



Universiteit
Leiden

The Netherlands

CelleKT: a chemical proteomics workflow to profile cellular target engagement of kinase inhibitors: a step-by-step guide

Gagestein, B.; Rüegger, J.; Janssen, A.P.A.; Kirchner, S.; Grether, U.; Rufer, A.C.; ... ; Tao, Z

Citation

Gagestein, B., Rüegger, J., Janssen, A. P. A., Kirchner, S., Grether, U., Rufer, A. C., & Stelt, M. van der. (2025). CelleKT: a chemical proteomics workflow to profile cellular target engagement of kinase inhibitors: a step-by-step guide. In Z. Tao (Ed.), *Methods in Molecular Biology* (Vol. 2921, pp. 231-253). New York: Humana New York. doi:10.1007/978-1-0716-4502-4_12

Version: Publisher's Version
License: [Leiden University Non-exclusive license](#)
Downloaded from: <https://hdl.handle.net/1887/4259114>

Note: To cite this publication please use the final published version (if applicable).

Methods in
Molecular Biology 2921

Springer Protocols

Zhipeng Tao *Editor*

Activity-Based Proteomics

Methods and Protocols



 Humana Press

METHODS IN MOLECULAR BIOLOGY

Series Editor

John M. Walker

School of Life and Medical Sciences

University of Hertfordshire

Hatfield, UK

For further volumes:

<http://www.springer.com/series/7651>

For over 35 years, biological scientists have come to rely on the research protocols and methodologies in the critically acclaimed *Methods in Molecular Biology* series. The series was the first to introduce the step-by-step protocols approach that has become the standard in all biomedical protocol publishing. Each protocol is provided in readily-reproducible step-by-step fashion, opening with an introductory overview, a list of the materials and reagents needed to complete the experiment, and followed by a detailed procedure that is supported with a helpful notes section offering tips and tricks of the trade as well as troubleshooting advice. These hallmark features were introduced by series editor Dr. John Walker and constitute the key ingredient in each and every volume of the *Methods in Molecular Biology* series. Tested and trusted, comprehensive and reliable, all protocols from the series are indexed in PubMed.

Activity-Based Proteomics

Methods and Protocols

Edited by

Zhipeng Tao

Texas Woman's University, Denton, TX, USA

 **Humana Press**

Editor

Zhipeng Tao
Texas Woman's University
Denton, TX, USA

ISSN 1064-3745

ISSN 1940-6029 (electronic)

Methods in Molecular Biology

ISBN 978-1-0716-4501-7

ISBN 978-1-0716-4502-4 (eBook)

<https://doi.org/10.1007/978-1-0716-4502-4>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2025

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Humana imprint is published by the registered company Springer Science+Business Media, LLC, part of Springer Nature.

The registered company address is: 1 New York Plaza, New York, NY 10004, U.S.A.

If disposing of this product, please recycle the paper.

Preface

Activity-based proteomics profiling, using chemical probes, is a technique designed to selectively detect reactive amino acids within complex proteomes. By employing probes that specifically target catalytically active enzymes, this method enables the rapid identification of a biological system's functional proteome. In recent years, it has been widely applied to globally profile various enzymes, including lipolytic enzymes, serine hydrolases, ubiquitinases/deubiquitinases, chitinases, and glycoside hydrolases. These enzymes have been identified across a range of biological systems, such as bacteria, viruses, and mammalian cells and tissues, under both normal and pathological conditions. This book will provide comprehensive protocols for using activity-based proteomics to investigate the biological functions of these proteins.

This volume aims to present a wide range of protocols and provides an overview of the application of activity-based proteomics profiling to study deISGylating enzymes, serine hydrolases, monooxygenases, ubiquitin E3 ligases, and reactive cysteine in various organisms, including bacteria, viruses, and mammalian cells and tissues. Additionally, it includes bioinformatics analysis for interpreting activity-based proteomics data and explores the use of this profiling method to identify potential small molecules that target specific proteins, with a focus on treatments for cancer and viral infections. This book provides updated, comprehensive protocols for utilizing activity-based proteomics profiling to discover new enzymes, detect novel modifications, and identify new drug therapy targets. We hope it will be a valuable resource for readers interested in applying cutting-edge technologies in activity-based proteomics. Additionally, this volume is designed to serve as a step-by-step guide for conducting experiments and troubleshooting throughout their research.

I would like to express my heartfelt gratitude to everyone who contributed to this issue. As many of us know, sharing lab protocols is both a personal sacrifice and an act of service for the broader scientific community, requiring immense dedication and humility. I am deeply thankful to each person who contributed to the chapters of this book. Without your efforts, this would not have been possible. I also want to extend my special thanks to my wife and son. Having a husband and father in science has meant you've missed out on much-needed support, and for that, I regret it. However, I am incredibly grateful for your patience, understanding, and unwavering support. Additionally, I want to thank my students for their assistance in the lab and in this work. Above all, I give thanks to God and my LORD Jesus Christ for providing me with the strength and guidance throughout the past year and a half. It has been nothing short of a miracle, and I owe the completion of this book to His divine support.

Denton, TX, USA

Zhipeng Tao

Contents

<i>Preface</i>	<i>v</i>
<i>Contributors</i>	<i>ix</i>
1 Activity-Based Protein Profiling (ABPP) of Cellular DeISGylating Enzymes and Inhibitor Screening	1
<i>Simeon D. Draganov, Hannah B. L. Jones, Qura Shah, Benedikt M. Kessler, and Adán Pinto-Fernández</i>	
2 Integrated Multi-omics Approaches for Studying Rare Genetic Diseases	31
<i>Refat M. Nimer, Lina A. Dahabiyeh, Reem AlMalki, Maha Al Mogren, Hicham Benabdelkamel, Assim A. Alfadda, Salini Scaria Joy, and Anas M. Abdel Rahman</i>	
3 Activity-Based Protein Profiling of Serine Hydrolases in Bacteria: Methods and Protocols	57
<i>Md Jalal Uddin, Jeanette S. Grunnvåg, and Christian S. Lentz</i>	
4 Profiling Ligand-Induced Changes in Nuclear Localization Using Proximity Labeling-Coupled Chemoproteomics	73
<i>Qianni Peng and Eranthie Weerapana</i>	
5 An Overview of Analytical Approaches to Cancer Proteogenomics	93
<i>Chad J. Creighton</i>	
6 Microplate-Based Enzymatic Activity Assay Protocol Powered by Activity-Based Probes	119
<i>Exequiel O. J. Porta, Karunakaran Kalesh, Jaime A. Isern, and Patrick G. Steel</i>	
7 Identifying Inhibitor Targets in Mycobacteria by Activity-Based Probe Profiling	139
<i>Neetika Jaisinghani, Isabel Sakarin, Hiren V. Patel, Michael Li, Mary L. Previti, and Jessica C. Seeliger</i>	
8 Application of Activity-Based Protein Profiling (ABPP) Strategy in the Identification of New Targets for Vascular Normalization and Binding Site Resolution	155
<i>Jie Zhang and Yanchen Li</i>	
9 Activity-Based Protein Profiling to Study Virus–Host and Viral-Induced MicroRNA–Host Interactions	181
<i>Parrish Evers and John Paul Pezacki</i>	
10 Activity-Based Protein Profiling of Active Host Cell Serine Hydrolases for Bioprocess Development of Therapeutic Proteins and Vaccines	201
<i>Xuanwen Li, Taku Tsukidate, Alyssa Q. Stiving, and Hillary A. Schuessler</i>	
11 Synthesis and Application of Activity-Based Probes for Serine Proteases	217
<i>Alba Ramos-Llorca, Valeria Parravicini, Koen Augustyns, and Michaela Prothiwa</i>	

12	CellEKT: A Chemical Proteomics Workflow to Profile Cellular Target Engagement of Kinase Inhibitors - A Step-by-Step Guide	231
	<i>Berend Gagestein, Joel Rüegger, Antonius P. A. Janssen, Stephan Kirchner, Uwe Grether, Arne C. Rufer, and Mario van der Stelt</i>	
13	Isolation of Exosome from Senescent Preadipocytes for Mass Spectrometric Analysis.....	255
	<i>Yao Wang</i>	
14	Identification of Anticancer ROS Targets by Cysteine Reactivity Protein Profiling.....	265
	<i>Junbing Zhang and Liron Bar-Peled</i>	
15	A Comprehensive Data Processing Pipeline for Phosphoproteomics in Viral Infection Studies	275
	<i>Sini Huuskonen, Xiaonan Liu, Kari Salokas, Antti Tuhkala, Salla Keskitalo, and Markku Varjosalo</i>	
16	Activity-Based Protein Profiling of Bacterial Monooxygenases.....	293
	<i>John B. Joyce and Michael R. Hyman</i>	
17	Activity-Based Profiling of Reactive Nucleophilic Cysteine Sites from Tissue Proteomes for Drug Discovery Applications	319
	<i>Hasini Kalpage, An Chi, and Smaranda Bodea</i>	
18	Activity-Based Protein Profiling for Functional Cysteines and Protein Target Identification	331
	<i>Tin-Yan Koo and Clive Yik-Sham Chung</i>	
19	System-Wide Profiling of Ubiquitin E3 Ligase Activities with Quantitative Proteomics and Bioinformatics Analysis.....	345
	<i>Yue Chen and Yao Gong</i>	
20	Recent Overview of Protein Palmitoylation and Profiling Methodologies	361
	<i>Changyan Jin, Qiaoling Yuan, and Zhipeng Tao</i>	
	<i>Index</i>	<i>371</i>

Contributors

- ANAS M. ABDEL RAHMAN • *Metabolomics Section, Precision Medicine Laboratory Department (PMLD) Genomic Medicine Center of Excellence (GMCoe), King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia; Department of Biochemistry and Molecular Medicine, College of Medicine, Alfaisal University, Riyadh, Saudi Arabia*
- MAHA AL MOGREN • *Metabolomics Section, Precision Medicine Laboratory Department (PMLD) Genomic Medicine Center of Excellence (GMCoe), King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia*
- ASSIM A. ALFADDA • *Proteomics Resource Unit, Obesity Research Center, College of Medicine, King Saud University, Riyadh, Saudi Arabia; Department of Medicine, College of Medicine, King Saud University, Riyadh, Saudi Arabia*
- REEM ALMALKI • *Metabolomics Section, Precision Medicine Laboratory Department (PMLD) Genomic Medicine Center of Excellence (GMCoe), King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia*
- KOEN AUGUSTYNS • *Laboratory of Medicinal Chemistry, Department of Pharmaceutical Sciences, Faculty of Pharmaceutical, Biomedical and Veterinary Sciences, Campus Drie Eiken, University of Antwerp, Antwerp, Belgium*
- LIRON BAR-PELED • *Krantz Family Center for Cancer Research, Massachusetts General Hospital Cancer Center, Charlestown, MA, USA; Department of Medicine, Harvard Medical School, Boston, MA, USA; Department of Medicine, Massachusetts General Hospital, Boston, MA, USA*
- HICHAM BENABDELKAMEL • *Proteomics Resource Unit, Obesity Research Center, College of Medicine, King Saud University, Riyadh, Saudi Arabia*
- SMARANDA BODEA • *Merck & Co., Inc., Cambridge, MA, USA*
- YUE CHEN • *University of Minnesota at Twin Cities, Minneapolis, MN, USA*
- AN CHI • *Merck & Co., Inc., Cambridge, MA, USA*
- CLIVE YIK-SHAM CHUNG • *School of Biomedical Sciences, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, People's Republic of China; Department of Pathology, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, People's Republic of China*
- CHAD J. CREIGHTON • *Dan L. Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, TX, USA; Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA; Department of Medicine, Baylor College of Medicine, Houston, TX, USA*
- LINA A. DAHABIYEH • *Department of Pharmaceutical Sciences, School of Pharmacy, The University of Jordan, Amman, Jordan*
- SIMEON D. DRAGANOV • *Chinese Academy of Medical Sciences Oxford Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK; Target Discovery Institute, Centre for Medicines Discovery, Nuffield Department of Medicine, University of Oxford, Oxford, UK*
- PARRISH EVERS • *Department of Chemistry and Biomolecular Sciences, University of Ottawa, Ottawa, ON, Canada*

- BEREND GAGESTEIN • *Department of Molecular Physiology, Leiden Institute of Chemistry, Leiden University, Leiden, The Netherlands; Oncode Institute, Amsterdam, The Netherlands; Department Medical Oncology, OncoProteomics Laboratory, Cancer Center Amsterdam, Amsterdam UMC, Location VUmc, Amsterdam, The Netherlands*
- YAO GONG • *University of Minnesota at Twin Cities, Minneapolis, MN, USA*
- UWE GRETHER • *Roche Pharma Research & Early Development, Roche Innovation Center Basel, Basel, Switzerland*
- JEANETTE S. GRUNNVÅG • *Department of Medical Biology and Centre for New Antibacterial Strategies (CANS), UiT The Arctic University of Norway, Tromsø, Norway*
- SINI HUUSKONEN • *Institute of Biotechnology, Helsinki Institute of Life Science HiLIFE, University of Helsinki, Helsinki, Finland*
- MICHAEL R. HYMAN • *Department of Plant and Microbial Biology, North Carolina State University, Raleigh, NC, USA*
- JAIME A. ISERN • *Department of Chemistry, Durham University, Durham, UK*
- NEETIKA JAISINGHANI • *Department of Pharmacological Sciences, Stony Brook University, Stony Brook, NY, USA*
- ANTONIUS P. A. JANSSEN • *Department of Molecular Physiology, Leiden Institute of Chemistry, Leiden University, Leiden, The Netherlands; Oncode Institute, Amsterdam, The Netherlands*
- CHANGYAN JIN • *Department of Nutrition and Food Sciences, Texas Woman's University, Denton, TX, USA*
- HANNAH B. L. JONES • *Target Discovery Institute, Centre for Medicines Discovery, Nuffield Department of Medicine, University of Oxford, Oxford, UK*
- SALINI SCARIA JOY • *Strategic Center for Diabetes Research, College of Medicine, King Saud University, Riyadh, Saudi Arabia*
- JOHN B. JOYCE • *Department of Plant and Microbial Biology, North Carolina State University, Raleigh, NC, USA*
- KARUNAKARAN KALESH • *School of Health and Life Sciences, Teesside University, Middlesbrough, UK; National Horizons Centre, Darlington, UK*
- HASINI KALPAGE • *Merck & Co., Inc., Boston, MA, USA*
- SALLA KESKITALO • *Institute of Biotechnology, Helsinki Institute of Life Science HiLIFE, University of Helsinki, Helsinki, Finland*
- BENEDIKT M. KESSLER • *Chinese Academy of Medical Sciences Oxford Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK; Target Discovery Institute, Centre for Medicines Discovery, Nuffield Department of Medicine, University of Oxford, Oxford, UK*
- STEPHAN KIRCHNER • *Roche Pharma Research & Early Development, Roche Innovation Center Basel, Basel, Switzerland*
- TIN-YAN KOO • *School of Biomedical Sciences, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, People's Republic of China*
- CHRISTIAN S. LENTZ • *Department of Medical Biology and Centre for New Antibacterial Strategies (CANS), UiT The Arctic University of Norway, Tromsø, Norway*
- MICHAEL LI • *Department of Microbiology and Immunology, Stony Brook University, Stony Brook, NY, USA*
- XUANWEN LI • *Analytical Research & Development Mass Spectrometry, Merck & Co., Inc., Rahway, NJ, USA*
- YANCHEN LI • *School of Pharmacy, Health Science Center, Xi'an Jiaotong University, Xi'an, China*

- XIAONAN LIU • *Institute of Biotechnology, Helsinki Institute of Life Science HiLIFE, University of Helsinki, Helsinki, Finland*
- REFAT M. NIMER • *Department of Medical Laboratory Sciences, Jordan University of Science and Technology, Irbid, Jordan*
- VALERIA PARRAVICINI • *Laboratory of Medicinal Chemistry, Department of Pharmaceutical Sciences, Faculty of Pharmaceutical, Biomedical and Veterinary Sciences, Campus Drie Eiken, University of Antwerp, Antwerp, Belgium*
- HIREN V. PATEL • *Department of Pharmacological Sciences, Stony Brook University, Stony Brook, NY, USA*
- QIANNI PENG • *Department of Chemistry, Boston College, Chestnut Hill, MA, USA*
- JOHN PAUL PEZACKI • *Department of Chemistry and Biomolecular Sciences, University of Ottawa, Ottawa, ON, Canada*
- ADÁN PINTO-FERNÁNDEZ • *Chinese Academy of Medical Sciences Oxford Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK; Target Discovery Institute, Centre for Medicines Discovery, Nuffield Department of Medicine, University of Oxford, Oxford, UK*
- ESEXUIEL O. J. PORTA • *Department of Chemistry, Durham University, Durham, UK*
- MARY L. PREVITI • *Department of Pharmacological Sciences, Stony Brook University, Stony Brook, NY, USA*
- MICHAELA PROTHIWA • *Laboratory of Medicinal Chemistry, Department of Pharmaceutical Sciences, Faculty of Pharmaceutical, Biomedical and Veterinary Sciences, Campus Drie Eiken, University of Antwerp, Antwerp, Belgium*
- ALBA RAMOS-LLORCA • *Laboratory of Medicinal Chemistry, Department of Pharmaceutical Sciences, Faculty of Pharmaceutical, Biomedical and Veterinary Sciences, Campus Drie Eiken, University of Antwerp, Antwerp, Belgium*
- JOEL RÜEGGER • *Department of Molecular Physiology, Leiden Institute of Chemistry, Leiden University, Leiden, The Netherlands; Oncode Institute, Amsterdam, The Netherlands*
- ARNE C. RUFER • *Roche Pharma Research & Early Development, Roche Innovation Center Basel, Basel, Switzerland*
- ISABEL SAKARIN • *Department of Microbiology and Immunology, Stony Brook University, Stony Brook, NY, USA*
- KARI SALOKAS • *Institute of Biotechnology, Helsinki Institute of Life Science HiLIFE, University of Helsinki, Helsinki, Finland*
- HILLARY A. SCHUESSLER • *Analytical Research & Development Mass Spectrometry, Merck & Co., Inc., Rahway, NJ, USA*
- JESSICA C. SEELIGER • *Department of Pharmacological Sciences, Stony Brook University, Stony Brook, NY, USA*
- QURA SHAH • *Life Sciences Solutions, Thermo Fisher Scientific, Paisley, UK*
- PATRICK G. STEEL • *Department of Chemistry, Durham University, Durham, UK*
- ALYSSA Q. STIVING • *Analytical Research & Development Mass Spectrometry, Merck & Co., Inc., Rahway, NJ, USA*
- ZHIPENG TAO • *Texas Woman's University, Denton, TX, USA*
- TAKU TSUKIDATE • *Analytical Research & Development Mass Spectrometry, Merck & Co., Inc., Rahway, NJ, USA*
- ANTTI TUHKALA • *Institute of Biotechnology, Helsinki Institute of Life Science HiLIFE, University of Helsinki, Helsinki, Finland*
- MD JALAL UDDIN • *Department of Medical Biology and Centre for New Antibacterial Strategies (CANS), UiT The Arctic University of Norway, Tromsø, Norway*

- MARIO VAN DER STELT • *Department of Molecular Physiology, Leiden Institute of Chemistry, Leiden University, Leiden, The Netherlands; Oncode Institute, Amsterdam, The Netherlands*
- MARKKU VARJOSALO • *Institute of Biotechnology, Helsinki Institute of Life Science HiLIFE, University of Helsinki, Helsinki, Finland*
- YAO WANG • *Diabetes Center, University of California, San Francisco, San Francisco, CA, USA*
- ERANTHIE WEERAPANA • *Department of Chemistry, Boston College, Chestnut Hill, MA, USA*
- QIAOLING YUAN • *Department of Nutrition and Food Sciences, Texas Woman's University, Denton, TX, USA*
- JIE ZHANG • *School of Pharmacy, Health Science Center, Xi'an Jiaotong University, Xi'an, China*
- JUNBING ZHANG • *Krantz Family Center for Cancer Research, Massachusetts General Hospital Cancer Center, Charlestown, MA, USA; Department of Medicine, Harvard Medical School, Boston, MA, USA; Shanghai Institute of Nutrition and Health, Chinese Academy of Sciences, Shanghai, China*



Chapter 1

Activity-Based Protein Profiling (ABPP) of Cellular DeISGylating Enzymes and Inhibitor Screening

Simeon D. Draganov, Hannah B. L. Jones, Qura Shah, Benedikt M. Kessler, and Adán Pinto-Fernández

Abstract

A detailed methodology platform is described for activity-based protein profiling (ABPP) of cellular deISGylating enzymes using a specific activity-based interferon-stimulated gene 15 (ISG15) probe. Manual and semi-automated workflows for medium- to high-throughput applications are outlined in this chapter, with western blotting and proteomics-based techniques as the main readouts. This methodology informs us of endogenous deISGylating enzyme expression and activity in a cellular context, including USP18, the type I interferon (IFN-I)-inducible deISGylase, and several constitutively expressed deubiquitinases (DUBs), such as USP5, USP14, USP16, and USP36, that exert cross-reactivity to ISG15. ISG15-ABPP also enables the identification and characterization of potent and selective deISGylating enzyme modulators.

Key words Activity-based probes, Activity proteomics, DeISGylating enzymes, Chemical biology, DeISGylation, Drug discovery, ISG15, Proteomics, Pharmacology

1 Introduction

Activity-based probes (ABPs) are useful tools for assessing the activity of enzymes and profiling the potency and selectivity of inhibitors against proteins of interest [1, 2]. A class of cysteine proteases, known as deISGylating enzymes (deISGylases), are of particular interest to profile and study due to their roles in the innate immune response [3–5], metabolism [6], and tumorigenesis/cancer cell metastasis [7–13]. The interferon-stimulated gene 15 (ISG15) is a 15 kDa ubiquitin-like protein that is induced by different classes of interferons [14], DNA damage [15], or upon viral [16, 17] or bacterial infection [18, 19], due to subsequent activation of the type I interferon (IFN-I) response. ISG15 plays an important role in regulating these processes, such as inducing an antiviral response in cells, through conjugation of ISG15 onto

proteins as a post-translational modification known as ISGylation [3, 4, 16, 17]. To fine-tune the cellular effects of ISGylation, deISGylating proteins [6, 20–22] cleave the ISG15 modification from ISG15-conjugated (ISGylated) proteins in a process known as deISGylation. A notable deISGylating protein is the Ubl carboxyl-terminal hydrolase 18 (USP18), which is a potent negative regulator of the IFN-I response [3–5]. Although several cellular deubiquitinases (DUBs) have been described to cross-react and cleave ISG15 from substrates, such as USP2, USP5, USP13, USP14, USP16, and USP36 [6, 20–22], their role in ISG15 biology and interferon pathways is not yet fully understood. Several viral proteases have been found to exert deISGylase activity with the aim of interfering with the host's antiviral response [23–27]. There is also an increasing amount of literature evidencing that overexpression of USP18, and other deISGylating proteins, may be involved in tumorigenesis and pathways critical for promoting cancer cell growth [7–13]. Thus, deISGylating enzymes are potential drug-gable therapeutic targets of interest in the context of pathogen diseases, immuno-oncology, and other pathologies involving inflammation.

In this chapter, we describe an activity-based protein profiling (ABPP) approach for cellular deISGylating enzymes and deISGylating enzyme inhibitors in a high-throughput manner (Fig. 1). We also describe a protocol for a standard, lower-throughput format. Our approach employs the use of a double-tagged, truncated (~10 kDa; containing the C-terminal ubiquitin-like domain, residues 79–156) ISG15 activity-based probe [6, 27] for the labeling of deISGylating enzymes in cell or tissue protein extracts, with readouts of immunoblotting or activity proteomics after isolation (enrichment) of probe–enzyme complexes followed by liquid chromatography tandem-mass spectrometry (LC-MS/MS) (mass spectrometry) analysis. Labeling is achieved through a covalent reaction of catalytic cysteines of ISG15 proteases with a propargylamine (PA) warhead present on the C-terminus of the probe, and protein pull-down can be successfully performed through a biotin tag present on the N-terminus of the probe. This methodology has been extensively described for Ub-based ABPs [28–31].

One of the main applications of this assay is the screening of compound libraries for potential inhibitors by western blot since using such a high-molecular weight probe results in an ~10 kDa shift of a labeled protein band, relative to the corresponding non-labeled protein band, and the extent of labeling can be quantified by measuring band intensity in the presence and absence of inhibitor treatment. In a competition assay, where a compound successfully binds to the active site of a particular deISGylase, less of that deISGylase is available to bind to the ISG15 probe. This results in reduced labeling compared to a probe-only (no compound) treatment control, and the decrease in labeling can be quantified by

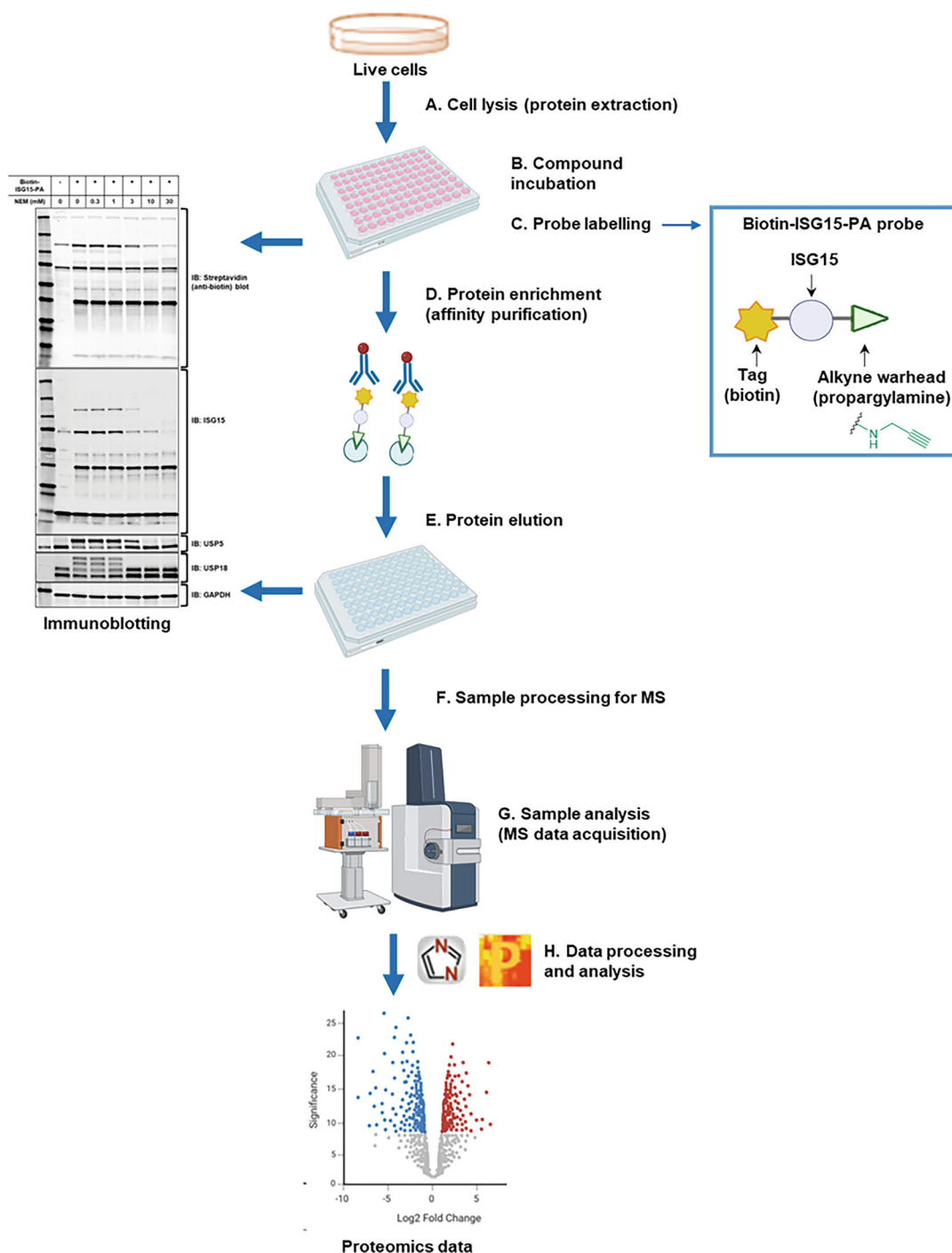


Fig. 1 Typical workflow for assessing potency and selectivity of a compound library toward deISGylases in a high-throughput manner. Cells are lysed and protein quantification performed (a). Lysate protein is initially incubated with compound(s) in a 96-well plate (b), followed by labeling with biotin-ISG15-PA probe (c). Labeling efficiency and inhibitor potency is evaluated by western blotting, and pull-down is subsequently

measuring the intensity of the probe-labeled band—the approach is thus useful for calculating IC_{50} values in compound titration assays. The method also has the advantage of allowing visualization of multiple probe-labeled deISGylases via western blot with streptavidin detection of the biotin tag and therefore allows for the assessment of whether a compound exhibits broad-range inhibition by observing changes in overall labeling following compound treatment. For initial compound screens, performing the activity-based probe profiling in cell lysates overcomes potential issues such as compound stability, permeability into cells, disruption of cellular pathways that may affect deISGylases induction, and compartmentalization of deISGylases. Subsequent validations are typically performed in cell/tissue protein extracts after compound treatment of cells, or living animals, and informs of target engagement in a more relevant physiological context [32].

In addition to immunoblotting, LC-MS/MS approaches can be combined with ISG15 probe-labeling assays and pull-down of probe-labeled proteins for ABPP (or activity proteomics) of deISGylases (Fig. 1). LC-MS/MS can be applied to assess the true coverage of potential broad-range inhibition due to potential sensitivity limitations of detection by western blot and the lack of information on deISGylase identity in cases where such inhibition is observed by western blot. In the ABPP approach, catalytic inhibition of a deISGylase with a compound will result in reduced labeling of that deISGylase with the probe, which will lead to a subsequent reduction in the enrichment of that protein. It should be noted, however, that depending on how an inhibitor is predicted to bind the catalytic cysteine (i.e., irreversible, reversible, covalent, noncovalent), additional optimization may be required, such as probe incubation times, to minimize/avoid outcompeting, for example, a reversible inhibitor with the ISG15 probe due to fast dissociation kinetics of the inhibitor, which will result in a lower apparent potency for that compound.

The described methodology in this chapter can thus be applied for profiling deISGylating enzymes in different conditions (i.e., disease vs. healthy) and assessing selectivity and potency of inhibitor compounds against cellular deISGylases, and we adapted the approach for high-throughput screening of compound libraries based on previously reported protocols for profiling deubiquitinases and DUB inhibitors by LC-MS/MS [33–35].

Fig. 1 (continued) performed (d). Enriched proteins are eluted from beads (e) and efficiency of enrichment is evaluated by western blotting, prior to MS sample processing (f). Processed samples are analyzed on a Tims-TOF LC-MS/MS system (g), and data processing and analysis performed on freely available software packages (DIA-NN, Perseus) (h). (Figure created with BioRender.com)

2 Materials

2.1 Biotin-ISG15-PA Probe Synthesis

1. pET21b-Avi-ISG15 (79–156)-Intein-CBD (chitin-binding domain) plasmid.
2. 20 μ L, 200 μ L, and 1 mL micropipette tips.
3. 1.5 mL microcentrifuge tubes.
4. 50 mL conical centrifuge tubes.
5. Competent BL21 (DE3) *Escherichia coli*.
6. Luria broth (LB) media containing 100 μ g/mL ampicillin.
7. LB media containing 100 μ g/mL ampicillin and 1.5% agar (w/v).
8. SOC media (sterile) containing 0.5% yeast extract (w/v), 2% tryptone (w/v), and 10 mM NaCl.
9. 100 mm Petri dishes.
10. 1 M isopropyl β -D-1-thiogalactopyranoside (IPTG) stock.
11. 60% glycerol stock (sterile).
12. Buffer A: 50 mM HEPES, 150 mM NaCl, 0.5 mM TCEP, pH 7.4 (adjustment with NaOH).
13. Chitin bead slurry.
14. Clamp stand.
15. Boss head.
16. Clamp.
17. Glass (liquid) chromatography column (1.5 $\times$ 10 cm).
18. Buffer B: 50 mM HEPES, 150 mM NaCl, 0.5 mM TCEP, 100 mM 2-mercaptoethanesulfonic acid sodium salt (MESNA), pH 7.4 (adjustment with NaOH).
19. 4x Laemmli buffer: 40% glycerol (v/v), 8% sodium dodecyl sulfate (SDS), 0.2% bromophenol blue (w/v), 50 mM dithiothreitol (DTT), 250 mM Tris base, pH 6.8 (adjustment with HCl).
20. Tris-glycine acrylamide gel.
21. Gel electrophoresis apparatus (Criterion Cell or alternative).
22. 1x Tris-glycine gel running buffer: 25 mM Tris base, 192 mM glycine, and 0.1% SDS (w/v).
23. Prestained protein molecular weight ladder.
24. Coomassie stain (instant blue or alternative).
25. PVDF (or nitrocellulose) membrane for western blot (pore size 0.45 μ M).
26. Thick blot filter paper (thickness 0.84 mm).

27. Western blot apparatus (Criterion Blotter or alternative), including foam pads and an ice pack.
28. Western blot transfer (Towbin) buffer: 25 mM Tris base, 192 mM glycine, 20% methanol (MeOH) (v/v), and 80% distilled H₂O (v/v).
29. Streptavidin (biotin detection reagent) or anti-biotin antibody.
30. 3 kDa MWCO spin concentrators.
31. 3.5 kDa MWCO dialysis devices.
32. Desalting columns (PD-10 with Sephadex G-25 resin).
33. Protein estimation kit (BCA or alternative).
34. 250 mM PA stock.
35. Enzymatic protein biotinylation kit (Merck, #CS0008 or alternative).

**2.2 Cell/Tissue Lysis
for Biotin-ISG15-PA
Probe Synthesis**

1. Buffer C: 50 mM Tris base, 5 mM MgCl₂, 0.5 mM EDTA, 2.5% glycerol (v/v), pH 7.5 (adjustment with HCl).
2. 1.25 M DTT (can be stored at -20 °C and reused).
3. Acid-washed glass beads ($\leq 106 \mu\text{m}$).
4. Protein estimation kit (BCA or alternative).

**2.3 Biotin-ISG15-PA
Pull-Down for LC-MS/
MS**

1. 5% SDS (w/v) stock.
2. 10% NP-40 (v/v) stock.
3. Buffer D: 50 mM Tris base, 0.5% NP-40 (v/v), 150 mM NaCl, 20 mM MgCl₂, pH 7.4 (adjust with HCl).
4. Streptavidin magnetic Sepharose beads slurry (Cytiva, #GE28-9857-99).
5. Magnetic separation rack (for 1.5/2 mL Eppendorf tubes).
6. 2x Laemmli buffer: 20% glycerol (v/v), 4% SDS (w/v), 0.1% bromophenol blue (w/v), 25 mM DTT, 125 mM Tris base, pH 6.8 (adjustment with HCl).
7. S-trap mini columns (100–300 μg) (ProtiFi, #C02-mini-80).
8. 96-well S-trap plate (100–300 μg per well) (ProtiFi, #C02-96well-1).
9. 200 mM DTT stock.
10. 200 mM iodoacetamide stock.
11. 12% phosphoric acid (w/v).
12. 50% acetonitrile (ACN) (v/v).
13. 0.2% formic acid (v/v).
14. 90% MeOH, 100 mM triethylammonium bicarbonate (TEAB), pH 7.6.
15. 50 mM TEAB, pH 8.5.

16. Trypsin powder.
17. Vacuum centrifuge (Speedvac).
18. Buffer E: 98% H₂O (v/v), 2% ACN (v/v), and 0.1% trifluoroacetic acid (TFA) (v/v).
19. Glass vials for mass spectrometry.
20. Caps for glass vials.

**2.4 Biotin-ISG15-PA
Pull-Down for High-
Throughput LC-MS/MS
Analysis: 96-Well Plate
Format**

1. 5% SDS (w/v) stock.
2. 10% NP-40 (v/v) stock.
3. Buffer D: 50 mM Tris base, 0.5% NP-40 (v/v), 150 mM NaCl, 20 mM MgCl₂, pH 7.4 (adjustment with HCl).
4. Reagent reservoir (55 mL).
5. 8/12-channel multichannel pipette.
6. Streptavidin magnetic Sepharose beads slurry (Cytiva, #GE28-9857-99).
7. Magnetic separation rack (for 1.5/2 mL Eppendorf tubes).
8. 96-well plate magnetic separation rack.
9. Thermomixer shaker (with a temperature range of at least 4–90 °C).
10. Plate adaptor (for thermomixer shaker).
11. 2x Laemmli buffer: 20% glycerol (v/v), 4% SDS (w/v), 0.1% bromophenol blue (w/v), 25 mM DTT, 125 mM Tris base, pH 6.8 (adjustment with HCl).
12. 96-well S-trap plate (100–300 µg per well) (ProtiFi, #C02-96well-1).
13. 200 mM DTT stock.
14. 200 mM iodoacetamide stock.
15. 12% phosphoric acid (w/v).
16. 0.2% formic acid (v/v).
17. 50% ACN (w/v).
18. 90% MeOH (v/v), 100 mM TEAB, pH 7.6.
19. 50 mM TEAB, pH 8.5.
20. Trypsin.
21. Vacuum centrifuge (Speedvac).
22. Buffer E: 98% H₂O (v/v), 2% ACN (v/v), 0.1% TFA.
23. 96-well plate.
24. 96-well deep well plate.
25. Plate sealing cover.

**2.5 Biotin-ISG15-PA
Pull-Down for High-
Throughput LC-MS/MS
Using High-
Throughput Thermo
KingFisher Apex
System**

1. 5% SDS (w/v) stock.
2. 10% NP-40 (v/v) stock.
3. Buffer D: 50 mM Tris base, 0.5% NP-40 (v/v), 150 mM NaCl, 20 mM MgCl₂, pH 7.4 (adjustment with HCl).
4. Reagent reservoir (55 mL).
5. 8/12-channel multichannel pipette.
6. Streptavidin magnetic Sepharose beads slurry (Cytiva, #GE28-9857-99).
7. 96-well plate magnetic separation rack.
8. 2x Laemmli buffer: 20% glycerol (v/v), 4% SDS (w/v), 0.1% bromophenol blue (w/v), 25 mM DTT, 125 mM Tris base, pH 6.8 (adjusted with HCl).
9. 96-well S-trap plate (100–300 µg per well) (ProtiFi, #C02-96well-1).
10. 200 mM DTT stock.
11. 200 mM iodoacetamide stock.
12. 12% phosphoric acid (w/v).
13. 0.2% formic acid (v/v).
14. 50% ACN (w/v).
15. 90% MeOH, 100 mM TEAB, pH 7.6.
16. 50 mM TEAB, pH 8.5.
17. Trypsin.
18. Vacuum centrifuge (Speedvac).
19. Buffer E: 98% H₂O (v/v), 2% ACN (v/v), 0.1% TFA (v/v).
20. 96-deep well plate.
21. KingFisher 96 tip comb for deep-well (DW) magnets (Thermo Scientific, #97002534).
22. KingFisher 96 deep-well plate, v bottom, polypropylene (Thermo Scientific, #95040450).

**2.6 LC-MS/MS
TimsTOF**

1. Evotips (Evosep, #EV2011).
2. Isopropanol (LC-MS grade, Supelco, #1027811000 or alternative).
3. Buffer F: ACN, 0.1% formic acid (v/v).
4. Buffer G: H₂O, 0.1% formic acid (v/v).
5. PepSep8 cm analytical C18 column (3 µm beads, 100 µm ID).

3 Methods

3.1 Biotin-ISG15-PA Probe Synthesis

3.1.1 Transform the pET21b-Avi-ISG15 (79–156)-Intein-CBD (Chitin-Binding domain) Plasmid (or a Suitable Alternative) into Competent BL21(DE3) E. coli Cells Using the Aseptic Technique

The pET21b vector contains an intein and a CBD at the C-terminal of ISG15, necessary for purification and subsequent reaction with the warhead.

1. Add 1 pg to 100 ng of the vector to 50 μ L of the competent *E. coli* and place on ice for 30 min.
2. Heat shock for 10 s at 42 °C.
3. Place back on ice for 5 min.
4. Add 950 μ L of sterile SOC media and shake at 37 °C, at 250 rpm for 1 h.
5. Mix the cells and carry out 5×1 in 10 serial dilutions.
6. Spread 100 μ L of each serial dilution onto separate sterile LB agar plates containing 100 μ g/mL ampicillin (or different antibiotic and concentration if an alternative selection gene is used) (*see Note 1*).

3.1.2 Liquid Culture of Transformed E. coli Using the Aseptic Technique

1. Following transformation, inoculate 40 mL of LB media containing 100 μ g/mL ampicillin in a 250 mL conical flask, with a single colony, and incubate at 37 °C, at 250 rpm for 18 h.
2. Optional: after overnight growth, glycerol stocks can be prepared in 20% glycerol and stored at –80 °C for future use.
3. Add 10 mL of the inoculated culture to 1 L of sterile LB in a 2 L conical flask containing 100 μ g/mL ampicillin and incubate at 37 °C, at 200 rpm until $OD_{600} = 0.8$ – 0.9 is measured (~3–4 h).
4. Induce protein expression by adding 0.4 mM IPTG and incubate cultures overnight at 17 °C, at 180 rpm.
5. Collect bacterial pellets by centrifugation at 6800 g for 12 min and discard the supernatant. Bacterial pellets can be stored at –80 °C (Fig. 1).

3.1.3 Biotin-ISG15-MESNA Purification from Biotin-ISG15-Intein-CBD

1. From the steps marked with *, 200 μ L of sample can be kept for subsequent analysis of purification efficiency by SDS-PAGE gel/western blot (Fig. 2).
2. Resuspend bacterial pellet in 20 mL of buffer A and sonicate at an amplitude of 35% on ice for 10×30 s, with 30 s breaks in-between each sonication cycle.
3. Centrifuge the lysate at 30000 g for 30 min at 4 °C.
4. Transfer the supernatant to a fresh tube*. Disinfect the tube with pellet.
5. Fill a 1.5×10 cm chromatography column with 10 mL of chitin bead slurry using a clamp stand.

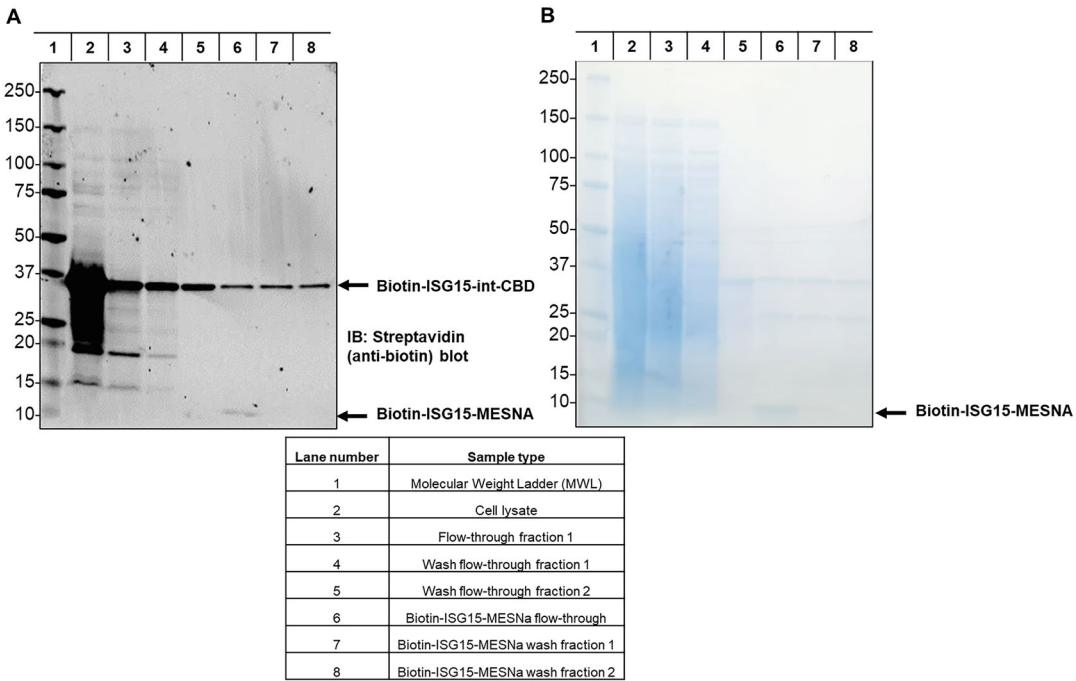


Fig. 2 Example (a) streptavidin western blot and (b) Coomassie stain SDS-PAGE gel for detecting biotin-ISG15-PA probe intermediates in the different stages of biotin-ISG15-PA probe purification. The included table shows sample identity for lanes 1–8

6. Wash the column with 10 mL of buffer A ($\times 4$).
7. Add the lysate supernatant to the column and collect the flow-through.
8. Reapply the flow-through to the column (to maximize probe-binding) and collect the flow-through again*. Label this flow-through fraction 1.
9. Wash the column with 10 mL of buffer A and collect the flow-through*. Label this wash flow-through fraction 1.
10. Wash the column with a further 30 mL of buffer A and collect the flow-through*. Label this wash flow-through fraction 2.
11. Store all flow-throughs and 200 μ L aliquots at -80°C .
12. Close the column and fill it with buffer B.
13. Seal the column with parafilm to prevent liquid from leaking out and incubate at 37°C at 180 rpm for 18 h.
14. Let the resin settle and collect flow-through*. This flow-through will contain biotin-ISG15-MESNA. Label this biotin-ISG15-MESNA flow-through.
15. Wash the column with 15 mL buffer A*. Label this biotin-ISG15-MESNA wash fraction 1.

16. Wash the column with a further 45 mL buffer A*. Label this biotin-ISG15-MESNA wash fraction 2.
17. Aliquot biotin-ISG15-MESNA-containing eluates and store at -80°C .

3.1.4 Assessing the Formation of Biotin-ISG15-MESNA by Western Blot and SDS-PAGE Gel

1. From the 200 μL aliquots from each step marked with *, take 40 μL and mix with 13.3 μL of 4x Laemmli buffer, and boil the samples at 90°C for 5 min.
2. Load 25 μL of each sample on an acrylamide gel $\times 2$, as shown in Fig. 1. One gel will be used for western blotting and the other will be used for Coomassie staining. Run the gels.
3. Assess biotin-ISG15-MESNA purity by Coomassie staining of the SDS-PAGE gel. The size of biotin-ISG15-MESNA is expected to be at ~ 10 kDa.
4. Confirm formation of biotin-ISG15-MESNA by performing a western blot with streptavidin (Fig. 2) (*see Note 2*).
5. Combine samples containing pure biotin-ISG15-MESNA (typically, these are the biotin-ISG15-MESNA sample and biotin-ISG15-MESNA wash1 flow-throughs) and concentrate ($\sim$ ten-fold) biotin-ISG15-MESNA by centrifugation using a 3 kDa MWCO filter.
6. Dialyze concentrated biotin-ISG15-MESNA in buffer A using a 3.5 kDa MWCO dialysis device at 4°C overnight.
7. Desalt biotin-ISG15-biotin using PD-10 desalting columns with buffer A.
8. Check biotin-ISG15-MESNA concentration using BCA or equivalent protein estimation kit.

3.1.5 Conversion of Biotin-ISG15-MESNA into Biotin-ISG15-PA

1. Dilute biotin-ISG15-MESNA stock to 1.5 mg/mL with buffer A.
2. Add 250 mM PA to biotin-ISG15-MESNA and incubate at room temperature (r.t.) at 1400 rpm for 1 h.
3. Desalt biotin-ISG15-PA using PD-10 desalting columns with buffer A.
4. Check biotin-ISG15-PA concentration using BCA or equivalent protein estimation kit.
5. Aliquot biotin-ISG15-PA probe and store at -80°C .
6. Optional: Perform enzymatic biotinylation of avi-ISG15-PA to biotin-ISG15-PA (*see Note 3* and Fig. 3). If enzymatic biotinylation is performed, subsequently desalt the biotinylated probe as described in **step 3** and concentrate to a desired volume.

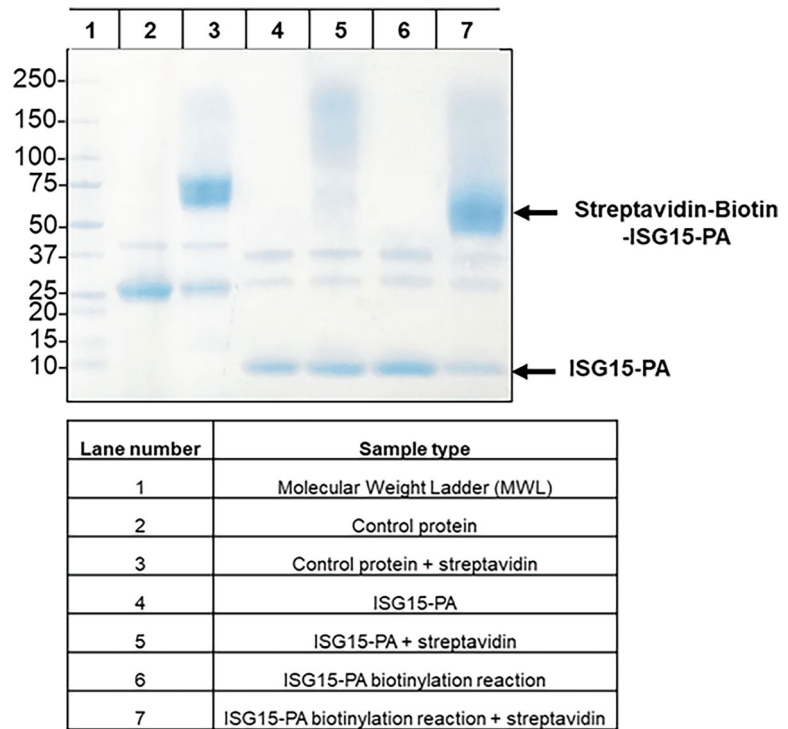


Fig. 3 Example streptavidin gel-shift assay Coomassie stain SDS-PAGE gel for assessing enzymatic biotinylation efficiency of the biotin-ISG15-PA probe. The included table shows sample identity for lanes 1–7

**3.2 Cell/Tissue Lysis
for Labeling with
Biotin-ISG15-PA**

1. Store lysate on ice.
2. Add DTT (frozen stock is good) to buffer C to a 1 mM final concentration.
3. Lyse cells by resuspending in appropriate volume of buffer C to obtain a final protein concentration of 2–10 mg/mL (*see Note 4*).
4. Add acid-washed glass beads to each lysate (~1:2 ratio (v/v) of glass beads to buffer C).
5. Vortex the lysate-beads mixture 10 × 30 s. After each 30 s vortex cycle, place the lysate on ice for 1 min (*see Note 5*).
6. Clear lysate by two-step centrifugation. Centrifuge at 600 g at 4 °C for 5 min and transfer ~80–90% of supernatant into a clean tube (*see Note 6*). Avoid disturbing the formed pellet.
7. Centrifuge the remaining lysate at 14000 g at 4 °C for 10 min and combine supernatant with the supernatant from the first centrifugation step. Avoid disturbing the formed pellet.
8. Perform protein quantification using BCA or equivalent protein estimation kit.

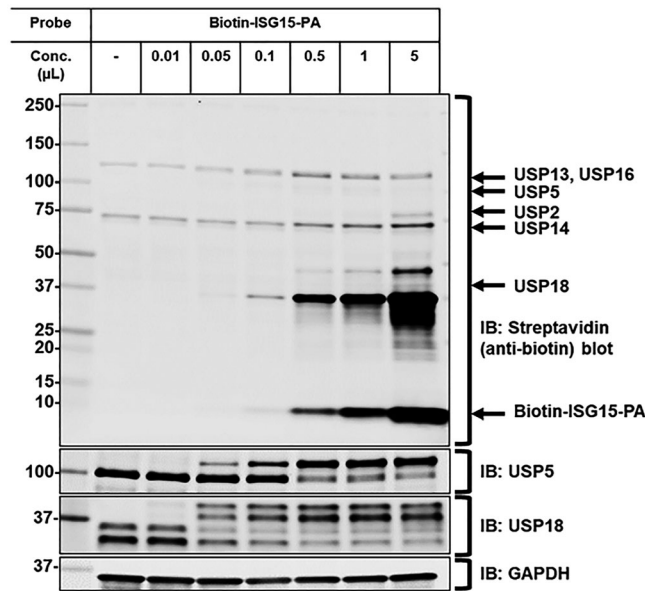


Fig. 4 Example western blots for titration of biotin-ISG15-PA probe to assess dose-dependent labeling of deISGylases in HAP1 lysates. 25 μ g of HAP1 lysate protein were loaded per lane. USP5, USP18, and streptavidin blots illustrate saturation of labeling with 0.5 μ L of biotin-ISG15-PA probe per 75 μ g of HAP1 lysate protein

3.3 Optimization of Biotin-ISG15-PA Labeling

1. For each batch of biotin-ISG15-PA, the ratio of biotin-ISG15-PA:lysate protein concentration needs to be optimized.
2. Titration of biotin-ISG15-PA (0.01–5 μ L) indicates that typically 0.5 μ L of probe are sufficient to fully label 75 μ g of protein lysate at 37 $^{\circ}$ C for 45 min (*see Note 7*). Include a negative control sample without probe. A streptavidin blot will show a band at ~10 kDa, corresponding to free biotin-ISG15-PA probe, indicating saturation of lysate labeling (Fig. 4).

3.4 Pull-Down of Biotin-ISG15-PA for LC-MS/MS Analysis: Standard

1. Perform cell lysis as outlined in Subheading 3.2 and aliquot 275 μ g of lysate per condition. Include a negative control sample without probe.
2. Dilute with buffer C to 147 μ L final volume [minus inhibitor volume (in DMSO), if used, and minus biotin-ISG15-PA volume (required volume specified in Subheading 3.3)].
3. Optional: If using an inhibitor, add 1.5 μ L of inhibitor stock in DMSO (final dilution 1:100) and treat negative control with 1.5 μ L of DMSO. Incubate at r.t. for 1 h.
4. Incubate with biotin-ISG15-PA at 37 $^{\circ}$ C for 45 min (*see Note 7*).

- 5. Quench labeling reaction with 0.4% SDS (w/v) (from a 5% stock) and 0.5% NP-40 (w/v) (from a 10% stock). This is easiest done by mixing 99 μ L of 5% SDS (w/v) stock with 62 μ L of 10% NP-40 stock and adding 21.5 μ L of this mixture to each labeling reaction (*see Note 8*).
- 6. Remove 15 μ L aliquot and label as “input” to check efficiency of probe labeling and check pull-down efficiency later.
- 7. Dilute the remaining ~150 μ L labeled sample with 250 μ L of buffer D.
- 8. Aliquot 20 μ L of streptavidin magnetic Sepharose beads slurry per sample using a cut tip (to avoid damage to the beads and ensure reproducible pipetting of slurry). Titration of streptavidin magnetic Sepharose beads (10–75 μ L) indicates that probe-labeled proteins are fully enriched with 20 μ L of beads, as indicated by depletion of supernatant fractions (Fig. 5).
- 9. Wash beads by adding 500 μ L of buffer D, inverting tubes 10 times, and place on a magnetic rack for ~30 s until all beads have pelleted toward the magnet.

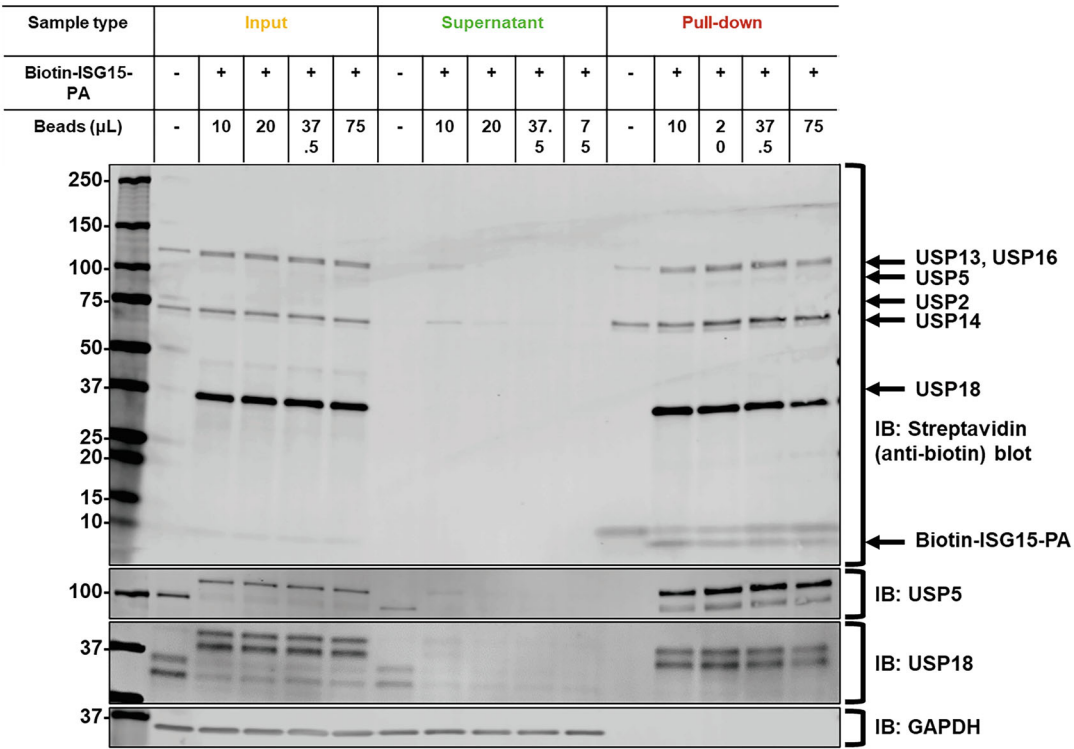


Fig. 5 Example western blots for titration of streptavidin magnetic Sepharose beads to assess pull-down efficiency of probe-labeled proteins. USP5, USP18, and streptavidin blots illustrate complete pull-down of labeled proteins with 20 μ L of streptavidin Sepharose magnetic beads per 250 μ g of HAP1 lysate protein, as evidenced by depletion of supernatant fractions and saturation in the eluate (pull-down) fractions

10. Discard supernatant and repeat wash step three more times.
11. After the removal of supernatant from the final wash step, add the biotin-ISG15-PA-labeled lysate to the beads and incubate for 16–18 h at 4 °C on an end-to-end rotation device.
12. Place beads on a magnetic rack for ~30 s until all beads have pelleted toward the magnet and collect supernatant. Label as “supernatant” to check pull-down efficiency by looking at depletion of probe-labeled proteins by western blot and store at –80 °C if pull-down has not been successful.
13. Wash beads by adding 500 µL of buffer D, inverting tubes 10 times, and place on a magnetic rack for ~30 s until all beads have pelleted toward the magnet.
14. Discard supernatant and repeat wash step three more times.
15. Add 60 µL of 2x Laemmli buffer, vortex briefly, and elute proteins by boiling at 90 °C for 5 min.
16. Place beads on a magnetic rack for ~30 s until all beads have pelleted toward the magnet and collect the eluate.
17. Repeat the elution step once more by adding 50 µL of 2x Laemmli buffer, vortexing briefly, and eluting proteins by boiling at 90 °C for 5 min.
18. Place beads on a magnetic rack for ~30 s until all beads have pelleted toward the magnet and collect the eluate. Combine with corresponding eluate from the first elution step.
19. Retain 10 µL of the “eluate” fractions to check pull-down efficiency by confirming the presence of probe-labeled proteins by western blot.
20. Pause point: samples can be stored at –80 °C for future use.
21. Run “input,” “supernatant,” and “eluate” fractions on a western blot. Blot with streptavidin to ensure lysate input fractions have been labeled with biotin-ISG15-PA, relative to the no probe control, to ensure that supernatant fractions have been completely depleted of probe-labeled proteins, and to confirm enrichment of probe-labeled proteins in the eluate fractions (Fig. 5).
22. Prepare samples for analysis by LC-MS/MS using S-trap mini-columns and following the manufacturer’s protocol for the reduction, alkylation, digestion, and peptide elution steps.
23. Dry samples on a vacuum centrifuge (Speedvac) and resuspend dry peptides in 20 µL of buffer E. Shake at r.t. at 800 rpm for 5 min, briefly vortex, and sonicate samples at r.t. for 5 min to ensure full resuspension of peptides.
24. Transfer samples to glass mass spectrometry vials.

3.5 Pull-Down of Biotin-ISG15-PA for High-Throughput LC-MS/MS Analysis: 96-Well Plate Format

1. Perform cell lysis as outlined in Subheading 3.2 and aliquot 275 μg of lysate per condition per well in a 96-well plate. Include a negative control sample without probe.
2. Dilute with buffer C to 64.8 μL final volume [minus inhibitor volume (in DMSO), if used, and minus biotin-ISG15-PA volume (required volume specified in Subheading 3.3)].
3. Optional: If using an inhibitor, add 1.5 μL of inhibitor stock in DMSO (final dilution $\sim 1:40$) and treat negative control with 1.5 μL of DMSO. Incubate at r.t. for 1 h.
4. Incubate with biotin-ISG15-PA at 37 $^{\circ}\text{C}$ for 45 min (*see Note 7*).
5. Quench labeling reaction with 0.4% SDS (w/v) (from a 5% stock) and 0.5% NP-40 (w/v) (from a 10% stock) using an 8-channel or 12-channel multichannel pipette. This is easiest done by mixing 99 μL of 5% SDS (w/v) stock with 62 μL of 10% NP-40 stock and adding 9.5 μL of this mixture to each labeling reaction (*see Note 8*).
6. Dilute each sample volume to 216 μL labeled samples with 141.7 μL of buffer D.
7. Remove 16 μL aliquot and label as “input” to check the efficiency of probe labeling/initial assessment of inhibition (if inhibitors are included) (Fig. 6) and check pull-down efficiency later.
8. Aliquot 20 μL of streptavidin magnetic Sepharose beads slurry per well in a 96-well plate using an 8-channel or 12-channel multichannel pipette (*see Note 9*).
9. Wash beads by adding 200 μL of buffer D, pipetting up and down several times, and place on a 96-well magnetic rack for ~ 30 s until all beads have pelleted toward the magnet.
10. Discard supernatant and repeat wash step four more times.
11. After the removal of supernatant from the final wash step, add the biotin-ISG15-PA-labeled lysate to the beads and incubate for 16–18 h on a thermomixer shaker, equipped with a plate adaptor, pre-set to 4 $^{\circ}\text{C}$, with shaking (*see Note 10*).
12. Place plate on a 96-well plate magnetic rack for ~ 30 s until all beads have pelleted toward the magnet and collect supernatant. Label as “supernatant” to check pull-down efficiency by looking at the depletion of probe-labeled proteins by western blot and store at -80°C if pull-down has not been successful.
13. Wash beads by adding 200 μL of buffer D, pipetting up and down several times, and place on a 96-well magnetic rack for ~ 30 s until all beads have pelleted toward the magnet.
14. Discard supernatant and repeat wash step four more times.

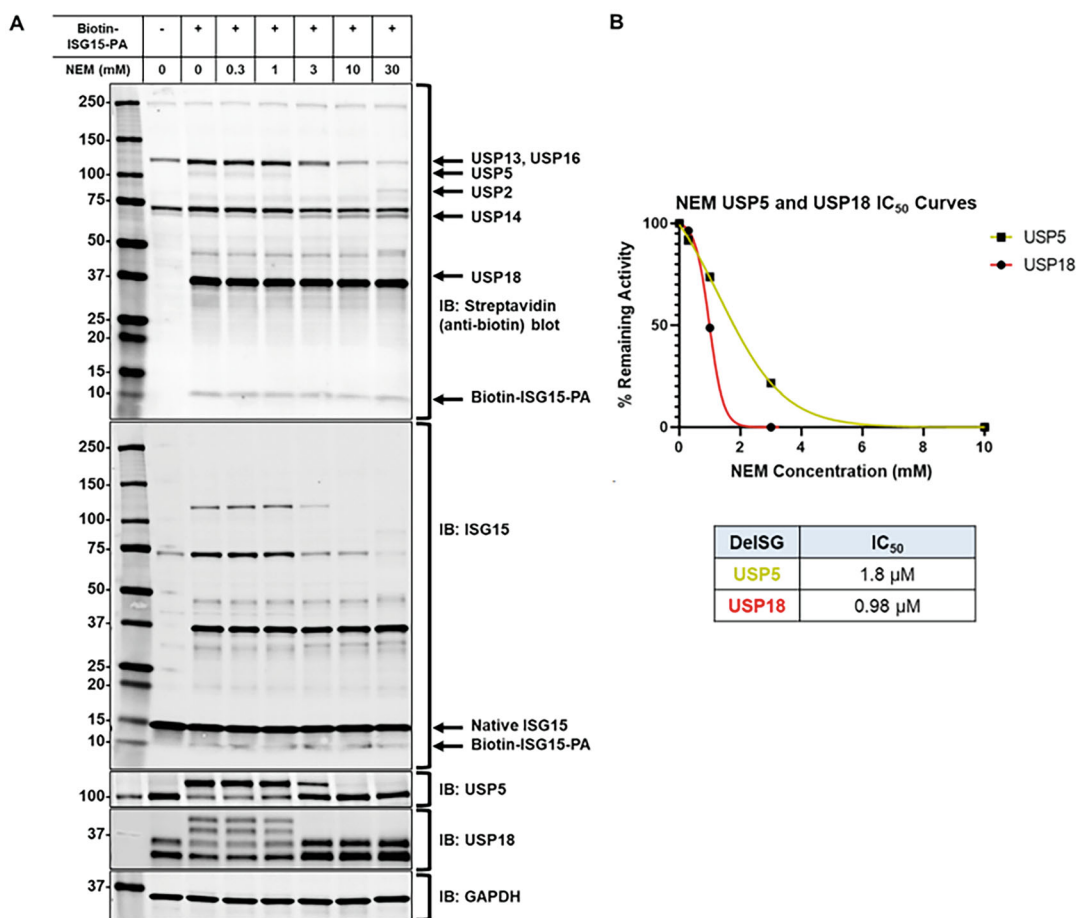


Fig. 6 Example western blots for screening of inhibitors against USP5 and USP18. **(a)** Immunoblotting for USP5 and USP18 shows dose-dependent inhibition of both enzymes with N-ethylmaleimide (NEM). The biotin and ISG15 blots show that NEM is a broad inhibitor of deISGylases, as evidenced by dose-dependent disappearance of multiple bands across the two blots, and also in the USP5 and USP18 blots. **(b)** IC₅₀ curves generated for NEM against USP5 and USP18 following quantification of probe-labeled USP5 and USP18 band intensities at the different NEM concentrations and normalization, as a %, relative to probe-labeled USP5 and USP18, respectively, in the probe-only positive control (= 100%). A table with the corresponding calculated IC₅₀ values for USP5 and USP18 is included

15. Add 100 μ L of 2x Laemmli buffer and elute proteins by boiling for 10 min on a thermomixer shaker, equipped with a plate adaptor, pre-set to 90 $^{\circ}$ C, with shaking (*see Note 10*).
16. Place plate on a 96-well plate magnetic rack for ~30 s until all beads have pelleted toward the magnet and collect the eluate.
17. Retain 10 μ L of the “eluate” fractions to check pull-down efficiency by confirming the presence of probe-labeled proteins by western blot.
18. Pause point: samples can be stored at -80° C for future use.

19. Run “input,” “supernatant,” and “eluate” fractions on a western blot for the no probe (negative control) and probe-only (positive control) samples. Blot with streptavidin to ensure lysate input fractions have been labeled with biotin-ISG15-PA, relative to the no probe control, to ensure that supernatant fractions have been completely depleted of probe-labeled proteins, and confirm enrichment of probe-labeled proteins in the eluate fractions.
20. Prepare samples for analysis by LC-MS/MS using a 96-well S-trap plate and following the manufacturer’s protocol for the reduction, alkylation, digestion, and peptide elution steps.
21. Dry samples on a vacuum centrifuge (Speedvac) and resuspend dry peptides in 20 μL of buffer E. Place plate on a thermomixer shaker at r.t. at 800 rpm for 5 min to ensure full resuspension of peptides.
22. Transfer samples to glass mass spectrometry vials or put a plate sealing cover if samples are to be analyzed directly from the 96-well plate.

3.6 Pull-Down of Biotin-ISG15-PA for High-Throughput LC-MS/MS Analysis Using the KingFisher Apex Purification System

1. Perform cell lysis as outlined in Subheading 3.2 and aliquot 275 μg of lysate per condition per well in a 96-deep well plate. Include a negative control sample without probe.
2. Dilute with buffer C to 147 μL final volume [minus inhibitor volume (in DMSO), if used, and minus biotin-ISG15-PA volume (required volume specified in Subheading 3.3)].
3. Optional: If using an inhibitor, add 1.5 μL of inhibitor stock in DMSO (final dilution 1:100) and treat negative control with 1.5 μL of DMSO. Incubate at r.t. for 1 h.
4. Incubate with biotin-ISG15-PA at 37 °C for 45 min (*see Note 7*).
5. Quench labeling reaction with 0.4% SDS (w/v) (from a 5% stock) and 0.5% NP-40 (w/v) (from a 10% stock) using an 8-channel or 12-channel multichannel pipette. This is easiest done by mixing 99 μL of 5% SDS (w/v) stock with 62 μL of 10% NP-40 stock and adding 21.5 μL of this mixture to each labeling reaction (*see Note 8*).
6. Remove 15 μL aliquot and label as “input” to check the efficiency of probe labeling/initial assessment of inhibition (if inhibitors are included) (Fig. 6) and check pull-down efficiency later.
7. Dilute the remaining ~150 μL labeled samples with 250 μL of buffer D.
8. Depending on the amount of beads required, aliquot 450 μL of streptavidin magnetic Sepharose beads slurry in 2 mL Eppendorf tubes. Make sure to mark the volume level of the slurry on the tube.

9. Wash beads by adding 1.6 mL of buffer D, inverting tubes 10 times, and place on a magnetic rack for ~30 s, until all beads have pelleted toward the magnet.
10. Discard supernatant and repeat wash step three more times.
11. Resuspend beads to their original volume by adding buffer D to the marked volume level on the tube (from **step 8**).
12. Prepare a script for the KingFisher Apex containing details for all steps regarding initial bead washing, sample incubation, and sample elution, as outlined in Subheading 3.6.1 (*see Note 11*).
13. Aliquot 20 μ L of streptavidin magnetic Sepharose beads slurry into each well of the sample plate diluted with buffer D (from **step 7**) using an 8-channel or 12-channel multichannel pipette (*see Note 9*).

3.6.1 Preparation and Running of KingFisher Apex Script

1. Open Thermo Scientific BindIx Software.
2. Click “Create protocol.”
3. Click “New protocol.”
4. In the “Protocol overview” tab, set the “Application type” to “IP/Co-IP” and provide details of “Kit used” and “Description” of the experiment.
5. In the “Magnet heads” tab, tick the “96 DW” magnet head option.
6. In the “Protocol steps” tab, a “Pick Up Tip” and “Leave Tip” steps should have been automatically added. If this is not the case, click on “Add tip” option. Both steps should now be visible, along with a “Define step” step in-between. For the “Pick Up Tip” and “Leave Tip” steps, leave all settings as default.
7. Click on “Define step.”
8. Under “Step Type” category, click on the dropdown menu and select “Bind,” tick “Automatic numbering,” and untick “Custom naming.”
9. Under “Step contents” category, set “Precollect beads” to “Off.”
10. Set “Release beads” to “On,” without setting a duration.
11. Set “Heating & Cooling” to “On,” click on the dropdown arrow, set the “Temperature” to 4 °C, and set “Preheat” to “On.”
12. Click on the dropdown arrow on “Mixing,” under “Mixing time 1” set to 4 h (04:00:00, hh:mm:ss format) and set “Mixing type” to “Medium” (*see Note 12*).
13. Set “Postmix” to “Off.”

14. Set “Collect beads” to “On,” click on the dropdown arrow, change “Count” to “5,” and change “Collect time (s)” to “3.”
15. Click on the tick option to save the settings and to be taken back to the “Protocol steps” tab.
16. Click on “Define plate” for that step.
17. Click on “New plate.”
18. Under “Plate Type” category, select “DW 96” from the left-hand dropdown menu, select “Sample Plate” from the right-hand dropdown menu, tick “Automatic numbering,” and untick “Custom naming.”
19. Under “Plate Contents” category, click “Add Content” and set “Volume [μ L]” to “400.” Leave the “Multiple plates” option unticked.
20. Click on the tick option to save the settings and to be taken back to the “Protocol steps” tab. This step will now appear as a “Bind” step in the “Protocol steps” workflow.
21. Click on “Add step” and click on “Define step.”
22. Under “Step Type” category, click on the dropdown menu and select “Wash,” tick “Automatic numbering,” and untick “Custom naming.”
23. Under “Step contents” category, set “Precollect beads” to “Off.”
24. Set “Release beads” to “On,” without setting a duration.
25. Set “Heating & Cooling” to “On,” click on the dropdown arrow, set the “Temperature” to 4 °C, and set “Preheat” to “On.”
26. Click on the dropdown arrow on “Mixing,” under “Mixing time 1” set to 30 seconds (00:00:30, hh:mm:ss format) and set “Mixing type” to “Medium” (*see Note 12*).
27. Set “Postmix” to “Off.”
28. Set “Collect beads” to “On,” click on the dropdown arrow, change “Count” to “5,” and change “Collect time (s)” to “3.”
29. Click on the tick option to save the settings and to be taken back to the “Protocol steps” tab. This step will now appear as a “Wash 1” step in the “Protocol steps” workflow.
30. Click on “Define plate.”
31. Click on “New plate.”
32. Under “Plate Type” category, select “DW 96” from the left-hand dropdown menu, select “Wash Plate” from the right-hand dropdown menu, tick “Automatic numbering,” and untick “Custom naming.”

33. Under “Plate Contents” category, click “Add Content” and set “Volume [μ L]” to “500.”
34. Click on the tick option to save the settings and to be taken back to the “Protocol steps” tab. This step will now appear as a “Wash 1” and “Wash Plate 1” step in the “Protocol steps” workflow.
35. Create three additional wash steps, “Wash 2,” “Wash 3,” and “Wash 4,” in the exact same manner as described for step “Wash 1.”
36. Click on “Add step” in the “Protocol steps” and click on “Define step.”
37. Under “Step Type” category, click on the dropdown menu and select “Elute,” tick “Automatic numbering,” and untick “Custom naming.”
38. Under “Step contents” category, set “Precollect beads” to “Off.”
39. Set “Release beads” to “On,” without setting a duration.
40. Set “Heating & Cooling” to “On,” click on the dropdown arrow, set the “Temperature” to 95 °C, and set “Preheat” to “On.”
41. Click on the dropdown arrow on “Mixing,” under “Mixing time 1” set to 8 minutes (00:08:00, hh:mm:ss format) and set “Mixing type” to “Slow” (*see* **Note 13**).
42. Set “Postmix” to “Off.”
43. Set “Collect beads” to “On,” click on the dropdown arrow, change “Count” to “5,” and change “Collect time (s)” to “3.”
44. Click on the tick option to save the settings and to be taken back to the “Protocol steps” tab. This step will now appear as a “Elute 1” step in the “Protocol steps” workflow.
45. Click on “Define plate.”
46. Click on “New plate.”
47. Under “Plate Type” category, select “DW 96” from the left-hand dropdown menu, select “Elution Plate” from the right-hand dropdown menu, tick “Automatic numbering,” and untick “Custom naming.”
48. Under “Plate Contents” category, click “Add Content” and set “Volume [μ L]” to “55.”
49. Click on the tick option to save the settings and to be taken back to the “Protocol steps” tab. This step will now appear as a “Elute 1” and “Elution plate 1” step in the “Protocol steps” workflow.

50. Create an additional elution step, “Elute 2” and “Elution plate 2” in the exact same manner as described for step “Elute 1” and “Elution plate 1,” with the only exception that in the “Mixing” category, “Mixing time 1” should be set to 5 min (00:05:00, hh:mm:ss format), instead of 8 min.
51. Click on “Protocol overview” tab and scroll down to the “Protocol overview” category toward the bottom of the page. Ensure that the “Tips” section states “1 × 96 DW tip comb,” the “Plates” section states “8 × DW 96,” and the “Steps” section states “9 × steps.”
52. Click “Save as” to save the script.
53. Prior to executing the script, prepare the “Pick Up Tip” plate (plate 1), the “Bind” sample plate (plate 2), the four “Wash” plates containing 500 µL buffer D, and the two “Elute” plates containing 55 µL of 2x loading buffer.
54. Click “Run” to execute the script and follow on-screen instructions.
55. Add plates one by one, as prompted on the screen of the instrument. The instrument prompts the user to add plates in reverse order—“Elution 2” plate at position 8 will be placed on the instrument first, and “Pick Up Tip” plate at position 1 will be placed on the instrument last.
56. Start the incubation.
57. After the script has finished running, remove all plates from the instrument and combine “Elution Plate” 1 and 2.
58. Retain 10 µL of the “eluate” fractions to check pull-down efficiency by confirming the presence of probe-labeled proteins by western blot.
59. Pause point: samples can be stored at −80 °C for future use.
60. Run “input,” “supernatant,” and “eluate” fractions on a western blot. Blot with streptavidin to ensure lysate input fractions have been labeled with biotin-ISG15-PA, relative to the no probe control, to ensure that supernatant fractions have been completely depleted of probe-labeled proteins, and to confirm enrichment of probe-labeled proteins in the eluate fractions.
61. Prepare samples for analysis by LC-MS/MS using a 96-well S-trap plate and following the manufacturer’s protocol for the reduction, alkylation, digestion, and peptide elution steps.
62. Dry samples on a vacuum centrifuge (Speedvac) and resuspend dry peptides in 20 µL of buffer E. Shake at r.t. at 800 rpm for 5 min, briefly vortex, and sonicate samples at r.t. for 5 min to ensure full resuspension of peptides.

3.7 LC-MS/MS High-Throughput Analysis Using TimsTOF

For high-throughput analysis of samples generated in the manner described in Subheadings 3.5 and 3.6, an Evosep Liquid Chromatography (LC) and TimsTOF MS/MS can be employed, which allows analysis of up to 100 samples per day. Details of the Evosep-TimsTOF setup employed for sample analysis generated using the methodology are outlined below.

1. Activate Evtips by soaking in 100 μ L isopropanol for at least 10 s.
2. Add 20 μ L of buffer F and centrifuge at 700 g for 1 min.
3. Soak tips in 100 μ L isopropanol again for at least 10 s.
4. Add 20 μ L of buffer G and centrifuge at 700 g for 1 min.
5. Load 200 ng of peptide in 20 μ L and centrifuge at 700 g for 1 min.
6. Add 20 μ L of buffer G again and centrifuge at 700 g for 1 min.
7. Add 100 μ L of buffer G once more and centrifuge at 700 g for 1 min.
8. For peptide separation, run samples on a PepSep 8 cm analytical C18 column (3 μ m beads, 100 μ m ID).
9. Perform sample analysis using the pre-set 100 samples per day method on the Evosep One instrument.
10. Acquire data using data-independent acquisition (DIA) PASEF mode in a mass range of 475–1000 m/z, with a 25 m/z width of each mass window isolated. Set the ion mobility window to 0.85–1.3 V and set the ion mobility-dependent collision energy to increase linearly from 27 eV to 45 eV for the 0.85–1.3 V window.

3.8 Data Processing

Several software packages are available as freeware that can be used for analysis of DIA proteomics data. For DIA, we used DIA-NN [36]. Below are outlined the settings used in DIA-NN for searching proteomics data acquired in DIA mode on the Evosep One TimsTOF platform. Unless specified below, all other settings should be left as default. Ensure that the latest UniProt FASTA database for the relevant species is downloaded and used.

1. Set the “Experiment name.”
2. Under “Input,” click on “diaPASEF.d” and select all the mass spectrometry acquired files that will be searched.
3. Click “Add FASTA” and add the relevant FASTA file.
4. Tick “FASTA digest for library-free search/library generation.”
5. Under “Precursor ion generation,” tick “Deep learning-based spectra, RTs and IMs prediction.”

6. Select the type of protease used for protein digestion from the dropdown menu.
7. Set “Missed cleavages” to “1”.
8. Set “Maximum number of variable modifications” to “1”.
9. Tick “N-term M excision” and “C carbamidomethylation”, and leave all other modifications unticked.
10. Set “Peptide length range” to “7 – 30”.
11. Set “Precursor charge range” to “1 – 4”.
12. Set “Precursor m/z range” to “300 – 1800”.
13. Set “Fragment ion m/z range” to “200 – 1800”.
14. Under “Output,” untick “Use existing quant files when available”.
15. Select the “Main output” folder, where all the data from the search will be saved.
16. Tick “Generate spectral library” and “Quantities matrices”.
17. Select the “Output library” folder. This can be the same folder as set for the “Main output” in **step 14**.
18. Untick “Generate Prosit input from FASTA or spectral library.”
19. Set “Precursor FDR (%)” to “1”.
20. Select the relevant number of “Threads” supported by the computer hardware. It is recommended to set the “Threads” number to 1 or 2 less than the total number of threads that the computer hardware supports—if the computer hardware supports 12 threads, set “Threads” to 10 or 11.
21. Tick “Generate PDF report” and set “Log level” to “1”.
22. Under “Algorithm,” set “Mass accuracy” to “10.0”.
23. Set “MS1 accuracy” to “10.0”.
24. Set “Scan window” to 0.
25. Untick “Unrelated runs”.
26. Tick “Use isotopologues,” “MBR,” “Heuristic protein inference,” and “No shared spectra” options.
27. For “Protein inference,” select “Genes” (from the dropdown menu).
28. For “Neural network classifier,” select “Single-pass mode.”
29. For “Quantification strategy,” select “Any LC (high accuracy).”
30. For “Cross-run normalisation,” select “On.”
31. For “Library generation,” select “Smart profiling.”
32. For “Speed and RAM usage,” select “Optimal results.”

33. Once all settings have been set, click “Run” to start the search. List of proteins identified in the LC-MS/MS data can be viewed in the generated “report.pg_matrix.tsv” file in the selected “Main output” folder.

3.9 Analysis of Western Blot and LC-MS/MS Data

1. Data from western blotting can be used to identify potential inhibitors against deISGylating proteins and assess inhibitor potency. This can be performed by measuring band intensities using a suitable imaging software (Image Studio, ImageJ) and comparison of intensity of probe-labeled bands in the absence of inhibitor with intensity of probe-labeled bands in the presence of inhibitor treatment. A decrease in intensity of probe-labeled protein bands is indicative of inhibition against that protein. The % of probe-labeled band intensity of inhibitor-treated sample at a given inhibitor concentration relative to the same band in the probe-only (no inhibitor) control can be plotted to generate a dose–response curve, in a software such as GraphPad, to determine an IC_{50} value for that inhibitor.
2. Data from LC-MS/MS can be used to identify (novel) deISGylating proteins by comparison of probe-treated samples versus no probe control and looking at proteins that are enriched in the probe-treated samples (Table 1). This implies that the proteins have been successfully enriched as a result of probe-labeling and not due to nonspecific interactions with the beads. In the probe-treated sample, the intensity of enriched deISGylating proteins can be assumed to represent 100% activity for

Table 1

List of known cysteine deISGylases identified in references [6, 20–22]. DeISGylases identified, by mass spectrometry, in biotin-ISG15-PA probe-treated versus no probe control samples derived from the HAP1 cellular proteome are highlighted in blue. Molecular weights for human deISGylases were taken from uniprot.com database

UniProt identifier	Gene	Molecular weight (kDa)	Reference
O75604	USP2	68.1	[20]
P45974	USP5	95.8	[6, 20, 22]
Q5BKP2	USP13	97.3	[20]
P54578	USP14	56.1	[6, 20]
Q9Y5T5	USP16	93.6	[6, 21]
Q9UMW8	USP18	41.0/43.0	[6, 20, 21]
Q9P275	USP36	122.9	[21]

those proteins. A decrease in intensity following treatment with inhibitor can be calculated as a % activity remaining relative to the probe-only control sample, and this can be plotted on a dose–response curve, in a software such as GraphPad, to determine an IC_{50} value for that inhibitor.

4 Notes

1. For the preparation of agar plates, mix LB media with 1.5% agar (w/v) and sterilize by autoclaving. Ampicillin should be added at the appropriate concentration once the bottle is sufficiently cool to touch (this avoids degradation of the antibiotic) but prior to the media solidifying. Once ampicillin is added, immediately pour into plates. Avoid formation of bubbles.
2. In some cases, endogenous biotinylation levels of avi-ISG15-MESNA are low, and the corresponding biotin-ISG15-MESNA band (~10 kDa, Fig. 2a) may be of low intensity or undetectable on a streptavidin blot. If the band for biotin-ISG15-MESNA is present on the Coomassie stain gel (~10 kDa, Fig. 2b), proceed with probe synthesis regardless of whether the band is seen by western blot with streptavidin.
3. Enzymatic biotinylation of avi-ISG15-PA to biotin-ISG15-PA is recommended in cases where endogenous biotinylation levels are low/undetectable by western blot. Successful biotinylation of the probe can be detected by a streptavidin gel-shift assay (Fig. 3).
4. Amount of protein obtained will vary between different cell lines due to differences in size; therefore, the volume of lysis buffer used must be adjusted so that protein concentration falls within the desired range.
5. For tissue analysis, tissue should be homogenized either using a glass dounce homogenizer, followed by glass bead lysis, or using a bead beater.
6. A two-step centrifugation, a low-speed step followed by a high-speed step, is recommended to maximize deISGylases coverage by isolating proteins from different-sized organelles.
7. The biotin-ISG15-PA probe labeling reaction typically works well at 37 °C; however, with older lysates degradation of probe-labeled USP18 has been observed at 37 °C during the reaction. If this is the case, the reaction can also be performed at r.t., resulting in nearly the same labeling efficiency as that of a new lysate.
8. Prepare the SDS and NP40 mixture fresh and add to the samples immediately after mixing to avoid potential quenching of SDS by the nonionic detergent.

9. Streptavidin magnetic Sepharose beads tend to settle very quickly; therefore, when transferring beads into wells using a multichannel pipette, mix the beads slurry by pipetting up and down several times between each transfer. This ensures homogeneous slurry and that the same amount of beads have been transferred into each well each time.
10. Shaking speed must be adjusted to high enough to ensure sufficient bead mixing and avoid pelleting, but not too high to avoid splashing of beads/samples into other wells. Adjust the mixing speed, monitor the mixing, and, if necessary, further adjust the mixing speed accordingly. Insufficient mixing will compromise the pull-down efficiency.
11. It is recommended to prepare the script for the KingFisher Apex prior to starting the actual experiment.
12. “Medium” setting should be used for mixing during sample incubation and bead washing steps since streptavidin magnetic Sepharose beads tend to settle very quickly and will pellet at the bottom of the wells if a “Slow” setting is used, which will compromise the pull-down efficiency and wash steps.
13. “Slow” setting is recommended for mixing during sample elution steps to avoid sample splashing onto well walls since a small elution volume is used. Following quality control by western blotting, if elution appears to be incomplete, mixing can be set to “Medium.”

Acknowledgments

We would like to thank Amber Truepenny in the Kessler Group and Helene Greenwood for their help with biotin-ISG15 probe synthesis and initial ABPP optimization, respectively. We would also like to thank Raphael Heilig for his help with initial DIA-PASEF method development and the Discovery Proteomics Facility at the University of Oxford for their help with sample analysis by mass spectrometry. Additionally, we would like to thank Thermo Fisher Scientific for providing us with the KingFisher Apex Purification System to test. S.D.D., B.M.K., and A.P.F. were supported by Pfizer and the Chinese Academy of Medical Sciences (CAMS) Innovation Fund for Medical Science (CIFMS), China (grant no. 2018-I2M-2-002). H.B.L.J. was supported by Bristol Myers Squibb.

References

1. Fang H et al (2021) Recent advances in activity-based probes (ABPs) and affinity-based probes (AfBPs) for profiling of enzymes. *Chem Sci* 12:8288–8310
2. Porta EOJ et al (2023) Activity-based protein profiling: a graphical review. *Curr Rev Pharmacol Drug Discov* 5:100164
3. Ritchie KJ et al (2004) Role of ISG15 protease UBP43 (USP18) in innate immunity to viral infection. *Nat Med* 10:1374–1378
4. Ketscher L et al (2015) Selective inactivation of USP18 isopeptidase activity in vivo enhances ISG15 conjugation and viral resistance. *Proc Natl Acad Sci USA* 10:1577–1582
5. Tang L et al (2023) USP18 promotes innate immune responses and apoptosis in influenza A virus-infected A549 cells via cGAS-STING pathway. *Virology* 585:240–247
6. Gan J et al (2023) USP16 is an ISG15 cross-reactive deubiquitinase that targets pro-ISG15 and ISGylated proteins involved in metabolism. *Proc Natl Acad Sci USA* 120: e2315163120
7. Huang W et al (2022) USP5 promotes breast cancer cell proliferation and metastasis by stabilizing HIF2 α . *J Cell Physiol* 237:2211–2219
8. Song C et al (2021) USP18 promotes tumor metastasis in esophageal squamous cell carcinomas via deubiquitinating ZEB1. *Exp Cell Res* 409:112884
9. Pinto-Fernandez A et al (2021) Deletion of the deISGylating enzyme USP18 enhances tumour cell antigenicity and radiosensitivity. *Br J Cancer* 124:817–830
10. Du XH et al (2023) USP14 promotes colorectal cancer progression by targeting JNK for stabilization. *Cell Death Dis* 14:56
11. Deng Q et al (2022) USP36 promotes tumor growth of non-small cell lung cancer via increasing KHK-A expression by regulating c-MYC-hnRNPH1/H2 axis. *Hum Cell* 35: 694–704
12. Qin K et al (2023) USP36 stabilizes nuclear Snail1 to promote ribosome biogenesis and cancer cell survival upon ribotoxic stress. *Nat Commun* 14:6473
13. Jove V et al (2023) Type I interferon regulation by USP18 is a key vulnerability in cancer. *iScience* 27:109593
14. Fernandes dos Santos P et al (2017) Beyond ISGylation: functions of free intracellular and extracellular ISG15. *J Interf Cytokine Res* 37: 246–253
15. Sandy Z et al (2020) More than meetings the ISG15: emerging roles in the DNA damage response and beyond. *Biomol Ther* 10:1557
16. Malakhova OA et al (2008) ISG15 inhibits Nedd4 ubiquitin E3 activity and enhances the innate antiviral response. *J Biol Chem* 283: 8783–8787
17. Perng YC et al (2018) ISG15 in antiviral immunity and beyond. *Nat Rev Microbiol* 16: 423–439
18. Radoshevich L et al (2015) ISG15 counteracts *Listeria monocytogenes* infection. *elife* 4: e06848
19. Zhang Y et al (2019) The in vivo ISGylome links ISG15 to metabolic pathways and autophagy upon *Listeria monocytogenes* infection. *Nat Commun* 10:5838
20. Catic A et al (2007) Screen for ISG15-crossreactive Deubiquitinases. *PLoS One* 2: e679
21. O'Dea R (2023) Molecular basis for ubiquitin/Fubi cross-reactivity in USP16 and USP36. *Nat Chem Biol* 19:1394–1405
22. Hemelaar J et al (2004) Specific and covalent targeting of conjugating and deconjugating enzymes of ubiquitin-like proteins. *Mol Cell Biol* 24:84–95
23. Frias-Staheli N et al (2007) Ovarian tumor domain-containing viral proteases evade ubiquitin- and ISG15-dependent innate immune responses. *Cell Host Microbe* 2:404–416
24. Akutsu M et al (2011) Molecular basis for ubiquitin and ISG15 cross-reactivity in viral ovarian tumor domains. *Proc Natl Acad Sci USA* 108:2228–2233
25. Shin D et al (2020) Papain-like protease regulates SARS-COV-2 viral spread and innate immunity. *Nature* 587:657–662
26. Klemm T et al (2020) Mechanism and inhibition of the papain-like protease, PLpro, of SARS-COV-2. *EMBO J* 39:e106275
27. Swatek KN et al (2018) Irreversible inactivation of ISG15 by a viral leader protease enables alternative infection detection strategies. *Proc Natl Acad Sci USA* 115:2371–2376
28. Sui X et al (2020) Development and application of ubiquitin-based chemical probes. *Chem Sci* 11:12633–12646
29. Ekkebus R (2013) On terminal alkynes that can react with active-site cysteine nucleophiles in proteases. *J Am Chem Soc* 135:2867–2870
30. Chan WC et al (2023) Accelerating inhibitor discovery for deubiquitinating enzymes. *Nat Commun* 14:686

31. Hewings DS et al (2017) Activity-based probes for the ubiquitin conjugation-deconjugation machinery: new chemistries, new tools, and new insights. *FEBS J* 284:1555–1576
32. Turnbull AP et al (2017) Molecular basis of USP7 inhibition by selective small-molecule inhibitors. *Nature* 550:481–486
33. Mahajan S et al (2022) Deubiquitinase vulnerabilities identified through activity-based protein profiling in non-small cell lung cancer. *ACS Chem Biol* 17:776–784
34. Jones HBL et al (2021) ABPP-HT—High-throughput activity-based profiling of deubiquitylating enzyme inhibitors in a cellular context. *Front Chem* 9:640105
35. Jones HBL et al (2022) ABPP-HT*—deep meets fast for activity-based profiling of deubiquitylating enzymes using advanced DIA mass spectrometry methods. *Int J Mol Sci* 23:3263
36. Demichev V et al (2020) DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods* 17:41–44



Chapter 2

Integrated Multi-omics Approaches for Studying Rare Genetic Diseases

Refat M. Nimer, Lina A. Dahabiyeh, Reem AlMalki, Maha Al Mogren, Hicham Benabdelkamel, Assim A. Alfadda, Salini Scaria Joy, and Anas M. Abdel Rahman

Abstract

Despite the transformation of genomics and genetics, DNA- and RNA-based information provides only a partial view of disease etiology and pathogenesis. This has increased awareness that genetic and gene expression data must be integrated with downstream product activity and cellular metabolite regulation to understand disease processes fully.

By simultaneously analyzing the genome, transcriptome, proteome, and metabolome, crucial molecular pathways and novel biomarkers associated with various genetic diseases have been identified using multi-omics approaches. A more comprehensive understanding of the complex interactions between genetic factors (genotype) and disease development (phenotype) has been enabled by these approaches.

This chapter describes multi-omics protocols for genetic diseases, emphasizing metabolomics and proteomics approaches.

Key words Genetic diseases, Proteomics, Metabolomics, Multi-omics

1 Introduction

Approximately 10,000 distinct genetic illnesses impact the lives of hundreds of millions of people [1]. Genetic disorders (GDs) include many diseases that arise from gene variations and mutations. On their own, GDs are very uncommon, but together, they make up more than 80% of all rare diseases [2]. Biofluids, cells, and tissues are all components of biological matrices, and GD is no exception; as with any disease, it produces unique, abnormal molecular profiles. Measuring proteins (gene product) and metabolites (most biomolecules reflecting the phenotype) may help identify novel biomarkers of various disorders and provide new insight into their molecular causes [3].

In recent years, multi-omics approaches, particularly metabolomics and proteomics, have gained increasing attention in studying GDs. These approaches allow for a comprehensive analysis of various biological molecules and their interactions, providing valuable insights into the underlying molecular mechanisms of GDs [4]. Researchers can better understand disease pathogenesis by integrating data from multiple omics platforms and discovering potential biomarkers and therapeutic targets [5]. This chapter explores applying multi-omics approaches to studying rare GDs, specifically focusing on metabolomics and proteomics. The genomics approach alone is inadequate for understanding the complexity of GDs' phenotypes. In addition to genomics, transcriptomics, proteomics, lipidomics, and metabolomics play crucial roles in bridging the gap between genotypes and phenotypes [6, 7]. Proteomics studies the entire complement of proteins in a biological system, including their structures, functions, and interactions [8]. These findings provide insights into the dynamic nature of proteins, their post-translational modifications (PTMs), and how they contribute to the development and progression of GDs [9, 10]. Metabolomics, on the other hand, focuses on the comprehensive analysis of small molecules, or metabolites, less than 1500 Daltons within cells, tissues, or biofluids, providing the most precise representation of the phenotype [11–13]. It provides a snapshot of the biochemical activity within a cell and can reveal important metabolic alterations associated with GDs [14, 15].

Furthermore, metabolomics and proteomics can provide valuable information about disease progression, classification, response to treatment, and potential therapeutic targets in a personalized healthcare setting [16–18].

Metabolomics and proteomics studies can be broadly categorized into untargeted and targeted approaches. Untargeted metabolomics involves comprehensively analyzing all detectable metabolites within a biological sample without prior knowledge of specific metabolites of interest. This approach allows for identifying novel metabolites and metabolic pathways that may be dysregulated in the GDs [19]. Untargeted metabolomics often begins with sample collection and metabolite extraction, then separating the extracted metabolites using chromatography (liquid or gas chromatography) [20]. Metabolites will be analyzed using mass spectrometry (MS) or nuclear magnetic resonance (NMR) spectroscopy. Raw data processing and metabolite identification will be performed using specialized software such as MS Dial, Progenesis, and XCMS [21–23]. Finally, uni- and multivariate statistical analyses will be used to identify potential novel biomarkers [24].

Untargeted proteomics involves various technical approaches and sample preparation methods such as sample collection, protein extraction, depletion, digestion, and MS-based experiments [25]. One useful strategy for untargeted proteomics is

two-dimensional difference gel electrophoresis (2D-DIGE). This technique uses cyanine fluorescent dyes to identify numerous protein samples that simultaneously comigrate in the same gel, speeding up the analytical and experimental processes. 2D-DIGE offers a more rapid and accurate method of gel matching, which reduces the effect of gel-to-gel variation and provides a wider dynamic range for quantitative comparisons than conventional post-staining 2D-polyacrylamide GE (PAGE) techniques such as colloidal Coomassie or silver nitrate. Using internal standards, it is feasible to standardize experimental differences in spot intensities and gel patterns [26].

Regarding the targeted approach, targeted metabolomics focuses on quantifying specific metabolites of interest, often based on prior knowledge of their involvement in disease pathways. Similarly, targeted proteomics aims to quantify particular proteins or PTMs that are known to be relevant to the pathogenesis of genetic diseases [27, 28].

In the following sections, we will outline the specific methodologies and technologies commonly used in untargeted approaches for metabolomic and proteomic studies.

2 Materials

2.1 Equipment

1. Waters ACQUITY ultra-performance liquid chromatography (UPLC) system (Waters Ltd.).
2. Xevo G2-S QToF mass spectrometer equipped with an electrospray ionization source (ESI) (Waters Ltd.).
3. ACQUITY UPLC (HSS T3 1.8 μm , 2.1 $\times$ 100 mm) column, (Waters).
4. A Q Exactive Plus Hybrid Quadrupole-Orbitrap mass spectrometer.
5. Matrix-Assisted Laser Desorption Ionization (MALDI) target (384 MTP 800 μm AnchorChip; Bruker Daltonik).
6. MALDI-TOF MS.
7. Vortex mixer.
8. ThermoMixer.
9. Centrifuge 5425 R.
10. SpeedVac.
11. Ettan IPGphor IEF unit (GE Healthcare).
12. Ettan Dalt Six device (GE Healthcare).
13. Sapphire biomolecular imager.

14. Eppendorf micropipettes and tips of different volumes (2000, 1500, 1000, 500, 100, 200, 10) μL .
15. LC-MS vials.
16. Electrophoresis Power Supply—EPS 601 (GE Healthcare).
17. Glass vial inserts (Waters).
18. Millipore water, deionized water (dH_2O) (should be kept at room temperature).
19. Dionex UltiMate 3000 nano-LC system, Dionex UltiMate3000 micro-LC system with a WPS-3000 autosampler (Thermo Fisher Scientific).
20. Analytical column (PepMap[™] C18, 50 cm $\times$ 75 μm) (Fisher Scientific).
21. Trapping column (C18 PepMap100, 5 μm , 100A)) (Fisher Scientific).
22. Freezer ($-30\text{ }^\circ\text{C}$) (SANYO MEDICOOL).
23. Analytical balance (KERN).
24. A UV/Visible spectrophotometer (Biochrom Ltd.).
25. Heating and Drying Ovens (Thermo Scientific).
26. Sonicator.
27. Micro Palm Centrifuge.
28. Automated DBS puncher instrument (PerkinElmer).
29. Cell scraper.
30. Cell culture dish (100 mm $\times$ 20 mm).
31. T25 digital ULTRA TURRAX homogenizer.

2.2 Chemicals (see Note 1)

1. LC-MS grade methanol (CH_3OH) (MeOH).
2. LC-MS grade acetonitrile (CH_3CN) (ACN).
3. LC-MS grade formic acid (FA) (HCOOH).
4. Sodium formate (HCOONa).
5. Isopropanol ($\text{C}_3\text{H}_8\text{O}$).
6. LockSpray [Leucine-enkephalin] ($\text{C}_{28}\text{H}_{37}\text{N}_5\text{O}_7$) (an external reference to the ion m/z 556.2771 in (ESI^+) and 554.2615 in (ESI^-)).
7. Protease inhibitor.
8. Depletion Spin Columns.
9. A 2D Quant kit.
10. IPG Buffer.
11. Tris.
12. Acrylamide ($\text{C}_3\text{H}_5\text{NO}$).

13. Methylenebisacrylamide ($C_7H_{10}N_2O_2$).
14. Urea (CH_4N_2O).
15. Ammonium hydrogen carbonate ($[NH_4]HCO_3$).
16. Thiourea ($SC(NH_2)_2$).
17. Dithiothetrol (DTT) ($C_4H_{10}O_2S_2$).
18. Iodoacetamide (IAA) (C_2H_4INO).
19. Coomassie® Brilliant blue G 250.
20. Bromophenol blue.
21. Lysine ($C_6H_{14}N_2O_2$).
22. *NN*-Dimethylformamide.
23. CyDye™ DIGE Fluor dyes.
24. CHAPS ($C_{32}H_{58}N_2O_7S$).
25. Ammonium hydrogen carbonate (NH_4HCO_3).
26. Sequencing grade modified trypsin.
27. Pierce™ C18 Spin Columns.
28. Dry Strips.
29. Dry Strips.
30. Trifluoroacetic acid ($C_2HF_3O_2$).
31. Phosphate-buffered saline (PBS).

2.3 Software

1. Masslynx™ V4.1 workstation.
2. Progenesis QI v.3.0 software from Waters.
3. Progenesis SameSpots software.
4. MetaboAnalyst v.6 (<http://www.metaboanalyst.ca>).
5. Sapphire capture system image analysis software.
6. Mascot search algorithm (version 2.0.04; Matrix Science Ltd.).
7. Flex Analysis software (version 2.4, Bruker Daltonics).
8. BioTools v3.2 (Bruker Daltonics).
9. Ingenuity Pathway Analysis (IPA) software (license of 2024, Ingenuity® Systems, <http://www.ingenuity.com>).
10. Mass Profiler Professional (MPP) v.15.0 software (Agilent).
11. Proteome Discoverer 3.0 (Thermo Fisher Scientific).
12. Universal Protein Resource (UniProt) database (<https://www.uniprot.org>).
13. PANTHER (protein analysis through evolutionary relationships <http://www.pantherdb.org>), Thomas lab at the University of Southern California.

3 Methods

3.1 General

Untargeted

Metabolomics Protocol

3.1.1 Metabolite

Extraction and Sample

Preparation

Metabolite Extraction from

Dry Blood Spot

Dried blood spot (DBS) is the most useful sample collected from newborns or babies with rare GDs for metabolomics analysis. Below is the main protocol developed in our lab for untargeted metabolomics analysis in studying these cases, as published earlier [19, 29, 30].

1. One punch, 3.2 mm in size, is used for DBS for each sample using an automated puncher.
2. Prepare quality control samples (QCs) by collecting a single 3.2 mm punch from each study sample and pooling it for extraction.
3. Then, add 300 μL of extraction solvent (20:40:40) (dH_2O : MeOH:ACN).
4. Vortex the samples in a ThermoMixer at 600 rpm, 25 $^{\circ}\text{C}$ for 2 h.
5. Spin the samples down at 16000 rpm for 5 min.
6. Transfer the supernatant to new Eppendorf tubes and discard the punches.
7. Dry the samples using a SpeedVac.
8. Store the samples at -80°C until LC-MS analysis.

Metabolite Extraction from

Plasma and Serum

Serum and plasma are the second most commonly used biofluids in metabolomics analysis; below is the main protocol developed in our lab for untargeted metabolomics analysis to study many conditions, such as hyperthyroidism, with treatment [31, 32].

1. Collect 50 μL of plasma or serum. Meanwhile, prepare QCs by mixing equal volumes (10 μL) of each study sample and pooling them for extraction.
2. Then, add 950 μL of extraction solvent (1:1) (v/v) (ACN: MeOH) to each study and QC sample.
3. Vortex the samples in a ThermoMixer at 600 rpm, 25 $^{\circ}\text{C}$ for 1 h.
4. Spin the samples down at 16000 rpm, 4 $^{\circ}\text{C}$ for 10 min.
5. Transfer the supernatant to a new Eppendorf tube.
6. Dry the samples using a SpeedVac.
7. Store the samples at -80°C until LC-MS analysis.

Metabolite Extraction from

Cells

Primary cell culture experiments are commonly used to study rare GDs that interact with metabolism at the cellular and systemic levels. Here, we present the main protocol developed in our lab for untargeted metabolomics analysis in studying these cases, as published earlier [33, 34].

Cell Pellets (Intracellular Metabolites)

1. Collect the conditioned media for secretome metabolite extraction.
2. Wash the cells quickly with cold 1× PBS.
3. Place the plate in liquid nitrogen to quench the cells.
4. Next, extract the metabolites by adding 500 µL of cold 80% (v/v) MeOH:dH₂O. Agitate the wells to ensure that all cells are covered.
5. Place the plates on liquid nitrogen, scrape the cell monolayer with a cell scraper, and collect the cells with 500 µL of cold 80% (v/v) MeOH:dH₂O in a labeled Eppendorf tube.
6. Prepare QCs by mixing equal volumes (10 µL) of each study sample and pooling them for extraction.
7. Submerge the sample tube in ice and place it in a ThermoMixer at 600 rpm, 4 °C for 1 h.
8. Spin the samples at 16000 rpm, 4 °C for 10 min, and transfer the supernatant to a newly labeled Eppendorf tube.
9. Dry the samples using a SpeedVac and store them at −80 °C until LC-MS analysis.

Secretome (Extracellular Metabolites)

1. Add 950 µL of extraction solvent (1:1) (v/v) (ACN:MeOH) to 50 µL of culture-conditioned media.
2. Meanwhile, prepare QCs by mixing equal volumes (10 µL) of each culture-conditioned media and pooling them for extraction.
3. Vortex the samples in a ThermoMixer at 600 rpm, 25 °C for 1 h.
4. Speed down the samples at 16000 rpm, 4 °C for 10 min.
5. Transfer the supernatant to a new Eppendorf tube.
6. Dry the samples using a SpeedVac.
7. Store the samples at −80 °C for LC-MS analysis.

*3.1.2 LC-MS/MS Using High-Resolution Mass Spectrometry (Q-TOF)**Reconstitution of the Samples for LC-MS Analysis*

1. Reconstitute the dry samples prepared in previous sections with 100 µL of (1:1) (v/v) mobile phase A:B [**A**: 0.1% FA in dH₂O, **B**: 0.1% FA in (1:1) (v/v) MeOH and ACN].
2. Vortex the samples for 2 min.
3. Centrifuge the samples in an Eppendorf tube for 5 min at 16,000 rpm, 4 °C.
4. Transfer the samples to glass vial inserts and place them in autosampler vials for LC-MS analysis.

LC-MS Conditions for Metabolomics Analysis (see Note 2)

1. Connect the analytical column (ACQUITY UPLC HSS T3 1.8 µm, 2.1× 100 mm) to the UPLC instrument.

2. Prepare the mobile phase solution (0.1% FA in dH₂O) as solvents A and B consisting of 0.1% FA in 1:1 (v/v) ACN: MeOH.
3. Open the inlet page from the Masslynx™ V4.1 workstation:
 - 3.1. Set the gradient elution schedule as follows: 0–16 min 95–5% A, 16–19 min 5% A, 19–20 min 5–95% A, and 20–22 min 5–95% A with a 300 µL/min flow rate.
 - 3.2. Set the run time to 22 min.
 - 3.3. Set the injector volume to 5 µL.
 - 3.4. Sample Manager, choose the following option:
 - a. Needle placement (from bottom) 4 mm
 - b. Offline loop: disabled.
 - c. Wash solvents: 0.1% FA in water.
 - d. Temperature: column (40 °C) and sample (4 °C).
4. Open the tuning page from the Masslynx™ V4.1 workstation and set up the ion source parameters as detailed in Table 1.
5. In MS^E mode, the collision energies of the low and high functions are set at off and 10 V to 50 V, respectively.
6. Calibrate the mass spectrometer with sodium formate (HCOONa) in 100–1200 Da using the Console page from the Masslynx™ V4.1 workstation.
7. From the tune page in the Masslynx™ V4.1 workstation, set LockSpray [Leucine-enkephalin] (C₂₈H₃₇N₅O₇) to be injected continuously, switching between the sample and reference every 45 and 60 s for ESI+ and ESI-, respectively, scan time is 0.5 seconds, the flow rate is 10 µL/min, the cone is 30 V, and collision energy is 4 V.

Table 1
The ion source parameters

Parameter	Value	
	ESI+	ESI–
Capillary voltage (kv)	3.0	2.5
Cone voltage (v)	40	40
Source temperature (°C)	150	150
Desolvation temperature (°C)	500	140
Desolvation gas (L/hr)	800	800
Cone gas (L/hr)	50	50
LockSpray capillary (kv)	3.0	2.49

8. In the Masslynx™ V4.1 workstation, set the instrument sample list to include five blanks and three QCs at the beginning of the analytical batch, followed by study samples. Ensure there is a QC after each of the randomized 10 study samples, and end the batch with three QCs and two blank samples.
9. Collect the data in continuum mode with a Masslynx™ V4.1 workstation.

3.1.3 Raw Data Processing

After completing the sample batch analysis, process the raw MS data following a standard pipeline starting from alignment based on the m/z value and the ion signal retention time, peak picking, and signal filtering based on the peak quality using Progenesis QI v.3.0 software from Waters.

3.1.4 Metabolite Identification

Using Progenesis QI v.3.0 software, the ion features are annotated using databases:

1. According to the Human Metabolome Database (HMDB), metabolites' chemical structures will be identified by acquiring their accurate precursor masses and theoretical MS/MS fragmentation tolerance values of a 5-ppm mass window and an isotopic similarity of 80%.
2. METLIN MS/MS (www.metlin.scripps.edu), with precursor masses and fragmentation patterns of 5 ppm, filters fragmentations in silico or empirically.
3. Kyoto Encyclopedia of Genes and Genomes (KEGG), with a precursor mass and fragmentation pattern of 5 ppm and an isotopic similarity of 80%.
4. LipidBlast, with precursor masses and fragmentation pattern of 5 ppm.
5. Lipid Map, with precursor masses and fragmentation pattern of 5 ppm and isotopic similarity of 80%.
6. Mass Bank, with precursor masses and fragmentation pattern of 5 ppm and isotopic similarity of 80%.
7. After annotating metabolites, pick the significant metabolites, keep endogenous metabolites, and exclude exogenous compounds such as drugs, food additives, and environmental compounds from the final list.

3.1.5 Statistical Analysis (see **Note 3**)

Multivariate Statistical Analysis

1. Multivariate analysis will be performed using MetaboAnalyst v.6 (<http://www.metaboanalyst.ca>).
2. The data sets (compound names and their raw abundances) will be median-normalized, Pareto-scaled, and log-transformed to maintain normal distribution.
3. Generate partial least squares-discriminant analysis (PLS-DA) to evaluate the quality of the experiment and the segregation between the study groups.

4. Generate an orthogonal partial least squares-discriminant analysis (OPLS-DA) model to evaluate the fitness of the models (R^2Y) and their predictive ability (Q^2) via permutation validation of 100 samples.

Univariate Analysis

1. Univariate analysis will be performed using Mass Profiler Professional (MPP) v.15.0 software.
2. One-way analysis of variance (ANOVA) [Tukey's post hoc $p < 0.05$] will be performed between groups (more than two groups).
3. Venn diagrams will be developed using MPP v.15.0 software to determine the overlap between groups.
4. Volcano plot representation, for the comparison between two groups, will identify significantly altered mass features based on a fold change (FC) cutoff of 2 or 1.5 and an FDR of $p < 0.05$.
5. Heatmap analysis will be performed for significantly endogenous metabolites using the Pearson distance measure according to the Pearson similarity test.
6. Pathway analysis will be performed for significantly endogenous metabolites.
7. Based on the significant list, a biomarker analysis linked the metabolites with the disorder is evaluated by receiver operating characteristic (ROC) curves model on MetaboAnalyst (v.6.0) by Monte Carlo cross-validation (MCCV) with balanced subsampling. During each MCCV iteration, two-thirds (2/3) of the samples are designated for assessing feature importance. The top-ranked features, as determined by the PLS-DA algorithm, are subsequently employed to construct classification models. These models will then be validated on the remaining one-third (1/3) of the samples. This procedure will be iteratively performed multiple times to ascertain each model's performance and confidence interval. The PLS-DA algorithm will be utilized for feature ranking with latent variables (LV) 2. The value is considered only if the provided LV count is within the number of features.

3.2 General Untargeted Proteomics Protocol

3.2.1 Highly Abundant Protein Depletion for Plasma and Serum Samples

Highly Abundant Protein Depletion for Plasma Samples-Spin Column Procedure for the Removal of the Top 12 Proteins

1. Equilibrate the depletion spin column to room temperature.
2. Remove the column screw cap and add 15 μ L of the sample directly to the resin slurry in the column.
3. Cap the column and invert until the resin is completely suspended in the solution.

4. Incubate the mixture in the column with gentle end-over-end mixing for 60 min at room temperature. Mix the sample with the resin during the incubation period. Alternatively, vortex the mixture every 5 min.
5. Twist off the bottom closure and loosen the cap. Place the column into a 1 mL collection tube and centrifuge at $1000\times g$ for 2 min.
6. Discard the column containing the resin.
7. Remove the top 12 proteins from the filtrate. Use for further processing or storage for later use (*see* **Note 4**).

3.2.2 Protein Extraction

Plasma/Serum

1. Incubate the depleted sample (*see* Subheading 3.2.1, **step 1**) (with ice-cold acetone four times the sample volume overnight at -20°C).
2. On the next day, centrifuge the sample at $14000\times g$ for 10 min.
3. Discard the supernatant and wash the protein pellet with ice-cold acetone (volume of acetone to cover the pellets).
4. Centrifuge at $14000\times g$ for 10 min and discard the supernatant.
5. Incubate pellets to evaporate all the remaining acetone.
6. Add lysis buffer containing 7 M urea, 2 M thiourea, 4% CHAPS, and 30 mM Tris-base (*see* **Notes 5–8**).
7. Sonicate 1 min (pulse three times) using Sonicator.
8. Vortex and centrifuge at $14000\times g$ for 10 min.
9. Remove the nondissolved particles in the tube if present.

Tissue

1. Weight 100 mg of tissue sample
2. Clean and remove blood clots and blood vessels in 1X PBS.
3. Homogenize the tissue samples with T25 digital ULTRA TURRAX homogenizer directly in lysis buffer containing 7 M urea, 2 M thiourea, 4% CHAPS, 30 mM Tris-base, and 10 μL of protease inhibitor (*see* **Notes 5–8**).
4. The suspension was shaken for 1 h at room temperature and then sonicated (pulse, two intervals of 1 min each, separated by a 1 min gap).
5. Centrifuge at $16000\times g$ at 4°C . Collect the supernatant.
6. Add four times ice-cold acetone and incubate overnight at -20°C .
7. After overnight incubation, centrifuge the protein extracts ($20,000\times g$, 40 min, 4°C).
8. Discard the supernatant and keep the pellet dry at the bottom.
9. Wash the protein pellet with ice-cold acetone (volume of acetone to cover the pellets).

10. Incubate pellets to evaporate all the remaining acetone.
11. The protein pellets should be dissolved in lysis buffer containing 7 M urea, 2 M thiourea, 4% CHAPS, and 30 mM Tris-base (*see* **Notes 5–8**).
12. Sonicate pulse three times.
13. Vortex and centrifuge ($20,000\times g$, 40 min, 4 °C).
14. Remove nondissolved particles in the tube if present.

Urine

1. Add 3 mL of urine sample to 3 mL of MeOH and 0.75 mL of chloroform.
2. Centrifuge the mixture ($5000\times g$, 5 min), remove the upper phase, and add another three volumes of MeOH.
3. Remix the sample and centrifuge ($16000\times g$, 10 min) to pellet the protein.
4. Dry the protein pellet at room temperature using a speed vac.
5. Dissolve in lysis buffer containing 7 M urea, 2 M thiourea, 4% CHAPS, and 30 mM Tris-base (*see* **Notes 5–8**).
6. Sonicate 1 min (pulse three times) using Sonicator.
7. Vortex and centrifuge ($16000\times g$, 10 min).
8. Remove nondissolved particles in the tube if present.

3.2.3 Protein
Quantification (2D Quant
Kit; GE Healthcare)

Standard Procedure

1. Set up six tubes and add standard solutions according to Table 2.
2. Prepare tubes containing 2 μ L of the sample in duplicate from step 2.
3. Next, add 500 μ L of precipitant to each tube (including the standard curve tubes). Vortex the tube briefly for 15 s and incubate for 2–3 min at room temperature.
4. Next, add 500 μ L of co-precipitant to each tube and mix by vortexing or inversion (15 s).
5. Centrifuge the tubes by keeping all tubes' cap hinge outward at $10,000\times g$ for 5 min.

Table 2

The procedure of preparing the calibration curve for protein quantification based on a standard BSA.

Reagents	Tube number					
	1	2	3	4	5	6
BSA (2 mg/mL) standard solution (μ L)	0	5	10	15	20	25
Protein quantity (μ g)	0	10	20	30	40	50

6. Remove the tubes from the centrifuge. A small pellet should be visible. Decant the supernatants. Rapidly proceed to the next step to avoid resuspension or dispersion of the pellets.
7. Reposition the tubes carefully in the microcentrifuge, with the cap-hinge and pellet facing outward. Centrifuge the tubes again to bring any remaining liquid to the bottom of the tube.
8. Using a micropipette, remove the remaining supernatant without disturbing the pellets. There should be no visible liquid remaining in the tubes.
9. Next, add 100 μL of copper solution and 400 μL of distilled or deionized water to each tube. Vortex the mixture briefly for 15 s to dissolve the precipitated protein.
10. Prepare an appropriate volume of working color reagent by mixing 100 parts of color reagent A with 1 part of color reagent B. Each assay requires 1 mL of working color reagent (*see* **Notes 9** and **10**).
11. Add 1 mL of working color reagent to each tube and mix by inversion.
12. Incubate the mixture at room temperature for 15–20 min.
13. Measure the absorbance (within 40 min of the addition of working color reagent) of each sample and standard at 480 nm using water as the reference.
14. Generate a standard curve by plotting the absorbance of the standards against the quantity of protein. Use this standard curve to determine the protein concentration of the samples.

*3.2.4 Two-Dimensional
Difference Gel
Electrophoresis (2D
DIGE)—Matrix-Assisted
Laser Desorption
Ionization–Time-of-Flight
Mass Spectrometry (MALDI
TOF)*

*Protein Labeling with
Cyanine Dyes*

*Dimensional Gel
Electrophoresis (1-DIGE)
(Isoelectric Focusing
Electrophoresis)*

1. Prepare stock rehydration buffer containing 7 M urea, 2 M thiourea, and 2% CHAPS (*see* **Note 11**).
 2. Prepare a working rehydration buffer (*see* **Note 12**).
 3. Take 50 μg of protein from each sample.
 4. Label each 50 μg of protein sample with 3 μL of Cy5 dye (CY3 and CY5) to differentiate different groups (CyDye™ DIGE Fluor dyes) and run on different gels.
 5. Add 1 μL of lysine to 50 μg of each sample.
 6. Make up the with working rehydration buffer (up to 450 μL).
 7. Pool equal quantities of all samples, label with Cy2, and utilize as an internal standard, standardized and matched across gels to eliminate gel-to-gel variance.
 8. In each gel, analyze two samples from two different groups.
1. Dilute 50 μg of protein (in μL) from each sample to 450 μL with a working rehydration buffer (*see* **Note 12**).
 2. Vortex well for 30 s.

3. Use a ceramic strip holder. Pipette the appropriate volume of rehydration solution into each holder. Deliver the solution at a central point in the strip holder channel away from the sample application wells. Remove any larger bubbles. Excess rehydration solution was not applied to ensure complete sample uptake.
4. Remove the protective cover from the IPG strip (note the barcode). Position it with the gel side down and the pointed (anodic) end of the strip directed toward the pointed end of the strip holder. Point the end first and lower the strip onto the solution. To help coat the entire strip, gently lift and lower the strip and slide it back and forth along the surface of the solution, tilting the strip holder slightly as needed to ensure complete and even wetting. Finally, lower the cathodic (square) end of the strip into the channel, ensuring that the IPG gel contacts the strip holder electrodes at each end. (The gel can be visually identified once the rehydration solution begins to dye the gel.) Take care not to trap bubbles under the strip.
5. Apply Immobiline Drystrip Cover Fluid to minimize evaporation and urea crystallization. Pipet the fluid dropwise into one end of the strip holder until one-half of the strip is covered. Then, pipet the fluid dropwise into the other end of the strip holder and add fluid until the entire IPG strip is covered.
6. Place the cover on the holder. Pressure blocks on the underside of the cover ensure that the strip maintains good contact with the electrodes as the gel rehydrates.
7. Rehydration can proceed on the benchtop or on the IPGphor unit platform. Ensure that the holder is on a level surface. A minimum of 10 h is required for rehydration; overnight rehydration is recommended. Alternatively, the rehydration period can be programmed as the first step of an IPGphor protocol.

IPG Strip Washing Step

1. Prepare the equilibration of the SDS stock solution (75 mM Tris, 6 M urea, 34.5% glycerol, 2% SDS) (*see* **Note 13**).
2. Then, divide 40 mL of the equilibrated SDS stock solution into two aliquots (20 mL each).
3. Add DTT (32 mM) (0.1 g) to 20 mL of equilibration SDS stock solution in the first tube and mix well until dissolved.
4. In the second tube, add IAA (0.9 g, 240 mM) to 20 mL of equilibration SDS stock solution and mix the solution well until it dissolves.
5. Transfer the equilibration SDS stock solution with 0.1 g of DTT to a 24 cm screw cap plastic tube. Place the IPG strips in this tube and shake for 20 min.

6. Discard the equilibration SDS stock solution with 0.1 g of DTT from the tube and add 20 mL of equilibration SDS stock solution with 0.9 g of IAA. Shake the tube with an IPG strip for 20 min.
7. Discard the 20 mL equilibration SDS stock solution and add 0.9 g of IAA.
8. Store the tube at -20°C until use.

Preparation of Two-Dimensional Gels

1. Prepare Monomer Stock solution (30% acrylamide and 0.8% N. N-methylenebisacrylamide) (*see* **Note 14**).
2. Prepare the 4X resolving gel buffer solution by weighing 181.7 g of Tris-base and placing it in a container. Add 750 milliliters of distilled water (dH_2O) to the Tris-base and stir until it completely dissolves. Next, adjust the pH of the solution to 8.8 using HCl, adding it drop by drop while stirring continuously. Use a pH meter or strips to check the pH level until it reaches 8.8. Once the pH is adjusted, pour enough distilled water to bring the total volume of the solution to 1000 milliliters. Stir the solution gently to ensure thorough mixing. Store the 4X resolving gel buffer solution in the refrigerator at 4°C until ready.
3. Following these steps, prepare the 2-D GEL preparation solution (six gels with a 12% concentration). Start by adding 250 mL of the monomer solution to 150 mL of 4x resolving gel buffer solution and 6 mL of 10% SDS solution [dissolve 1 gram in 10 mL of distilled water (dH_2O) to make 6 mL]. Then, add 190.46 mL of dH_2O to the mixture, bringing the total volume to 600 mL. Add 3.33 mL of 10% ammonium persulfate (501 k) solution (dissolve 0.5 grams in 5 mL of dH_2O). Just before gel stacking, add 166 μL of TEMED.
4. Prepare the butanol spray by mixing 40 mL pure butanol with 10 mL dH_2O
5. Gel stacking apparatus (for six gels).
6. Clean and dry the glass plates and spacers.
7. Assemble the *gel* plates by placing 1.5 mm spacers *vertically* between long and short spacers.
8. Position clamps on each side of the gel sandwich over the spacers and place them on the casting stand. Align the plates and spacers properly, then tighten the clamps and cams to obtain a leak-proof seal. The plates are adjusted so that they are level and vertical.
9. Pour the 2-D GEL preparation solution through the top of the gel plate. Use the gel plates to remove the bubbles. Fill the gel sandwich to 2 cm below the top of the plate.

10. Spray the prepared butanol on top of the gel.
11. Place the lid and allow the gels to polymerize overnight.
12. The following day, remove each gel plate and clean it properly. Place the wet tissues between the gel plates. Store the gel horizontally in closed boxes at 4 °C until use.

Two-Dimensional-Gel
Electrophoresis (2D-SDS-
PAGE)

1. Prepare the agarose sealing buffer by mixing 50 mL of 10x Tris-glycine SDS running buffer, 0.25 g of agarose, and 15–20 µL of bromophenol blue.
2. Use a heated stirrer with maximum heat until boiling and check. Distribute each 1 mL in a tube and store at 4 °C.
3. Prepare the 10X Tris-glycine SDS running buffer: 250 mM Trizma base, 1.92 M glycine, 1% SDS (*see Note 15*).
4. Prepare the 2X Tris-glycine SDS running buffer: dissolve 500 mL of 10X Tris-glycine SDS running buffer up to 5000 mL of dH₂O.
5. Prepare the 2X Tris-glycine SDS running buffer: dissolve 100 mL of 10X Tris-glycine SDS running buffer up to 1000 mL of dH₂O.
6. Remove the IPG strips at –20 °C. Allow the mixture to stand at room temperature for 30 min.
7. Remove the prepared gels from 4 °C and wipe with tissue to remove moisture.
8. Place the agarose sealing buffer in a heating mantle for melting.
9. Remove the IPG strips from the tube using tweezers, dip them in 2X Tris-glycine SDS running buffer, and place them at the top of the gel. Allow all sides of the IPG strip to touch the gel properly.
10. Add melted agarose sealing buffer to the IPG strips, which are then sealed.
11. Transfer the gels to a gel holder and cover with a lid.
12. The ¾ of the Ettan DALTsix electrophoresis unit with 2X Tris-glycine SDS running buffer.
13. Place the gel holder with gels into the Ettan DALTsix electrophoresis unit. Then, 1X Tris-glycine SDS Rewning buff to the outer layer until the maximum filling line reaches the electrophoresis unit.
14. Then, 2X Tris-glycine SDS Rewing Buffer is added to the inner layer until the gel reaches the max fill line in the electrophoresis unit.
15. Place the cover and switch on an Electrophoresis Power Supply—EPS 601.
16. Run the gel as detailed in Table 3 as needed Watt for a single gel.

Table 3
The parameters for running the gel

For single gel		
Step	Watt/gel	Time
P1	1	1 h
P2	2	5 min
	5	10 min
	10	10 min
	15	10 min
	17	4–5 h

- Switch off the electrophoresis power supply when the “dye front” reached the bottom. Remove the gels from the electrophoresis unit and remove the clamps.
- Scan the gels with a Sapphire biomolecular imager.
- Digitalize the samples with Sapphire capture system image analysis software.

**2D-DIGE Statistical
Analyses**

- Upload the 2D-DIGE gel images to Progenesis SameSpots software and analyze them using an automated spot detection method.
- Although automatic analysis is performed to detect all the spots across all gel images, each selected spot is verified and manually analyzed wherever necessary.
- Use Normalized volumes to identify the differentially expressed spots. Consider a cutoff ($p \leq 0.05$ and fold change ≥ 1.5) as significant.

**Preparatory Gel Preparation
and Colloidal Coomassie
Blue Staining**

- Prepare Colloidal Coomassie blue G-250 dye stock solution. 10% ammonium sulfate, 1% w/w phosphoric acid, 0.1% Coomassie blue G-250 up to 500 mL with dH₂O.
- Prepare the colloidal Coomassie blue G-250 dye working solution as shown in the table below. 80% Colloidal Coomassie blue G-250 stock solution, 20% MeOH up to 500 with dH₂O.
- Prepare the gel-fixing solutions according to the table below. Ethanol 40%, acetic acid 10% up to 400 mL with dH₂O.
- Obtain total protein (1 mg) from a pool of equal amounts of protein from all the samples.
- Separate the protein samples according to the first and second dimensions under the same conditions described in the DIGE section.
- Remove the gels from the electrophoresis unit and then remove the clamps.

7. Pry the glass plates apart with a spatula.
8. Place the gel in a gel-fixing solution for 1–3 h.
9. Discard the gel-fixing solution and wash the samples with dH₂O.
10. Change the water every 30 minutes.
11. Place the gel in colloidal Coomassie blue G-250 dye working solution and leave for 5 days with shaking.
12. Discard the solution and wash the cells with dH₂O.
13. Change the water every 30 minutes.
14. Store the gel at 4 °C until protein spot excision.

Protein Spot Excision

1. Place the gel over a white plastic sheet.
2. Analyze the gel images using an automated spot detection method with Progenesis SameSpots software.
3. Excise the differentially expressed spots using wide-mouth micropipette tips with a cutoff ratio of ≥ 1.5 -fold and p-value < 0.05 from the gel.
4. Transfer the spots to a 96-well plate filled with distilled water.

In-Gel Digestion

Note: The 96-well plate containing the gel spots will be used in this experiment:

1. Prepare a 1:1 volume from 100 mM ammonium hydrogen carbonate extra pure (AB) and ACN up to 20 mL. Remove all water from each well of the plate. Then, add 100 μ L of AB/ACN to the gel spots and place on a vortex incubator for 30 min at 400 rpm/25 °C. Repeat this process two times until the blue color disappears.
2. Remove (AB/ACN).
3. Then, add 100 μ L of ACN and vortex the mixture for 10 min at 400 rpm/25 °C.
4. Remove the ACN and allow the mixture to stand at room temperature for 5–10 min or until the gel dried (*see Note 16*).
5. Prepare trypsin by adding 1300 μ L of 50 mM AB to 20 μ L of trypsin (1 μ g/ μ L trypsin stored at –80 °C) (*see Note 17*). Then, add 15–20 μ L of each sample to the gel spots, which spin for 1 min and incubate for 1 h at 4 °C.
6. Remove the extra trypsin, add 10 μ L of 50 mM AB, spin the plate for 1 min, fix the plate properly to avoid evaporation, and incubate the plate overnight in a 37 °C incubator.
7. The next day, remove the plate, leave it at room temperature for a while, and then quickly spin it for 1 min.

8. Add 1 μL of 1% trifluoroacetic acid (TFA) (10 μL of TFA in 990 μL of dH_2O) to stop the reaction and to extract the peptides.
9. Spin the mixture for 1 min and vortexed for 1 h at 400 rpm/25 $^{\circ}\text{C}$.
10. Ready for MS/MS Analysis.
11. Spot 0.8 μL of samples onto a MALDI target (384 MTP Anchorchip) (800 μm Anchorchip), leave to dry or seal properly, and store at -20°C if used later or the next day.

3.2.5 MALDI TOF MS/MS Analysis

1. Obtain the spectra using an UltraflexTerm time-of-flight (TOF) mass spectrometer equipped with an IR-laser device at the reflector and detector voltages of 21 kV and 17 kV, respectively.
2. Calibrate the Peptide Mass Finger Printing (PMF) experiments against a standard peptide calibration standard II.
3. Assess the PMFs using Flex Analysis software (version 2.4).
4. Interpret the MS data using BioTools v3.2.
5. Search the peptide masses against the Mascot search.
6. Screen the identified proteins for Mascot scores higher than 56 and $p < 0.05$.

Bioinformatics Analysis

1. Construct the PPI network of differentially expressed proteins using the STRING database (<https://string-db.org/>) and use it to analyze protein interaction networks and the functions of the proteins differentially expressed in samples.
2. Employ Ingenuity pathway analysis (IPA) v9.0 to analyze the interaction of protein networks and the functions of proteins that were differentially expressed in samples.
3. Both databases allow UniProt IDs into the ingenuity knowledge base, the largest manually curated resource combining information from all published scientific studies.
4. This software aids in determining the functions and pathways that are most strongly associated with the MS-generated protein list by overlaying the experimental expression data onto networks constructed from published interactions.
5. Additionally, classify the identified proteins into different categories according to their function and location using the PANTHER classification system.
6. For the Gene Ontology (GO) enrichment analysis, classify proteins by GO annotation into three categories—biological process, cellular compartment, and molecular function—according to the UniProt-GOA database.

3.3 Label-Free Quantification Protocol (LC-MS/MS)

3.3.1 Sample Preparation

1. Remove plasma samples from -80°C .
2. Thaw the samples at $25-30^{\circ}\text{C}$.
3. Vortex the sample for 1–3 min.
4. Centrifuge the samples at $2000\times g$ for 2–3 min.
5. Transfer 100 μL of each sample to labeled Eppendorf tubes for further experiments.
6. Then, add 0.6 μL of protease inhibitor (25 $\times$) to 15 μL of sample (plasma/serum) (*see* **Notes 18** and **19**).
7. Carry out the depletion step (*see* Subheading **3.2.1**) (*see* **Notes 20** and **21**).
8. Determine the protein concentration (*see* Subheading **3.2.3**). The protein concentration should be 50 μg . Take volume accordingly.
9. The total protein concentration in each sample is 50 μg . Dry the samples completely on a speed vac. Store the samples at -20°C for further experiments.

3.3.2 In-Solution Protein Digestion

1. Add 10 μL of 6 M urea to 50 μg of protein (*see* **Notes 22** and **23**).
2. The reduction step is as follows: add (1 μL) 200 mM DTT, vortex, and then incubate the sample at 60°C for 30 min with shaking at 350 rpm (*see* **Note 24**).
3. Alkylation step: cool the sample (**step 11**) and add 1 μL of 400 mM (IAA) to a final concentration of 10 mM. Vortex the sample and then incubate at room temperature in the dark for 30 min (*see* **Note 25**).
4. pH stabilization: add 65 μL of 50 mM ammonium hydrogen carbonate (ABC) (*see* **Note 26**).
5. Trypsin addition step: add to each sample (**step 13**), 2.5 μL of trypsin (1 $\mu\text{g}/\text{mL}$) at a ratio of 1:20 (trypsin: protein) (*see* **Note 27**), and then incubate the sample overnight at 37°C with slow shaking.
6. RapiGest Removal Step: add to the peptide sample (**step 5**), 10% FA (7 μL) (to stop the reaction) and check the pH [**1**, **2**].

3.3.3 Peptides Desalting [Desalt the Peptides Using Pierce™ C18 Spin Columns (89873; 50 Spin Columns, Each Containing 8 mg of C18 resin)]

Solution Preparation

1. Activation solution: for each sample, 400 μL of activation solution (50% MeOH) is needed (*see* **Note 28**).
2. Equilibration solution: 400 μL of equilibration solution (0.5% TFA in 5% ACN) is needed for each sample (*see* **Note 29**).
3. Wash solution: for each sample, add 400 μL of wash solution (0.5% TFA in 5% ACN). (*see* **Note 30**).
4. Elution buffer: for each sample, 40 μL of elution buffer (70% ACN + 0.1% FA/TFA) is needed (*see* **Notes 31** and **32**).

For sample preparation, collect 3 μL of plasma sample and add 1 μL of sample buffer (see **Notes 33** and **34**).

Protocol

1. Trap the column to settle the resin. Remove the top and bottom cap. Then, place the column into a receiver tube.
2. Then, add 200 μL of activation solution to rinse the walls of the spin column and to wet the resin.
3. Centrifuge at $1500\times g$ for 1 min. Discard the flow-through. Repeat **steps 2** and **3**.
4. Next, add 200 μL of equilibration solution. Centrifuge at $1500\times g$ for 1 min. Discard the flow-through. Repeat **step 4**.
5. Load the prepared sample on top of the resin bed.
6. Place the column into a receiver tube. Centrifuge at $1500\times g$ for 1 min.
7. To ensure complete binding, recover the flow-through. Repeat **steps 5** and **6**.
8. Place the column into a receiver tube. Add 200 μL of wash solution to the column, then centrifuge at $1500\times g$ for 1 min. Discard the flow-through. Repeat **step 8**.
9. Place the column in a new receiver tube. Then, add 20 μL of elution buffer to the top of the resin bed. Centrifuge at $1500\times g$ for 1 min. Repeat **step 9**. Obtain a total of 40 μL of sample.
10. Transfer 40 μL to new Eppendorf tubes (500 μL). Air-dry and store all the samples at -20°C . Each tube contained $\sim 50\ \mu\text{g}$ of protein.
11. Pierce Quantitative Colorimetric Peptide Assay, 500x.
12. Gently dry the sample in a vacuum evaporator. For LC-MS/MS analysis, suspend the sample carefully in 1–2 μL of matrix solution prepared just before use. For LC-ESI applications, suspend the sample in 0.1% FA or the appropriate buffer.

3.3.4 Liquid Chromatography-Coupled to Tandem Mass Spectrometry (LC-MS/MS)

1. The peptides will be dissolved in 0.1% (v/v) FA and then separated by a Dionex UltiMate 3000 nano LC system, with a WPS-3000 autosampler.
2. The peptide will be loaded onto a trapping column (C18 PepMap100, 5 μm , 100A) coupled to the analytical column (PepMap™ C18, 50 $\text{cm} \times 75\ \mu\text{m}$).
3. This trapping nano LC setting achieves high-capacity sample loading, Peptide separation will be carried out at 300 nL/min and delivery Mobile Phase A consists of 0.1% FA and 2% ACN, and mobile phase B consists of 0.1% FA and 90% ACN.
4. After separation, the peptides will have injected into the NSI source for ionization and then analyzed by an Q Exactive Plus Hybrid Quadrupole-Orbitrap mass spectrometer operating in positive ion mode, with nano electrospray (nESI) potential at 1800 V with a maximal duty cycle of 3 s.

5. The scanning range for primary MS will be set to 350–1550 m/z , and the scanning resolution will be set to 70,000; the fixed starting point of the scanning range for secondary MS will be 100 m/z , and the scanning resolution for secondary MS will be set to 15,000. Data acquisition mode will use a data-dependent scanning (DIA).

3.3.5 Data Analysis

1. Data analysis by using Proteome Discoverer 1.3, using Sequest as search engine, and the HUMAN-refprot-isoforms.fasta as sequence database.
2. Parameters applied for the analysis will be the following: The search parameters included
 - (a) Pre-cursor/fragment mass tolerance: 15 ppm/0.02 Da
 - (b) Maximal missed cleavages: 2
 - (c) Enzyme name: trypsin (full)
 - (d) Dynamic modifications: peptide N-terminal acetylation, methionine oxidation
 - (e) Static modification: cysteine carbamidomethylation
 1. Peptide-spectrum match (PSM) filtering.
 2. The protein and peptide FDRs will be set at 1, with a minimum of two unique peptides per protein. The filtered PSM list will be exported and manually formatted in Excel.

3.3.6 Bioinformatics Analysis

For bioinformatics analysis see Subheading [3.2.5.1](#).

4 Notes

1. All reagents should be labeled with the reagent name, concentration, preparation and expiration dates, storage conditions, and initials of the preparing person. People unfamiliar with the hazards associated with the following chemicals should read the relevant material safety data sheets.
2. Before starting a run, ensure that the LC-MS is calibrated.
3. Statistical analysis highly depends on the study design and the aim of the project.
4. The filtrate obtained is ~300 μL . The depleted samples (300 μL) will undergo a speed vac. Reduce the volume to 70 μL . This can be stored for later use. The depleted sample will be in 10 mM PBS, 0.15 M NaCl, and 0.02% azide, pH 7.4.
5. Prepare 1.5 M Tris-base pH 8.8: dissolve 181.7 g of Tris-base, and 27 mL of 37% HCL (to adjust pH = 8.8) in up to 1000 mL dH_2O .

6. $g = V(L) * M.W * M$ (molarity).
7. To prepare lysis buffer: dissolve 10.5 g of urea, 3.8 g of thio-urea, 1 g of CHAPS, and 0.5 mL of Tris-base in up to 25 mL dH₂O.
8. Protease inhibitor mix: 10 μ L of 100x Protease inhibitor mix added to 1 mL lysis buffer if required later.
9. Then, add 50 mL of color reagent (49.5 mL of A and 500 μ L of B).
10. While calculating the amount of color reagent, you should include the number of standards.
11. To prepare 25 mL stock rehydration buffer: Dissolve 10.5 g of urea, 3.8 g of thiourea, and 0.5 g of CHAPS in up to 25 mL dH₂O. Distribute 1 mL in Eppendorf tubes and store it at -20°C .
12. To prepare a working rehydration buffer: mix 1 mL of stock rehydration buffer, 5 μ L of IPG buffer (pH 3–11), 2 μ L of bromophenol blue, and 0.006 gm of DTT (40 mM).
13. To prepare equilibration SDS stock solution: Dissolve 20 mL of Tris (1.5 M pH 8.8), 144.14 g of urea, 8 g of SDS, 10–13 μ L of bromophenol blue, 138 mL of glycerol (add to the end), fill it with up to 400 mL of dH₂O. Distribute each 40 mL in a tube and store it at -20°C .
14. To prepare Monomer STOCK solution: dissolve 300 g of acrylamide and 8 g of *N,N*-methylenebisacrylamide in up to 1000 mL dH₂O. Store at 4°C .
15. Prepare the 10X Tris glycine SDS RUNNING BUFFER: Dissolve 30.28 g of Trizma base (Tris), 144.1 g of glycine, 10 g of SDS, in up to 1000 mL of dH₂O. Keep at room temperature.
16. At this point, the gel will look like gelatin and be ready to absorb the trypsin and digest the proteins.
17. To prepare 50 mM AB, dissolve 0.039 g AB in 10 mL dH₂O.
18. Add 4 μ L of 25X protease inhibitor to 100 μ L of sample (plasma/serum). Adjust the volume of the protease inhibitor according to the sample volume.
19. Preparation of working 25X (200 μ L) protease inhibitors from 100X stock protease inhibitors (GE 28955972): dilute 50 μ L of 100X PI in 150 μ L of dH₂O.
20. The concentration of serum/plasma proteins can vary widely depending on the origin of the sample. Adjust each 810 μ L aliquot of antibody resin slurry to 10 μ L (600 μ g) of serum or plasma. For the best results, optimize the sample to slurry volume ratio for each specific application.

21. We optimized the method to use 15 μL of plasma for the experiment.
22. To prepare 6 M urea, dissolve 3.6 g of urea in 10 mL of dH_2O .
23. 6 M urea disrupts hydrogen bonds and hydrophobic interactions.
24. To prepare 200 mM DTT, 0.030 g in 1 mL of dH_2O .
25. To prepare 400 mM IAA, dissolve 0.074 gm in 1 mL of dH_2O .
26. To prepare 50 mM ABC, dissolve 0.039 gm in 10 mL of dH_2O .
27. $\rightarrow\text{X}$ (volume of trypsin required to be added to the sample):
 $50\text{ }\mu\text{g (our sample)} \rightarrow x = 50\text{ }\mu\text{g}/20 = 2.5\text{ }\mu\text{g}.$
28. To prepare 50% MeOH, add 25 mL of MeOH to 25 mL of dH_2O .
29. To prepare the equilibration solution (50 mL), add 250 μL of TFA to 2.5 mL of ACN to a final volume of 50 mL.
30. To prepare the wash solution (50 mL), add 250 μL of TFA to 2.5 mL of ACN to a final volume of 50 mL.
31. To prepare the elution buffer (10 mL), add 5 mL of ACN and 10 μL of FA, and adjust the volume to 10 mL with water.
32. For ESI-MS analysis, replace TFA with 0.1% FA for the best results.
33. To prepare 2000 μL of sample buffer, add 40 μL of TFA in 400 μL of ACN to 2000 μL of dH_2O .
34. From **step 15**, add 79.5 μL of total sample to 26.5 μL of sample buffer for each sample.

References

1. Boycott KM, Hartley T, Biesecker LG, Gibbs RA, Innes AM, Riess O et al (2019) A diagnosis for all rare genetic diseases: the horizon and the next frontiers. *Cell* 177(1):32–37
2. Jackson M, Marks L, May GHW, Wilson JB (2018) The genetic basis of disease. *Essays Biochem* 62(5):643–723
3. Al-Qahtani W, Abdel Jabar M, Masood A, Jacob M, Nizami I, Dasouki M et al (2020) Dried blood spot-based metabolomic profiling in adults with cystic fibrosis. *J Proteome Res* 19(6):2346–2357
4. Gloerich J, Wevers RA, Smeitink JAM, van Engelen BG, van den Heuvel LP (2007) Proteomics approaches to study genetic and metabolic disorders. *J Proteome Res* 6(2):506–512
5. Graw S, Chappell K, Washam CL, Gies A, Bird J, Robeson MS et al (2021) Multi-omics data integration considerations and study design for biological systems and disease. *Molecular omics* 17(2):170–185
6. Smirnov D, Konstantinovskiy N, Prokisch H (2023) Integrative omics approaches to advance rare disease diagnostics. *J Inherit Metab Dis* 46(5):824–838
7. Sebaa R, Harper M-E, Al-Tassan R, Al-Owain M, Abdel Rahman AM (2023) Metabolomics in the study of human mitochondrial diseases. In: Abdel Rahman AM (ed) *Clinical metabolomics applications in genetic diseases*. Springer Nature Singapore, Singapore, pp 147–171
8. Aebersold R, Mann M (2016) Mass-spectrometric exploration of proteome structure and function. *Nature* 537(7620):347–355
9. Nimer RM, Sumaily KM, Almuslat A, Abdel Jabar M, Sabi EM, Al-Muhaizea MA et al (2022) Dystrophin protein quantification as a

- duchenne muscular dystrophy diagnostic biomarker in dried blood spots using multiple reaction monitoring tandem mass spectrometry: a preliminary study. *Molecules* 27(12)
10. Sumaily KM, Nimer R, Alzahrani M, Abdel Jabar M, Alodib A, Sabi EM et al (2022) CFTR protein quantification as a cystic fibrosis diagnostic biomarker in dried blood spots using multiple reaction monitoring tandem mass spectrometry. *J Pharm Biomed Anal* 216:114801
 11. Johnson CH, Ivanisevic J, Siuzdak G (2016) Metabolomics: beyond biomarkers and towards mechanisms. *Nat Rev Mol Cell Biol* 17(7):451–459
 12. Jaber MA, Benabdelkamel H, Dahabiyeh LA, Masood A, AlMalki RH, Musambil M et al (2022) The metabolomics approach revealed a distinctive metabolomics pattern associated with hyperthyroidism treatment. *Front Endocrinol* 13:1050201
 13. Dahabiyeh LA, Nimer RM, Sumaily KM, Alabdajabar MS, Jacob M, Sabi EM et al (2023) Metabolomics profiling distinctively identified end-stage renal disease patients from chronic kidney disease patients. *Sci Rep* 13(1):6161
 14. Alodaib AN, Nimer RM, Alhumaidy R, Alhenaky A, Abdel Jabar M, AlMalki RH et al (2023) Biomarker discovery in galactosemia: metabolomics with UPLC/HRMS in dried blood spots. *Front Mol Biosci*:10
 15. Masood A, Malkawi A, Abdel Rahman AM, Siaj M (2023) Metabolomics of rare endocrine, genetic disease: a focus on the pituitary gland. In: Abdel Rahman AM (ed) *Clinical metabolomics applications in genetic diseases*. Springer Nature Singapore, Singapore, pp 173–187
 16. Pinu FR, Beale DJ, Paten AM, Kouremenos K, Swarup S, Schirra HJ et al (2019) Systems biology and multi-omics integration: viewpoints from the metabolomics research community. *Meta* 9(4):76
 17. Nice EC (2016) From proteomics to personalized medicine: the road ahead. *Expert Rev Proteomics* 13(4):341–343
 18. Dahabiyeh LA, Mahmoud NN, Al-Natour MA, Safo L, Kim DH, Khalil EA et al (2021) Phospholipid-gold nanorods induce energy crisis in MCF-7 cells: cytotoxicity evaluation using LC-MS-based metabolomics approach. *Biomol Ther* 11(3)
 19. Sabi EM, AlMogren M, Sebaa R, Sumaily KM, AlMalki R, Mujammami AH et al (2024) Comprehensive metabolomics analysis reveals novel biomarkers and pathways in falsely suspected glutaric aciduria Type-1 newborns. *Clin Chim Acta* 557:117861
 20. Manor J, Elsea SH (2023) Untargeted metabolomics in newborn screening. In: Abdel Rahman AM (ed) *Clinical metabolomics applications in genetic diseases*. Springer Nature Singapore, Singapore, pp 97–116
 21. Dahabiyeh LA, Nimer RM, Rashed M, Wells JD, Fiehn O (2023) Serum-based lipid panels for diagnosis of idiopathic Parkinson's disease. *Meta* 13(9):990
 22. Jacob M, Gu X, Luo X, Al-Mousa H, Arnaout R, Al-Saud B et al (2019) Metabolomics distinguishes DOCK8 deficiency from atopic dermatitis: towards a biomarker discovery. *Meta* 9(11):274
 23. Aleidi SM, Alnehmi EA, Alshaker M, Masood A, Benabdelkamel H, Al-Ansari MM et al (2021) A distinctive human metabolomics alteration associated with osteopenic and osteoporotic patients. *Meta* 11(9)
 24. Dahabiyeh LA, Nimer RM (2023) Metabolomics: a pipeline for biomarker discovery in genetic diseases. In: Abdel Rahman AM (ed) *Clinical metabolomics applications in genetic diseases*. Springer Nature Singapore, Singapore, pp 43–69
 25. Nimer RM, Alfaqih MA, Shehabat ER, Mujammami M, Abdel Rahman AM (2023) Label-free quantitative proteomics analysis for type 2 diabetes mellitus early diagnostic marker discovery using data-independent acquisition mass spectrometry (DIA-MS). *Sci Rep* 13(1):20880
 26. Pasquali M, Serchi T, Planchon S, Renaut J (2017) 2D-DIGE in proteomics. *Methods Mol Biol* 1654:245–254
 27. Jacob M, Masood A, Shinwari Z, Abdel Jabbar M, Al-Mousa H, Arnaout R et al (2021) Proteomics profiling to distinguish DOCK8 deficiency from atopic dermatitis. *Front Allergy* 2:774902
 28. Nimer RM, Abdel Rahman AM (2023) Recent advances in proteomic-based diagnostics of cystic fibrosis. *Expert Rev Proteomics* 20(7–9):151–169
 29. Sebaa R, AlMalki RH, Alseraty W, Abdel Rahman AM (2023) A distinctive metabolomics profile and potential biomarkers for very long Acylcarnitine dehydrogenase deficiency (VLCADD) diagnosis in newborns. *Meta* 13(6):725
 30. Sebaa R, AlMogren M, Alseraty W, Abdel Rahman AM (2023) Untargeted metabolomics identifies biomarkers for MCADD neonates in dried blood spots. *Int J Mol Sci* 24(11)

31. Jaber MA, Benabdelkamel H, Dahabiyeh LA, Masood A, AlMalki RH, Musambil M et al (2022) The metabolomics approach revealed a distinctive metabolomics pattern associated with hyperthyroidism treatment. *Front Endocrinol (Lausanne)* 13:1050201
32. Benabdelkamel H, Jaber MA, Dahabiyeh LA, Masood A, AlMalki RH, Musambil M et al (2023) Metabolomic profile of patients on levothyroxine treatment for hypothyroidism. *Eur Thyroid J* 12(4)
33. AlMalki RH, Jaber MA, Al-Ansari MM, Sumaily KM, Al-Alwan M, Sabi EM et al (2023) Metabolic alteration of MCF-7 cells upon indirect exposure to E. Coli Secretome: a model of studying the microbiota effect on Human breast tissue. *Meta* 13(8)
34. AlMalki RH, Sebaa R, Al-Ansari MM, Al-Alwan M, Alwehaibi MA, Rahman AMA (2023) E. coli secretome metabolically modulates MDA-MB-231 breast cancer cells' energy metabolism. *Int J Mol Sci* 24(4):4219



Chapter 3

Activity-Based Protein Profiling of Serine Hydrolases in Bacteria: Methods and Protocols

Md Jalal Uddin, Jeanette S. Grunnvåg, and Christian S. Lentz

Abstract

Activity-based protein profiling (ABPP) is a chemoproteomic technique that uses functionalized covalent enzyme inhibitors known as activity-based probes (ABPs) to detect, quantify, and identify active enzymes in a biological sample. Here, we describe experimental protocols for profiling of enzyme activities in bacterial cultures using both fluorescent and biotinylated ABPs that allow for the gel-based detection and resolution of target enzymes, or their enrichment and mass spectrometry-based identification, respectively. These protocols have been optimized for profiling of serine hydrolases using fluorophosphonate probes in both Gram-positive and Gram-negative bacteria, and they can easily be adapted for the application of probes targeting other enzyme families of interest in bacteria under diverse growth conditions.

Key words Activity-based probe, Serine hydrolases, Bacteria, Chemoproteomics, Fluorophosphonate

Abbreviations

ABP	Activity-based probe
ABPP	Activity-based protein profiling
Ambic	Ammonium bicarbonate
BHI broth/agar	Brain–Heart Infusion broth/agar
Bis-Tris	Bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane
Cy3	Cyanine 3
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
FP	Fluorophosphonate
IAA	Iodoacetamide
LC	Liquid chromatography
MS	Mass spectrometry
OD	Optical density
PBS	Phosphate-buffered saline

Md Jalal Uddin and Jeanette S. Grunnvåg contributed equally with all other contributors.

PMSF	Phenylmethylsulfonyl fluoride
PMT	Photomultiplier
rpm	Rotations per minute
RT	Room temperature
SDS-PAGE	Sodium dodecylsulfate polyacrylamide gel-electrophoresis
TEAB	Triethylammonium bicarbonate
TMR	Tetramethylrhodamine
Tris	Tris(hydroxymethyl)aminomethane

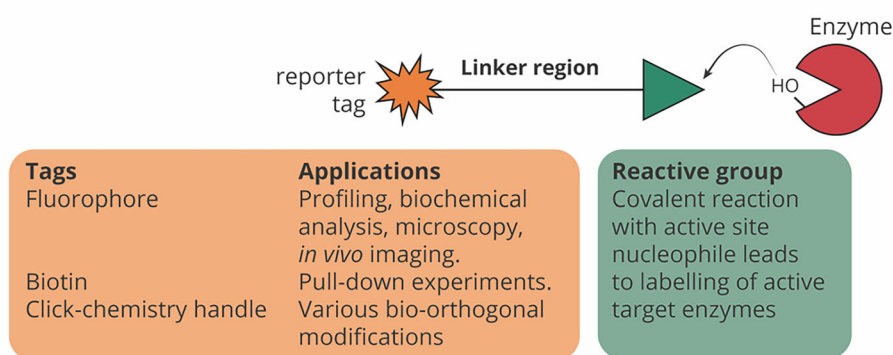
1 Introduction

Activity-based protein profiling (ABPP) is a chemoproteomic technique that employs specialized small molecules, termed activity-based probes (ABPs), for the detection, quantification, and identification of active species of a particular enzyme class within a biological sample. ABPs are functionalized covalent enzyme inhibitors that are typically comprised of three main parts (Fig. 1a). (1) A reactive electrophile (also referred to as warhead) that leads to covalent, irreversible attachment to the active site nucleophile of a target enzyme. (2) A reporter tag such as a fluorophore, biotin tag, or click-chemistry handle depending on the downstream application of choice. (3) A linker region connecting electrophile and reporter tag that also contributes to the target selectivity of the probe.

Fluorescent ABPs are commonly used in gel-based ABPP (Fig. 1b), where labeled target enzymes within proteomes are separated according to size by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by fluorescence scanning. This method allows for rapid and efficient detection of the most abundant targets and is well suited for comparative assessment of activity profiles across diverse experimental conditions or in different strains. ABPP can also be performed in a competitive setup to identify inhibitors or validate enzyme–inhibitor interactions. In competitive gel-based ABPP, active-site-directed binding of unlabeled inhibitors will prevent binding of the fluorescent ABP in a competitive way, which can be detected as a decrease in fluorescent intensity of the target band(s). This allows to screen for or validate potential inhibitors against multiple enzyme targets in parallel in live bacteria, without the need for recombination and expression of purified enzymes [1–4].

For the identification of enzymes, however, mass spectrometry-based approaches are needed. These techniques typically rely on the use of ABPs functionalized with biotin or desthiobiotin tags that allow for selective enrichment of target enzymes via streptavidin beads. Post-enrichment, the target enzymes can be processed and identified by liquid chromatography–tandem mass spectrometry

a)



b)

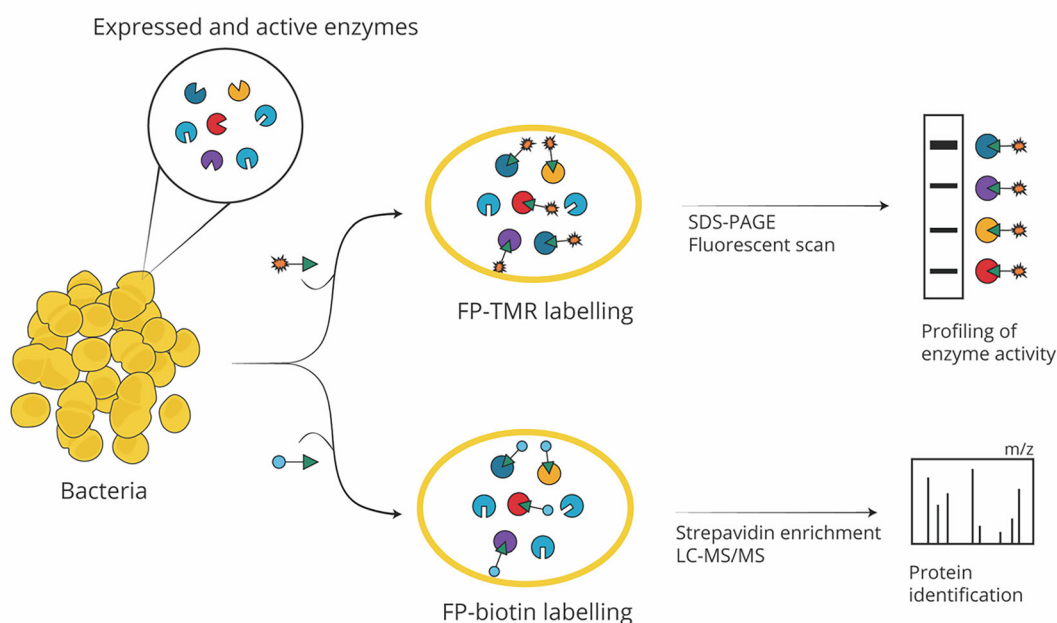


Fig. 1 ABPP of serine hydrolases in bacteria. **(a)** Overview of the structure of an activity-based probe (ABP), with a reporter tag indicated in orange, linker region and the reactive group (green) that binds to the active site of serine hydrolases. **(b)** Schematic overview of the workflow of ABPP to target expressed and active enzymes in any bacterial species, either for the purpose of profiling using a fluorescent probe like FP-TMR, or for identification of protein targets using FP-desthiobiotin

(LC-MS/MS) (Fig. 1b). Alternatively, probes with small “click-chemistry” handles such as azide or alkyne groups can be used in both gel-based and mass spectrometry approaches in a two-step protocol. In the first step, these probes are added to the biological sample to label target enzymes, whereas a suitable detection or enrichment tag is attached subsequently.

Over the last two decades, probes have been developed to target a broad variety of enzyme classes such as glycoside hydrolases [5, 6], proteases [7], serine hydrolases [8], or oxido-reductases [9]. One of the most widely used groups of ABPs is fluorophosphonates, which target the family of serine hydrolases [10, 11] and has been applied on diverse biological samples, including mammalian cell lines and tissues [10, 12, 13], fungi [14], parasites [15, 16], and archaea [17]. In bacteria, fluorophosphonate (FP) probes have been used in recent years to profile enzyme activities both in bacterial culture samples [1, 18–21, 3], at the host–pathogen interface in cecal fluid [22] and in the gut microbiome [23].

Here, we present our routine protocols for gel-based and mass spectrometry-based ABPP in live bacterial cultures that were established for profiling of serine hydrolases using FP probes. We have used these protocols successfully for both Gram-positive (*Enterococcus faecium*, *Staphylococcus aureus*) and Gram-negative bacteria (*Klebsiella pneumoniae*) as well as with probes targeting other enzyme families [24, 25]. The described protocols for gel-based ABPP include labeling of bacterial cultures with ABPs lysis of bacterial samples, SDS-PAGE analysis, and competitive ABPP in the presence of unlabeled inhibitors. Protocols for MS-based ABPP include labeling of bacterial cultures, cell lysis, enrichment of biotinylated targets using streptavidin beads, and sample preparation for MS analysis. These protocols may serve as a blueprint that can easily be adapted for diverse combinations of bacteria and ABPs of interest.

2 Materials

2.1 Reagents for Labeling of Bacterial Cultures

1. Bacterial liquid or solid culture media [e.g., brain–heart infusion (BHI) agar and BHI broth] suitable for the cultivation of bacteria of interest.
2. Inoculation loops (1 μ L and 10 μ L).
3. 37 °C incubator for overnight cultivation.
4. 1.5 mL Eppendorf tubes.
5. Thermo Scientific™ ActivX™ TAMRA-FP Serine Hydrolase Probe (FP-TMR) or Thermo Scientific™ ActivX™ Desthiobiotin-FP Serine Hydrolase (FP-desthiobiotin).
6. Phenylmethylsulfonyl fluoride (PMSF) or other nonfluorescent inhibitors for competitive ABPP.
7. Dimethyl sulfoxide (DMSO).
8. Thermomixer.
9. BioSpec products 0.1 mm glass beads (BioSpec Products).
10. Heat block.

11. Centrifuge with rotor for 1.5 mL tubes.
12. Precellys[®] Evolution homogenizer (Bertin Technologies).

2.2 SDS-PAGE

Analysis and

Fluorescence

Scanning

1. 4x Lysis/SDS-PAGE loading buffer (40% glycerol, 240 mM Tris/HCl at pH 6.8, 8% sodium dodecyl sulfate (SDS), 0.04% bromophenol blue, and 5% beta-mercaptoethanol).
2. Invitrogen[™] NuPAGE[®] Novex[®] Gel System (Thermo Fisher Scientific).
3. Invitrogen[™] NuPAGE[®] Bis-Tris gradient gel (Thermo Fisher Scientific).
4. Invitrogen[™] NuPAGE[®] MES SDS Running Buffer (20X).
5. Amersham ECL Plex rainbow marker.
6. Power supply.
7. Amersham Typhoon[™] 5 scanner (Cytiva) or another fluorescent scanner.
8. Invitrogen[™] SimplyBlue[™] SafeStain (Thermo Fisher Scientific).

2.3 Streptavidin

Enrichment of

Biotinylated Targets

2.3.1 Reagents and Solutions

1. SDS solution.
2. Protease Inhibitor Cocktail Powder (Sigma, P2714) for target enrichment and MS-based ABPP.
3. Triton-X 100 (Sigma, T8787).
4. Disposable PD-10 desalting columns (VWR) (optional).
5. Phosphate-buffered saline (PBS).
6. BCA protein assay kit (Bio-Rad) or alternative kit for measuring of protein concentrations.
7. 1 M KCl.
8. 0.1 M Na₂CO₃ (this solution should be prepared freshly before use).
9. 2 M urea in 10 mM Tris-HCl (pH 8.0) (this solution should be prepared freshly before use).
10. 4 M Urea Wash Buffer: 4 M urea in 50 mM Tris-HCl (pH 7.5). Alternatively use 4 M urea in 50 mM ammonium bicarbonate (Ambic), or 4 M urea in 50 mM triethylammonium bicarbonate (TEAB) (this solution should be prepared freshly before use).
11. 50 mM Tris-HCl (pH 7.5) or 50 mM Ambic or 50 mM TEAB.
12. 1 mM CaCl₂.
13. RIPA lysis buffer [50 mM Tris-HCl, 150 mM NaCl, 0.1% (wt/vol) SDS, 0.5% (wt/vol) sodium deoxycholate, 1% (vol/vol) Triton X-100] in dH₂O and adjusted to pH to 7.5. This solution can be stored at 4 °C for many months.

14. TPM buffer (50 mM Tris-HCl, 1 mM potassium phosphate, 8 mM MgSO₄) in dH₂O and adjust the pH to 7.6. Can be stored for several months at room temperature (RT).
15. Dithiothreitol (DTT).
16. Iodoacetamide (IAA).
17. Trypsin solution (0.5 mg/mL), Sequencing Grade Modified Trypsin dissolved in Trypsin Resuspension Buffer.
18. Formic acid.

2.3.2 Materials and Instruments

1. 5 mL Eppendorf and Lo-Bind Eppendorf tubes.
2. Rotator for microtubes.
3. Pierce[™] Streptavidin Magnetic Beads.
4. Magnetic rack for use with 1.5 mL microtubes.
5. OMIX C18 tip (A5700310, Varian).
6. Nanodrop spectrophotometer (Nanodrop ONE, Thermo Fisher Scientific).
7. Vacuum concentrator.
8. EASY-nLC 1200 system (by Thermo Fisher Scientific) linked to either an Exploris 480 Orbitrap or an Orbitrap Fusion Lumos mass spectrometer.

3 Methods

If not stated otherwise, perform all steps at RT.

3.1 Gel-Based Analysis of Bacterial Serine Hydrolase Activities Using the Fluorescent Probe FP-TMR

3.1.1 Probe Labeling of Bacterial Cultures and Cell Lysis

1. Streak out bacterial strain from frozen permanent culture onto the preferred culture media agar plates (*see Note 1*). Place for ~16–18 h at 37 °C.
2. Harvest bacteria from the plate with a 10 µL inoculation loop and streak it onto a fresh agar plate for lawn growth (*see Note 2*). Place for ~16–18 h at 37 °C.
3. Harvest bacteria from plate with an inoculation loop and make a homogeneous bacterial suspension in liquid culture media (vortex well). Adjust to an OD₆₀₀ ≈ 20 (*see Note 3*).
4. Aliquot 150 µL (*see Note 4*) of the bacterial suspension to Eppendorf tubes and add 1.5 µL of the FP-TMR solution at 100 µM (in DMSO) for a final concentration of 1 µM (*see Note 5*).
5. Incubate for 30–60 min at 37 °C, 300 rpm in a thermomixer (*see Note 6*).
6. Transfer samples to screw-top tubes containing ~100 µL 0.1 mm glass beads and 50 µL 4X Lysis/SDS-PAGE loading

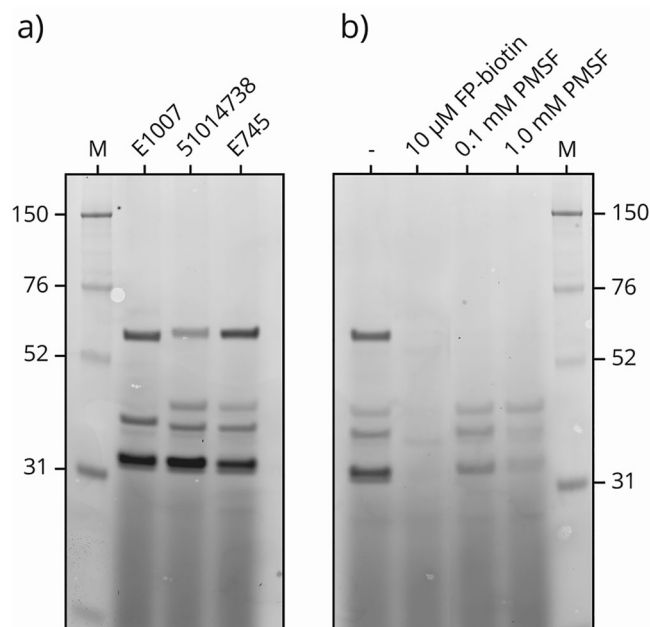


Fig. 2 Profiling of serine hydrolases in Gram-positive *Enterococci*. **(a)** Enzyme activity profiles of targeted serine hydrolases by the ABP FP-TMR in three strains; E1007 belonging to *Enterococcus lactis*, 51014738, which is a commensal *Enterococcus faecium*, and E745, a vancomycin-resistant *E. faecium* clinical strain. **(b)** Competitive gel-based ABPP in *E. faecium* E745; samples were first treated with the inhibitor, either FP-desthiobiotin (FP-biotin) (final concentration 10 μ M), PMSF (final concentrations 0.1 and 1 mM), or negative control (–)(DMSO) and incubated for 45 min, followed by labeling with FP-TMR (final concentration 1 μ M) in all samples for 45 min. The images show gels scanned for fluorescence in the Cy3 channel after SDS-PAGE analysis

buffer. Lyse by bead-beating (3×30 s at full speed with 1 min in-between).

7. Centrifuge samples at 4000–5000 g for 5 min at 4 °C.
8. Boil samples at 95 °C in a heating block for 10 min. The samples can be stored at –20 °C or loaded directly on a gel after cooling down (see Subheading 3.1.3). (See Fig. 2a for exemplary results of gel-based profiling of serine hydrolase activities across three *Enterococcus* strains.)

3.1.2 Competitive ABPP with Unlabeled Inhibitors

1. Grow bacteria either on plate or in liquid culture by following steps 1–3 from Subheading 3.1.1.
2. Aliquot 150 μ L of the bacterial suspension to Eppendorf tubes.
3. Add 1.5 μ L of PMSF stock solution at 1 mM, 100 μ M, and 10 μ M concentration (in DMSO) or alternative unlabeled inhibitor (see Note 7). Incubate for 30–60 min at 37 °C, 300 rpm in a thermomixer (see Note 6).

4. Add 1.5 μL of the probe solution at 100 μM (in DMSO). Incubate for 30–60 min at 37 °C, 300 rpm in a thermomixer (*see* **Note 6**).
5. Follow protocol **steps 6–8** from Subheading 3.1.1. (*See* Fig. 2b for exemplary results of a competitive gel-based ABPP experiment.)

3.1.3 SDS-PAGE and Fluorescence Scan Analysis

1. Use a precast gradient gel like NuPAGE 4–12% Bis-Tris gel (or cast one, *see* **Note 8**).
2. Assemble the gel chamber and fill to cover the gel(s) with 1X NuPAGE MES SDS running buffer (precast NuPAGE gradient gel). Remove the comb.
3. The gel is ready for loading samples. Load 3 μL Amersham ECL Rainbow Marker (ladder), then load 20 μL per sample in the wells (*see* **Note 9**).
4. Run gel at 120 V for 1 h 50 min (for pre-cast mini gels) (*see* **Note 10**).
5. Analyze the gel on the Typhoon 5 Scanner in the Cy3 channel to detect FP-TMR-labeled bands (most often PMT 700) (Fig. 2).
6. To control for loading of the gel, carefully place the gel in a tray with MilliQ water and continue with the staining protocol for Invitrogen[™] SimplyBlue[™] SafeStain provided by the manufacturer (*see* **Note 11**). For best results, destain overnight.

3.2 MS-Based ABPP Using FP-Desthiobiotin

This protocol describes the enrichment of bacterial serine hydrolases labeled with a (desthio)biotin-ABP using magnetic streptavidin beads and subsequent sample preparation for LC-MS/MS-based detection. To distinguish between specific ABP-labeled and nonspecifically enriched background proteins, it is essential to include a negative control (vehicle-treated instead of ABP-treated) and prepare all samples as three biological replicates allowing for a quantitative comparison of the two datasets (*see* **Note 12**).

3.2.1 Labeling of Bacterial Cells with FP-Desthiobiotin for Target Enrichment

1. Grow bacteria in appropriate culture media using three independent biological culture replicates.
2. Determine the optical density (OD) at 600 nm (OD₆₀₀) of the bacterial cultures using a spectrophotometer and adjust to an OD₆₀₀ of ≈ 2.0 .
3. Add 10 μL FP-desthiobiotin probe (200–300 μM stock in DMSO) or an equal volume of DMSO (negative control) to 1 mL of bacterial culture (2–3 μM final concentration) and incubate at 37 °C for 60 min at 37 °C in a rotator (or alternatively shaking at 700 rpm in a thermomixer) (*see* **Note 13**).

4. Centrifuge the samples at $4,500\times g$ for 5 min at 4 °C and remove the supernatant.
5. Optional: wash cell pellet in 1 mL of TPM buffer (*see* **Note 14**).
6. Resuspend the pellets in 1.2 mL RIPA lysis buffer and transfer the samples to a 2 mL screw-cap tube filled with ca. 200 μ L of 0.1 mm glass beads.
7. Optional: add 1/10 volume (120 of μ L) of Protease Inhibitor cocktail (*see* **Note 15**).
8. Place the samples on ice until they are placed into a Precellys[®] Evolution homogenizer and lysed by bead-beating (3×30 s, 6500 rpm, with 60 s pause in-between). Perform this process two to three times with intervals of 1 minute on ice between each session to prevent overheating and ensure efficient lysis.
9. Pellet insoluble cell debris and beads by centrifuging at $10,000\times g$ for 5 min and transfer supernatant into a new 1.5 mL tube (*see* **Note 16**).
10. Measure the protein concentration (e.g., with BCA protein assay, Bio-Rad).
11. OPTIONAL: labeling with FP-desthiobiotin can be detected on SDS-PAGE by western blotting (*see* **Note 17**).
12. Store samples at -80 °C.

**3.2.2 Enrichment of
Biotinylated Probes-
Labeled Proteins and
Sample Preparation for
LC-MS**

1. Equilibrate streptavidin magnetic beads (*see* **Note 18**):
 - (a) Transfer 40 μ L streptavidin magnetic beads to a 1.5 mL protein Lo-Bind tube per sample.
 - (b) Wash twice with 1 mL RIPA lysis buffer using a magnetic rack.
 - (c) Resuspend beads in 500 μ L RIPA lysis buffer.
2. Add 20 μ L of 10% SDS and boil at 95 °C for 8 min in a Thermomixer. Cool down to RT (optional).
3. Mix 500 μ L of proteins (corresponding to a protein concentration of 1 mg/mL) with 500 μ L of equilibrated beads in a 1.5 mL protein Lo-Bind Eppendorf tube (*see* **Note 19**) and incubate for at least 1 h on a rotary shaker at 18 rpm at 4 °C (*see* **Note 20**).
4. Place tube in a magnetic rack, remove the supernatant from the beads, and discard the supernatant.
5. Wash the beads twice with 1 mL of RIPA lysis buffer for 2 min.
6. Wash beads once with 1 mL of 1 M KCl for 2 min.
7. Wash beads once with 1 mL of 0.1 M Na_2CO_3 for approximately 10 s.

8. Wash beads once with 1 mL of 2 M urea in 10 mM Tris-HCl (pH 8.0) for approximately 10 s.
9. Wash beads twice with 1 mL of RIPA lysis buffer for 2 min per wash (*see* **Note 21**).
10. Transfer the beads into fresh tubes containing 1 mL of RIPA lysis buffer, mix, place tube in magnetic rack, and remove supernatant.
11. Wash the beads with 500 μ L of 4 M Urea Wash Buffer.
12. Shake for 5 min at 500 rpm.
13. Remove the supernatant and add 100 μ L of 4 M Urea Wash Buffer.
14. Transfer the beads to a new tube and add additional 400 μ L of 4 M Urea Wash Buffer. (This tube change is to ensure that any remnants of RIPA buffer with detergents on the tube walls/lid are removed.)
15. Shake for 5 min at 500 rpm, place tube in magnetic rack, and remove supernatant.
16. Repeat **steps 13–15** once more.
17. Wash the beads twice (or more) with 500 μ L of 50 mM Tris-HCl (pH 7.5), or alternatively, with 50 mM Ambic or 50 mM TEAB. Shake for 5–7 min at 650 rpm. Transfer the beads to a new tube between washes to ensure the complete removal of all detergents and nonspecific proteins (*see* **Note 22**).
18. Add 150 μ L of 50 mM TEAB (or Ambic), 1 mM CaCl_2 , and 1 μ g of trypsin (i.e., 2 μ L of trypsin solution). Incubate for 1 h at RT on a shaker at 800 rpm (*see* **Note 23**).
19. Reduction: Add DTT to achieve a final concentration of 5 mM (according to the buffer volume; 150 μ L). Shake for 45 min at RT.
20. Alkylation: Add IAA to a final concentration of 15 mM. Shake at RT for 30 min. Incubate in the **DARK** since IAA is light-sensitive.
21. Add DTT again to achieve a final concentration of 5 mM DTT (*see* **Note 24**).
22. Add 3 μ g of trypsin (i.e., 6 μ L of trypsin solution). Gently vortex the sample and incubate overnight at 37 °C on a shaker at 800 rpm.
23. After overnight incubation, transfer the supernatant containing peptides to a fresh protein Lo-Bind tube.
24. Wash the beads with 70 μ L of 50 mM TEAB (or 50 mM Ambic), shaking for 20 min at RT. Pool this wash with the first extraction.

25. Acidify the pooled sample by adding formic acid to achieve a concentration of approximately 1% (vol/vol).
26. Desalt and concentrate the samples using an OMIX C18 tip according to the manufacturer's instructions. The purified peptide samples are first dried using a vacuum concentrator and then reconstituted in 10–20 μ L of 0.1% formic acid.
27. Determine the concentration of peptides using a Nanodrop spectrophotometer at a wavelength of 205 nm.
28. Store sample in the protein Lo-Bind tubes at -20°C if LC–MS/MS analysis is not conducted on the same day.

3.2.3 LC-MS Analysis

This protocol was established for use of an EASY-nLC 1200 system linked to either an Exploris 480 Orbitrap or an Orbitrap Fusion Lumos mass spectrometer using a data-dependent acquisition (DDA) MS strategy to facilitate label-free quantification of protein. Due to the wide variety of available LC–MS/MS analytical setups, essential parameters can substantially differ depending on the HPLC and MS systems employed (*see Note 25*).

1. Load 0.5 μ g of concentrated peptides mixtures in 0.1% formic acid onto an EASY-nLC1200 system equipped with an EASY-Spray column (C18, 2 μ m, 100 $\text{\AA}$, 50 μ m $\times$ 50 cm).
2. Fractionate the peptides using a gradient of 5–80% acetonitrile in 0.1% formic acid at a flow rate of 300 nL/min over 60 min.
3. Analyze the fractionated peptides with an Orbitrap Exploris 480 or Orbitrap Fusion Lumos mass spectrometer employing a DDA mode using a Top20 method.
4. Analyze MS raw data, for example, using an appropriate software for label-free quantification such as MaxQuant (*see Note 26*).

4 Notes

1. In our work with *E. faecium*, we have used BHI agar as standard growth medium. However, choice of growth media and conditions will depend on the bacterial species being investigated, and the overarching research aims. Choose media best suited for the bacteria being studied. It is also possible to use ABPP to compare the enzymatic activity across samples of different growth conditions.
2. An alternative to growing the bacteria on agar plates is to use broth media. Pick one colony from the plate and resuspend in 3 mL media broth (for instance, BHI) in 14 mL falcon round-bottom tubes and incubate at 37°C for ~16–18 h in a shaking incubator (~200 rpm).

3. Remember the linear range of the spectrophotometer when measuring optical density (OD). An OD₆₀₀ of 20 is beyond this linear range, and the sample therefore must be diluted before the OD measurement. For instance, dilute sample 1:40 in clear culture media by adding 25 μ L of the bacterial suspension to 975 μ L media. The measured value should be $\sim$ 0.5.
4. It is possible to adjust the volumes used and the assay can be miniaturized to 50 μ L bacterial suspension and 0.5 μ L probe solution, if needed. Remember to adjust the volumes in **step 6** accordingly that is if using 50 μ L bacterial suspension, reduce the volumes of reagents and glass beads in **step 6** by 1/3.
5. Probe concentration can be varied. When testing a probe for the first time for a new organism of interest, we suggest to first determine the potency and selectivity of the probe in a titration experiment (e.g., testing the probe at final concentrations of 1 μ M, 300 nM, 100 nM, 30 nM). With decreasing concentrations of probe, a dose-dependent *decrease in the number and intensity of detected bands is expected*.
6. The incubation time of probe (or inhibitor in competitive ABPP) with the bacterial culture may be varied and a time-course experiment may be performed (e.g., comparing 15 min, 30 min, 60 min) to optimize this step for the bacterial species and growth condition of interest. Longer labeling times may give better labeling outcomes but also increase the risk of inducing artifacts.
7. The competitive effects of unlabeled inhibitors in the gel-based ABPP setup are dose-dependent and depend on the potency and selectivity of the inhibitor. We generally recommend to test inhibitors in dose-dependent fashion in a 10-fold dilution series (e.g., 10 μ M to 10 nM final concentration).
8. Commercial pre-cast gradient gels give very good separation, and their use is convenient, avoids handling of the toxic chemical acrylamide, but is also very expensive. In most cases, homemade gels give similar performance, but are much more cost-efficient. The separation may be optimized depending on the molecular weight of the target proteins of interest by tuning the acrylamide percentage in the gel, but as a starting point we recommend a polyacrylamide percentage of 10%. The use of homemade wide or triple-wide gel systems can significantly enhance the throughput allowing to run dozens of samples on the same gel (30–102, depending on system and comb size).
9. The commercial precast NuPAGE gels come in different sizes, with different number of wells (for instance, 10 well, 12 well, and 15 well). The volume of each well depends on the number

of wells on the gel and is stated on the webpage of the manufacturer. The same principle is true for a homemade gel, which can be cast with different combs. Be aware of the maximum possible loading volume before starting to load the gels.

10. The time and voltage might need to be optimized to gain the best separation of proteins on the gel, which is usually achieved once the dye at the running front has just run out of the gel.
11. It is possible to stain the gel with any other staining solution, for example, using Coomassie Blue or Silver stain, to determine the total protein content and control the loading of the gel.
12. Significant enrichment of some non-serine hydrolase targets (that may have been enriched nonspecifically) is not uncommon when comparing the FP-desthiobiotin-labeled and vehicle-treated control dataset. As an additional control to identify those targets that have been labeled by the ABP in an active-site-directed mechanism, datasets can be included that have been pretreated with non-biotinylated inhibitors such as FP-TMR in a competitive ABPP setup may be considered.
13. A vehicle (DMSO)-treated control is crucial to distinguish specifically enriched proteins that have been labeled with FP-desthiobiotin from endogenously biotinylated and non-specifically enriched proteins. Our standard labeling protocol uses lower concentrations of FP-desthiobiotin compared to previous studies [1]. In the live cell labeling setup, this reduction in probe concentration facilitates removal of undesired free probe after labeling by simple centrifugation rendering additional clean-up using desalting column unnecessary.
14. Extensive washing is used to ensure efficient removal of free probe that could interfere with the enrichment of biotinylated proteins.
15. Addition of protease inhibitors can be included to reduce unwanted proteolytic degradation in the sample. We have performed this protocol with and without addition of protease inhibitors with similar results.
16. Desalting columns can be used to eliminate any free probe remaining in the lysate. Rather than employing a spin protocol, desalting columns can also be used with a gravity protocol for a more efficient removal of free probe. We recommend using desalting column when directly labeling in protein extracts rather than in live cells.
17. Labeling with FP-desthiobiotin can also be analyzed by SDS-PAGE/western blotting using standard protocols and detection by streptavidin-horse radish peroxidase (HRP) or fluorescent streptavidin conjugates. However, the high abundance of endogenously biotinylated proteins in bacterial

proteomes can make the detection of probe-labeled targets on the background of these proteins by western blot challenging.

18. The use of magnetic streptavidin beads allows for facile removal of nonspecific proteins and excess detergents using magnetic racks. It is also possible to use non-magnetic streptavidin resins (e.g., agarose-based) with filtration-based wash protocols.
19. In our experience, using 1.5 mL tubes facilitates easier separation of magnetic beads.
20. The incubation can also be performed overnight at 4 °C. Extending the incubation time enhances the binding and solubilization of a greater quantity of proteins.
21. Avoid letting the beads sit in Na₂CO₃ or urea for extended periods, as prolonged exposure can denature the beads and lead to aggregation.
22. Extensive washing is performed to ensure removal of nonspecifically bound proteins and removal of detergent is essential before addition of trypsin for efficient on-bead digest.
23. This pre-digestion step is intended to improve accessibility of disulfide bridges for reduction and alkylation in the next steps.
24. Addition of excess DTT at this step will quench nonreacted IAA.
25. Efficient detergent removal in the sample preparation protocol (through extensive washing facilitated by magnetic beads) eliminates the need for a biphasic column before the (LC-) LC-MS/MS setup.
26. For data analysis using MaxQuant, the bacterial protein sequence database from UniProt in FASTA format is used to match against fragmentation spectra. The search parameters are configured according to the MaxQuant reference setup for “Label-free quantification” mode [26]. Access the MaxQuant output in the “.txt” folder and utilize the “LFQ intensities” from the “proteinGroups.txt” file for further analysis. The output is further analyzed using the Perseus software [27] for label-free protein quantification. The significance and degree of protein enrichment compared to control strains are calculated. Utilizing defined enrichment criteria (e.g., log₂ ratio ≥ 1; *p*-value < 0.05), proteins of interest are identified as significantly enriched by probe treatment.

Acknowledgments

This work was financed by the Northern Norway Regional Health Authority Medical Research Programme grant (HNF1570-21) to C.S.L. Mass spectrometry-based proteomic analyses were

performed by UiT Proteomics and Metabolomics Core Facility (PRiME). This facility is a member of the National Network of Advanced Proteomics Infrastructure (NAPI), which is funded by the Research Council of Norway INFRASTRUKTUR-program (project number: 295910). We thank Toril Anne Grønset and Jack-Ansgar Bruun for their support in this work.

References

1. Lentz CS, Sheldon JR, Crawford LA, Cooper R, Garland M, Amieva MR, Weerapana E, Skaar EP, Bogyo M (2018) Identification of a *S. aureus* virulence factor by activity-based protein profiling (ABPP). *Nat Chem Biol* 14(6):609–617. <https://doi.org/10.1038/s41589-018-0060-1>
2. Chen L, Keller LJ, Cordasco E, Bogyo M, Lentz CS (2019) Fluorescent triazole urea activity-based probes for the single-cell phenotypic characterization of *Staphylococcus aureus*. *Angew Chem Int Ed Engl* 58(17):5643–5647. <https://doi.org/10.1002/anie.201900511>
3. Li M, Patel HV, Cognetta AB 3rd, Smith TC 2nd, Mallick I, Cavalier JF, Previti ML, Canaan S, Aldridge BB, Cravatt BF, Seeliger JC (2022) Identification of cell wall synthesis inhibitors active against *Mycobacterium tuberculosis* by competitive activity-based protein profiling. *Cell. Chem Biol* 29(5):883–896 e885. <https://doi.org/10.1016/j.chembiol.2021.09.002>
4. Babin BM, Keller LJ, Pinto Y, Li VL, Eneim AS, Vance SE, Terrell SM, Bhatt AS, Long JZ, Bogyo M (2022) Identification of covalent inhibitors that disrupt *M. tuberculosis* growth by targeting multiple serine hydrolases involved in lipid metabolism. *Cell. Chem Biol* 29(5):897–909 e897. <https://doi.org/10.1016/j.chembiol.2021.08.013>
5. Wu L, Armstrong Z, Schroder SP, de Boer C, Artola M, Aerts JM, Overkleeft HS, Davies GJ (2019) An overview of activity-based probes for glycosidases. *Curr Opin Chem Biol* 53:25–36. <https://doi.org/10.1016/j.cbpa.2019.05.030>
6. McGregor NGS, Overkleeft HS, Davies GJ (2022) Detecting and identifying glycoside hydrolases using cyclophellitol-derived activity-based probes. *Methods Enzymol* 664:103–134. <https://doi.org/10.1016/bs.mie.2022.01.007>
7. Chakrabarty S, Kahler JP, van de Plassche MAT, Vanhoutte R, Verhelst SHL (2019) Recent advances in activity-based protein profiling of proteases. *Curr Top Microbiol Immunol* 420:253–281. https://doi.org/10.1007/82_2018_138
8. Faucher F, Bennett JM, Bogyo M, Lovell S (2020) Strategies for tuning the selectivity of chemical probes that target serine hydrolases. *Cell Chem Biol* 27(8):937–952. <https://doi.org/10.1016/j.chembiol.2020.07.008>
9. Fuerst R, Breinbauer R (2021) Activity-based protein profiling (ABPP) of oxidoreductases. *Chembiochem* 22(4):630–638. <https://doi.org/10.1002/cbic.202000542>
10. Liu Y, Patricelli MP, Cravatt BF (1999) Activity-based protein profiling: the serine hydrolases. *Proc Natl Acad Sci USA* 96(26):14694–14699. <https://doi.org/10.1073/pnas.96.26.14694>
11. Kidd D, Liu Y, Cravatt BF (2001) Profiling serine hydrolase activities in complex proteomes. *Biochemistry* 40(13):4005–4015. <https://doi.org/10.1021/bi00679j>
12. Nomura DK, Leung D, Chiang KP, Quistad GB, Cravatt BF, Casida JE (2005) A brain detoxifying enzyme for organophosphorus nerve poisons. *Proc Natl Acad Sci USA* 102(17):6195–6200. <https://doi.org/10.1073/pnas.0501915102>
13. Jessani N, Niessen S, Wei BQ, Nicolau M, Humphrey M, Ji Y, Han W, Noh DY, Yates JR 3rd, Jeffrey SS, Cravatt BF (2005) A streamlined platform for high-content functional proteomics of primary human specimens. *Nat Methods* 2(9):691–697. <https://doi.org/10.1038/nmeth778>
14. Schmerling C, Sewald L, Heilmann G, Witfeld F, Begerow D, Jensen K, Brasen C, Kaschani F, Overkleeft HS, Siebers B, Kaiser M (2022) Identification of fungal lignocellulose-degrading biocatalysts secreted by *Phanerochaete chrysosporium* via activity-based protein profiling. *Commun Biol* 5(1):1254. <https://doi.org/10.1038/s42003-022-04141-x>
15. Child MA, Hall CI, Beck JR, Ofori LO, Albrow VE, Garland M, Bowyer PW, Bradley PJ, Powers JC, Boothroyd JC, Weerapana E, Bogyo M (2013) Small-molecule inhibition of

- a depalmitoylase enhances *Toxoplasma* host-cell invasion. *Nat Chem Biol* 9(10):651–656. <https://doi.org/10.1038/nchembio.1315>
16. Davison D, Howell S, Snijders AP, Deu E (2022) Activity-based protein profiling of human and plasmodium serine hydrolases and interrogation of potential antimalarial targets. *iScience* 25(9):104996. <https://doi.org/10.1016/j.isci.2022.104996>
 17. Zweerink S, Kallnik V, Ninck S, Nickel S, Verheyen J, Blum M, Wagner A, Feldmann I, Sickmann A, Albers SV, Brasen C, Kaschani F, Siebers B, Kaiser M (2017) Activity-based protein profiling as a robust method for enzyme identification and screening in extremophilic Archaea. *Nat Commun* 8:15352. <https://doi.org/10.1038/ncomms15352>
 18. Keller LJ, Lentz CS, Chen YE, Metivier RJ, Weerapana E, Fischbach MA, Bogoy M (2020) Characterization of serine hydrolases across clinical isolates of commensal skin bacteria *Staphylococcus epidermidis* using activity-based protein profiling. *ACS Infect Dis* 6(5): 930–938. <https://doi.org/10.1021/acinfecdis.0c00095>
 19. Keller LJ, Nguyen TH, Liu LJ, Hurysz BM, Lakemeyer M, Guerra M, Gelsinger DJ, Chanin R, Ngo N, Lum KM, Faucher F, Ipock P, Niphakis MJ, Bhatt AS, O'Donoghue AJ, Huang KC, Bogoy M (2023) Chemoproteomic identification of a DPP4 homolog in *Bacteroides thetaiotaomicron*. *Nat Chem Biol* 19(12):1469–1479. <https://doi.org/10.1038/s41589-023-01357-8>
 20. Ortega C, Anderson LN, Frando A, Sadler NC, Brown RW, Smith RD, Wright AT, Grundner C (2016) Systematic survey of serine hydrolase activity in *Mycobacterium tuberculosis* defines changes associated with persistence. *Cell Chem Biol* 23(2):290–298. <https://doi.org/10.1016/j.chembiol.2016.01.003>
 21. Tallman KR, Levine SR, Beatty KE (2016) Small-molecule probes reveal esterases with persistent activity in dormant and reactivating *Mycobacterium tuberculosis*. *ACS Infect Dis* 2(12):936–944. <https://doi.org/10.1021/acinfecdis.6b00135>
 22. Hatzios SK, Abel S, Martell J, Hubbard T, Sasabe J, Munera D, Clark L, Bachovchin DA, Qadri F, Ryan ET, Davis BM, Weerapana E, Waldor MK (2016) Chemoproteomic profiling of host and pathogen enzymes active in cholera. *Nat Chem Biol* 12(4):268–274. <https://doi.org/10.1038/nchembio.2025>
 23. Thuy-Boun PS, Wang AY, Crissien-Martinez A, Xu JH, Chatterjee S, Stupp GS, Su AI, Coyle WJ, Wolan DW (2022) Quantitative metaproteomics and activity-based protein profiling of patient fecal microbiome identifies host and microbial serine-type endopeptidase activity associated with ulcerative colitis. *Mol Cell Proteomics* 21(3):100197. <https://doi.org/10.1016/j.mcpro.2022.100197>
 24. Uddin MJ, Overkleeft HS, Lentz CS (2023) Activity-based protein profiling in methicillin-resistant *Staphylococcus aureus* reveals the broad reactivity of a carmofur-derived probe. *Chembiochem* 24(21):e202300473. <https://doi.org/10.1002/cbic.202300473>
 25. Grunnvåg JS, Hegstad K, Lentz CS (2024) Activity-based protein profiling of serine hydrolases and penicillin-binding proteins in *Enterococcus faecium*. *FEMS Microbes* 5:xtae015. <https://doi.org/10.1093/femsmc/xtae015>
 26. Tyanova S, Temu T, Cox J (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* 11(12):2301–2319. <https://doi.org/10.1038/nprot.2016.136>
 27. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* 13(9):731–740. <https://doi.org/10.1038/nmeth.3901>



Profiling Ligand-Induced Changes in Nuclear Localization Using Proximity Labeling-Coupled Chemoproteomics

Qianni Peng and Eranthie Weerapana

Abstract

The nuclear proteome encompasses diverse proteins that regulate critical cellular functions, including histone modifications, chromatin structure, and transcription. Mutations to many of these nuclear proteins correlate with the onset of diseases such as cancer. Due to the disease relevance of nuclear proteins, drug development efforts have focused on identifying small-molecule modulators of nuclear protein function. Covalent ligands provide a promising strategy to therapeutically target nuclear proteins that lack distinct substrate binding pockets. In particular, chemoproteomic strategies have enabled the identification of ligandable sites within the proteome, with a particular emphasis on covalent targeting of cysteine residues. Nuclear proteins are typically poorly represented in chemoproteomic workflows that utilize whole-cell lysates due to the low abundance of these proteins and the localization of nuclear proteins in multiple cellular compartments. To specifically focus on the nuclear proteome, we coupled proximity labeling using a histone-TurboID construct with chemoproteomics. Notably, this platform can be utilized to identify ligandable sites within the nuclear proteome, and monitor changes in nuclear localization and chromatin association upon exposure to covalent ligands. Here, we describe the steps required to generate histone-TurboID expressing cell lines, and apply tandem mass tag (TMT)-based quantitative proteomics to monitor protein localization changes induced by covalent ligands. Together, this methodology provides a streamlined approach toward identifying covalent ligands that regulate nuclear protein function.

Key words TurboID, Proximity labeling, Chemoproteomics, Nuclear cysteine ligandability, Nuclear localization, Chromatin association

1 Introduction

The nucleus is a unique organelle in eukaryotic cells responsible for storing genetic materials and regulating cell proliferation, differentiation, and apoptosis [1–3]. Most nuclear proteins exert their functions by interacting with chromatin for epigenetic regulation, including transcription factors, chromatin remodelers, and histone modifying proteins [4–6]. The interactions between chromatin and these chromatin-associated proteins are well-regulated to ensure cellular homeostasis, while genetic mutations or external stimuli

that cause aberrant protein–protein interactions or chromatin association can lead to the onset of disease [7]. For example, the tumor suppressor gene TP53 is the most frequently mutated gene in adenocarcinoma [8]. Common missense mutations of p53 disrupt its DNA binding ability, affecting cell cycle regulation, DNA damage repair, and apoptosis induction [9]. Many nuclear proteins are challenging to therapeutically target with small-molecule ligands due to the presence of unstructured protein regions and shallow binding pockets. Covalent ligands can more effectively target proteins that lack prototypical ligand binding sites, and additionally exhibit enhanced potency, selectivity, and prolonged duration of action [10]. Cysteine residues have been the nucleophile of choice for covalent targeting due to their high nucleophilicity [11], and clinically approved cysteine-targeted covalent therapeutics include Ibrutinib (Bruton’s tyrosine kinase, BTK), Afatinib (epidermal growth factor receptor, EGFR), and Sotorasib (KRAS G12C) [12–14].

To broaden the protein targets amenable to covalent liganding, chemoproteomic platforms, including the competitive isotopic tandem orthogonal proteolysis-activity-based protein profiling (isoTOP-ABPP) method, have been developed [15–17]. Proteome-wide screening of highly promiscuous cysteine-reactive scout fragments has unveiled ligandable cysteine residues on proteins that were previously considered to be “undruggable.” Standard isoTOP-ABPP workflows for screening covalent ligands utilize whole-cell proteomic analyses, which provide poor coverage of low-abundance proteins localized within specific subcellular organelles, including the nucleus. Additionally, many nuclear proteins are distributed across various cellular locations [18], and therefore determining the liganding capacity of the nuclear pool of a given protein is challenging. These limitations can be circumvented by using proximity labeling to enrich nuclear proteomes prior to competitive isoTOP-ABPP analysis. We demonstrated the utility of this platform to identify ligandable cysteines within the nuclear proteome and reveal liganding events that modulate nuclear localization and chromatin association [19].

Here, we provide a detailed protocol for analyzing ligand-induced changes in nuclear localization. This platform utilizes the biotin ligase TurboID fused to the C-terminus of histone H3.3 (His-TID), a histone variant that undergoes extensive post-translational modifications and localizes at gene loci with active transcription [20, 21]. The His-TID construct allows for the enrichment of nuclear proteins for mass spectrometry (MS) analysis (Fig. 1). Detailed workflows are provided for the following steps in the protocol; (1) generation of a monoclonal cell line that stably expresses His-TID by lentiviral transduction and monoclonal selection (Fig. 2); and (2) the use of His-TID cells in a tandem mass tag (TMT)-MS workflow (Fig. 3) to identify proteins

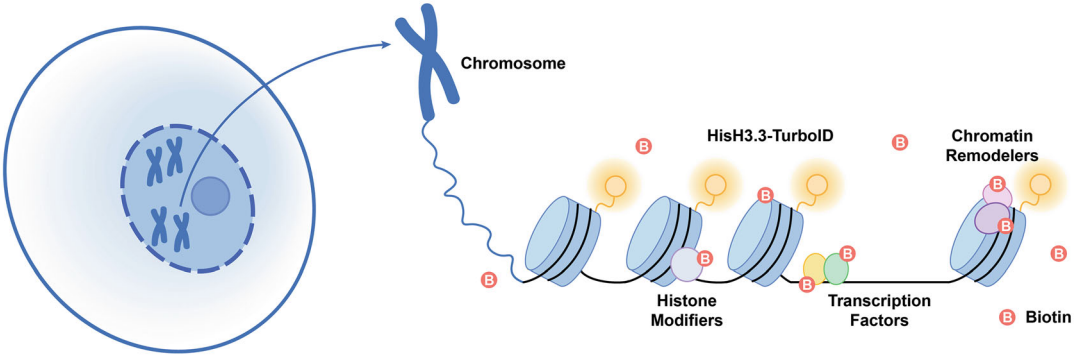


Fig. 1 Schematic illustration of a His-TID-based proximity labeling platform to biotinylate nuclear proteins

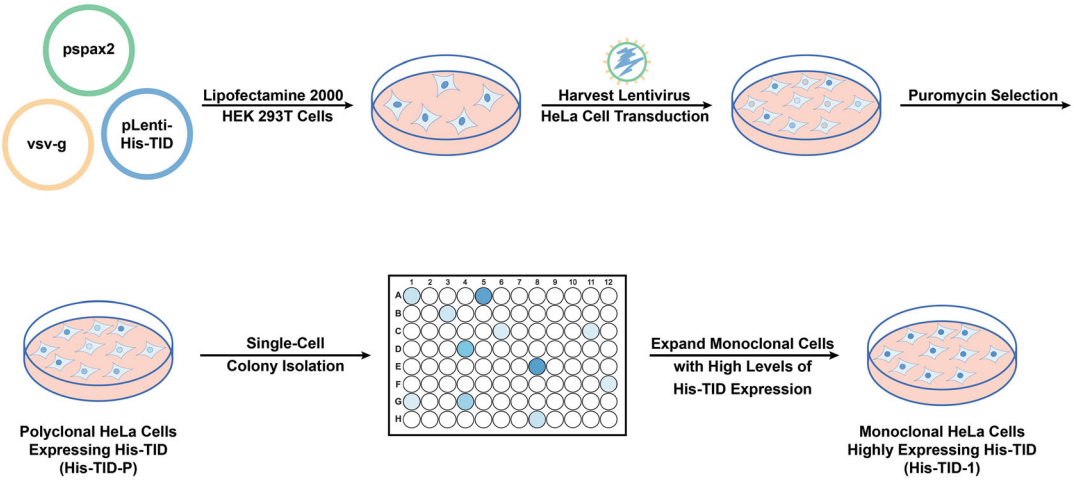
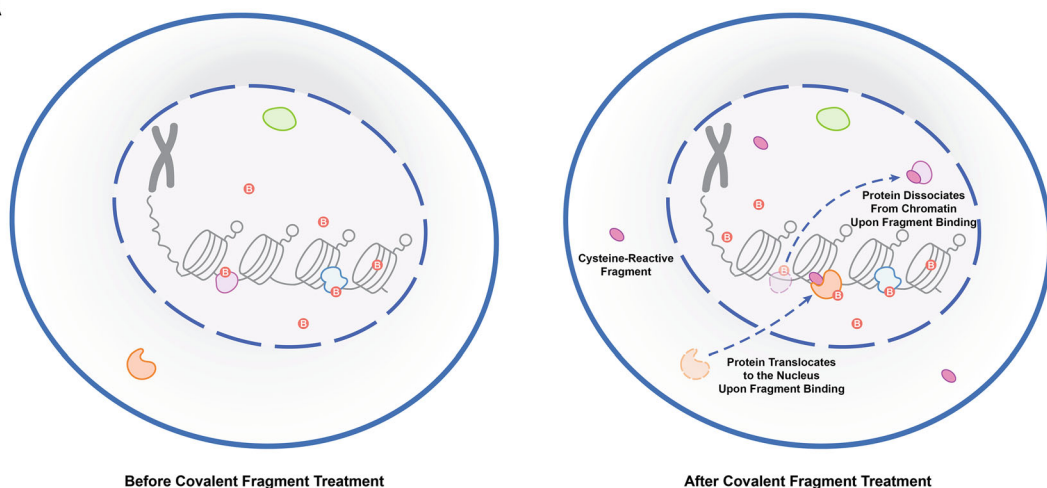


Fig. 2 Workflow to generate polyclonal HeLa cells stably expressing His-TID (His-TID-P) and single-cell colony selection to isolate and expand monoclonal populations with high expression levels of His-TID (His-TID-1)

that display changes in nuclear localization and chromatin enrichment upon ligand treatment. We show prototypical data for the cysteine-reactive scout fragments KB02 and KB05 [17] (Fig. 4), where 1657 nuclear proteins, including 296 transcription factors, were identified [19].

In summary, we present a platform for enriching nuclear proteins, and its application in investigating liganding-mediated changes in protein nuclear localization. We highlight the use of this platform to identify covalent ligands that regulate the nuclear localization and chromatin association of proteins. However, the platform is highly versatile and can be applied to investigate the effects of other classes of small molecules, or cellular stresses, on the nuclear proteome.

A



B

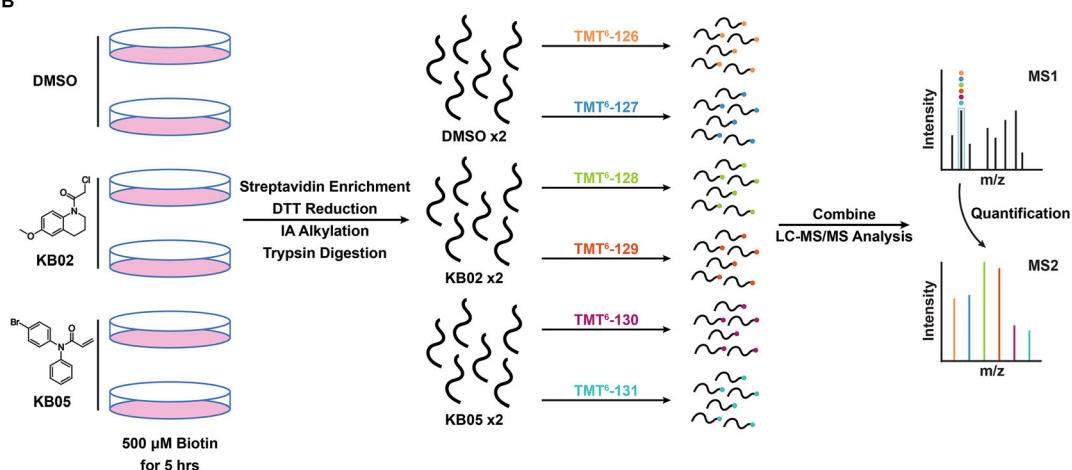


Fig. 3 Profiling changes in protein nuclear localization and chromatin association mediated by covalent fragment treatment. **(a)** Illustration of potential effects of fragments on protein chromatin association and its subsequent biotinylation level. **(b)** Proximity labeling and TMT workflow for profiling changes in protein biotinylation levels mediated by fragment treatment

2 Materials

All buffers and solutions are prepared with ultrapure MilliQ water unless indicated otherwise.

2.1 Generation of a Polyclonal HeLa Cell Line Stably Expressing His-TID

1. HEK293T cells (ATCC); HeLa cells (ATCC).
2. DMEM with 10% heat-inactivated fetal bovine serum (HIS): heat 50 mL fetal bovine serum (FBS) at 56 °C for 30 min with inverting every 10 min and combine it with 500 mL of DMEM.
3. Serum-free DMEM.

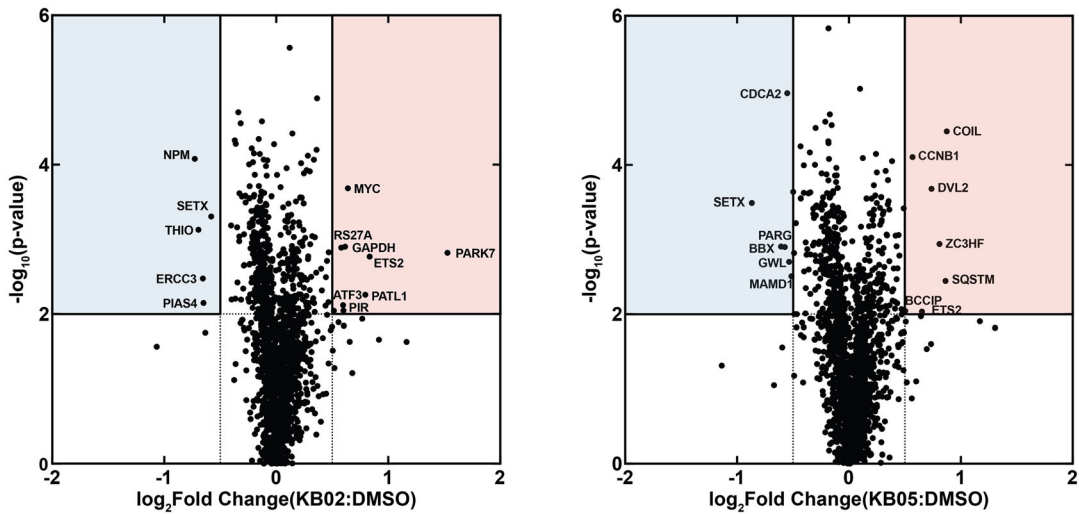


Fig. 4 Volcano plots of changes in protein biotinylation upon KB02 or KB05 treatment. Proteins with significant changes in biotinylation were highlighted in blue ($\log_2\text{Fold Change} < -0.5$) or red ($\log_2\text{Fold Change} > 0.5$), $p < 0.01$. Data were presented as the average values of three technical replicates. A two-tailed t -test was used to calculate p -values

4. Lipofectamine 2000.
5. Envelope plasmid pCMV-VSV-G (Addgene) [22]; packaging plasmid psPAX2 (Addgene); pLenti-His-TID (made in-house) [19]; and positive control for lentiviral transduction pLenti-eGFP (made in-house).
6. 10 mg/mL polybrene.
7. Luer Lok disposable syringes without needles, 10 mL.
8. PES syringe filter, 0.45 μm , 30 mm, sterile.
9. Complete DMEM medium: DMEM with 10% FBS, 100 units/mL of penicillin, 100 $\mu\text{g}/\text{mL}$ of streptomycin, and 0.25 $\mu\text{g}/\text{mL}$ of amphotericin B.
10. 0.25% Trypsin.
11. Puromycin dihydrochloride.

2.2 Generation of Monoclonal HeLa Cell Lines Stably Expressing His-TID

1. Cell counting slides.
2. TC20 automated cell counter.
3. 96-Well tissue culture plates (Chemglass); 24-well tissue culture plates; 6-well tissue culture plates.

2.3 Screening of Monoclonal HeLa Cell Lines for His-TID Expression

1. DPBS.
2. Ultrasonic Tip Sonicator.
3. DC Protein Assay Kit.

4. 2x SDS-PAGE loading dye: 100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 0.2% (w/v) bromophenol blue, 20% (v/v) glycerol, 200 mM dithiothreitol (DTT). For 20 mL of 2x SDS-PAGE loading dye, dissolve 0.24 g Tris, 0.8 g SDS, 40 mg bromophenol blue, 4 mL glycerol, and 0.62 g DTT in 15 mL H₂O. Use 2 M HCl to adjust the pH to 6.8, and add H₂O to bring the volume to 20 mL. Aliquot the loading dye into microcentrifuge tubes and store at -20 °C.
5. 12.5% SDS-PAGE Gel: For the separating gel, mix 6.26 mL H₂O, 8.34 mL of 30% acrylamide/bis solution, 5 mL of 1.5 M Tris-HCl (pH 8.8), 200 µL of 10% SDS, 200 µL of 10% APS, and 10 µL TEMED and add to a gel-casting module. Add methanol to the top of the separating gel to flatten the front line and leave it at room temperature to solidify. For the stacking gel, mix 5 mL H₂O, 900 µL of 30% acrylamide/bis solution, 867.8 µL of 1.5 M Tris-HCl (pH 6.8), 69 µL of 10% SDS, 34.4 µL of 10% APS, and 6.9 µL TEMED. When the separating gel is solidified, pour off the methanol and add the stacking gel on top of the separating gel, gently insert a 10-well comb and wait for the gel to solidify at room temperature.
6. Nitrocellulose membranes.
7. Tris-buffered saline Tween 20 (TBST): Dissolve 60.6 g Tris-base and 116.9 g NaCl in ~750 mL H₂O. Adjust the pH to 7.4 with concentrated HCl and bring the total volume up to 1 L with H₂O to make 1 L of 10x Tris-buffered saline (10x TBS). To prepare 1x TBST, mix 100 mL of 10x TBS and 1 mL Tween 20 with 900 mL H₂O.
8. 5% Milk/TBST: Dissolve 2.5 g of nonfat dry milk in 50 mL of TBST to make 5% milk blocking solution.
9. Antibody solution: Dilute primary or secondary antibodies according to product usage information for western blotting in either 5% milk/TBST or 5% bovine serum albumin (BSA)/TBST. Store at 4 °C.
10. Antibodies: anti-HA-Tag rabbit mAb, anti-GAPDH rabbit mAb, and anti-rabbit IgG, HRP-linked antibody.
11. SuperSignal™ West Pico PLUS chemiluminescent substrate.
12. ChemiDoc MP Imaging System.

2.4 Confirmation of Nuclear Localization of the His-TID Construct

1. 35 mm glass-bottom dishes.
2. 4% formaldehyde: dilute 16% methanol-free formaldehyde four-fold in warm DPBS and prepare it fresh.
3. Blocking buffer: 1x DPBS with 5% normal goat serum and 0.3% Triton X-100. Add 0.5 mL normal goat serum to 9.5 mL DPBS. Add 30 µL Triton X-100 and mix well.

4. Antibody dilution buffer: 1x DPBS with 1% BSA and 0.3% Triton X-100. Dissolve 0.1 g BSA in 10 mL DPBS. Add 30 μ L Triton X-100 and mix well. Dilute primary antibodies or fluorochrome-conjugated secondary antibodies according to product usage information for immunofluorescence.
5. Antibodies: anti-HA-Tag mouse mAb, anti-mouse IgG with AlexaFluor[®] 488 Conjugate.
6. ProLong[®] Gold Antifade Reagent with DAPI.
7. Cover glasses.
8. Leica TCS SP5 Scanning Confocal Microscope.

2.5 Optimization of Biotinylating Conditions

1. D-Biotin.
2. Ice-cold DPBS.
3. DC Protein Assay Kit.
4. Antibodies: anti-HA-Tag rabbit mAb, anti-GAPDH rabbit mAb, streptavidin HRP, and anti-rabbit IgG, HRP-linked antibody.
5. SuperSignal[™] West Pico PLUS chemiluminescent substrate.
6. ChemiDoc MP Imaging System.

2.6 Optimization of Covalent Fragment Treatment Conditions

1. 2-Chloro-1-(6-Methoxy-1,2,3,4-Tetrahydroquinolin-1-Yl) Ethan-1-One (KB02) [17].
2. N-(4-Bromophenyl)-N-Phenylacrylamide (KB05) [17].
3. Serum-Free DMEM.

2.7 Mass Spectrometry Sample Preparation for Investigating Changes in Chromatin Association Mediated by Fragment Treatment

1. 100% (w/v) trichloroacetic acid (TCA) solution: dissolve 2.2 g TCA in 1 mL DPBS.
2. Ice-cold acetone.
3. 1.2% SDS/DPBS: dissolve 0.6 g sodium dodecyl sulfate (SDS) in 50 mL DPBS.
4. Pierce streptavidin agarose.
5. 0.2% SDS/DPBS: dilute 1.2% SDS/DPBS six-fold in DPBS.
6. Screw cap micro tubes.
7. 6 M urea/DPBS: dissolve 1.8 g urea in 5 mL DPBS and prepare it fresh.
8. 200 mM DTT solution: dissolve 30 mg DTT in 1 mL DPBS and make it fresh.
9. 400 mM iodoacetamide (IA) solution: dissolve 74 mg IA in 1 mL DPBS and make it fresh.
10. 100 mM triethylammonium bicarbonate (TEAB) buffer: dilute 1 M TEAB, pH 8.5 ± 0.1 , ten-fold in H₂O.

11. Premixed trypsin solution: 200 μ L of 100 mM TEAB with 2 μ L of 100 mM CaCl_2 and 4 μ L of 0.5 mg/mL trypsin. Dissolve 11 mg CaCl_2 in 1 mL H_2O to make a 100 mM CaCl_2 solution. For a vial that contains 20 μ g lyophilized sequencing-grade modified trypsin, add 40 μ L of 100 mM TEAB and vortex to resuspend. Add 2 μ L of 100 mM CaCl_2 and 4 μ L of 0.5 mg/mL trypsin to 200 μ L of 100 mM TEAB, multiplying the volume of each reagent according to the number of samples to be digested.
12. EPPS buffer with 30% acetonitrile: add 300 μ L acetonitrile, Optima LC/MS grade, to 700 μ L of 200 mM EPPS buffer, pH 8.5, to make 1 mL of EPPS buffer with 30% acetonitrile.
13. TMTsixplex isobaric tag: resuspend each of the 0.8 mg TMTsixplex isobaric label reagent in 40 μ L of LC/MS grade acetonitrile to achieve a final concentration of 20 mg/mL. Aliquot the TMT label reagent into Eppendorf tubes, dispensing 4 μ L per tube, and store at -20°C to prevent freeze-thaw cycles.
14. 5% hydroxylamine: dilute 50% hydroxylamine buffer ten-fold with 100 mM TEAB buffer
15. 99.0 + % formic acid, Optima™ LC/MS grade.
16. 1.5 mL protein LoBind® microcentrifuge tubes.
17. Low-pH desalting buffer A: 95% H_2O , 5% acetonitrile, and 0.1% formic acid. Add 5 mL acetonitrile and 0.1 mL formic acid to 95 mL H_2O ; all three are Optima™ LC/MS grade.
18. Sep-Pak C18 Classic Cartridge.
19. Acetonitrile, Optima™ LC/MS grade.
20. Low-pH desalting buffer B: 20% H_2O , 80% acetonitrile, and 0.1% formic acid. Add 20 mL H_2O and 0.1 mL formic acid to 80 mL acetonitrile; all three are LC/MS grade.
21. 2 mL protein LoBind® microcentrifuge tubes.

**2.8 Liquid
Chromatography-
Tandem Mass
Spectrometry (LC-MS/
MS) Analysis**

1. 9 mm autosampler vial with screw thread caps.
2. Pierce® Quantitative Colorimetric Peptide Assay.
3. Orbitrap Exploris 240 Mass Spectrometer.
4. Dionex Ultimate 3000 RSLCnano System.
5. Xcalibur v4.4 Software Package.
6. Low-pH MS buffer A: 100% H_2O , 0.1% formic acid. Add 0.1 mL formic acid to 100 mL of H_2O ; both are Optima LC/MS grade.
7. Low-pH MS buffer B: 20% H_2O , 80% acetonitrile, and 0.1% formic acid. Add 20 mL H_2O and 0.1 mL formic acid to 80 mL of acetonitrile; all three are Optima LC/MS grade.

8. Acclaim PepMap 100 Trap Column.
9. Acclaim PepMap RSLC Analytical Column.

2.9 Analysis of LC-MS/MS Data

1. Proteome Discoverer v2.4 (Thermo Fisher Scientific).
2. GraphPad Prism 9.

3 Methods

3.1 Generation of a Polyclonal HeLa Cell Line Stably Expressing His-TID

1. Seed HEK 293 T cells under passage 15 in 2×10 -cm tissue culture dishes at 35% confluence in DMEM with 10% HIS for transfection the next day (*see Note 1*).
2. For each construct, prepare two Eppendorf tubes with 600 μ L of serum-free DMEM. Add 24 μ L Lipofectamine 2000 to one tube and wait for 5 min. Add 2 μ g VSV-G, 2 μ g psPAX2, and 2 μ g pLenti-His-TID or pLenti-eGFP to the other tube (*see Note 2*). Combine the two tubes by adding the plasmid-containing medium to the medium with Lipofectamine 2000 and wait for 30–40 min at room temperature. Gently add the mixture to HEK 293 T cells at 70–80% confluence.
3. Aspirate the media 24 h after transfection and replenish HEK 293T cells with 5 mL DMEM with 10% HIS (*see Note 3*). Verify eGFP fluorescence for the pLenti-eGFP transfected plate to confirm successful transfection. Bleach the aspirated lentivirus-containing media with 30% bleach.
4. Seed HeLa cells at 10% confluence in DMEM in 3×10 -cm tissue culture dishes using DMEM with 10% HIS for transduction the next day (*see Note 4*).
5. Add 5 mL of DMEM with 10% HIS and 10 μ L of 10 mg/mL polybrene to a 15 mL conical tube. Collect and carefully filter the lentivirus-containing media with a 10 mL Luer Lok syringe with a 0.45 μ m filter into the corresponding 15 mL conical tube. Add the filtered media to HeLa cells at 20% confluence. Bleach the HEK cell dishes and all cell culture supplies that have come into contact with lentivirus using a 30% bleach solution.
6. 24 h after the transduction, swap lentivirus-containing media with complete DMEM and 1 μ g/mL puromycin (*see Note 5*). Grow transduced HeLa cells in complete media with puromycin for 3 days or until HeLa cells in the negative control plate are completely inviable. Verify eGFP fluorescence in the plate transduced with lentivirus carrying pLenti-eGFP to confirm successful lentiviral transduction and puromycin selection. Replace the media with fresh puromycin if necessary.

3.2 Generation of Monoclonal HeLa Cell Lines Stably Expressing His-TID

1. Resuspend polyclonal HeLa cells stably expressing His-TID (His-TID-P) by trypsinization and pass cell suspension through a serological pipet several times to break any cell clumps. Measure the cell density with a cell counter and dilute the cell suspension to 1000 and 5 cells/mL by serial dilution. Add 100 μ L of the 1000 cells/mL cell suspension to the first well of each 96-well plate. Add 100 μ L of the 5 cells/mL cell suspension to the remaining wells in 96-well plates (*see Note 6*).
2. 7–14 days after seeding the cells, check each well for signs of a single-cell colony (*see Note 7*). For wells with a single-cell colony, trypsinize and transfer the cells into a 24-well plate when they become fully confluent (*see Note 8*). Gradually expand the monoclonal cells into 6-well plates, 10-cm and 15-cm cell culture dishes.

3.3 Selection of Monoclonal HeLa Cell Lines with High His-TID Expression

1. Seed monoclonal HeLa cells in 10-cm cell culture dishes and harvest when the cells are fully confluent. Resuspend the cell pellets in DPBS and lyse the cells at 75% amplitude for 45 pulses with an ultrasonic tip sonicator. Chill cell lysates on ice every 15 pulses to avoid heat.
2. Centrifuge cell lysates at 14,000 rpm for 10 min at 4 °C and transfer the supernatants to new Eppendorf tubes. Measure protein concentrations with a DC protein assay kit and normalize cell lysates to 2 mg/mL with DPBS. Mix normalized cell lysates with 2x SDS-PAGE loading dye.
3. Run samples of monoclonal His-TID cell lysates together with the polyclonal cell lysate sample on a 12.5% SDS-PAGE gel at 165 V for 1 h. Transfer the proteins from the gel to a nitrocellulose membrane at 75 V for 100 min (*see Note 9*).
4. Block the membrane with 5% milk/TBST blocking solution for 1 h at room temperature. Wash the membranes with TBST three times, each time for 5 min. Incubate the membranes with anti-HA and anti-GAPDH rabbit primary antibody solutions overnight at 4 °C, respectively (*see Note 10*). Wash the membranes with TBST three times, each time for 5 min. Incubate the membranes with anti-rabbit secondary antibody solutions for 1 h at room temperature. Wash the membranes with TBST three times. Image the membranes with chemiluminescent substrates with a ChemiDoc MP imaging system.

3.4 Confirmation of Nuclear Localization of the His-TID Construct

1. Seed wild-type HeLa cells and monoclonal HeLa cells expressing His-TID in 35-mm dishes with glass coverslips (*see Note 11*). When cells reach 60–70% confluence, aspirate the media and add 2 mL of 4% formaldehyde diluted in warm DPBS into each dish. Incubate cells with the fixative solution for 15 min at room temperature.

2. Aspirate the fixative solution, rinse cells with 2 mL DPBS three times, each time for 5 min. Incubate cells with 2 mL blocking buffer for 1 h at room temperature. Aspirate the blocking buffer, incubate cells with anti-HA mouse primary antibody solution at 4 °C overnight.
3. Aspirate the primary antibody solution and rinse cells with DPBS three times. Incubate cells with AlexaFluor[®] 488-conjugated anti-mouse secondary antibody solution for 2 h at room temperature in the dark.
4. Rinse cells with DPBS three times and add 100 µL of Pro-Long[®] Gold Antifade Reagent with DAPI to the coverslip before applying cover glasses (*see Note 12*). Incubate cells with the mounting reagent overnight at room temperature in the dark. Seal coverslips with nail polish for long-term storage at 4 °C and protect from light.
5. Image cell fluorescence using Leica TCS SP5 scanning confocal microscopy.

3.5 Optimization of Biotinylating Conditions

1. Seed wild-type HeLa cells, His-TID-P, and monoclonal cells with the highest His-TID expression (His-TID-1) in 10-cm cell culture dishes. When fully confluent, treat cells with 500 µM biotin for different periods of time (0, 1, 2, 3, 4, 5 h) or with different concentrations of biotin (0, 50, 100, 500, 1,000 µM) for 1 h (*see Note 13*).
2. Wash cells with ice-cold DPBS three times and harvest by scraping the cells off the plates. Lyse the cells in ice-cold DPBS with a tip sonicator at 75% amplitude for 45 pulses, with chilling on ice every 15 pulses. Measure the protein concentration using a DC protein assay kit and normalize the sample concentrations to 2 mg/mL with DPBS.
3. Conduct western blot analysis according to the aforementioned procedures in Subheading 3.3. Notably for biotin blot, block the membrane with 5% milk/TBST blocking solution overnight at 4 °C instead of 1 h. Incubate the biotin blot with streptavidin-HRP solution in 5% BSA/TBST for 1 h at room temperature. Wash the blot three times before imaging with a ChemiDoc MP imaging system.

3.6 Optimization of Covalent Fragment Treatment Conditions

1. Seed His-TID-1 cells in 10-cm cell culture dishes. When fully confluent, treat cells with various concentrations of the scout fragment, KB02 or KB05, (0, 10, 20, 30, 50, 75, 100 µM) for 1 h in serum-free DMEM (*see Note 14*). Check cell morphology and viability using a microscope to decide the treatment conditions for MS samples.

3.7 Mass Spectrometry Sample Preparation for Investigating Changes in Chromatin Association Mediated by Fragment Treatment

1. Treat His-TID-1 cells with 30 μ M KB02, KB05, or DMSO for 1 h in serum-free media at 37 °C. Aspirate the media and replenish cells with complete DMEM containing 500 μ M biotin. Incubate the cells with biotin for 5 h at 37 °C.
2. Aspirate the media and wash cells with ice-cold DPBS three times. Harvest cells by scraping them off the plates. Resuspend the cell pellets in PBS and lyse at 75% amplitude for 45 pulses with a tip sonicator, chilling the samples on ice every 15 pulses. Centrifuge the cell lysates at 14,000 rpm for 10 min at 4 °C and transfer the supernatants to new tubes. Measure the protein concentrations for the supernatants with a DC protein assay kit.
3. Dilute lysates from different labeling conditions to 2 mg/mL using DPBS. Precipitate 1 mL of 2 mg/mL lysates (*see Note 15*) by mixing with 100 μ L of 100% (w/v) TCA solution (*see Note 16*), vortex vigorously and keep at –80 °C for 1 h or overnight. Prepare two samples for each labeling condition (DMSO, KB02, and KB05).
4. Thaw samples on ice and pellet down the proteins by centrifuging at 14,000 rpm for 10 min at 4 °C. Discard the supernatant and resuspend the protein pellet in 500 μ L ice-cold acetone by gentle tip sonicating at 60% amplitude. Centrifuge the samples at 5,000 rpm for 10 min at 4 °C.
5. Discard the acetone supernatant and let the protein pellet air dry to remove traces of remaining acetone (*see Note 17*). Add 1 mL of 1.2% SDS/DPBS to each sample and resolubilize the protein pellet by repeated rounds of heating at 65 °C for 5 min and tip sonication (*see Note 18*).
6. Aliquot 170 μ L of streptavidin agarose slurry into a 15 mL conical tube for each sample (*see Note 19*). Wash the beads with 5 mL DPBS three times by centrifuging at 1,400 g for 3 min.
7. After the final wash, add 5 mL DPBS to each tube and combine it with the 1 mL solubilized protein solution to make the protein solution a final volume of 6 mL in 0.2% SDS/DPBS. Incubate the samples with streptavidin beads overnight at 4 °C with rotating.
8. Rotate the samples at room temperature for 2–3 h to re-solubilize the SDS. Spin down the beads at 1,400 g for 3 min and discard the supernatant.
9. Add 5 mL of 0.2% SDS/DPBS to each sample and rotate at room temperature for 10 min. Spin down the samples at 1,400 g for 3 min and discard the supernatant. Further wash the beads three times each with 5 mL of DPBS and H₂O (*see Note 20*).

10. Transfer the beads to a 1.5 mL tube with a screw cap and rinse the 15 mL tube with an additional 500 μ L DPBS to transfer the residual beads (*see Note 21*). Spin at 1,400 g for 3 min, discard the supernatant, and add 500 μ L of freshly made 6 M urea/DPBS.
11. Add 25 μ L of freshly made 200 mM DTT solution (30 mg/mL in DPBS) to each sample. Heat the reaction at 65 °C for 20 min and resuspend the beads every 10 min (*see Note 22*).
12. Wait until the sample cools down to room temperature. Add 25 μ L of freshly made 400 mM IA (74 mg/mL in DPBS) to each sample and incubate at 37 °C for 30 min with rotating (*see Note 23*).
13. Add 950 μ L DPBS to each sample and spin at 1,400 g for 3 min to discard the supernatant. Further wash the beads with 1 mL of 100 mM TEAB three times to get rid of residual urea (*see Note 24*).
14. Add 206 μ L of premixed trypsin solution (200 μ L of 100 mM TEAB in H₂O, 2 μ L of 100 mM CaCl₂, and 4 μ L of 0.5 mg/mL trypsin in 100 mM TEAB) to each sample (*see Note 25*). Incubate at 37 °C overnight with rotating.
15. Spin at 1,400 g for 3 min to pellet the beads and transfer the supernatant that contains digested peptides to a new tube. Wash the beads with 50 μ L of 100 mM TEAB three times and combine the TEAB washes with the previous supernatant to make a final volume of ~350 μ L. Dry the peptide samples on SpeedVac.
16. Resuspend the peptide samples with 100 μ L 200 mM EPPS buffer (pH 8.0) with 30% acetonitrile by rehydrating for 30 min and bath sonicating for 1 min. Add 3 μ L of the corresponding TMT tag (20 mg/mL in acetonitrile) to each sample (*see Note 26*). Vortex and incubate the samples at room temperature for 1.5 h in the dark with vortexing every 45 min.
17. Quench the TMT labeling reaction by adding 5 μ L of 5% hydroxylamine to each sample. Vortex and incubate at room temperature for 15 min.
18. Add 5 μ L formic acid to each sample. Vortex and combine the six samples into a 1.5 mL protein LoBind[®] tube. Put the samples on SpeedVac to dryness overnight.
19. Resuspend peptide samples in 500 μ L of low-pH desalting buffer A. Allow samples to rehydrate for 30 min and bath sonicate for 1 min.
20. Attach a 1 mL syringe without a plunger to a C18 Sep-Pak cartridge. Add 1 mL LC/MS grade acetonitrile to the syringe and let it flow through the cartridge by gravity three times to rehydrate the C18 column.

21. Equilibrate the column with 3×1 mL low-pH desalting buffer A.
22. Add the resuspended peptide sample to the syringe and let it flow through the C18 cartridge by gravity three times to allow peptides to bind.
23. Wash the column three times with 1 mL low-pH desalting buffer A and elute the peptides off the cartridge into a 2 mL LoBind[®] tube with 3×0.5 mL low-pH desalting buffer B.
24. Put the samples on SpeedVac to dryness overnight and store the samples at -20 °C until LC-MS/MS analysis.

**3.8 Liquid
Chromatography-
Tandem Mass
Spectrometry (LC-MS/
MS) Analysis**

1. Resuspend the desalted peptide sample in 50 μ L low-pH MS buffer A. Vortex vigorously and allow the sample to rehydrate for 30 min on the benchtop with 1 min bath sonication. Measure the peptide concentration using Pierce[®] Quantitative Colorimetric Peptide Assay. Dilute the sample to a desired concentration with low-pH MS buffer A and centrifuge at 14,000 rpm for 10 min (*see Note 27*). Transfer the supernatant to an autosampler vial.
2. Perform LC-MS/MS analysis on an Orbitrap Exploris 240 mass spectrometer with Xcalibur v4.4 coupled to a Dionex Ultimate 3000 RSLCnano system. 5 μ L of each sample is loaded onto an Acclaim PepMap loading column for further peptide separation on an Acclaim PepMap RSLC. The peptide separation is conducted with a 2-h gradient method ranging from 5% to 25% low-pH MS buffer B (20% H₂O, 80% acetonitrile, and 0.1% formic acid) in low-pH MS buffer A at a flow rate of 0.3 μ L/min with a spray voltage at 2.1 kV.
3. A full MS scan is obtained every 2 s with a resolution of 120,000 and within the m/z range between 350 and 1,800. The parameters for MS scan are set as follows: RF lens of 65%, automatic gain control of 300%, automatic maximum injection time, 30% desired apex window, and with profile mode, dynamic exclusion, and isotopic exclusion enabled.
4. A variable number of MS2 scans with the resolution of 15,000 is obtained between each MS1 scan based on highest precursor masses and filtered for monoisotopic peak determination, theoretical precursor isotopic envelope fit, intensity ($5E4$), and charge state (2–6). The parameters for MS2 scans are set as follows: 75% automatic gain control, maximum injection time of 100 ms, and with centroid mode enabled. MS2 scan analysis is conducted by isolating precursor ions with the isolation window of 2 m/z , followed by higher-energy collision dissociation with 30% collision energy. Three injections are conducted for each TMT sample to serve as three technical replicates for one biological sample.

3.9 Analysis of LC-MS/MS Data

1. LC-MS/MS data is analyzed using the Thermo Proteome Discoverer v2.4 software package. Search the data with SequestHT and Percolator algorithms against the UniProt Protein Knowledgebase (UniProtKB) database of the *Homo sapiens* proteome [23, 24]. Specify trypsin as the protease with a maximum of two missed cleavages. Set the peptide precursor mass tolerance as 10 ppm with a fragment mass tolerance of 0.02 Da. Set methionine oxidation (+15.995 Da), protein N-terminus acetylation (+42.011 Da), and methionine loss (−131.040 Da) as dynamic modifications. Specify cysteine carbamidomethylation (+57.021 Da) and TMT sixplex modification on lysine residues or protein N-termini (+229.163 Da) as static modifications. Set the strict target false discovery rate for high-confident peptide hits as 0.01. Set TMT channel 126 as the control channel to generate protein abundance ratios between each fragment labeling conditions.
2. Search identified proteins against the UniProtKB database for subcellular localization and filter to retain only nuclear proteins [24]. Calculate the median value for the TMT ratios in each channel, then normalize all the ratios in that channel to its corresponding median (*see* **Note 28**). Next, quantify the average ratio values of the three technical replicates for the KB02 and KB05 labeling conditions and calculate the $\log_2$ values of these averaged ratios. Then, determine the p -values for changes in protein biotinylation levels using a two-tailed t -test for the ratios quantified in the three technical replicates and calculate the $-\log_{10}$ values of the p -values. Finally, use GraphPad Prism to plot the $\log_2$ ratio values on the x-axis and $-\log_{10} p$ -values on the y-axis, which generates a volcano plot showing changes in protein biotinylation levels after fragment labeling (Fig. 4). Proteins with significant increase or decrease in biotinylation upon fragment treatment are highlighted in red or blue, respectively ($\log_2(\text{fold change}) < -0.5$ or > 0.5 , $p < 0.01$, $-\log_{10}(p\text{-value}) > 2$).

4 Notes

1. The optimal cell density for transfection is 70–80% confluence as cells that are actively dividing are better at taking up foreign nucleic acids than quiescent cells. Excessive cell density can lead to contact inhibition, while insufficient cell density can enhance the toxicity effect of Lipofectamine on cells.
2. pLenti-His-TID is the construct of interest, while pLenti-eGFP serves as a positive control for lentiviral transfection.

3. HEK cells can easily detach from cell culture dishes, so carefully add media along the side of the dish to avoid cell detachment.
4. For the three plates of HeLa cells, two are for lentiviral transduction and the other one is for verifying successful antibiotic selection in the following steps.
5. 1 $\mu\text{g}/\text{mL}$ puromycin is the concentration that can kill wild-type HeLa cells in 3 days to select transduced cells. If transducing a different cell line, or if the lentiviral plasmid carries a different antibiotic resistance gene, test various antibiotic concentrations beforehand to determine the appropriate concentration for selection.
6. Seeding more cells in the first well of each plate helps adjust the focus of the microscope. For the remaining wells, seeding 100 μL of 5 cells/mL results in 0.5 cells per well, ensuring that each well has the chance to receive either 0 or 1 cell. Invert the cell suspension every few minutes to avoid cells settling down at the bottom of the conical tube.
7. Cells tend to settle in the corner of the well, so look thoroughly within each well when checking for single-cell colony. Discard the wells that contain multiple colonies.
8. Check if cells have detached off the plate after trypsinization using a microscope. Be careful not to cross-contaminate individual single-cell colonies when expanding the cell populations.
9. The His-TID construct is ~ 52 kD, so a 12.5% SDS-PAGE gel and 100 min transfer time are appropriate for immunoblotting analysis of this construct. If expressing a construct of a different molecular weight, adjust the acrylamide concentration of SDS-PAGE gels and the membrane transfer time accordingly.
10. The His-TID construct harbors an HA-tag on its C-terminus, and GAPDH serves as a loading control.
11. For cells that can easily detach from the dishes, such as HEK 293 T cells, coat the dishes with poly-D-lysine before seeding them to avoid cell loss during extensive washes.
12. To avoid the formation of bubbles in the mounting reagent, do not shake or invert the antifade reagent bottle during the thawing process. Slowly cover the coverslip onto the mounting reagent.
13. Make a biotin stock solution of 200 mM in DMSO and calculate the volume needed for each plate. A concentration that is too high will lead to biotin crashing out, while a concentration that is too low can lead to high DMSO percentage in cell media.

14. Complete DMEM is rich in protein, which can result in protein labeling in the media instead of inside the cells by the scout fragments.
15. The protein input amount for this study is 2 mg per sample. The optimal protein input varies between different constructs in different cell lines and needs to be tested beforehand.
16. Given the extra volume brought up by TCA, add 1 mL DPBS to 2.2 g TCA to make 100% (w/v) TCA solution.
17. Air dry the sample until the acetone smell is gone. Completely drying the pellet will cause difficulty in re-solubilizing the protein pellet.
18. If the pellet is allowed to dry completely, this will take up to 5–6 cycles to heating and sonication.
19. Invert the bottle of streptavidin agarose to mix the settled beads. Do not put the pipet tip all the way down to the bottom when taking up the slurry as it will block the beads from being aspirated into the tip. 170 μ L of slurry should result in $\sim$ 100 μ L of streptavidin agarose.
20. Be careful not to accidentally aspirate off the streptavidin beads during the wash steps.
21. Urea will leak out of an Eppendorf tube with normal snap cap during overnight incubation, so it is recommended to use tubes with screw caps instead.
22. Incubation with DTT is to reduce any reversible cysteine oxidation or disulfide bonds. Resuspend the beads by inverting the tubes or flicking them with your fingers. Avoid vortexing streptavidin beads to prevent bead fragmentation.
23. The reaction with IA is to alkylate all reduced cysteines residues, which prevents cysteine re-oxidation and re-formation of disulfide bonds, as this process can hamper subsequent peptide identification by MS.
24. It is necessary to remove urea before TMT labeling since the primary amine groups on urea can react with TMT labeling reagents. Moreover, it is important to be careful not to remove any of the beads during the extensive wash steps.
25. Resuspend the lyophilized trypsin in 100 mM TEAB instead of the supplied trypsin resuspension buffer if the digested samples undergo TMT labeling directly, to prevent any interference with the subsequent TMT labeling.
26. Keep the TMT isobaric tags on ice after taking out from -20°C to prevent acetonitrile evaporation. Vortex and spin the tubes briefly before use to bring down the reagent to the bottom of the tube.

27. The optimal peptide concentration for MS analysis varies among different instruments and MS workflows. Loading too few peptides can lead to low protein coverage, while overloading can result in interference in MS2 fragmentation. If the peptide sample is overdiluted, re-dry the sample on SpeedVac and resuspend in a smaller volume of low-pH MS buffer A.
28. The overall levels of protein biotinylation should remain constant between DMSO, and the fragment-treated samples even though certain proteins show significant changes in biotinylation upon fragment treatment. The median value of the TMT ratios in each channel should be close to 1. Normalizing TMT ratios to the median values of their respective channels help eliminate human errors during the MS sample preparation.

Acknowledgements

This work was supported by NIH [R35GM134964](#) to E.W.

References

1. Lammerding J (2011) Mechanics of the nucleus. *Compr Physiol* 1(2):783–807. <https://doi.org/10.1002/cphy.c100038>
2. Macara IG (2001) Transport into and out of the nucleus. *Microbiol Mol Biol R* 65(4): 570–594. <https://doi.org/10.1128/mmr.65.4.570-594.2001>
3. Guo M, Hay BA (1999) Cell proliferation and apoptosis. *Curr Opin Cell Biol* 11(6): 745–752. [https://doi.org/10.1016/S0955-0674\(99\)00046-0](https://doi.org/10.1016/S0955-0674(99)00046-0)
4. Bentley DL (2014) Coupling mRNA processing with transcription in time and space. *Nat Rev Genet* 15(3):163–175. <https://doi.org/10.1038/nrg3662>
5. Clapier CR, Iwasa J, Cairns BR, Peterson CL (2017) Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* 18(7): 407–422. <https://doi.org/10.1038/nrm.2017.26>
6. Soutourina J (2018) Transcription regulation by the mediator complex. *Nat Rev Mol Cell Biol* 19(4):262–274. <https://doi.org/10.1038/nrm.2017.115>
7. Lu HY, Zhou QD, He J, Jiang ZL, Peng C, Tong RS, Shi JY (2020) Recent advances in the development of protein-protein interactions modulators: mechanisms and clinical trials. *Signal Transduct Target Ther* 5(1):213. <https://doi.org/10.1038/s41392-020-00315-3>
8. Mendiratta G, Ke E, Aziz M, Liarakos D, Tong M, Stites EC (2021) Cancer gene mutation frequencies for the U.S. population. *Nat Commun* 12(1):5961. <https://doi.org/10.1038/s41467-021-26213-y>
9. Chen XH, Zhang TT, Su W, Dou ZH, Zhao DP, Jin XD, Lei HW, Wang J, Xie XD, Cheng B, Li Q, Zhang H, Di CX (2022) Mutant p53 in cancer: from molecular mechanism to therapeutic modulation. *Cell Death Dis* 13(11):974. <https://doi.org/10.1038/s41419-022-05408-1>
10. Singh J, Petter RC, Baillie TA, Whitty A (2011) The resurgence of covalent drugs. *Nat Rev Drug Discov* 10(4):307–317. <https://doi.org/10.1038/nrd3410>
11. Dalton SE, Campos S (2020) Covalent small molecules as enabling platforms for drug discovery. *Chembiochem* 21(8):1080–1100. <https://doi.org/10.1002/cbic.201900674>
12. Davids MS, Brown JR (2014) Ibrutinib: a first in class covalent inhibitor of Bruton's tyrosine kinase. *Future Oncol* 10(6):957–967. <https://doi.org/10.2217/fon.14.51>
13. Harvey RD, Adams VR, Beardslee T, Medina P (2020) Afatinib for the treatment of mutation-positive NSCLC: a review of clinical findings. *J Oncol Pharm Pract* 26(6):1461–1474. <https://doi.org/10.1177/1078155220931926>

14. Skoulidis F, Li BT, Dy GK, Price TJ, Falchook GS, Wolf J, Italiano A, Schuler M, Borghaei H, Barlesi F, Kato T, Curioni-Fontecedro A, Sacher A, Spira A, Ramalingam SS, Takahashi T, Besse B, Anderson A, Ang A, Tran Q, Mather O, Henary H, Ngarmchamnanrith G, Friberg G, Velcheti V, Govindan R (2021) Sotorasib for lung cancers with KRAS p.G12C mutation. *N Engl J Med* 384(25):2371–2381. <https://doi.org/10.1056/NEJMoa2103695>
15. Weerapana E, Wang C, Simon GM, Richter F, Khare S, Dillon MBD, Bachovchin DA, Mowen K, Baker D, Cravatt BF (2010) Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* 468(7325):790–795. <https://doi.org/10.1038/nature09472>
16. Wang C, Weerapana E, Blewett MM, Cravatt BF (2014) A chemoproteomic platform to quantitatively map targets of lipid-derived electrophiles. *Nat Methods* 11(1):79–85. <https://doi.org/10.1038/nmeth.2759>
17. Backus KM, Correia BE, Lum KM, Forli S, Horning BD, González-Páez GE, Chatterjee S, Lanning BR, Teijaro JR, Olson AJ, Wolan DW, Cravatt BF (2016) Proteome-wide covalent ligand discovery in native biological systems. *Nature* 534(7608):570–574. <https://doi.org/10.1038/nature18002>
18. Thul PJ, Åkesson L, Wiking M, Mahdessian D, Geladaki A, Blal HA, Alm T, Asplund A, Björk L, Breckels LM, Bäckström A, Danielsson F, Fagerberg L, Fall J, Gatto L, Gnann C, Hober S, Hjelmare M, Johansson F, Lee S, Lindskog C, Mulder J, Mulvey CM, Nilsson P, Oksvold P, Rockberg J, Schutten R, Schwenk JM, Sivertsson Å, Sjöstedt E, Skogs M, Stadler C, Sullivan DP, Tegel H, Winsnes C, Zhang C, Zwahlen M, Mardinoglu A, Pontén F, von Feilitzen K, Lilley KS, Uhlén M, Lundberg E (2017) A subcellular map of the human proteome. *Science* 356(6340):eaal3321. <https://doi.org/10.1126/science.aal3321>
19. Peng Q, Weerapana E (2024) Profiling nuclear cysteine ligandability and effects on nuclear localization using proximity labeling-coupled chemoproteomics. *Cell Chem Biol* 31(3):550–564.e9. <https://doi.org/10.1016/j.chembiol.2023.11.010>
20. Rodriguez-Paredes M, Esteller M (2011) Cancer epigenetics reaches mainstream oncology. *Nat Med* 17(3):330–339. <https://doi.org/10.1038/nm.2305>
21. Chen P, Zhao J, Wang Y, Wang M, Long H, Liang D, Huang L, Wen Z, Li W, Li X, Feng H, Zhao H, Zhu P, Li M, Wang QF, Li G (2013) H3.3 actively marks enhancers and primes gene transcription via opening higher-ordered chromatin. *Genes Dev* 27(19):2109–2124. <https://doi.org/10.1101/gad.222174.113>
22. Stewart SA, Dykxhoorn DM, Palliser D, Mizuno H, Yu EY, An DS, Sabatini DM, Chen IS, Hahn WC, Sharp PA, Weinberg RA, Novina CD (2003) Lentivirus-delivered stable gene silencing by RNAi in primary cells. *RNA* 9(4):493–501. <https://doi.org/10.1261/rna.2192803>
23. Kall L, Canterbury JD, Weston J, Noble WS, MacCoss MJ (2007) Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat Methods* 4(11):923–925. <https://doi.org/10.1038/nmeth1113>
24. UniProt C (2021) UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res* 49(D1):D480–D489. <https://doi.org/10.1093/nar/gkaa1100>



An Overview of Analytical Approaches to Cancer Proteogenomics

Chad J. Creighton

Abstract

The molecular landscape of human cancers involves multiple omics layers of complexity, from genome to proteome and beyond. Cancer proteogenomics involves the integration of protein expression patterns with somatic DNA alterations. Recently, advances in mass spectrometry-based proteomic profiling technologies have enabled the generation of combined proteomic and multi-omic data for thousands of human tumors across dozens of studies. These data in the public domain can be utilized to give us a more complete picture of cancer-specific pathways and processes and identify gene candidates for therapeutic targeting. Many proteogenomic studies are ongoing involving various cancer types according to tissue or cell of origin, including studies to predict response to therapy. In addition, pan-cancer analyses across multiple studies can identify molecular commonalities, differences, and emergent themes across tumor lineages. Data integration can determine which gene alterations at the transcriptome level are translated to the protein level. A wealth of knowledge and analytical approaches developed historically to integrate gene transcription with genomic data can be readily applied to proteogenomic analyses. Here is provided an overview of higher-level analyses of proteogenomic datasets. Such analyses include defining proteomic subtypes of cancer, exploring the impact of somatic mutations and epigenetic modifications on protein expression, cataloging proteomic correlates of more aggressive disease or drug response, and identifying enriched pathways.

Key words Proteomics, Proteogenomics, Multi-omics, Cancer, Data integration

1 Introduction

Cancer as a disease involves multiple layers of molecular complexity, from the genome and epigenome to the regulation of gene expression. Gene expression can be regulated at both transcriptional and post-transcriptional levels [1], and transcript levels across cancers only partially predict the corresponding protein levels for most genes [2–6]. Major technological advances over the past 30 years have enabled researchers to examine several different biological omics layers of cancer cells and how the layers interact [7–10]. Begun in the wake of the completion of the Human Genome Project [11, 12], The Cancer Genome Atlas (TCGA) was a large-

scale collaborative effort to systematically characterize the genomic changes that occur in over 10,000 human cancers [13–15]. TCGA data included somatic single-nucleotide variants (SNVs) and insertions-deletions (INDELs) by whole-exome sequencing, copy number alterations (CNAs) by SNP arrays, DNA methylation by Illumina arrays, gene transcription by RNA-sequencing (RNA-seq), and protein expression by reverse-phase protein arrays (RPPA). The RPPA platform used for TCGA involved antibody-based techniques to profile on the order of 200–300 predefined protein features [16, 17], including total protein levels or post-translational modifications. In contrast, mass spectrometry (MS)-based proteomics captures around 10,000–15,000 total protein features (which could include different protein isoforms) and 30,000–50,000 phosphorylation features [18]. Taking advantage of recent technological advancements in MS-based proteomics technologies [18], the Clinical Proteomic Tumor Analysis Consortium (CPTAC), a large-scale collaborative effort initiated while TCGA data generation efforts were winding down, has generated MS-based proteomic profiling data combined with corresponding multi-omics data on over 1000 human tumors to date [5, 19].

Proteogenomics represents the intersection of proteomics and genomics. One application of proteogenomics involves using customized protein sequence databases generated from genomic and transcriptomic information to help identify novel peptides from MS-based proteomic data [20]. Cancer is a disease characterized by widespread somatic genomic alterations [21]. The application of proteogenomics to the study of cancer can involve integrating somatic genomic data with protein expression data to catalog what somatic alterations are reflected in the cancer proteome [2, 4, 5, 22, 23]. Ideally, cancer proteogenomics studies should involve ~100 tumor samples of a given cancer type as defined by tissue or cell of origin [19], as cancer involves extensive genetic heterogeneity both across different patients and within individual tumors [24, 25]. Where TCGA helped pioneer the concept of “pan-cancer” genomic analyses, whereby data from tumors of various tissues of origin are analyzed together to identify molecular commonalities, differences, and emergent themes across tumor lineages [13, 14, 17], pan-cancer studies have been subsequently carried out for MS-based proteomics by both CPTAC and independent groups [19, 23, 26–32]. Mirroring the eventual democratization of DNA microarray and DNA sequencing technologies, recent advancements in MS-based proteomics technologies are making the platform accessible to more and more laboratories and core facilities [33]. This development has resulted in an explosion of cancer proteogenomics studies and datasets in recent years (Fig. 1) [3–5, 34–67]. Public databases, such as the Proteomic Data Commons (PDC, <https://pdc.cancer.gov>) [19] and LinkedOmics [68], make MS-based proteogenomics data from dozens of

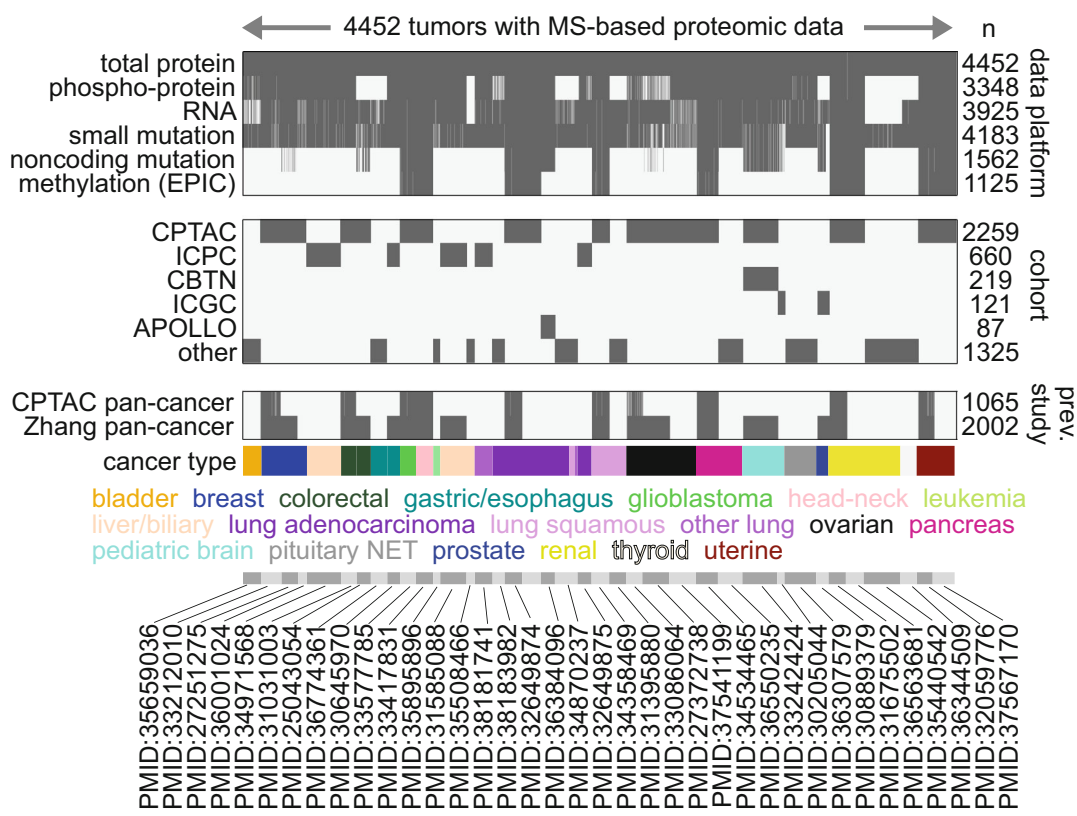


Fig. 1 Snapshot of the availability of cancer proteogenomic datasets in the public domain. Based on a survey in January 2024 of datasets available via public data repository or associated publication [3–5, 34–67]. Available data involve 4452 human tumors with mass spectrometry (MS)-based proteomic data combined with data from other omics platforms. The PubMed identifiers for the original publications associated with each dataset are indicated at the bottom. Sample and data availability numbers are subject to change and were based here on original reports and associated data repositories. CPTAC, Clinical Proteomic Tumor Analysis Consortium; ICPC, International Cancer Proteogenome Consortium; CBTN, Children’s Brain Tumor Network; ICGC, International Cancer Genomics Consortium; APOLLO, Applied Proteogenomics Organizational Learning and Outcomes. “CPTAC pan-cancer” from Ref. [19]; “Zhang pan-cancer” from Ref. [2]

published studies available for scientists to download and reanalyze with additional questions in mind. Additional proteogenomics data from other studies can be obtained from the supplementary data of the associated published papers [2, 69].

With the advent of DNA microarray technologies allowing for the simultaneous profiling of thousands of mRNAs in a sample [9, 10], analytical methods for extracting meaningful information from gene expression profiling data at the transcription level were developed [70, 71]. As additional data platforms representing other omics levels were implemented, studies integrating these additional levels with expression profiling data were readily carried out [72, 73]. Over several years, TCGA and other multi-omic cancer studies provided ample opportunities for data integration across

platforms to obtain novel insights [14, 74]. The wealth of knowledge and analytical approaches developed over the years to integrate gene transcription with genomic data can be readily applied to proteogenomic analyses. A sampling of studies from both independent groups and large-scale proteogenomics initiatives—including CPTAC, the International Cancer Proteogenome Consortium (ICPC), and the Applied Proteogenomics Organization Learning and Outcomes (APOLLO) Network—provides lots of examples of effective proteogenomic analyses and the types of questions that such analyses can address [69]. This chapter explores higher-level analyses of proteogenomic datasets, drawing examples from recent studies. Higher-level analyses would include defining proteomic subtypes of cancer, exploring the impact of somatic mutations and epigenetic modifications on protein expression, cataloging proteomic correlates of more aggressive disease or drug response, and identifying enriched pathways. Low-level analyses [75], including analyses of raw MS data, are not covered here, though many established tools exist to analyze raw MS-based proteomic data to produce identified and quantified proteins, phosphosites, acetyl sites, and ubiquityl sites [18].

2 Data Normalization

Low-level analysis of liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) data can generate a data table of protein features by tumor samples, with protein feature levels quantified in each sample. Depending on the MS-based proteomics platform utilized, protein levels may be represented as absolute values above zero or as a ratio to a reference. In either case, applying a log transformation to the data is recommended as this typically brings the data more in alignment with the assumptions for linear modeling or t -tests [76]. In practice, not all protein features would be detected for all samples. In contrast to RNA sequencing datasets, one should not assume that missing protein values denote no protein or very low protein levels. For proteomics datasets, missing values simply denote that we have no measurements for the feature in the given sample. With proteomic datasets from CPTAC and ICPC, about 80% of protein features represented have detected levels in at least 75% of samples profiled, and about 10% of features have detected levels in fewer than half of samples [2]. For some analysis approaches or software previously designed for the RNA-seq platform, missing values can be an issue. One way to deal with missing is by imputation [77], for example, by replacing any missing value with the mean of the protein feature for the other samples. However, imputation can attenuate observed correlations or associations and complicate analyses to identify protein expression outliers in cancer [22].

The different cancer studies carried out as part of CPTAC or ICPC involve different cancer types, with the proteomic data generated by different labs at different times, often using different platforms. These differences lead to “batch” effects in the proteomics datasets [19]. Batch effects lead to patterns in the data being revealed that involve technical rather than biological phenomena. Within each study dataset, batch effects should not be an issue. However, comparing protein feature levels between two different studies—for example, comparing renal and ovarian tumors—would reveal widespread differences in protein levels. Thus, when analyzing pan-cancer proteomics data across different studies, one must carefully adjust for batch effects across different cancer types. For RNA-seq data, the Combat algorithm [78, 79], now included as part of the Bioconductor SVA package, is one effective method for addressing batch effects. Given an expression dataset with batches (e.g., defined by time or laboratory) and comparison groups specified for each sample, Combat can generate an expression dataset with batch effects removed. However, for proteomics data, missing values would be an issue for methods like Combat, so data imputation is needed here. Also, the batches should each represent the various comparison groups involved. Otherwise, batch effects correction could remove the biological differences of interest. Other approaches to dealing with batch effects could include fitting a regression model to study the dependence of expression levels on other attributes, including batch effect indicators as covariates to account for batch-specific mean expression values [19].

With proteomics data now available for thousands of tumor samples across dozens of different studies (Fig. 1), there is utility in combining these individual datasets for pan-cancer analyses [2, 19]. My group has utilized a straightforward approach for combining all these datasets into one unified proteomics dataset [2, 17, 22, 28, 80, 81]. Within each independent dataset, we normalize expression values for each gene or protein feature across samples to standard deviations from the median so that no feature would have any real differences in relative levels between datasets. By making the expression values unitless, neither tissue-dominant differences nor inter-laboratory batch effects would drive analyses across the combined datasets. In addition, for some datasets, the average protein expression values across all protein features may vary substantially from sample to sample. One way to adjust for this is to normalize expression values within each proteomic profile to standard deviations from the median. Quantile normalization is another approach for between-sample normalization [82], though data imputation may be needed to apply this to proteomic data. With a compendium dataset comprising individual datasets with inter-dataset differences removed as described above, protein-level comparisons by tissue-based cancer type are not possible, since all cancer proteogenomics studies to date have each involved a single

cancer type. However, protein-level comparisons by cancer type are often not of primary interest as there would be widespread expression differences between any two tissues due to cell differentiation but not representing disease drivers. At the same time, as described below, other comparisons that would cut across cancer types and molecular subtype discovery can be readily carried out on a proteomic compendium dataset.

3 Molecular Subtype Discovery

Cancer is a heterogeneous disease, and multiple distinct disease entities usually exist within a given cancer type as defined by tissue or cell of origin [8, 81, 83]. Cancer subtypes can be defined using molecular profiling data from any omics platform. For many cancer types, initial studies defined molecular subtypes using gene transcription profiling data based on expression arrays [8]. Subsequent multi-omic studies using TCGA data could define cancer subtypes for different omics levels. TCGA studies helped pioneer a “cluster-of-clusters” approach, whereby the different platform-level subtypes are all combined to define a set of consensus subtypes reflecting contributions from each of the omics platforms [74, 84–86]. Across 9716 tumors, TCGA consortium defined 28 pan-cancer subtypes by multi-omic consensus clustering [84]. These pan-cancer subtypes closely followed the cancer tissue of origin, which would not be surprising given the widespread molecular differences involving cell and tissue type, although a TCGA pan-cancer squamous subtype would span multiple tissues of origin [74]. In contrast, we applied a different approach to defining TCGA pan-cancer subtypes [81], whereby we first normalized mRNA expression patterns within each cancer type to standard deviations from the median and then used the combined normalized datasets to define 10 pan-cancer subtypes spanning tissue and cell type. Except for one pan-cancer subtype representing basal-like breast cancer, our subtypes each involved multiple tissue-based cancer types. In contrast to subtypes reflecting the cell of origin, our pan-cancer subtypes represented pathways, processes, and tumor microenvironmental influences transcending tumor lineage.

Many analytical methods for molecular subtyping of tumors developed for gene transcription data can be applied to proteomic data. Subtyping methods include hierarchical clustering and k-means clustering [87, 88]. Standard aggregative hierarchical clustering represents the data in the shape of a binary tree, where the most similar patterns are clustered in a hierarchy of nested subsets [8]. Non-hierarchical methods such as k-means clustering start with a predefined number of clusters and, by iterative reallocation of cluster members, minimize the overall intra-cluster dispersion [87]. The number of subtypes best represented in a given dataset

can be an open question. Consensus Clustering, originally developed for mRNA expression data to aid subtype discovery, is a resampling-based method for class discovery and visualization [89]. Consensus Clustering begins by subsampling a proportion of items and a proportion of features from a data matrix, and each subsample is then partitioned into up to k groups by a user-specified clustering algorithm, with the process repeated for a specified number of repetitions. ConsensusClusterPlus, which implements the Consensus Cluster method in R, is a popular subtyping tool used in TCGA and CPTAC studies [39, 74, 90]. ConsensusClusterPlus generates several graphics illustrating the degree of consensus for each k , from which the user can select an optimal k beyond which there is no substantial improvement in the overall degree of consensus. One practical issue with ConsensusClusterPlus is that it may not scale up to very large datasets with more than 2000 tumor samples [81]. In this case, a random sampling of profiles from the larger dataset might be used to first define the subtypes, with a gene classifier based on the subtypes then applied to the other sample profiles.

In recent cancer proteomics studies from my group, we combined data from individual studies involving different cancer types (normalizing the data as described above) to define pan-cancer molecular subtypes [2, 28]. Our earlier study [28] classified 532 cancers, representing six tissue-based types (breast, colon, ovarian, renal, uterine), into 10 proteome-based, pan-cancer subtypes. In a subsequent study [2], we assembled a compendium dataset of MS-based proteomics data of 2002 primary tumors from 14 cancer types and 17 studies, which tumors we classified into 11 pan-cancer subtypes. The 11 subtypes from our second study largely recapitulated the subtype-associated patterns of our first study, with additional proteome-based subtypes also identified using the larger dataset involving more samples and cancer types. Some molecular subtypes we identified at the proteome level were not manifested at the mRNA level. These included an immune-based subtype featuring complement proteins and additional subtypes involving proteins related to Golgi apparatus, hemoglobin complex, and endoplasmic reticulum [28]. Here, in this overview, we assembled another expanded proteomics dataset of 4334 tumors from over 35 studies and then classified each tumor into one of the 11 subtypes from the 2002 tumor study (Fig. 2). As we can observe, each proteome-based subtype involves multiple tissue-based cancer types, and the additional proteomic profiles show differential expression values consistent with their assigned subtype. With larger proteomic compendium datasets, additional pan-cancer subtypes might be discoverable, representing future work.

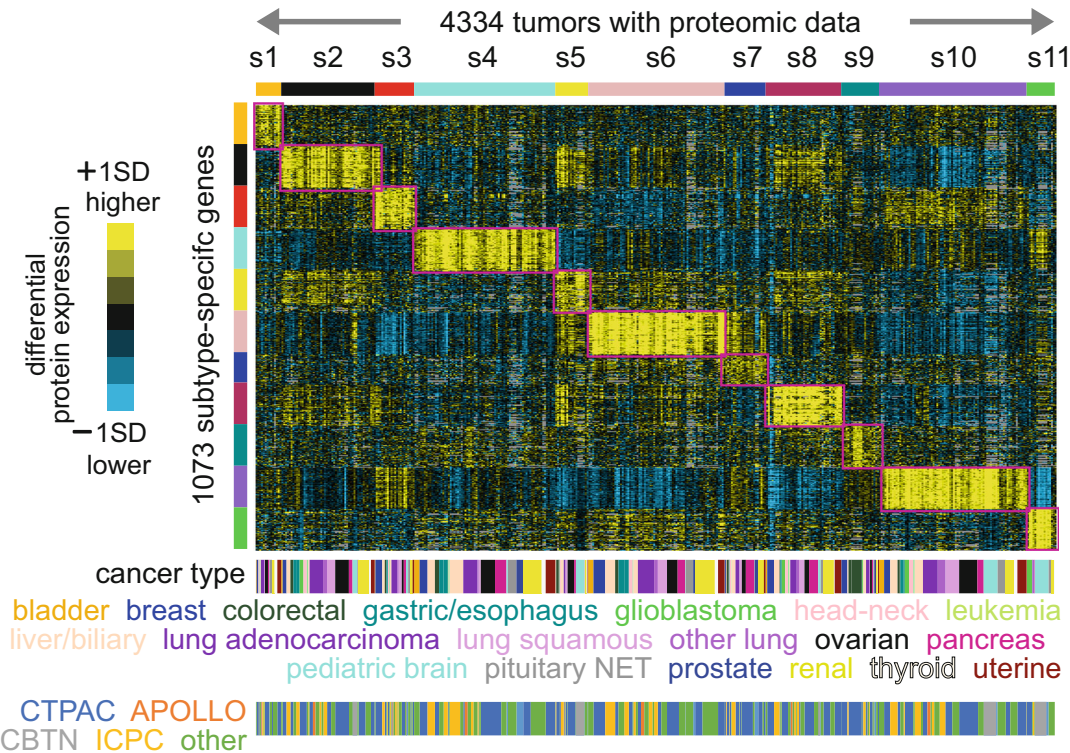


Fig. 2 Pan-cancer proteomic subtypes. In a previous study [2], we used ConsensusClusterPlus [90] to call 11 molecular subtypes based on MS-based proteomic profiling of 2002 human tumors. A set of 1073 total protein features that best distinguished the subtypes were previously found. Here, we examine the 11 subtypes and associated differential proteins in an extended proteomics compendium dataset of 4334 human tumor proteomic profiles taken from 37 studies [3–5, 34–67]. Like what we carried out for our previously reported 2002 tumor proteomics compendium dataset, we normalized the protein expression values within each dataset to standard deviations (SDs) from the median for each protein across tumor profile, thereby removing any laboratory batch-specific and cancer type-specific patterns. We then classified each additional tumor profile according to one of the 11 subtypes using classification methods previously described [2, 28]. Similar to previous observations, the pan-cancer subtypes each span multiple cancer types and multiple independent studies. The 118 tumors in the CPTAC lung adenocarcinoma “Confirmatory” dataset, referenced in Fig. 1 as part of the 4452 tumors with available proteomics data, did not have proteomics data publicly released at the time of writing, though these data were pending

4 Differential Expression Analyses

Individual proteins or protein groups with differential levels between tumor comparison groups may provide clues as to the underlying biological phenomena that distinguish the groups. In practice, protein-level differential analyses can be generated using a *t*-test on log2-transformed protein values (or values normalized from log2-transformed values). For correlating protein expression with a continuous variable across tumors, a Pearson’s correlation using log2-transformed values would suffice. It is a given that

parametric tests, such as *t*-tests and Pearson's correlations, work off some assumptions of the data, for example, that the data are normally distributed and that the comparison groups have equal variances. The log₂ transform would bring the protein data more in line with model assumptions [76]. When the data do not fit well within the model assumptions, it can be more difficult for the model to identify significance patterns. At the same time, *t*-tests do not require normally distributed data to perform reasonably well in most circumstances [91], though they often may not detect significance in the case of extreme outliers, which the log₂-transform can alleviate. Nonparametric tests do not rely upon model assumptions or make minimal assumptions. At the same time, one notable drawback with nonparametric tests is that they need much more data and are underpowered in situations where a parametric test can perform well [91]. Also, nonparametric methods such as Wilcoxon rank sum would not factor in the amount of change between groups. Where many data points are available, *t*-tests and Wilcoxon rank sum tests tend to give similar answers.

Multiple testing comes into play when testing up to tens of thousands of protein features for differential levels or correlation associations. For example, when testing 1000 features, we expect about 10 to have a nominal *p*-value < 0.01, representing 1% of the features tested having a significant *p*-value by chance alone. Therefore, we cannot rely solely on the nominal *p*-value when attaching significance to genes but instead must consider the false discovery rate (FDR). The protein-level FDR is not a *p*-value but rather would apply to the entire set of proteins meeting a given statistical cutoff. The FDR denotes the percentage of the top proteins likely to be false positives. For example, FDR < 10% indicates that ~10% of the proteins at that cutoff would be false positives, but on the other hand, ~90% should represent true positives. One commonly used method to estimate FDR in expression profiling studies [70] uses the following formula for each feature: [(nominal *p*-value) X (total number of features tested)] / (total number of features that were significant at the given *p*-value)]. What FDR cutoff to use would be a balance of limiting both false negatives and false positives. A stricter FDR cutoff may be preferred for global molecular associations involving a single analysis using one dataset. However, when bringing additional information to evaluate top protein features, a more relaxed cutoff can be employed to limit false negatives. Additional information might include another set of top genes or proteins from an independent source. A biologist can also use domain knowledge to note top proteins of likely relevance to the system under study.

5 Pathway Enrichment Analyses

Top proteins arising from differential analyses can be searched for enriched pathways or functional gene groups. Such enrichment analyses can provide clues that might explain pathways or processes that underlie the system under study. Curated gene sets for carrying out enrichment analysis would include Gene Ontology [92]—which uses a controlled vocabulary to annotate each gene according to its molecular function, location, and associated biological processes—and the Molecular Signature Database (MSigDB) [93]—which curates gene sets of potential interest based on the literature and other sources such as wikiPathways [94]. Two commonly used approaches to enrichment analysis include one-sided Fisher’s exact tests and Gene Set Enrichment Analysis (GSEA). One-sided Fisher’s exact tests compute the significance of overlap between two independently derived gene sets. For this method, a statistical cutoff would be applied to the expression data to define top genes. In contrast, GSEA does not rely on a strict cutoff but instead involves ranking all genes in the expression dataset based on some metric (e.g., differential comparison statistic or correlation with a continuous variable). The GSEA method can identify significant enrichment patterns even when the absolute expression changes are subtle and differential gene sets involved do not have low FDRs [95]. User-friendly software for carrying out GSEA is available [93]. Software for carrying out pathway enrichment by one-sided Fisher’s exact tests include DAVID [96] and SigTerms [97].

6 Gene Signature Analyses

A gene signature represents an entire set of genes that are relationally associated with each other in some way. A gene signature may be defined in any number of ways, including gene annotation (e.g., by Gene Ontology [92]), curated pathway-associated gene sets (e.g., by MSigDB [93] or wikiPathways [94]), or using molecular profiling datasets (e.g., determining what genes go up or down in response to treatment in an experimental model, using a specified selection cutoff). Early transcriptomic profiling experiments of cancer could define gene expression signatures of worse patient outcomes, whereby a top set of genes with associations with outcome in a “discovery” dataset could also predict patient outcomes in one or more “validation” datasets not used to derive the signature [98, 99]. In applying gene signatures to molecular profiling datasets, the relative direction of each gene in the signature (e.g., up versus down, or positive versus negative association) should be taken into account. In contrast, many early transcriptomic studies

used hierarchical clustering to apply gene expression signatures to external expression datasets [100], where such unsupervised clustering approaches do not utilize gene signature direction. With unsupervised approaches, a randomly selected set of genes may show the same type of associations in the external dataset as what the gene signature of interest may show [101].

When applying gene signatures to a proteomic dataset, a good practice is for the protein feature levels to be normalized across samples, for example, to standard deviations from the sample median (for each protein value, subtract the median and then divide by the standard deviation). This normalization would correct for differences in overall protein levels between datasets from different laboratories or platforms and give each protein feature an equal weight in the signature scoring. For gene signatures where all the genes are supposed to be moving together in the same direction, the average of the signature genes in each profile of the external dataset can represent the signature score. For gene signatures where some genes are supposed to move in one direction and the other genes to move in the opposite direction, one way to compute the signature score for each profile would be the average of the “up” genes minus the average of the “down” genes [102]. Another way to compute the signature score is a “*t*-score” method utilized in several of our studies [2, 103–107]. Our *t*-score metric is defined as the two-sided *t*-statistic comparing, within each tumor expression profile, the average of the “up” signature genes with the average of the “down” signature genes. For a given tumor profile, the *t*-score is high when both the “up” genes in the signature are high and the “down” genes are low.

7 Proteome–Genome Integration

Proteogenomics represents the integration between proteome and genome. For example, somatic copy number alterations in genes are usually associated with corresponding changes in the measured protein levels of the genes, particularly where the correlation between protein and mRNA levels is high [2, 4, 5]. We have consistently found that somatic alterations of cancer-associated pathways are reflected in both the cancer transcriptome and proteome, whereby tumors with somatic alteration involving genes in a pathway tend to show higher levels of expression-based signature scoring for that pathway [2, 81, 108]. For example, our multiplatform-based survey of the PI3K/AKT/mTOR pathway across over 10,000 human cancers in TCGA cohort [108] involved the RPPA platform for which PI3K/AKT/mTOR signaling proteins were well-represented [16]. For each tumor in TCGA, we summarized RPPA signatures for PI3K/AKT and mTOR into activity scores. Using domain knowledge of PI3K/AKT/mTOR,

we regarded mTOR signaling as a separate pathway from PI3K/AKT, as proteins independent of PI3K/AKT can act upon mTOR. Proteomic scoring based on the PI3K/AKT and mTOR signatures was entirely distinct from transcriptomic scoring based on mRNA data, as, for example, the proteomic signatures included phosphoproteins denoting protein activation that would not be measurable by mRNA. In TCGA cohort, somatic mutations in genes in the canonical PI3K/AKT/mTOR pathway were strongly associated with PI3K/AKT/mTOR activation as measured at the protein level. For example, copy number loss, insertions-deletions (indels), and single-nucleotide variants (SNVs) in *PTEN* were each associated with higher PI3K/AKT and mTOR signaling, while *STK11* mutations were associated with higher mTOR but not PI3K/AKT signaling.

A subset of tumors may show high protein-level activation of a given pathway without harboring any somatic mutations involving canonical pathway genes. These tumors could have mutations in other genes not normally associated with the pathway, suggesting additional, unexplained, or underappreciated mechanisms of pathway activation. In our above-noted PI3K/AKT/mTOR proteogenomic study [108], we focused on the tumors with high phospho-AKT but lacking a DNA alteration classically associated with PI3K/AKT activation, and we found that these tumors were highly enriched for *IDH1/2* and *VHL* mutations. Notably, when surveying the literature, VHL was recently found to suppress AKT activity directly [109], and the generation of 2-hydroxyglutarate by mutated *IDH1/2* was found to lead to the activation of mTOR [110]; our study highlighted the importance of both of the above relationships in the setting of human cancer. In another study [2], we combined MS-based proteomics data and somatic mutation data on 1597 tumors to identify tumors with high MYC proteomic signature scores but lacking amplification of MYC family genes. Looking at somatic SNV/indel data, we identified 12 genes enriched for mutation events in these tumors, seven of these genes having a demonstrated role for their loss leading to increased MYC expression or activity, based on published experimental studies. While one or two published studies can support a pathway-specific role for a gene, such findings often need help gaining traction. Thus, our cancer proteogenomic surveys can highlight additional genes for consideration as part of the “canon” of what would be recognized to constitute the core standard model of well-studied pathways such as PI3K/AKT/mTOR and MYC.

Noncoding somatic mutations—as discoverable using whole genome sequencing or WGS—can also impact gene regulation. Regarding somatic SNVs, the cancer genome harbors relatively fewer “hotspot” noncoding mutations impacting gene expression, though these would include *TERT* promoter SNVs involving

multiple cancer types [111, 112]. Somatic structural variation in cancer, stemming from genomic rearrangements, is one major driver of altered transcription through CNAs, gene fusions, and altered cis-regulation [22, 113–116]. However, not all genes deregulated at the mRNA level would be translated at the protein level. In our proteogenomics study of 1307 tumors with combined MS-based proteomics data and WGS [22], we determined the extent to which somatic structural variant (SV) breakpoint patterns impact protein expression of nearby genes. In this cohort, 657 genes had a significant SV association with altered gene transcription, involving SV breakpoints within 1 Mb of the gene. Of these 657 genes, just 26% had a corresponding association at the protein level. While this overlap between mRNA- and protein-based results was highly nonrandom, this study underscores the importance of integrating protein with mRNA data into genomics studies. However, missing values involving MS-based proteomics can be an issue, particularly when looking for tumor expression outliers showing very high levels for a particular gene [22]. We have developed software called “SVExpress” [117] for integrating SV with gene expression data, which can be applied to proteomic data in the same way as mRNA data. Other open software for integrating SV with expression data include cis-X [118] and SV-HotSpot [119].

Tumor-specific somatic mutations could result in the production of mutated peptides presented on the cell surface by the MHC molecules [120]. Such neoantigens represent a promising target for immunotherapy. Several available bioinformatics pipelines can generate a list of neoantigens from combined whole-exome sequencing and RNA-seq data. High-throughput MS is another option for directly identifying MHC-bound peptides [121, 122]. Limitations of MS include low sensitivity and reproducibility, particularly for low-abundance peptides including tumor-specific neoantigens. MS-based identification of predicted neoantigens from mutation and RNA-seq data should involve immunoprecipitation of MHC-peptide complexes for greater sensitivity. Pipelines for neoantigen prediction workflow integrating genomic and MS data include ProGeo-neo [123].

8 Prognostic and Predictive Markers

Molecular correlates of more aggressive cancers can provide clues as to the pathways and processes involved. As part of TCGA, we carried out a multi-omic integration of proteomic, transcriptomic, and DNA methylation data to define molecular correlates of patient overall survival in clear cell renal cell carcinoma. When viewed in the context of metabolism, these molecular survival correlates highlighted a widespread metabolic shift toward a “Warburg effect”-like

state for aggressive cancers [106]. For human tumor molecular datasets that include long-term follow-up data for the patients, the association of each molecular feature with time-to-adverse event (e.g., death or relapse) can be determined by univariate Cox. Multivariate Cox can incorporate additional covariates of interest to determine if a molecular association with survival cannot be explainable by other factors. For many proteomic studies, the patient follow-up times may be limited such that there is insufficient statistical power for Cox or log-rank tests. In these instances, other clinical variables of aggressive disease can be utilized, such as tumor grade (the degree of differentiation of the cancer cells) or tumor stage (indicating how extensively the tumor has spread outside of its site of origin). In one study [27], we defined differentially expressed proteins and mRNAs associated with higher tumor grade or higher tumor stage for each of seven cancer types. We found widespread molecular associations with tumor grade and relatively fewer associations with tumor stage. For pan-cancer analyses, another approach to dealing with limited patient survival data is multivariate Cox analyses incorporating cancer type [22, 108], which benefits from the increased patient numbers while identifying survival associations that would not be explained by tissue-of-origin or histology alone.

In addition to prognostic markers (i.e., markers associated with better or worse patient outcomes), there is interest in identifying therapeutically predictive markers. Therapeutically predictive markers can predict to what extent the patient may benefit from a particular therapy. Well-known therapeutically predictive protein markers include estrogen receptor and HER2 in breast cancer [8]. Currently, there is high interest in identifying predictive biomarkers for immunotherapy or chemotherapy [124, 125]. A recent proteogenomics study identified protein markers of chemotherapy resistance and response in triple-negative breast cancer [34]. Another recent study identified a proteomic signature of resistance to platinum-based therapy in ovarian cancer [38]. In addition, multi-omics data on cancer cell lines, including MS-based proteomics and drug sensitivity data, can be leveraged to identify molecular markers of drug response [69, 126, 127]. Gene signatures and molecular subtypes, as defined using human tumor data, can be applied to cancer cell line proteomic datasets to identify associations with drug response [22, 80]. Other relevant resources to integrate with cell line proteomic data include the DepMap datasets, involving CRISPR assays to determine which genes are essential for cell proliferation in which cell lines [22, 69, 128]. There is a need for additional molecular datasets from human tumors with treatment data, including long-term response data. The most robust therapeutically predictive proteomic markers

would be those arising from integrating human tumor data and experimental data [27]. Top prognostic or predictive molecular correlates can be examined for enriched pathways and applied to additional molecular datasets, as described above.

9 Concluding Remarks

Analytical approaches developed historically to integrate gene transcription with genomic data can be readily applied to proteogenomic analyses. Publicly available software tools, many of which have been incorporated into Bioconductor in R [129], can be utilized for data integration. Bioinformatics analysis tools can also include Excel for integrating results between data tables [75]. The use of effective visualization of multi-omic results is critical [16] and can include use of software for generating heat maps [130, 131] and Adobe Illustrator (or equivalent freeware such as Inkscape). Effective proteogenomics analysis requires dedicated bioinformatics experts with extensive experience in applied data analysis and statistics.

Acknowledgments

This work was supported by the National Institutes of Health (NIH) grant P30CA125123 (C. Creighton).

References

1. Delgado M, León J (2006) Gene expression regulation and cancer. *Clin Transl Oncol* 8(11):780–787
2. Zhang Y, Chen F, Chandrashekar D, Varambally S, Creighton C (2022) Proteogenomic characterization of 2002 human cancers reveals pan-cancer molecular subtypes and associated pathways. *Nat Commun* 13(1):2669. <https://doi.org/10.1038/s41467-022-30342-3>
3. Zhang H, Liu T, Zhang Z, Payne S, Zhang B, McDermott J, Zhou J, Petyuk V, Chen L, Ray D, Sun S, Yang F, Chen L, Wang J, Shah P, Cha S, Aiyetan P, Woo S, Tian Y, Gritsenko M, Clauss T, Choi C, Monroe M, Thomas S, Nie S, Wu C, Moore R, Yu K, Tabb D, Fenyö D, Bafna V, Wang Y, Rodriguez H, Boja E, Hiltke T, Rivers R, Sokoll L, Zhu H, Shih I, Cope L, Pandey A, Zhang B, Snyder M, Levine D, Smith R, Chan D, Rodland K, CPTAC Investigators (2016) Integrated proteogenomic characterization of human high-grade serous ovarian cancer. *Cell* 166(3):755–765
4. Mertins P, Mani D, Ruggles K, Gillette M, Clauser K, Wang P, Wang X, Qiao J, Cao S, Petralia F, Kawaler E, Mundt F, Krug K, Tu Z, Lei J, Gatza M, Wilkerson M, Perou C, Yellapantula V, Huang K, Lin C, McLellan M, Yan P, Davies S, Townsend R, Skates S, Wang J, Zhang B, Kinsinger C, Mesri M, Rodriguez H, Ding L, Paulovich A, Fenyö D, Ellis M, Carr S, NCI CPTAC (2016) Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature* 534(7605):55–62
5. Zhang B, Wang J, Wang X, Zhu J, Liu Q, Shi Z, Chambers M, Zimmerman L, Shaddox K, Kim S, Davies S, Wang S, Wang P, Kinsinger C, Rivers R, Rodriguez H, Townsend R, Ellis M, Carr S, Tabb D, Coffey R, Slebos R, Liebler D, NCI CPTAC (2014) Proteogenomic characterization of human colon and rectal cancer. *Nature* 513(7518):382–387

6. Chen G, Gharib T, Huang C, Taylor J, Misek D, Kardias S, Giordano T, Iannettoni M, Orringer M, Hanash S, Beer D (2002) Discordant protein and mRNA expression in lung adenocarcinomas. *Mol Cell Proteomics* 1(5):304–313. <https://doi.org/10.1074/mcp.m200008-mcp200>
7. Raufaste-Cazavieille V, Santiago R, Droit A (2022) Multi-omics analysis: paving the path toward achieving precision medicine in cancer treatment and immuno-oncology. *Front Mol Biosci* 9:962743
8. Perou C, Sørli T, Eisen M, van de Rijn M, Jeffrey S, Rees C, Pollack J, Ross D, Johnsen H, Akslen L, Fluge O, Pergamenschikov A, Williams C, Zhu S, Lønning P, Børresen-Dale A, Brown P, Botstein D (2000) Molecular portraits of human breast tumours. *Nature* 406(6797):747–752. <https://doi.org/10.1038/35021093>
9. Lockhart D, Dong H, Byrne M, Follett M, Gallo M, Chee M, Mittmann M, Wang C, Kobayashi M, Horton H, Brown E (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* 14(13):1675–1680
10. Schena M, Shalon D, Davis R, Brown P (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270(5235):467–470
11. International_Human_Genome_Sequencing_Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* 409(6822):860–921
12. Venter J, Adams M, Myers E, Li P, Mural R, Sutton G, Smith H, Yandell M, Evans C, Holt R, Gocayne J, Amanatides P, Ballew R, Huson D, Wortman J, Zhang Q, Kodira C, Zheng X, Chen L, Skupski M, Subramanian G, Thomas PJZ, Gabor Miklos G, Nelson C, Broder S, Clark A, Nadeau J, McKusick V, Zinder N, Levine A, Roberts R, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley EBK, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R (2001) The sequence of the human genome. *Science* 291(5507):1304–1351
13. Ding L, Bailey M, Porta-Pardo E, Thorsson V, Colaprico A, Bertrand D, Gibbs D, Weerasinghe A, Huang K, Tokheim C, Cortés-Ciriano I, Jayasinghe R, Chen F, Yu L, Sun S, Olsen C, Kim J, Taylor A, Cherniack A, Akbani R, Suphailai C, Nagarajan N, Stuart J, Mills G, Wyczalkowski M, Vincent B, Hutter C, Zenklusen J, Hoadley K, Wendl M, Shmulevich L, Lazar A, Wheeler D, Getz G, Cancer Genome Atlas Research Network (2018) Perspective on oncogenic processes at the end of the beginning of cancer genomics. *Cell* 173(2):305–320
14. Cancer Genome Atlas Research Network, Weinstein J, Collisson E, Mills G, Shaw K, Ozenberger B, Ellrott K, Shmulevich I, Sander C, Stuart J (2013) The cancer genome atlas Pan-cancer analysis project. *Nat Genet* 45(10):1113–1120
15. Cancer Genome Atlas Research Network (2008) Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 455(7216):1061–1068
16. Creighton C, Huang S (2015) Reverse phase protein arrays in signaling pathways: a data integration perspective. *Drug Des Devel Ther* 9:3519–3527
17. Akbani R, Ng P, Werner H, Shahmoradgoli M, Zhang F, Ju Z, Liu W, Yang J, Yoshihara K, Li J, Ling S, Seviour E, Ram P, Minna J, Diao L, Tong P, Heymach J, Hill S, Dondelinger F, Städler N, Byers L, Meric-Bernstam F, Weinstein J, Broom B, Verhaak R, Liang H, Mukherjee S, Lu Y, Mills G (2014) A pan-cancer proteomic perspective on The Cancer Genome Atlas. *Nat Commun* 5(1):3887
18. Mani D, Krug K, Zhang B, Satpathy S, Clauser K, Ding L, Ellis M, Gillette M, Carr S (2022) Cancer proteogenomics: current impact and future prospects. *Nat Rev Cancer* 22(5):298–313. <https://doi.org/10.1038/s41568-022-00446-5>
19. Li Y, Dou Y, Da Veiga Leprevost F, Geffen Y, Calinawan A, Aguet F, Akiyama Y, Anand S, Birger C, Cao S, Chaudhary R, Chilappagari P, Cieslik M, Colaprico A, Zhou D, Day C, Domagalski M, Esai Selvan M, Fenyö D, Foltz S, Francis A, Gonzalez-Robles T, Gümüş Z, Heiman D, Holck M, Hong R, Hu Y, Jaehnig E, Ji J, Jiang W, Katsnelson L, Ketchum K, Klein R, Lei J, Liang W, Liao Y, Lindgren C, Ma W, Ma L, MacCoss M, Martins Rodrigues F, McKerrow W, Nguyen N, Oldroyd R, Pilozi A, Pugliese P, Reva B, Rudnick P, Ruggles K, Rykunov D, Savage S, Schnaubelt M, Schraink T, Shi Z, Singhal D, Song X, Storrs E, Terekhanova N, Thangudu R, Thiagarajan M, Wang L, Wang J, Wang Y, Wen B, Wu Y,

- Wyczalkowski M, Xin Y, Yao L, Yi X, Zhang H, Zhang Q, Zuhl M, Getz G, Ding L, Nesvizhskii A, Wang P, Robles A, Zhang B, Payne S, Clinical Proteomic Tumor Analysis Consortium (2023) Proteogenomic data and resources for pan-cancer analysis. *Cancer Cell* 41(8):1397–1406
20. Nesvizhskii A (2014) Proteogenomics: concepts, applications and computational strategies. *Nat Methods* 11(11):1114–1125
 21. Hanahan D, Weinberg R (2011) Hallmarks of cancer: the next generation. *Cell* 144(5):646–674
 22. Chen F, Zhang Y, Chandrashekar D, Varambally S, Creighton C (2023) Global impact of somatic structural variation on the cancer proteome. *Nat Commun* 14(1):5637
 23. Li Y, Porta-Pardo E, Tokheim C, Bailey M, Yaron T, Stathias V, Geffen Y, Imbach K, Cao S, Anand S, Akiyama Y, Liu W, Wyczalkowski M, Song Y, Storrs E, Wendl M, Zhang W, Sibai M, Ruiz-Serra V, Liang W, Terekhanova N, Rodrigues F, Clauser K, Heiman D, Zhang Q, Aguet F, Calinawan A, Dhanasekaran S, Birger C, Satpathy S, Zhou D, Wang L, Baral J, Johnson J, Huntsman E, Pugliese P, Colaprico A, Iavarone A, Chheda M, Ricketts C, Fenyö D, Payne S, Rodriguez H, Robles A, Gillette M, Kumar-Sinha C, Lazar A, Cantley L, Getz G, Ding L, Clinical Proteomic Tumor Analysis Consortium (2023) Pan-cancer proteogenomics connects oncogenic drivers to functional states. *Cell* 186(18):3921–3944
 24. Gerlinger M, Rowan A, Horswell S, Larkin J, Endesfelder D, Gronroos E, Martinez P, Matthews N, Stewart A, Tarpey P, Varela I, Phillimore B, Begum S, McDonald N, Butler A, Jones D, Raine K, Latimer C, Santos C, Nohadani M, Eklund A, Spencer-Dene B, Clark G, Pickering L, Stamp G, Gore M, Szallasi Z, Downward J, Futreal P, Swanton C (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 366(10):883–892
 25. Lawrence M, Stojanov P, Mermel C, Robinson J, Garraway L, Golub T, Meyerson M, Gabriel S, Lander E, Getz G (2014) Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 505(7484):495–501
 26. Chen F, Chandrashekar D, Scheurer M, Varambally S, Creighton C (2022) Global molecular alterations involving recurrence or progression of pediatric brain tumors. *Neoplasia* 24(1):22–33. <https://doi.org/10.1016/j.neo.2021.11.014>
 27. Monsivais D, Vasquez Y, Chen F, Zhang Y, Chandrashekar D, Faver J, Masand R, Scheurer M, Varambally S, Matzuk M, Creighton C (2021) Mass-spectrometry-based proteomic correlates of grade and stage reveal pathways and kinases associated with aggressive human cancers. *Oncogene* 40(11):2081–2095. <https://doi.org/10.1038/s41388-021-01681-0>
 28. Chen F, Chandrashekar D, Varambally S, Creighton C (2019) Pan-cancer molecular subtypes revealed by mass-spectrometry-based proteomic characterization of more than 500 human cancers. *Nat Commun* 10(1):5679. <https://doi.org/10.1038/s41467-019-13528-0>
 29. McKerrow W, Wang X, Mendez-Dorantes C, Mita P, Cao S, Grivainis M, Ding L, LaCava J, Burns K, Boeke J, Fenyö D (2022) LINE-1 expression in cancer correlates with p53 mutation, copy number alteration, and S phase checkpoint. *Proc Natl Acad Sci USA* 119(8):e2115999119
 30. Cheng P, Zhao X, Katsnelson L, Camacho-Hernandez E, Mermerian A, Mays J, Lippman S, Rosales-Alvarez R, Moya R, Shwetar J, Grun D, Fenyö D, Davoli T (2022) Proteogenomic analysis of cancer aneuploidy and normal tissues reveals divergent modes of gene regulation across cellular pathways. *elife* 11:e75227
 31. Geffen Y, Anand S, Akiyama Y, Yaron T, Song Y, Johnson J, Govindan A, Babur Ö, Li Y, Huntsman E, Wang L, Birger C, Heiman D, Zhang Q, Miller M, Maruvka Y, Haradhvala N, Calinawan A, Belkin S, Kerelsky A, Clauser K, Krug K, Satpathy S, Payne S, Mani D, Gillette M, Dhanasekaran S, Thiagarajan M, Mesri M, Rodriguez H, Robles A, Carr S, Lazar A, Aguet F, Cantley L, Ding L, Getz G, Clinical Proteomic Tumor Analysis Consortium (2023) Pan-cancer analysis of post-translational modifications reveals shared patterns of protein regulation. *Cell* 186(18):3945–3967
 32. Liang W, Lu R, Jayasinghe R, Foltz S, Porta-Pardo E, Geffen Y, Wendl M, Lazcano R, Kolodziejczak I, Song Y, Govindan A, Demicco E, Li X, Li Y, Sethuraman S, Payne S, Fenyö D, Rodriguez H, Wiznerowicz M, Shen H, Mani D, Rodland K, Lazar A, Robles A, Ding L, Clinical Proteomic Tumor Analysis Consortium (2023) Integrative multi-omic cancer profiling reveals DNA methylation patterns associated with therapeutic vulnerability and cell-of-origin. *Cancer Cell* 41(9):1567–1585

33. Macklin A, Khan S, Kislinger T (2020) Recent advances in mass spectrometry based clinical proteomics: applications to cancer research. *Clin Proteomics* 17:17. <https://doi.org/10.1186/s12014-020-09283-w>
34. Anurag M, Jaehnig E, Krug K, Lei J, Bergstrom E, Kim B, Vashist T, Huynh A, Dou Y, Gou X, Huang C, Shi Z, Wen B, Korchina V, Gibbs R, Muzny D, Doddapaneni H, Dobrolecki L, Rodriguez H, Robles A, Hiltke T, Lewis M, Nangia J, Nemati Shafae M, Li S, Hagemann I, Hoog J, Lim B, Osborne C, Mani D, Gillette M, Zhang B, Echeverria G, Miles G, Rimawi M, Carr S, Ademuyiwa F, Satpathy S, Ellis M (2022) Proteogenomic markers of chemotherapy resistance and response in triple-negative breast cancer. *Cancer Discov* 12(11):2586–2605
35. Archer T, Ehrenberger T, Mundt F, Gold M, Krug K, Mah C, Mahoney E, Daniel C, LeNail A, Ramamoorthy D, Mertins P, Mani D, Zhang H, Gillette M, Clauser K, Noble M, Tang L, Pierre-François J, Silterra J, Jensen J, Tamayo P, Korshunov A, Pfister S, Kool M, Northcott P, Sears R, Lipton J, Carr S, Mesirov J, Pomeroy S, Fraenkel E (2018) Proteomics, post-translational modifications, and integrative analyses reveal molecular heterogeneity within Medulloblastoma subgroups. *Cancer Cell* 34(3):396–410
36. Cao L, Huang C, Cui Zhou D, Hu Y, Lih T, Savage S, Krug K, Clark D, Schnaubelt M, Chen L, da Veiga Leprevost F, Eguez R, Yang W, Pan J, Wen B, Dou Y, Jiang W, Liao Y, Shi Z, Terekhanova N, Cao S, Lu R, Li Y, Liu R, Zhu H, Ronning P, Wu Y, Wyczalkowski M, Easwaran H, Danilova L, Mer A, Yoo S, Wang J, Liu W, Haibe-Kains B, Thiagarajan M, Jewell S, Hostetter G, Newton C, Li Q, Roehrl M, Fenyö D, Wang P, Nesvizhskii A, Mani D, Omenn G, Boja E, Mesri M, Robles A, Rodriguez H, Bathe O, Chan D, Hruban R, Ding L, Zhang B, Zhang H, Clinical Proteomic Tumor Analysis Consortium (2021) Proteogenomic characterization of pancreatic ductal adenocarcinoma. *Cell* 184(19):5031–5052
37. Chen Y, Roumeliotis T, Chang Y, Chen C, Han C, Lin M, Chen H, Chang G, Chang Y, Wu C, Lin M, Hsieh M, Wang Y, Chen Y, Jonassen I, Ghavidel F, Lin Z, Lin K, Chen C, Sheu P, Hung C, Huang K, Yang H, Lin P, Yen T, Lin Y, Wang J, Raghav L, Lin C, Chen Y, Wu P, Lai C, Weng S, Su K, Chang W, Tsai P, Robles A, Rodriguez H, Hsiao Y, Chang W, Sung T, Chen J, Yu S, Choudhary J, Chen H, Yang P, Chen Y (2020) Proteogenomics of non-smoking lung cancer in East Asia delineates molecular signatures of pathogenesis and progression. *Cell* 182(1):226–244
38. Chowdhury S, Kennedy J, Ivey R, Murillo O, Hosseini N, Song X, Petralia F, Calinawan A, Savage S, Berry A, Reva B, Ozbek U, Krek A, Ma W, da Veiga Leprevost F, Ji J, Yoo S, Lin C, Voytovich U, Huang Y, Lee S, Bergan L, Lorentzen T, Mesri M, Rodriguez H, Hoofnagle A, Herbert Z, Nesvizhskii A, Zhang B, Whiteaker J, Fenyö D, McKerrow W, Wang J, Schürer S, Stathias V, Chen X, Barcellos-Hoff M, Starr T, Winterhoff B, Nelson A, Mok S, Kaufmann S, Drescher C, Cieslik M, Wang P, Birrer M, Paulovich A (2023) Proteogenomic analysis of chemo-refractory high-grade serous ovarian cancer. *Cell* 186(16):3476–3498
39. Clark D, Dhanasekaran S, Petralia F, Pan J, Song X, Hu Y, da Veiga Leprevost F, Reva B, Lih T, Chang H, Ma W, Huang C, Ricketts C, Chen L, Krek A, Li Y, Rykunov D, Li Q, Chen L, Ozbek U, Vasaikar S, Wu Y, Yoo S, Chowdhury S, Wyczalkowski W, Ji J, Schnaubelt M, Kong A, Sethuraman S, Avtonomov D, Ao M, Colaprico A, Cao S, Cho K, Kalayci S, Ma S, Liu W, Ruggles K, Calinawan A, Gümüş Z, Geizler D, Kawaler E, Teo G, Wen B, Zhang Y, Keegan S, Li K, Chen F, Edwards N, Pierorazio P, Chen X, Pavlovich C, Hakimi A, Brominski G, Hsieh J, Antczak A, Omelchenko T, Lubinski J, Wiznerowicz M, Linehan W, Kinsinger C, Thiagarajan M, Boja E, Mesri M, Hiltke T, Robles A, Rodriguez H, Qian J, Fenyö D, Zhang B, Ding L, Schadt E, Chinnaiyan A, Zhang Z, Omenn G, Cieslik M, Chan D, Nesvizhskii A, Wang P, Zhang H, Clinical Proteomic Tumor Analysis Consortium (2019) Integrated proteogenomic characterization of clear cell renal cell carcinoma. *Cell* 179(4):964–983
40. Dong L, Lu D, Chen R, Lin Y, Zhu H, Zhang Z, Cai S, Cui P, Song G, Rao D, Yi X, Wu Y, Song N, Liu F, Zou Y, Zhang S, Zhang X, Wang X, Qiu S, Zhou J, Wang S, Zhang X, Shi Y, Figeys D, Ding L, Wang P, Zhang B, Rodriguez H, Gao Q, Gao D, Zhou H, Fan J (2022) Proteogenomic characterization identifies clinically relevant subgroups of intrahepatic cholangiocarcinoma. *Cancer Cell* 40(1):70–87
41. Dou Y, Kawaler E, Cui Zhou D, Gritsenko M, Huang C, Blumenberg L, Karpova A, Petyuk V, Savage S, Satpathy S, Liu W, Wu Y, Tsai C, Wen B, Li Z, Cao S, Moon J, Shi Z, Cornwell M, Wyczalkowski M, Chu R, Vasaikar S, Zhou H, Gao Q, Moore R, Li K,

- Sethuraman S, Monroe M, Zhao R, Heiman D, Krug K, Clauser K, Kothadia R, Maruvka Y, Pico A, Oliphant A, Hoskins E, Pugh S, Beecroft S, Adams D, Jarman J, Kong A, Chang H, Reva B, Liao Y, Rykunov D, Colaprico A, Chen X, Czeakański A, Jędryka M, Matkowski R, Wiznerowicz M, Hiltke T, Boja E, Kinsinger C, Mesri M, Robles A, Rodriguez H, Mutch D, Fuh K, Ellis M, DeLair D, Thiagarajan M, Mani D, Getz G, Noble M, Nesvizhskii A, Wang P, Anderson M, Levine D, Smith R, Payne S, Ruggles K, Rodland K, Ding L, Zhang B, Liu T, Fenyő D, Clinical Proteomic Tumor Analysis Consortium (2020) Proteogenomic characterization of endometrial carcinoma. *Cell* 180(4):729–748. <https://doi.org/10.1016/j.cell.2020.01.026>
42. Dou Y, Katsnelson L, Gritsenko M, Hu Y, Reva B, Hong R, Wang Y, Kolodziejczak I, Lu R, Tsai C, Bu W, Liu W, Guo X, An E, Arend R, Bavarva J, Chen L, Chu R, Czeakański A, Davoli T, Demicco E, DeLair D, Devereaux K, Dhanasekaran S, Dottino P, Dover B, Fillmore T, Foxall M, Hermann C, Hiltke T, Hostetter G, Jędryka M, Jewell S, Johnson I, Kahn A, Ku A, Kumar-Sinha C, Kurzawa P, Lazar A, Lazcano R, Lei J, Li Y, Liao Y, Lih T, Lin T, Martignetti J, Masand R, Matkowski R, McKerrow W, Mesri M, Monroe M, Moon J, Moore R, Nestor M, Newton C, Omelchenko T, Omenn G, Payne S, Petyuk V, Robles A, Rodriguez H, Ruggles K, Rykunov D, Savage S, Schepmoes A, Shi T, Shi Z, Tan J, Taylor M, Thiagarajan M, Wang J, Weitz K, Wen B, Williams C, Wu Y, Wyczalkowski M, Yi X, Zhang X, Zhao R, Mutch D, Chinnaiyan A, Smith R, Nesvizhskii A, Wang P, Wiznerowicz M, Ding L, Mani D, Zhang H, Anderson M, Rodland K, Zhang B, Liu T, Fenyő D, Clinical Proteomic Tumor Analysis Consortium (2023) Proteogenomic insights suggest druggable pathways in endometrial carcinoma. *Cancer Cell* 41(9):1586–1605
 43. Gao Q, Zhu H, Dong L, Shi W, Chen R, Song Z, Huang C, Li J, Dong X, Zhou Y, Liu Q, Ma L, Wang X, Zhou J, Liu Y, Boja E, Robles A, Ma W, Wang P, Li Y, Ding L, Wen B, Zhang B, Rodriguez H, Gao D, Zhou H, Fan J (2019) Integrated Proteogenomic characterization of HBV-related hepatocellular carcinoma. *Cell* 179(2):561–577. <https://doi.org/10.1016/j.cell.2019.08.052>
 44. Gillette M, Satpathy S, Cao S, Dhanasekaran S, Vasaikar S, Krug K, Petralia F, Li Y, Liang W, Reva B, Krek A, Ji J, Song X, Liu W, Hong R, Yao L, Blumenberg L, Savage S, Wendl M, Wen B, Li K, Tang L, MacMullan M, Avanesian S, Kane M, Newton C, Cornwell M, Kothadia R, Ma W, Yoo S, Mannan R, Vats P, Kumar-Sinha C, Kawaler E, Omelchenko T, Colaprico A, Geffen Y, Maruvka Y, da Veiga Leprevost F, Wiznerowicz M, Gümüş Z, Veluswamy R, Hostetter G, Heiman D, Wyczalkowski M, Hiltke T, Mesri M, Kinsinger C, Boja E, Omenn G, Chinnaiyan A, Rodriguez H, Li Q, Jewell S, Thiagarajan M, Getz G, Zhang B, Fenyő D, Ruggles K, Cieslik M, Robles A, Clauser K, Govindan R, Wang P, Nesvizhskii A, Ding L, Mani D, Carr S, Clinical Proteomic Tumor Analysis Consortium (2020) Proteogenomic characterization reveals therapeutic vulnerabilities in lung adenocarcinoma. *Cell* 182(1):200–225
 45. Hu Y, Pan J, Shah P, Ao M, Thomas S, Liu Y, Chen L, Schnaubelt M, Clark D, Rodriguez H, Boja E, Hiltke T, Kinsinger C, Rodland K, Li Q, Qian J, Zhang Z, Chan D, Zhang H, Clinical Proteomic Tumor Analysis Consortium (2020) Integrated proteomic and Glycoproteomic characterization of human high-grade serous ovarian carcinoma. *Cell Rep* 33(3):108276
 46. Huang C, Chen L, Savage S, Eguev R, Dou Y, Li Y, da Veiga Leprevost F, Jaehnig E, Lei J, Wen B, Schnaubelt M, Krug K, Song X, Cieślik M, Chang H, Wyczalkowski M, Li K, Colaprico A, Li Q, Clark D, Hu Y, Cao L, Pan J, Wang Y, Cho K, Shi Z, Liao Y, Jiang W, Anurag M, Ji J, Yoo S, Zhou D, Liang W, Wendl M, Vats P, Carr S, Mani D, Zhang Z, Qian J, Chen X, Pico A, Wang P, Chinnaiyan A, Ketchum K, Kinsinger C, Robles A, An E, Hiltke T, Mesri M, Thiagarajan M, Weaver A, Sikora A, Lubinski J, Wierzbicka M, Wiznerowicz M, Satpathy S, Gillette M, Miles G, Ellis M, Omenn G, Rodriguez H, Boja E, Dhanasekaran S, Ding L, Nesvizhskii A, El-Naggar AC, DW, Zhang H, Zhang B, Clinical Proteomic Tumor Analysis Consortium (2021) Proteogenomic insights into the biology and treatment of HPV-negative head and neck squamous cell carcinoma. *Cancer Cell* 39(3):361–379
 47. Hyeon D, Nam D, Han Y, Kim D, Kim G, Kim D, Bae J, Back S, Mun D, Madar I, Lee H, Kim S, Kim H, Hyun S, Kim C, Choi S, Kim Y, Jeong J, Jeon S, Choo Y, Lee K, Kwon W, Choi S, Goo T, Park T, Suh Y, Kim H, Ku J, Kim M, Paek E, Park D, Jung K, Back S, Jang J, Hwang D,

- Lee S (2023) Proteogenomic landscape of human pancreatic ductal adenocarcinoma in an Asian population reveals tumor cell-enriched and immune-rich subtypes. *Nat Cancer* 4(2):290–307
48. Kramer M, Zhang Q, Sprung R, Day R, Erdmann-Gilmore P, Li Y, Xu Z, Helton N, George D, Mi Y, Westervelt P, Payton J, Ramakrishnan S, Miller C, Link D, DiPersio J, Walter M, Townsend R, Ley T (2022) Proteomic and phosphoproteomic landscapes of acute myeloid leukemia. *Blood* 140(13):1533–1548
 49. Krug K, Jaehnig E, Satpathy S, Blumenberg L, Karpova A, Anurag M, Miles G, Mertins P, Geffen Y, Tang L, Heiman D, Cao S, Maruvka Y, Lei J, Huang C, Kothadia R, Colaprico A, Birger C, Wang J, Dou Y, Wen B, Shi Z, Liao Y, Wiznerowicz M, Wyczalkowski M, Chen X, Kennedy J, Paulovich A, Thiagarajan M, Kinsinger C, Hiltke T, Boja E, Mesri M, Robles A, Rodriguez H, Westbrook T, Ding L, Getz G, Clauser K, Fenyo D, Ruggles K, Zhang B, Mani D, Carr S, Ellis M, Gillette M, Clinical Proteomic Tumor Analysis Consortium (2020) Proteogenomic landscape of breast cancer tumorigenesis and targeted therapy. *Cell* 183(5):1436–1456
 50. Lehtiö J, Arslan T, Siavelis I, Pan Y, Socciarelli F, Berkovska O, Umer H, Mermelakas G, Pirmoradian M, Jönsson M, Brunnström H, Brustugun O, Purohit K, Cunningham R, Foroughi Asl H, Isaksson S, Arbajian E, Aine M, Karlsson A, Kotevska M, Gram Hansen C, Drageset Haakensen V, Heland Å, Tamborero D, Johansson H, Branca R, Planck M, Staaf J, Orre L (2021) Proteogenomics of non-small cell lung cancer reveals molecular subtypes associated with specific therapeutic targets and immune evasion mechanisms. *Nat Cancer* 2(11):1224–1242
 51. Li Y, Lih T, Dhanasekaran S, Mannan R, Chen L, Cieslik M, Wu Y, Lu R, Clark D, Kolodziejczak I, Hong R, Chen S, Zhao Y, Chugh S, Caravan W, Naser Al Deen N, Hosseini N, Newton C, Krug K, Xu Y, Cho K, Hu Y, Zhang Y, Kumar-Sinha C, Ma W, Calinawan A, Wyczalkowski M, Wendl M, Wang Y, Guo S, Zhang C, Le A, Dagar A, Hopkins A, Cho H, Leprevost F, Jing X, Teo G, Liu W, Reimers M, Pachynski R, Lazar A, Chinnaiyan A, Van Tine B, Zhang B, Rodland K, Getz G, Mani D, Wang P, Chen F, Hostetter G, Thiagarajan M, Linehan W, Fenyo D, Jewell S, Omenn G, Mehra R, Wiznerowicz M, Robles A, Mesri M, Hiltke T, An E, Rodriguez H, Chan D, Ricketts C, Nesvizhskii A, Zhang H, Ding L, Clinical Proteomic Tumor Analysis Consortium (2023) Histopathologic and proteogenomic heterogeneity reveals features of clear cell renal cell carcinoma aggressiveness. *Cancer Cell* 41(1):139–163
 52. Li S, Yuan L, Xu Z, Xu J, Chen G, Guan X, Pan G, Hu C, Dong J, Du Y, Yang L, Ni M, Jiang R, Zhu X, Lv H, Xu H, Zhang S, Qin J, Cheng X (2023) Integrative proteomic characterization of adenocarcinoma of esophago-gastric junction. *Nat Commun* 14(1):778
 53. Liu Q, Zhang J, Guo C, Wang M, Wang C, Yan Y, Sun L, Wang D, Zhang L, Yu H, Hou L, Wu C, Zhu Y, Jiang G, Zhu H, Zhou Y, Fang S, Zhang T, Hu L, Li J, Liu Y, Zhang H, Zhang B, Ding L, Robles A, Rodriguez H, Gao D, Ji H, Zhou H, Zhang P (2024) Proteogenomic characterization of small cell lung cancer identifies biological insights and subtype-specific therapeutic strategies. *Cell* 187(1):184–203
 54. Mun D, Bhin J, Kim S, Kim H, Jung J, Jung Y, Jang Y, Park J, Kim H, Jung Y, Lee H, Bae J, Back S, Kim S, Kim J, Park H, Li H, Hwang K, Park Y, Yook J, Kim B, Kwon S, Ryu S, Park D, Jeon T, Kim D, Lee J, Han S, Song K, Park D, Park J, Rodriguez H, Kim J, Lee H, Kim K, Yang E, Kim H, Paek E, Lee S, Lee S, Hwang D (2019) Proteogenomic characterization of human early-onset gastric cancer. *Cancer Cell* 35(1):111–124. <https://doi.org/10.1016/j.ccell.2018.12.003>
 55. Ng C, Dazert E, Boldanova T, Coto-Llerena-M, Nuciforo S, Ercan C, Suslov A, Meier M, Bock T, Schmidt A, Ketterer S, Wang X, Wieland S, Matter M, Colombi M, Piscuoglio S, Terracciano L, Hall M, Heim M (2022) Integrative proteogenomic characterization of hepatocellular carcinoma across etiologies and stages. *Nat Commun* 13(1):2436
 56. Petralia F, Tignor N, Reva B, Koptyra M, Chowdhury S, Rykunov D, Krek A, Ma W, Zhu Y, Ji J, Calinawan A, Whiteaker J, Colaprico A, Stathias V, Omelchenko T, Song X, Raman P, Guo Y, Brown M, Ivey R, Szpyt J, Thakurta S, Gritsenko M, Weitz K, Lopez G, Kalayci S, Gümüş Z, Yoo S, da Veiga Leprevost F, Chang H, Krug K, Katsnelson L, Wang Y, Kennedy J, Voytovich U, Zhao L, Gaonkar K, Ennis B, Zhang B, Baubet V, Tauhid L, Lilly J, Mason J, Farrow B, Young N, Leary S, Moon J, Petyuk V, Nazarian J, Adappa N, Palmer J, Lober R,

- Rivero-Hinojosa S, Wang L, Wang J, Broberg M, Chu R, Moore R, Monroe M, Zhao R, Smith R, Zhu J, Robles A, Mesri M, Boja E, Hiltke T, Rodriguez H, Zhang B, Schadt E, Mani D, Ding L, Iavarone A, Wiznerowicz M, Schürer S, Chen X, Heath A, Rokita J, Nesvizhskii A, Fenyő D, Rodland K, Liu T, Gygi S, Paulovich A, Resnick A, Storm P, Rood B, Wang P, Children's Brain Tumor Network and Clinical Proteomic Tumor Analysis Consortium (2020) Integrated Proteogenomic characterization across major histological types of pediatric brain cancer. *Cell* 183(7):1962–1985. <https://doi.org/10.1016/j.cell.2020.10.044>
57. Qu Y, Feng J, Wu X, Bai L, Xu W, Zhu L, Liu Y, Xu F, Zhang X, Yang G, Lv J, Chen X, Shi G, Wang H, Cao D, Xiang H, Li L, Tan S, Gan H, Sun M, Qiu J, Zhang H, Zhao J, Ye D, Ding C (2022) A proteogenomic analysis of clear cell renal cell carcinoma in a Chinese population. *Nat Commun* 13(1):2052
 58. Satpathy S, Krug K, Jean Beltran P, Savage S, Petralia F, Kumar-Sinha C, Dou Y, Reva B, Kane M, Avanesian S, Vasaikar S, Krek A, Lei J, Jaehnig E, Omelchenko T, Geffen Y, Bergstrom E, Stathias V, Christianson K, Heiman D, Cieslik M, Cao S, Song X, Ji J, Liu W, Li K, Wen B, Li Y, Gümüş Z, Selvan M, Soundararajan R, Visal T, Raso M, Parra E, Babur Ö, Vats P, Anand S, Schraink T, Cornwell M, Rodrigues F, Zhu H, Mo C, Zhang Y, da Veiga Leprevost F, Huang C, Chinnaiyan A, Wyczalkowski M, Omenn G, Newton C, Schurer S, Ruggles K, Fenyő D, Jewell S, Thiagarajan M, Mesri M, Rodriguez H, Mani S, Udeshi N, Getz G, Suh J, Li Q, Hostetter G, Paik P, Dhanasekaran S, Govindan R, Ding L, Robles A, Clauser K, Nesvizhskii A, Wang P, Carr S, Zhang B, Mani D, Gillette M, Clinical Proteomic Tumor Analysis Consortium (2021) A proteogenomic portrait of lung squamous cell carcinoma. *Cell* 184(16):4348–4371
 59. Shi X, Sun Y, Shen C, Zhang Y, Shi R, Zhang F, Liao T, Lv G, Zhu Z, Jiao L, Li P, Xu T, Qu N, Huang N, Hu J, Zhang T, Gu Y, Qin G, Guan H, Pu W, Li Y, Geng X, Zhang Y, Chen T, Huang S, Zhang Z, Ge S, Wang W, Xu W, Yu P, Lu Z, Wang Y, Guo L, Wang Y, Guo T, Ji Q, Wei W (2022) Integrated proteogenomic characterization of medullary thyroid carcinoma. *Cell Discov* 8(1):120
 60. Sinha A, Huang V, Livingstone J, Wang J, Fox N, Kurganovs N, Ignatchenko V, Fritsch K, Donmez N, Heisler L, Shiah Y, Yao C, Alfaro J, Volik S, Lapuk A, Fraser M, Kron K, Murison A, Lupien M, Sahinalp C, Collins C, Tetu B, Masoomian M, Berman D, van der Kwast T, Bristow R, Kislinger T, Boultros P (2019) The Proteogenomic landscape of curable prostate cancer. *Cancer Cell* 35(3):414–427. <https://doi.org/10.1016/j.ccell.2019.02.005>
 61. Soltis A, Bateman N, Liu J, Nguyen T, Franks T, Zhang X, Dalgard C, Viollet C, Somiari S, Yan C, Zeman K, Skinner W, Lee J, Pollard H, Turner C, Petricoin E, Meerzaman D, Conrads T, Hu H, APOLLO_Research_Network, Shriver C, Moskaluk C, Browning RJ, Wilkerson M (2022) Proteogenomic analysis of lung adenocarcinoma reveals tumor heterogeneity, survival determinants, and therapeutically relevant pathways. *Cell Rep Med* 3(11):100819. <https://doi.org/10.1016/j.xcrm.2022.100819>
 62. Stewart P, Welsh E, Slebos R, Fang B, Izumi V, Chambers M, Zhang G, Cen L, Pettersson F, Zhang Y, Chen Z, Cheng C, Thapa R, Thompson Z, Fellows K, Francis J, Saller J, Mesa T, Zhang C, Yoder S, DeNicola G, Beg A, Boyle T, Teer J, Ann Chen Y, Koomen J, Eschrich S, Haura E (2019) Proteogenomic landscape of squamous cell lung cancer. *Nat Commun* 10(1):3578
 63. Vasaikar S, Huang C, Wang X, Petyuk V, Savage S, Wen B, Dou Y, Zhang Y, Shi Z, Arshad O, Gritsenko M, Zimmerman L, McDermott J, Clauss T, Moore R, Zhao R, Monroe M, Wang Y, Chambers M, Slebos R, Lau K, Mo Q, Ding L, Ellis M, Thiagarajan M, Kinsinger C, Rodriguez H, Smith R, Rodland K, Liebler D, Liu T, Zhang B, Clinical Proteomic Tumor Analysis Consortium (2019) Proteogenomic analysis of human colon cancer reveals new therapeutic opportunities. *Cell* 177(4):1035–1049. e1019
 64. Wang L, Karpova A, Gritsenko M, Kyle J, Cao S, Li Y, Rykunov D, Colaprico A, Rothstein J, Hong R, Stathias V, Cornwell M, Petralia F, Wu Y, Reva B, Liang W, Song X, Dou Y, Wendl M, Caravan W, Liu W, Cui Zhou D, Ji J, Tsai C, Petyuk V, Moon J, Ma W, Chu R, Weitz K, Moore R, Monroe M, Zhao R, Yang X, Yoo S, Krek A, Demopoulos A, Zhu H, Wyczalkowski M, McMichael J, Henderson B, Lindgren C, Boekweg H, Lu S, Baral J, Yao L, Stratton K, Bramer L, Zink E, Couvillion S, Bloodsworth K, Satpathy S, Sieh W, Boca S, Schürer S,

- Chen F, Wiznerowicz M, Ketchum K, Boja E, Kinsinger C, Robles A, Hiltke T, Thiagarajan M, Nesvizhskii A, Zhang B, Mani D, Ceccarelli M, Chen X, Cottingham S, Li Q, Kim A, Fenyö D, Ruggles K, Rodriguez H, Mesri M, Payne S, Resnick A, Wang P, Smith R, Iavarone A, Chheda M, Barnholtz-Sloan J, Rodland K, Liu T, Ding L, Clinical Proteomic Tumor Analysis Consortium (2021) Proteogenomic and metabolomic characterization of human glioblastoma. *Cancer Cell* 39(4):509–528
65. Xu N, Yao Z, Shang G, Ye D, Wang H, Zhang H, Qu Y, Xu F, Wang Y, Qin Z, Zhu J, Zhang F, Feng J, Tian S, Liu Y, Zhao J, Hou J, Guo J, Hou Y, Ding C (2022) Integrated proteogenomic characterization of urothelial carcinoma of the bladder. *J Hematol Oncol* 15(1):76
 66. Zhang F, Zhang Q, Zhu J, Yao B, Ma C, Qiao N, He S, Ye Z, Wang Y, Han R, Feng J, Wang Y, Qin Z, Ma Z, Li K, Zhang Y, Tian S, Chen Z, Tan S, Wu Y, Ran P, Wang Y, Ding C, Zhao Y (2022) Integrated proteogenomic characterization across major histological types of pituitary neuroendocrine tumors. *Cell Res* 32(12):1047–1067
 67. Zhang Y, Fu F, Zhang Q, Li L, Liu H, Deng C, Xue Q, Zhao Y, Sun W, Han H, Gao Z, Guo C, Zheng Q, Hu H, Sun Y, Li Y, Ding C, Chen H (2024) Evolutionary proteogenomic landscape from pre-invasive to invasive lung adenocarcinoma. *Cell Rep Med* 5(1):101358
 68. Vasaikar S, Straub P, Wang J, Zhang B (2018) LinkedOmics: analyzing multi-omics data within and across 32 cancer types. *Nucleic Acids Res* 46(D1):D956–D963
 69. Creighton C (2022) Clinical proteomics towards multiomics in cancer. *Mass Spectrom Rev* 10:e21827. <https://doi.org/10.1002/mas.21827>
 70. Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci USA* 100:9440–9445. <https://doi.org/10.1073/pnas.1530509100>
 71. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 95:14863–14868
 72. Pollack J, Sørlie T, Perou C, Rees C, Jeffrey S, Lonning P, Tibshirani R, Botstein D, Børresen-Dale A, Brown P (2002) Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc Natl Acad Sci USA* 99(20):12963–12968
 73. Garraway L, Widlund H, Rubin M, Getz G, Berger A, Ramaswamy S, Beroukhi R, Milner D, Granter S, Du J, Lee C, Wagner S, Li C, Golub T, Rimm D, Meyerson M, Fisher D, Sellers W (2005) Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nature* 436(7047):117–122
 74. Hoadley K, Yau C, Wolf D, Cherniack A, Tamborero D, Ng S, Leiserson M, Niu B, McLellan M, Uzunangelov V, Zhang J, Kandoth C, Akbani R, Shen H, Omberg L, Chu A, Margolin A, Van't Veer L, Lopez-Bigas N, Laird P, Raphael B, Ding L, Robertson A, Byers L, Mills G, Weinstein J, Van Waes C, Chen Z, Collisson E, Cancer Genome Atlas Research Network, Benz C, Perou C, Stuart J (2014) Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. *Cell* 158(4):929–944
 75. Creighton C (2018) Making use of cancer genomic databases. *Curr Protoc Mol Biol* 121:19.14.11–19.14.13. <https://doi.org/10.1002/cpmb.49>
 76. West R (2022) Best practice in statistics: the use of log transformation. *Ann Clin Biochem* 59(3):162–165
 77. Aittokallio T (2010) Dealing with missing values in large-scale studies: microarray data imputation and beyond. *Brief Bioinform* 11(2):253–264
 78. Johnson W, Rabinovic A, Li C (2007) Adjusting batch effects in microarray expression data using Empirical Bayes methods. *Biostatistics* 8(1):118–127
 79. Leek J, Johnson W, Parker H, Jaffe A, Storey J (2012) The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 28(6):882–883
 80. Zhang Y, Chen F, Creighton C (2023) Pan-cancer molecular subtypes of metastasis reveal distinct and evolving transcriptional programs. *Cell Rep Med* 4(2):100932. <https://doi.org/10.1016/j.xcrm.2023.100932>
 81. Chen F, Zhang Y, Gibbons D, Deneen B, Kwiatkowski D, Ittmann M, Creighton C (2018) Pan-cancer molecular classes transcending tumor lineage across 32 cancer types, multiple data platforms, and over 10,000 cases. *Clin Cancer Res* 24(9):2182–2193
 82. Bolstad B, Ra I, Astrand M, Speed T (2003) A comparison of normalization methods for high density oligonucleotide Array data

- based on bias and variance. *Bioinformatics* 19(2):185–193
83. Martínez E, Yoshihara K, Kim H, Mills G, Treviño V, Verhaak R (2015) Comparison of gene expression patterns across 12 tumor types identifies a cancer supercluster characterized by TP53 mutations and cell cycle defects. *Oncogene* 34(21):2732–2740
 84. Hoadley K, Yau C, Hinoue T, Wolf D, Lazar A, Drill E, Shen R, Taylor A, Cherniack A, Thorsson V, Akbani R, Bowlby R, Wong C, Wiznerowicz M, Sanchez-Vega F, Robertson A, Schneider B, Lawrence M, Noushmehr H, Malta T, Cancer_Genome_Atlas_Network, Stuart J, Benz C, Laird P (2018) Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. *Cell* 173(2):291–304
 85. Chen F, Zhang Y, Şenbabaoğlu Y, Ciriello G, Yang L, Reznik E, Shuch B, Micevic G, De Velasco G, Shinbrot E, Noble M, Lu Y, Covington K, Xi L, Drummond J, Muzny D, Kang H, Lee J, Tamboli P, Reuter V, Shelley C, Kaiparettu B, Bottaro D, Godwin A, Gibbs R, Getz G, Kucherlapati R, Park P, Sander C, Henske E, Zhou J, Kwiatkowski D, Ho T, Choueiri T, Hsieh J, Akbani R, Mills G, Hakimi A, Wheeler D, Creighton C (2016) Multilevel genomics-based taxonomy of renal cell carcinoma. *Cell Rep* 14(10):2476–2489
 86. The Cancer Genome Atlas Research_Network (2015) Comprehensive molecular characterization of papillary renal cell carcinoma. *N Engl J Med* <https://doi.org/10.1056/NEJMoa1505917>
 87. Tamames J, Clark D, Herrero J, Dopazo J, Blaschke C, Fernández J, Oliveros J, Valencia A (2002) Bioinformatics methods for the analysis of expression arrays: data clustering and information extraction. *J Biotechnol* 98(2–3):269–283
 88. Sherlock G (2000) Analysis of large-scale gene expression data. *Curr Opin Immunol* 12(2):201–205
 89. Monti S, Tamayo P, Misirov J, Golub T (2003) Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Mach Learn* 52:91–118
 90. Wilkerson M, Hayes D (2010) ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26(12):1572–1573
 91. Lumley T, Diehr P, Emerson S, Chen L (2002) The importance of the normality assumption in large public health data sets. *Annu Rev Public Health* 23:151–169
 92. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT (2000) Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat Genet* 25:25–29
 93. Subramanian A, Tamayo P, Mootha V, Mukherjee S, Ebert B, Gillette M, Paulovich A, Pomeroy S, Golub T, Lander E, Mesirov J (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 102(43):15545–15550
 94. Slenter D, Kutmon M, Hanspers K, Riutta A, Windsor J, Nunes N, Mélius J, Cirillo E, Coort S, Digles D, Ehrhart F, Giesbertz P, Kalafati M, Martens M, Miller R, Nishida K, Rieswijk L, Waagmeester A, Eijssen L, Evelo C, Pico A, Willighagen E (2018) WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. *Nucleic Acids Res* 46(D1):D661–D667
 95. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E (2003) PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 34:267–273
 96. Sherman B, Hao M, Qiu J, Jiao X, Baseler M, Lane H, Imamichi T, Chang W (2022) DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res* 50(W1):W216–W221
 97. Creighton C, Nagaraja A, Hanash S, Matzuk M, Gunaratne P (2008) A bioinformatics tool for linking gene expression profiling results with public databases of microRNA target predictions. *RNA* 14(11):2290–2296
 98. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ (2002) A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 347:1999–2009
 99. van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT (2002) Gene expression profiling predicts clinical

- outcome of breast cancer. *Nature* 415:530–536. <https://doi.org/10.1038/415530a>
100. Chang HY, Sneddon JB, Alizadeh AA, Sood R, West RB, Montgomery K, Chi JT, van de Rijn M, Botstein D, Brown PO (2004) Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. *PLoS Biol* 2:E7
 101. Venet D, Dumont J, Detours V (2011) Most random gene expression signatures are significantly associated with breast cancer outcome. *PLoS Comput Biol* 7(10):e1002240
 102. Gibbons D, Creighton C (2018) Pan-cancer survey of epithelial-mesenchymal transition markers across The Cancer Genome Atlas. *Dev Dyn* 247(3):555–564
 103. Creighton C, Hernandez-Herrera A, Jacobsen A, Levine D, Mankoo P, Schultz N, Du Y, Zhang Y, Larsson E, Sheridan R, Xiao W, Spellman P, Getz G, Wheeler D, Perou C, Gibbs R, Sander C, Hayes D, Gunaratne P, Cancer Genome Atlas Research Network (2012) Integrated analyses of microRNAs demonstrate their widespread influence on gene expression in high-grade serous ovarian carcinoma. *PLoS One* 7(3):e34546. <https://doi.org/10.1371/journal.pone.0034546>
 104. Creighton C, Li X, Landis M, Dixon J, Neumeister V, Sjolund A, Rimm D, Wong H, Rodriguez A, Herschkowitz J, Fan C, Zhang X, He X, Pavlick A, Gutierrez M, Renshaw L, Larionov A, Faratian D, Hilsenbeck S, Perou C, Lewis M, Rosen J, Chang J (2009) Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. *Proc Natl Acad Sci USA* 106(33):13820–13825
 105. Creighton C, Casa A, Lazard Z, Huang S, Tsimelzon A, Hilsenbeck S, Osborne C, Lee A (2008) Insulin-like growth factor I (IGF-I) activates gene transcription programs strongly associated with poor breast cancer prognosis. *J Clin Oncol* 26(25):4078–4085
 106. The Cancer Genome Atlas Research Network (2013) Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* 499(7456):43–49. <https://doi.org/10.1038/nature12222>
 107. Cancer Genome Atlas Research Network (2011) Integrated genomic analyses of ovarian carcinoma. *Nature* 474(7353):609–615
 108. Zhang Y, Kwok-Shing Ng P, Kucherlapati M, Chen F, Liu Y, Tsang Y, de Velasco G, Jeong K, Akbani R, Hadjipanayis A, Pantazi A, Bristow C, Lee E, Mahadeshwar H, Tang J, Zhang J, Yang L, Seth S, Lee S, Ren X, Song X, Sun H, Seidman J, Luquette L, Xi R, Chin L, Protopopov A, Westbrook T, Shelley C, Choueiri T, Ittmann M, Van WC, Weinstein J, Liang H, Henske E, Godwin A, Park P, Kucherlapati R, Scott K, Mills G, Kwiatkowski D, Creighton C (2017) A pan-cancer Proteogenomic atlas of PI3K/AKT/mTOR pathway alterations. *Cancer Cell* 31(6):820–832
 109. Guo J, Chakraborty A, Liu P, Gan W, Zheng X, Inuzuka H, Wang B, Zhang J, Zhang L, Yuan M, Novak J, Cheng J, Toker A, Signoretti S, Zhang Q, Asara J, Kaelin WJ, Wei W (2016) pVHL suppresses kinase activity of Akt in a proline-hydroxylation-dependent manner. *Science* 353(6302):929–932
 110. Carbonneau M, Gagné L, Lalonde M, Germain M, Motorina A, Guiot M, Secco B, Vincent E, Tumber A, Hulea L, Bergeman J, Oppermann U, Jones R, Laplante M, Topisirovic I, Petrecca K, Huot M, Mallette F (2016) The oncometabolite 2-hydroxyglutarate activates the mTOR signalling pathway. *Nat Commun* 7:12700
 111. Chen F, Zhang Y, Creighton C (2021) Systematic identification of non-coding somatic single nucleotide variants associated with altered transcription and DNA methylation in adult and pediatric cancers. *NAR Cancer* 3(1):zcab001
 112. Rheinbay E, Nielsen MM, Abascal F, Tiao G, Hornshøj H, Hess JM, Pedersen RI, Feuerbach L, Sabarinathan R, Madsen T, Kim J, Mularoni L, Shuai S, Lanzós A, Herrmann C, Maruvka YE, Shen C, Amin SB, Bertl J, Dhingra P, Diamanti K, Gonzalez-Perez A, Guo Q, Haradhvala NJ, Isaev K, Juul M, Komorowski J, Kumar S, Lee D, Lochovsky L, Liu EM, Pich O, Tamborero D, Umer HM, Uusküla-Reimand L, Wadelius C, Wadi L, Zhang J, Boroevich KA, Carlevaro-Fita J, Chakravarty D, Chan CWY, Fonseca NA, Hamilton MP, Hong C, Kahles A, Kim Y, Lehmann K-V, Johnson TA, Kahraman A, Park K, Saksena G, Sieverling L, Sinnott-Armstrong NA, Campbell PJ, Hobolth A, Kellis M, Lawrence MS, Raphael B, Rubin MA, Sander C, Stein L, Stuart J, Tsunoda T, Wheeler DA, Johnson R, Reimand J, Gerstein MB, Khurana E, López-Bigas N, Martincorena I, Pedersen JS, Getz G (2020)

- Analyses of non-coding somatic drivers in 2,658 cancer whole genomes. *Nature* 578(7793):102–111
113. Zhang Y, Yang L, Kucherlapati M, Chen F, Hadjipanayis A, Pantazi A, Bristow C, Lee E, Mahadeshwar H, Tang J, Zhang J, Seth S, Lee S, Ren X, Song X, Sun H, Seidman J, Luquette L, Xi R, Chin L, Protopopov A, Li W, Park P, Kucherlapati R, Creighton C (2018) A Pan-cancer compendium of genes deregulated by somatic genomic rearrangement across more than 1,400 cases. *Cell Rep* 24(2):515–527
 114. Davis C, Ricketts C, Wang M, Yang L, Cherniack A, Shen H, Buhay C, Kang H, Kim S, Fahey C, Hacker K, Bhanot G, Gordenin D, Chu A, Gunaratne P, Biehl M, Seth S, Kaiparettu B, Bristow C, Donehower L, Wallen E, Smith A, Tickoo S, Tamboli P, Reuter V, Schmidt L, Hsieh J, Choueiri T, Hakimi A, Cancer Genome Atlas Research Network, Chin L, Meyerson M, Kucherlapati R, Park W, Robertson A, Laird P, Henske E, Kwiatkowski D, Park P, Morgan M, Shuch B, Muzny D, Wheeler D, Linehan W, Gibbs R, Rathmell W, Creighton C (2014) The somatic genomic landscape of chromophobe renal cell carcinoma. *Cancer Cell* 26(3):319–330
 115. Creighton C, Beer D, Hanash S (2003) Gene expression patterns define pathways correlated with loss of differentiation in lung adenocarcinomas. *FEBS Lett* 540:167–170
 116. Weischenfeldt J, Dubash T, Drinas A, Mardin B, Chen Y, Stütz A, Waszak S, Bosco G, Halvorsen A, Raeder B, Efthymiopoulos T, Erkek S, Siegl C, Brenner H, Brustugun O, Dieter S, Northcott P, Petersen I, Pfister S, Schneider M, Solberg S, Thunissen E, Weichert W, Zichner T, Thomas R, Peifer M, Helland A, Ball C, Jechlinger M, Sotillo R, Glimm H, Korbel J (2017) Pan-cancer analysis of somatic copy-number alterations implicates *IRS4* and *IGF2* in enhancer hijacking. *Nat Genet* 49(1):65–74
 117. Zhang Y, Chen F, Creighton C (2021) SVExpress: identifying gene features altered recurrently in expression with nearby structural variant breakpoints. *BMC Bioinf* 22(1):135
 118. Liu Y, Li C, Shen S, Chen X, Szlachta K, Edmonson M, Shao Y, Ma X, Hyle J, Wright S, Ju B, Rusch M, Liu Y, Li B, Macias M, Tian L, Easton J, Qian M, Yang J, Hu S, Look A, Zhang J (2020) Discovery of regulatory noncoding variants in individual cancer genomes by using cis-X. *Nat Genet* 52(8):811–818
 119. Eteleeb A, Quigley D, Zhao S, Pham D, Yang R, Dehm S, Luo J, Feng F, Dang H, Maher C (2020) SV-HotSpot: detection and visualization of hotspots targeted by structural variants associated with gene expression. *Sci Rep* 10(1):15890
 120. Gopanenko A, Kosobokova E, Kosorukov V (2020) Main strategies for the identification of Neoantigens. *Cancers (Basel)* 12(10):2879
 121. Chen R, Fulton K, Twine S, Li J (2021) Identification of MHC peptides using mass spectrometry for Neoantigen discovery and cancer vaccine development. *Mass Spectrom Rev* 40(2):110–125
 122. Creech A, Ting Y, Goulding S, Sauld J, Barthelme D, Rooney M, Addona T, Abelin J (2018) The role of mass spectrometry and Proteogenomics in the advancement of HLA epitope prediction. *Proteomics* 18(12):e1700259
 123. Li Y, Wang G, Tan X, Ouyang J, Zhang M, Song X, Liu Q, Leng Q, Chen L, Xie L (2020) ProGeo-neo: a customized proteogenomic workflow for neoantigen prediction and selection. *BMC Med Genet* 13:52
 124. Alban T, Chan T (2021) Immunotherapy biomarkers: the long and winding road. *Nat Rev Clin Oncol* 18(6):323–324
 125. Creighton C (2023) Gene expression profiles in cancers and their therapeutic implications. *Cancer J* 29(1):9–14. <https://doi.org/10.1097/PPO.0000000000000638>
 126. Nusinow D, Szpyt J, Ghandi M, Rose C, McDonald ER, Kalocsay M, Jané-Valbuena J, Gelfand E, Schweppe D, Jedrychowski M, Golji J, Porter D, Rejtar T, Wang Y, Kryukov G, Stegmeier F, Erickson B, Garraway L, Sellers W, Gygi S (2020) Quantitative proteomics of the cancer cell line encyclopedia. *Cell* 180(2):387–402. <https://doi.org/10.1016/j.cell.2019.12.023>
 127. Gonçalves E, Poulos R, Cai Z, Barthorpe S, Manda S, Lucas N, Beck A, Bucio-Noble D, Dausmann M, Hall C, Hecker M, Koh J, Lightfoot H, Mahboob S, Mali I, Morris J, Richardson L, Seneviratne A, Shepherd R, Sykes E, Thomas F, Valentini S, Williams S, Wu Y, Xavier D, MacKenzie K, Hains P, Tully B, Robinson P, Zhong Q, Garnett M, Reddel R (2022) Pan-cancer proteomic map of 949 human cell lines. *Cancer Cell* 40(8):835–849. <https://doi.org/10.1016/j.ccell.2022.06.010>
 128. Dempster J, Boyle I, Vazquez F, Root D, Boehm J, Hahn W, Tsherniak A, McFarland J (2021) Chronos: a cell population dynamics model of CRISPR experiments that improves inference of gene fitness effects. *Genome Biol*

- 22(1):343. <https://doi.org/10.1186/s13059-021-02540-7>
129. Gentleman R, Carey V, Huber W, Irizarry R, Dudoit S (2005) Bioinformatics and computational biology solutions using R and bioconductor. Springer Science+Business Media, New York
130. Saldanha AJ (2004) Java Treeview—extensible visualization of microarray data. *Bioinformatics* 20:3246–3248
131. Pavlidis P, Noble W (2003) Matrix2png: a utility for visualizing matrix data. *Bioinformatics* 19(2):295–296



Microplate-Based Enzymatic Activity Assay Protocol Powered by Activity-Based Probes

Exequiel O. J. Porta, Karunakaran Kalesh, Jaime A. Isern, and Patrick G. Steel

Abstract

Identifying new enzymes and gaining insights into their mode of action (and inhibition) is a critical step in drug discovery. In this chapter, we describe a novel methodology that accelerates this process by combining the precision of activity-based protein profiling (ABPP) with the efficiency of microplate assay technology. Central to this approach is the use of competitive ABPP, which utilizes an activity-based probe (ABP) to directly observe and characterize the interactions between inhibitors and an enzyme. In this study, we use an in-house fluorophosphonate (FP)-based probe, as an example of an electrophilic probe, and pig liver esterase (PLE), as a model enzyme. Simple chemical modification of PLE to introduce a streptavidin tag enables its immobilization and presentation on a biotinylated assay plate. Subsequent competitive ABPP experiments using FP probes and a commercially derived compound library can be run in parallel. This setup allows for the identification of new inhibitors and the estimation of IC_{50} values in a single operation. This chapter provides a comprehensive protocol for the synthesis of ABPP probes, preparation of enzyme solutions, and execution of plate-based assays. It offers a robust platform for discovering and characterizing novel inhibitors of enzymes with unknown structures or substrate specificities. By applying this technique, we not only expedite the identification of potential drug candidates but also gain deeper insights into their modes of action, thereby enhancing our understanding of their therapeutic potential.

Key words Activity-Based Protein Profiling, Activity-Based Probes, Chemical Probes, Drug Discovery, Inhibitors, Microplates, Plate-Based Assay, Screening

1 Introduction

Enzymes, as pivotal catalysts of biochemical reactions, are at the heart of metabolic pathways and are, thus, prime targets for therapeutic intervention [1]. The modulation of enzyme activity through inhibitors is a cornerstone of both medicinal chemistry and industrial biotechnology [2, 3]. These inhibitors are broadly categorized into reversible and irreversible types, depending on their binding dynamics and the nature of their interaction with the enzyme [4]. Given this, the identification and characterization

of enzyme inhibitors are critical goals for multiple applications. Despite being capable of running in ultra-high-throughput screening formats with millions of compounds at very low concentrations, traditional screening methods, such as enzyme-inhibition assays and fluorescence-based approaches, fail to provide a fully comprehensive understanding of enzyme inhibition profiles [5]. Moreover, reaching the final stage of this process involves several steps: identifying and isolating the enzyme, understanding and characterizing the chemistry it catalyzes, and designing an assay. Only then can it be assembled into a format that enables the screening for inhibitors.

In this chapter, we describe the development and application of a method for accelerating this process by enabling enzyme discovery and the identification and characterization of enzyme inhibitors to be directly coupled. Central to this endeavor is activity-based protein profiling (ABPP), a technique that leverages small-molecule chemical probes (activity-based probes or ABPs) to decode the interactions between potential drugs and their protein targets [6, 7]. Since the first reports, multiple forms of ABPP have been developed that not only identify the protein targets of small molecules but also their activity, thereby offering a window into the functional state of proteins within complex biological matrices [8–10]. Collectively ABPP methodologies have driven the development of modern chemical proteomics, which provides deeper insights into the proteome-wide landscape of drug targets [11, 12].

As the name suggests, ABPP focuses on the enzyme's function rather than its structure, requiring only that the enzyme be probed in its functionally active form. In a simple ABPP experiment, the enzyme can be probed either in a lysate or within the context of an entire cell. An ABP contains three basic components: a warhead (a reactive group) designed to label the catalytically active residues of specific proteins or protein families irreversibly and selectively, a tag to enable isolation and downstream identification of the protein(s), and a linker unit to connect these. Most commonly, the reactive group is an electrophile that is complementary to a nucleophilic residue in the active site [13, 14]. ABPs specifically react with related classes of enzymes that share a common mechanism of action, such as the catalytic triad of serine hydrolases [15], as discussed later (*vide infra*). ABP workflows lead to the identification of proteins related by function. Further characterization then requires isolating the protein, either by cloning, or, more commonly now, through *de novo* gene synthesis, and expression. Even then, identifying and characterizing inhibitors requires formatting an assay, which can be complicated by a lack of knowledge about the enzyme's structure and substrate requirements [16].

A solution to these problems can be realized through the application of competitive ABPP, a variant of ABPP [17, 18]. This involves pre-incubating the proteome or a target protein with compounds of interest, followed by the application of the cognate

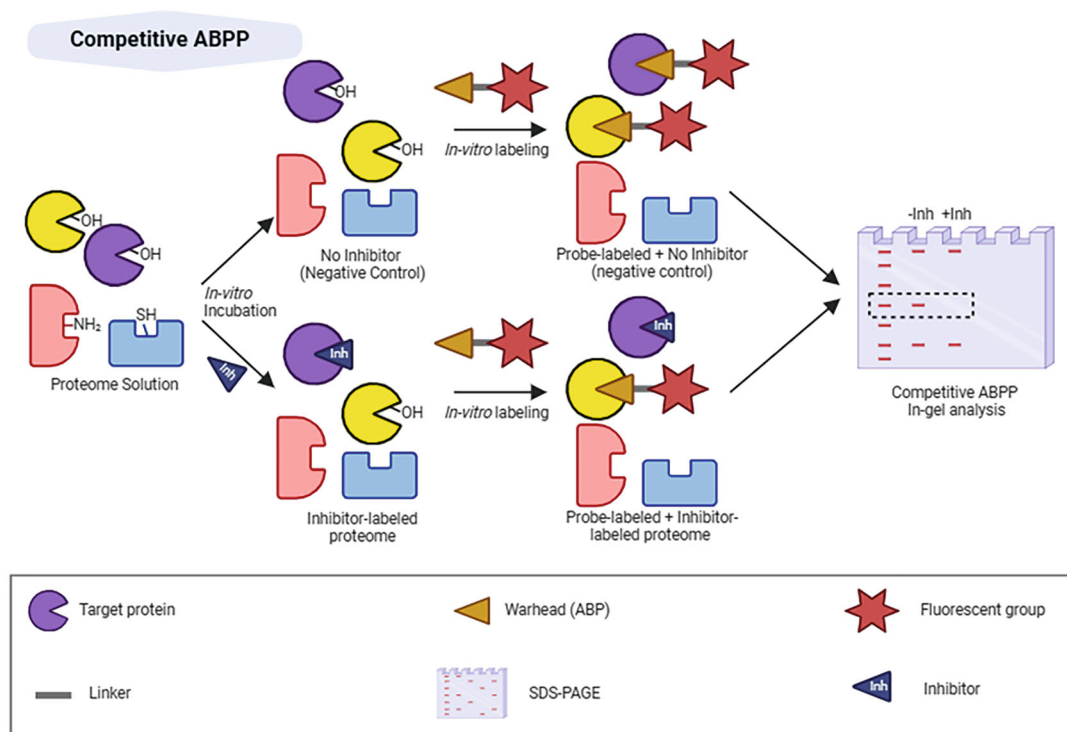


Fig. 1 Competitive ABPP experiments aim to identify inhibitors, using probes that target active-site residues; in our case, the nucleophilic serine found in serine hydrolases. Initially, the proteome is preincubated with the inhibitor, followed by incubation with the fluorescent probe. If the inhibitor effectively blocks the probe's binding site, this results in a reduction or complete absence of fluorescence intensity for the targeted protein. (Figure adapted from 6)

ABP used in the original identification process. If the compound interacts with the active site, it decreases the probe's binding with the target protein, by competing for the common binding site (Fig. 1), reducing the signal (response) obtained from the probe [19]. This setup facilitates the direct detection of compound–protein interactions by exploiting the competition for binding sites, thus providing invaluable insights into the mechanism of action of potential drugs. Importantly, as the probe reports based on enzyme function, this experiment does not require knowledge of the enzyme structure in order to identify an inhibitor. Therefore, this technology has a significant potential for high-throughput screening of targets with poorly characterized substrate or biological functions, for which traditional screening methods are not applicable [6].

Building on these processes, here we introduce a novel screening methodology that couples competitive ABPP with plate-based assays to afford a new method for identifying and mapping the inhibition profiles of enzyme inhibitors. This combination and integration offer a robust, efficient, rapid, cost-effective, and

versatile platform for the identification and characterization of enzyme inhibitors, even in the absence of knowledge of enzyme structure and substrate selectivity, thereby facilitating the acceleration of drug discovery and development.

To demonstrate the technique, we have selected a simple, commercially readily available enzyme pig liver esterase (PLE). PLE is a typical serine hydrolase, a diverse and widely distributed class of enzymes that play crucial roles in various biological processes [20]. Serine hydrolases are characterized by their serine-based mechanism of action, where a serine residue acts as a nucleophile in catalyzing the hydrolysis of ester, thioester, and amide bonds in substrates [21]. This functional versatility renders them essential for a myriad of physiological functions, including digestion, metabolism, signal transduction, and the regulation of post-translational modifications. Their ubiquity and significance in nature are underscored by their involvement in pathophysiological conditions, making them attractive targets for therapeutic intervention [22]. The importance of serine hydrolases extends beyond their biological functions; they also serve as markers for various diseases and are pivotal in the development of diagnostic and therapeutic strategies [23, 24].

Recognizing this, there is a plethora of electrophilic groups that selectively and covalently react with the nucleophilic serine at the active sites of serine hydrolases (e.g., fluorophosphonate, sulfonyl fluoride, β -lactams, coumarin, etc.) [25]. All of these have been found in ABPs offering diverse levels of selectivity. However, the most common (and the warhead of the probe employed in this study) are the fluorophosphonates (FP). FPs are favored due to their exquisite balance between selectivity and reactivity and their low rate of off-target labeling [26]. They are routinely utilized in competitive ABPP experiments to assess the selectivity and activity of serine hydrolase inhibitors within complex proteomes [24, 27–31]. Importantly for this demonstration, FPs have been used to characterize and quantify serine hydrolases, including PLE, in various assays [32]. This makes them an ideal example of the cognate ABP used for initial target protein identification.

Although the demonstration shown below uses a well-established serine hydrolase model (PLE), this methodology can accommodate a broad range of protein of interest (POI), including those that are not well-defined. The only requirement is the ability for the POI (enzyme) to respond (be labeled by) to an ABP. Moreover, the methodology is compatible with automation, and this, combined with the potential for miniaturization, represents an opportunity to significantly increase throughput. Although in our case we used a synthetically accessible in-house cell-permeable FP probe, it is worth noting that the assay process can be further streamlined by using probes directly conjugated with fluorescent

markers (e.g., the commercially available TAMRA-FP), eliminating (bypassing) the need for the additional click-chemistry steps and thereby simplifying the procedure.

In summary, the capability to conduct these assays in a high-throughput format, using 96-well microplates or potentially even higher-density formats, along with the previously highlighted potential for automation and miniaturization, establishes this protocol as a powerful tool for drug discovery. It allows for the rapid screening of extensive libraries of inhibitors against a diverse set of enzymes, including those not yet well-characterized, facilitating the identification and characterization of novel enzyme inhibitors coupled with insights into their modes of action. Through the application of this innovative ABPP microplate-based enzymatic activity assay protocol, we anticipate significant developments in the discovery of next-generation therapeutics.

2 Materials

Prepare all solutions using Milli-Q ultrapure water (obtained by purifying deionized water to achieve a resistivity of 18 M Ω /cm at 25 °C) and analytical or biological grade reagents. Prepare and store all reagents at room temperature (RT), unless indicated otherwise. Follow all waste disposal regulations when disposing of waste materials.

Unless otherwise indicated, all chemical reagents have been purchased from Sigma-Aldrich (Merck). Thin-layer chromatography (TLC) was carried out on Merck TLC silica gel 60 F₂₅₄ aluminum sheets and spots visualized by UV light (λ = 254 nm). Flash column chromatography was performed on fine spherical silica gel, 20–40 μ m particle size (Teledyne ISCO).

2.1 Reagents and Chemicals for Organic Synthesis and Assay Setup

1. 1 N hydrochloric acid (HCl).
2. 4-Bromo-1-butyne.
3. Brine.
4. Bifunctional TAMRA-N₃ (TAMRA-PEG₃-N₃, Biopharma PEG, catalog number: 10685).
5. Diethylaminosulfur trifluoride (DAST).
6. Dimethyl sulfoxide (DMSO) (Hybri-Max™, sterile-filtered, BioReagent, purity $\geq$ 99.7%, Sigma-Aldrich, catalog number: D2650).
7. Dry dichloromethane (DCM).
8. Ethanol (EtOH).
9. Ethyl acetate (EtOAc).
10. Inhibitors (OpnMe Library, Boehringer Ingelheim).

11. Phenylmethylsulfonyl fluoride (PMSE, Roche, Sigma-Aldrich, catalog number: 10837091001).
12. Phosphate-buffered saline (PBS, pH 7.2). Add about 300 mL of Milli-Q water to a 1 L graduated cylinder or a glass beaker. Dissolve 4 g of NaCl (137 mM), 100 mg of KCl (2.7 mM), 720 mg of Na₂HPO₄ (10 mM), and 120 mg of KH₂PO₄ (1.8 mM). Mix until fully dissolved. Verify and adjust pH to 7.2 if necessary. Bring the final volume to 500 mL with Milli-Q water.
13. Pig liver esterase (PLE, lyophilized, powder, ≥50 U/mg, Sigma-Aldrich, catalog number: 46058).
14. Sodium hydroxide (NaOH).
15. Streptavidin Conjugation Kit (Lightning-Link[®] from Abcam, catalog number: ab102921, containing streptavidin mix, modifier reagent and quencher reagent).
16. Triethylphosphate [P(OEt)₃].

2.2 Consumables

1. 96-well microplates (Pierce[™] Biotin Coated Plates, Clear, 8-Well Strip, Thermo Scientific[™], catalog number: 15151).
2. Aluminum foil (for covering and protecting light-sensitive reactions).
3. Eppendorf tubes.
4. Glass vials and Falcon tubes.
5. Pipettes and pipette tips.
6. RediSep Gold[®] Silica Gel Disposable Flash Columns (Teledyne ISCO).
7. Round-bottom flasks of various sizes.

2.3 Equipment

1. Flash column chromatography setup with fine spherical silica gel (20–40 μm particle size). Teledyne ISCO CombiFlash NextGen 300+ Chromatography System.
2. Fluorescent Scanner (Typhoon FLA9500 phosphor imager, GE Healthcare Life Sciences).
3. Laboratory Magnetic Stirring Heating apparatus capable of maintaining 150 °C.
4. Microplate mixer.
5. Microplate reader capable of measuring absorbance and/or fluorescence (Thermo Scientific[™] Multiskan[™] GO Microplate Spectrophotometer).
6. pH meter.
7. pH strips (pH range from 1 to 14, from Merck).
8. Rotary evaporator for removing volatiles.

9. TLC setup with Merck TLC silica gel 60 F₂₅₄ aluminum sheets.
10. UV light for TLC visualization ($\lambda = 254$ nm).
11. Vacuum filtration system equipped with a 0.22- μ m filter (Millipore[®] Steritop[®] Vacuum Bottle Top Filter).
12. Vortex mixer.

2.4 Click-Chemistry Solutions

1. *t*-Butanol–DMSO Solution Preparation (*t*-BuOH:DMSO 1:4): Heat a bottle of *t*-butanol in a water bath set to 40 °C. Occasionally, agitate the bottle over 20 minutes to facilitate melting. In parallel, measure 40 mL of DMSO into a 50 mL Falcon tube. Once the *t*-butanol has completely melted, use a pre-warmed 10 mL glass pipette to transfer 10 mL of *t*-butanol into the Falcon tube containing DMSO. Securely close the Falcon tube and mix the solution by inverting several times. Store this *t*-BuOH:DMSO solution at RT (*see* **Note 1**). 20 mM TBTA solution: Weigh 200 mg of TBTA directly into a 50 mL Falcon tube. Add 18.8 mL of the prepared *t*-BuOH:DMSO solution to the tube and mix, using a vortex, until fully dissolved. This might take some time. Once the solution is homogeneous and clear, aliquot into 500 μ L volumes and store at –20 °C for up to 1 year.
2. 100 mM TCEP solution: Weigh 100 mg of TCEP into a 4 mL brown glass screw-cap bottle. Add 3.5 mL of Milli-Q water and mix until completely dissolved, which may take some time. After achieving a clear solution, aliquot into 500 μ L volumes and freeze at –20 °C. Use within 6 months.
3. 100 mM CuSO₄: Weigh 625 mg of CuSO₄·5 H₂O (or 400 mg of CuSO₄ anhydrous) into a 50 mL Falcon tube. Dissolve in 25 mL of Milli-Q water, mixing well until the solution is clear and pale blue. Aliquot the solution into 500 μ L volumes. This solution can be stored at –20 °C indefinitely.
4. 20 mM bifunctional TAMRA-Azide Solution (TAMRA-N₃ in DMSO): This reporter molecule (TAMRA-PEG₃-N₃) includes a fluorescent TAMRA group for detection and an azide group for click-chemistry. Weigh 10 mg of TAMRA-PEG₃-N₃ into a 4 mL brown glass screw-cap bottle. Add 792 μ L of DMSO and mix until completely dissolved. After achieving a red clear solution, aliquot into 200 μ L volumes and freeze at –20 °C. Use within 6 months (*see* **Note 2**).

2.5 Streptavidin-Conjugated Protein of Interest (POI) Solution: Pig Liver Esterase (PLE)

1. 10 mM borate buffer (pH 8.0): Pour 80 mL of Milli-Q water into your container (*see* **Note 3**). Add 0.062 g of boric acid to the water (10 mM). Introduce 0.439 g of sodium chloride (75 mM) into the mixture. Use a stir bar and apply low heat to help dissolve the boric acid and sodium chloride completely (*see* **Note 4**). Add more Milli-Q water until the total volume of

the solution reaches 100 mL. Check the pH of the buffer. If it is not at the desired pH of 8.0, carefully add NaOH to adjust the pH upward. Use a pH meter or pH strips to monitor the pH accurately. Sterilize the solution by vacuum filtration using a 0.22- μ m filter (*see Note 5*).

2. POI solution—pig liver esterase (PLE) in borate buffer: Dissolve 1 mg of PLE in 1 mL of 10 mM borate buffer (pH 8.0) to ensure a final solution concentration of 1 mg/mL (*see Note 6*). Gently stir the solution until complete dissolution of PLE (*see Note 7*).
3. Preparation of POI (PLE)—streptavidin conjugated: Ensure all materials and prepared reagents are equilibrated to RT before use (*see Note 8*). Add 1 μ L of modifier reagent to every 10 μ L of POI to be labeled and gently mix (*see Notes 9–11*). Remove the cap from the vial containing Streptavidin Conjugation Mix and pipette the POI sample (with added modifier reagent) directly onto the lyophilized material. Resuspend gently by withdrawing and re-dispensing the liquid once or twice using a pipette. Replace the cap on the vial and allow it to stand for 3 h in the dark at RT (20–25 °C). Longer incubation times, such as overnight, do not have a detrimental effect on the conjugation. After incubating for 3 h (or more), add 1 μ L of Quencher reagent for every 10 μ L of POI used and gently mix. The conjugate can be used after 30 min. Purification of the conjugates is not necessary. Conjugate storage: Typically, the POI conjugate can be stored at 4 °C for up to 18 months. For longer storage, the conjugate can be stored at –20 °C with a cryoprotectant like 50% glycerol. Always store the POI conjugate in a dark environment.

3 Methods

3.1 *Synthesis of Diethyl (but-3-yn-1-yl) Phosphonate (FP-C4 Intermediate)*

1. Gather all required reagents, solvents, and equipment.
2. In a suitable reaction round-bottom flask, mix together 1.0 mL (6 mmol, 1 equivalent) of P(OEt)₃ with 1.2 mL (12.5 mmol, 2.1 equivalents) of 4-bromo-1-butyne.
3. Heat the reaction mixture at 150 °C for 2.5 h under an inert atmosphere. Ensure the reaction mixture is well contained and that the heating apparatus is set up to maintain a steady temperature.
4. After the heating period, allow the reaction mixture to cool down to RT. This may be done by simply removing the heat source and allowing natural cooling or employing a cooling bath to speed up the process.
5. Once the mixture has cooled to RT, remove volatile solvents. This can typically be done using a rotary evaporator. The aim is

to concentrate the reaction mixture by removing the excess of unreacted starting material.

6. After the removal of volatiles, the desired intermediate, compound **FP-C4 Intermediate**, should be obtained as a yellow oil. From the reaction, you should expect to obtain approximately 412 mg of compound **FP-C4 Intermediate**, which corresponds to a yield of 38% (*see* **Note 12**). **STOP POINT.**

3.2 Synthesis of Ethyl (but-3-yn-1-yl) Phosphonofluoridate (FP-C4 Probe)

1. Mix 100 mg of compound **FP-C4 Intermediate** (0.53 mmol, 1 equivalent) in a solution of 2 mL water and 2 mL ethanol in a round-bottom flask.
2. Add 64 mg of NaOH (1.60 mmol, 3 equivalents) to the solution, stirring to dissolve.
3. Heat the reaction mixture at 85 °C for 8 h, ensuring constant stirring.
4. After the reaction time, concentrate the mixture to remove ethanol.
5. Dilute the concentrated mixture with 1 mL water, then acidify with 1 mL of 1 N HCl.
6. Perform an extraction using three portions of 25 mL ethyl acetate, combining the organic layers after each extraction.
7. Wash the combined ethyl acetate extracts once with 5 mL of brine.
8. Concentrate the washed organic extracts under vacuum. **STOP POINT.**
9. Dissolve the crude product in 4 mL dry DCM, cool the solution to −78 °C, and add 0.146 mL of DAST (1.05 mmol, 2 equivalents) dropwise, stirring for 3 h at −78 °C under an inert atmosphere.
10. After stirring, quench the reaction with water.
11. Remove volatiles under vacuum, then extract the residue with 25 mL ethyl acetate, washing the organic layer once with 5 mL water and once with 5 mL brine.
12. Dry the organic layer and purify it using flash chromatography (Teledyne ISCO RediSep Gold[®] Silica Gel Disposable Flash Columns), applying a hexane:ethyl acetate 0–100% gradient. You should obtain approximately 60 mg of alkyl fluorophosphonate **FP-C4 Probe** (*see* **Note 13**) as a yellow liquid, which corresponds to a yield of 69% (Fig. 2) (*see* **Notes 14** and **15**). **STOP POINT.**

3.3 Plate-Based Preparation and Immobilization

1. Gather all required reagents, buffers, and equipment. Ensure all materials and prepared reagents are equilibrated to RT before use.

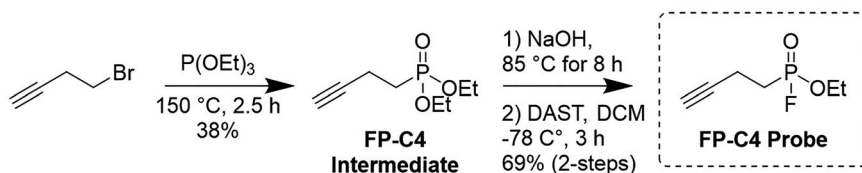


Fig. 2 In-house synthesis of **FP-C4 Probe**

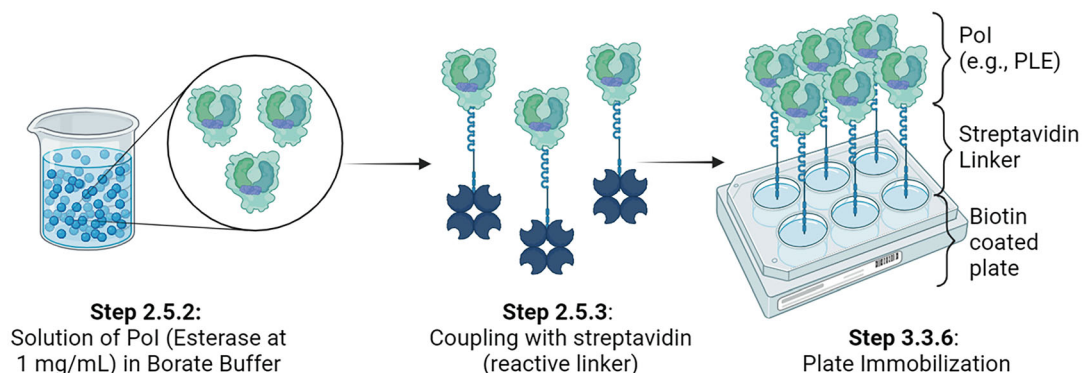


Fig. 3 Diagram of the preparation of the microplate containing the immobilized protein of interest

2. Rinse each well of the plate three times using 200 μL of PBS (*see Note 16*).
3. For an initial concentration of 1 mg/mL of POI (PLE)-streptavidin (1000 $\mu\text{g/mL}$), dilute the conjugate 1/100 by adding 1 μL of the POI-streptavidin conjugate to 99 μL of the borate buffer (per experiment).
4. Add 100 μL of the POI-streptavidin conjugate to each well.
5. Allow the plate to incubate at RT for 30 min (*see Note 17*).
6. After incubation, wash each well three times with 200 μL of borate buffer to remove any unbound conjugate (*see Note 18*). Now the microplate is ready for assays (Fig. 3). **STOP POINT.**

3.4 Protocol to Calculate the Percentage of Inhibition at a Fixed Concentration of Inhibitor: Activation and Procedure for Plate-Based Competitive ABPP Experiment

1. Ensure all materials and prepared reagents are equilibrated to RT before use.
2. Add 100 μL of borate buffer per well (Fig. 4) (*see Note 19*).
3. Add 0.5 μL of inhibitor solution (prepared at 20 mM concentration in DMSO) per well to achieve a final concentration of 100 μM . For control samples, use 0.5 μL of DMSO without inhibitor as the vehicle (Fig. 4) (*see Note 20*).
4. Incubate for 60 minutes at RT ($\sim 20\text{ }^{\circ}\text{C}$) (*see Note 21*).
5. Add 0.5 μL of **FP-C4 Probe** (from a 10 mM stock solution in DMSO) to achieve a final concentration of 50 μM in the

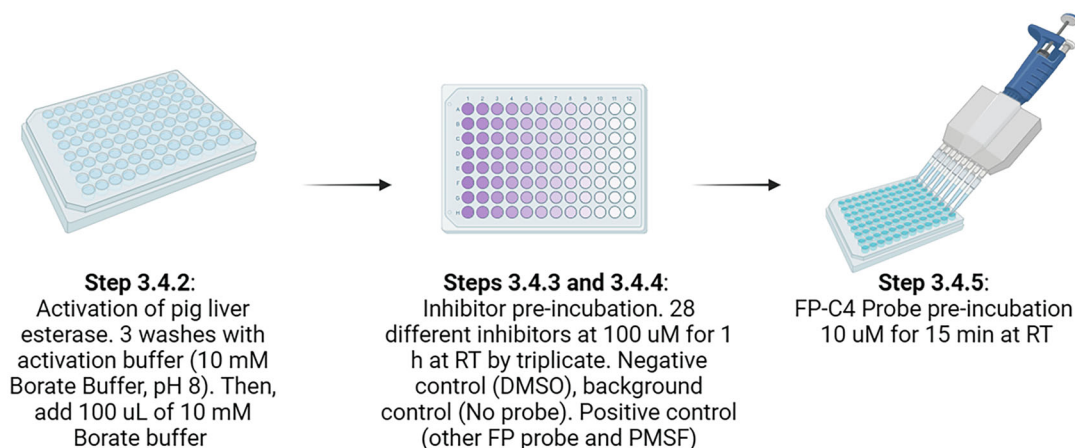


Fig. 4 Diagram of the core workflow containing the competitive ABPP experiment conducted in microplates

sample. Mix and incubate the samples for 15 minutes at RT (Fig. 4) (*see* **Notes 22** and **23**).

6. Introduce 0.5 μ L of commercial bifunctional TAMRA- N_3 solution (20 mM stock in DMSO) to each sample to achieve a final concentration of 100 μ M in the sample. Mix for even distribution.
7. Add 1 μ L of tris(carboxyethyl)phosphine (TCEP) from a 100 mM stock solution in water to achieve a final concentration of 1 mM in the sample. Mix the sample gently.
8. Incorporate 0.5 μ L of TBTA ligand from a 20 mM stock solution in a *t*-BuOH:DMSO mixture for a final concentration of 100 μ M. Mix the microplate gently.
9. To initiate the click-chemistry reaction, add 1 μ L of $CuSO_4$ (from a 100 mM stock solution in water) to achieve a final concentration of 1 mM in the sample. Mix gently to ensure uniform distribution of components.
10. Incubate the samples at RT for 1 h in the dark under constant mixing, if possible; otherwise, mix every 15 min to ensure homogeneity (Fig. 5) (*see* **Note 24**).
11. After incubation, wash each well three times with 200 μ L of borate buffer or PBS to remove any unbound TAMRA- N_3 reporter. Then, add 100 μ L of the same buffer to each well (*see* **Note 25**).
12. Measure absorbance using a Thermo Scientific™ Multiskan™ GO Microplate Spectrophotometer, targeting the TAMRA- N_3 excitation maximum (Ex = 546 nm) (Fig. 5). Alternatively, use an equivalent microplate reader that meets these specifications (*see* **Note 26**). Additionally, measure fluorescence using a

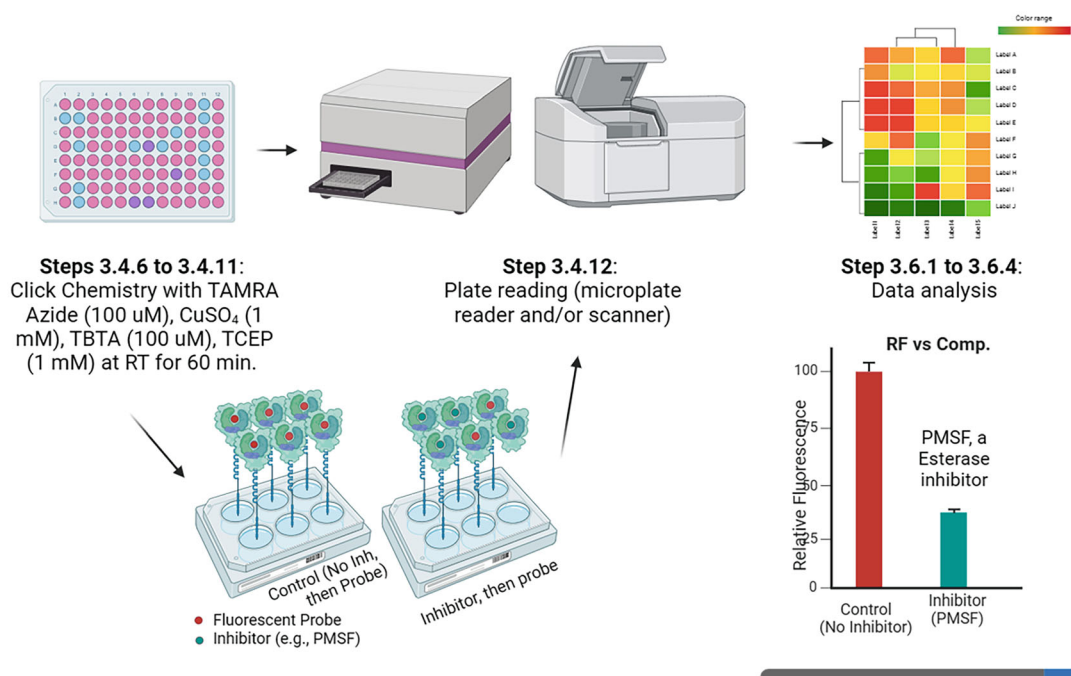


Fig. 5 Diagram illustrating the procedure for generating, collecting, and processing data from our assay

Typhoon FLA9500 phosphor imager, GE Healthcare Life Sciences, targeting the TAMRA- N_3 emission maximum (Ex = 579 nm) (*see Note 27*). **STOP POINT.**

3.5 Protocol to Determine the IC_{50} of Inhibitors: Activation and Procedure for Plate-Based Competitive ABPP Experiment

1. Prepare a series of dilutions from the 20 mM stock solution of the inhibitor to cover a concentration range from 200 μ M to 0.1 μ M (or lower). Utilize DMSO to maintain the solubility of the inhibitor across all dilutions. Prepare at least 8–10 concentrations in a twofold dilution series to accurately define the IC_{50} curve (*see Note 28*).
2. Repeat steps 3.4.1–3.4.2.
3. Introduce the inhibitor solutions to the wells to achieve the desired final concentrations, adjusting the volume of inhibitor solution added according to the dilution calculations (ideally it would be recommended to add ~ 0.5 μ L per experiment). For control samples, use 0.5 μ L of DMSO without inhibitor.
4. Repeat steps 3.4.4–3.4.12.

3.6 Data Analysis for This Assay

1. Normalize enzyme activity: Use the intensity of the wells treated with the vehicle (DMSO) as the reference point, setting it at 100%. Treat the condition with DMSO and no probe as the baseline, setting it at 0% (*see Note 29*).

2. Normalize absorbance or fluorescence intensity: For the remaining wells, normalize the absorbance or fluorescence intensity by using the well intensity of the vehicle-treated proteins to 100% and the condition with DMSO and no probe as the baseline (0%). Then, adjust this normalized value to account for the total protein loaded in each well, as determined in the previous **step 1**.
3. Transfer data: Input the normalized and corrected wells intensity data into a GraphPad Prism (or other comparable software) data table for further analysis (Fig. 5). **STOP POINT.**
4. For IC₅₀ calculations: Use the LOG [inhibitor] vs. normalized response setting with a variable slope to fit an IC₅₀ curve. This calculation provides an apparent IC₅₀ value, which is valuable for comparing inhibitors evaluated under identical assay conditions (*see* **Note 30**). **STOP POINT.**

Using our innovative protocol, we successfully screened 32 experimental conditions on a single 96-well plate, comprising 1 negative control (DMSO), 1 background control (no probe), and 30 inhibitors at single concentration (100 μ M). This set of inhibitors included two positive controls (a synthetic FP probe and PMSF) along with 28 molecules from the Boehringer Ingelheim OpnMe library. Each condition was tested in triplicate (Fig. 6a). Additionally, we established a dose–response curve (Fig. 6b) for PMSF (Fig. 6c) against the PLE, evaluating 18 distinct concentrations, also in triplicate and, as an internal control, experiments employing classical competitive in-gel ABPP have been conducted to validate our methodology (Fig. 6d).

As a result, we have showcased the versatility and efficacy of a novel microplate-based enzymatic activity assay protocol that employs ABPP methods for the rapid and sensitive screening of enzyme inhibitors. Central to our approach is the application of competitive ABPP, a method that provides precise insights into the interaction dynamics between inhibitors and targeted enzymes. Nonetheless, traditional (classical) competitive ABPP experiments often depend on multiple gels or quantitative chemoproteomics mass spectrometry, requiring substantial material, incurring high costs, and involving cumbersome procedures. These factors collectively hinder the ability to use ABPP to conduct large-scale screenings for inhibitor discovery [33]. For comparison, to illustrate the efficiency and scalability of this method, conducting the same analysis using traditional in-gel methods would have necessitated 10 individual gels for the inhibitor screening and another 6 gels for the dose–response evaluation. In response to these limitations, we have developed this new methodology that combines the sensitivity and precision of competitive ABPP with the efficiency of microplate assays, enabling the simultaneous operation of numerous assays.

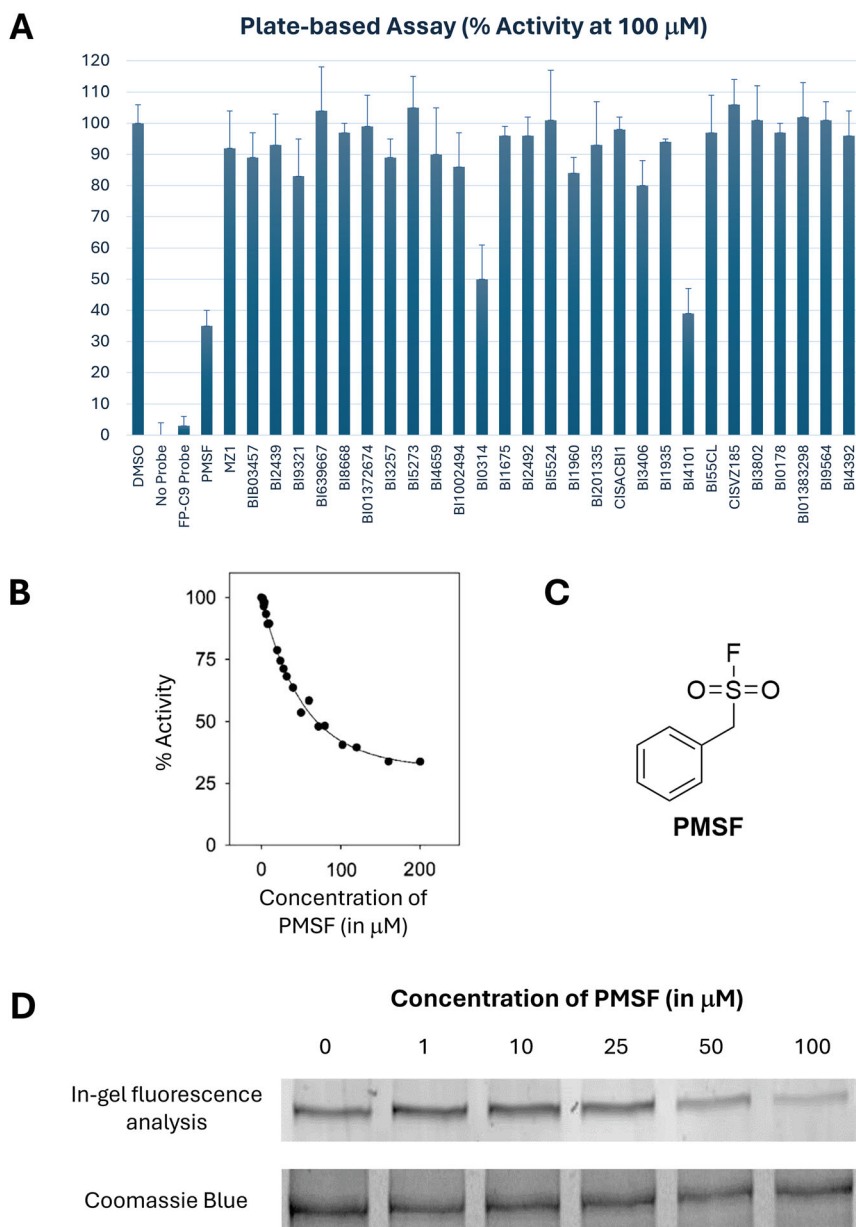


Fig. 6 Comprehensive analysis of inhibitor screening and dose–response evaluation using our 96-plate-based assay. **(a)** Results of the experimental setup showcasing 32 conditions including 1 negative control (DMSO), 1 background control (No probe), and 30 inhibitors at a single concentration (100 μ M) of which 28 are from the Boehringer Ingelheim OpnMe library, tested in triplicate. As a result, we have identified BI4101 as a potential inhibitor of PLE and are currently working on further optimization. **(b)** Dose–response curve established for PMSF against PLE using our plate-based assay. Calculation of PMSF's IC_{50} value against PLE, determined to be 64 ± 3 μ M. **(c)** Chemical structure of PMSF. **(d)** Validation of the methodology through fluorescent competitive in-gel ABPP as an internal control. **(d, top)** In-gel fluorescence analysis (emission at 560 nm) of the competitive ABPP between increasing concentration of PMSF and FP-C4 Probe (50 μ M) in pig liver esterase (1 mg/mL). Calculation of PMSF's IC_{50} value against PLE, determined to be 57 ± 8 μ M. **(d, bottom)** Coomassie blue protein stain

4 Notes

1. Ensure the glass pipette is preheated to avoid rapid cooling and solidification of the *t*-butanol.
2. Although commercial versions of this molecule are available, it can also be synthesized in the laboratory.
3. This preliminary volume makes it easier to dissolve the solids before reaching the final volume.
4. The goal is to have a clear solution without any undissolved particles.
5. This step is essential for removing any microorganisms and particulates, making the buffer suitable for our assay. This buffer can be stored at 4 °C for several months.
6. Ideally, use a kit to calculate the protein concentration.
7. Avoid vortex. This solution can be stored at –20 °C for several months.
8. Buffers containing amines should be avoided.
9. It is crucial to adhere to the specified fixed volumes. Modifying the reaction volume may lead to suboptimal conjugates.
10. POI can be diluted using either Milli-Q water, borate buffer, or PBS.
11. Adding less than the recommended amount of POI may result in unbound label post-conjugation. Excess label will be deactivated by the quencher and subsequently removed during the initial wash step of any application.
12. Diethyl (but-3-yn-1-yl)phosphonate: $\nu_{\max}$ (ATR) 3311 (C-H alkyne), 2244 (C $\equiv$ C), 1242 (P=O) cm^{-1} . δH NMR (400 MHz, CDCl_3) 4.19–4.01 (4H, m, $-\text{OCH}_2\text{CH}_3$), 2.47 (2H, m, 1-H₂), 2.05–1.93 (3H, m, 4-H, 2-H₂), 1.32 (6H, t, J 7.0, $-\text{OCH}_2\text{CH}_3$). δC NMR (101 MHz, CDCl_3) 82.8 (d, $^3J_{\text{C-P}}$ 20.6, C3), 69.2 (d, $^4J_{\text{C-P}}$ 1.8, C-4), 61.9 (d, $^2J_{\text{C-P}}$ 6.5, $-\text{OCH}_2\text{CH}_3$), 25.34 (d, $^1J_{\text{C-P}}$ 142.0, C-1), 16.6 (d, $^3J_{\text{C-P}}$ 6.0, $-\text{OCH}_2\text{CH}_3$), 12.7 (d, $^2J_{\text{C-P}}$ 3.6, C-2). δP NMR (162 MHz, CDCl_3) 28.9. m/z LCMS (ES^+) 191.2 $[\text{M} + \text{H}]^+$.
13. Once the probe is synthesized, it is strongly recommended to conduct preliminary LC-MS and proteomics experiments to verify that the probe selectively reacts with the nucleophilic active residue of the POI.
14. Ethyl (but-3-yn-1-yl)phosphonofluoridate: $\nu_{\max}$ (ATR) 3305 (C-H alkyne), 2248 (C $\equiv$ C), 1677 (C=O), 1272 (P=O), 1043 (CO) cm^{-1} . δH NMR (599 MHz, CDCl_3) 4.35–4.23 (2H, m, $-\text{OCH}_2\text{CH}_3$), 2.60–2.51 (2H, m, 2-H₂), 2.22–2.12 (2H, m, 1-H₂), 2.05 (1H, t, J 2.7, 4-H), 1.39 (3H, t, J 7.1,

–OCH₂CH₃). ¹³C NMR (151 MHz, CDCl₃) 81.4 (d, ³J_{C-P} 20.1, C-3), 70.0 (d, ⁴J_{C-P} 1.8, C4), 63.8 (d, ²J_{C-P} 7.1, –OCH₂CH₃), 24.2 (dd, ¹J_{C-P} 145.8, ¹J_{C-F} 24.0, C-1), 16.5 (d, ³J_{C-P} 5.6, –OCH₂CH₃), 12.1 (d, ²J_{C-P} 3.5, C-2). ³¹P NMR (162 MHz, CDCl₃) 28.1 (d, ¹J_{P-F} 1070). ¹⁹F NMR (376 MHz, CDCl₃) –63.6 (d, ¹J_{F-P} 1070). HRMS (ES⁺) found [M + H]⁺ 165.0476; C₆H₁₀FO₂P requires M 165.0481.

15. Product and intermediate can be stored at –20 °C for up to 6 months. While this compound has not been individually assessed, fluorophosphonates are potentially highly toxic and can cause irreversible effects. Appropriate caution should be taken when handling them.
16. This step ensures that any residual substances are removed, and the wells are ready for the enzyme conjugates.
17. This step is crucial for the conjugates to bind properly to the well surfaces.
18. Proper washing is critical for reducing background noises in the assay.
19. Volume adjustments are feasible; however, any changes require corresponding modifications to the volumes of added reagents to preserve the reagent concentrations in each experiment.
20. Ensure that each experiment is conducted in triplicate to guarantee reproducibility and reliability of the results.
21. If available, it is recommended to use a microplate mixer for this procedure. Utilizing a microplate mixer ensures uniform mixing of the samples, which is crucial for the consistency and reliability of the experiment's results.
22. Set up a control sample that includes neither the **FP-C4** Probe nor inhibitors, using only DMSO as the component.
23. Although FP-TAMRA is a viable commercial alternative, it is often more expensive and less versatile compared to the **FP-C4** Probe [34].
24. This step allows for the click-chemistry reaction to proceed, attaching the TAMRA-N₃ tag to the labeled POI-streptavidin conjugate.
25. Proper washing is critical for reducing background noise in the assay. The volume of the final buffer addition before measuring absorbance or fluorescence can be adjusted to control signal intensities. This adjustment should be based on the linearity of the measurement functions and the sensitivity of the detecting equipment.
26. If your microplate reader is equipped with a fluorescence detection unit, configure it to detect at the maximum emission wavelength of TAMRA (Em = 579 nm).

27. Fluorescence profiles and peak volume quantitation for labeled proteins are generated using the instrument's ImageQuant TL v2005 software.
28. For example, to prepare the first dilution at 200 μM from a 20 mM stock, mix 10 μL of the stock solution with 990 μL of DMSO. For subsequent dilutions, take 500 μL from the previous dilution and mix with 500 μL of DMSO.
29. Each experiment must have been performed at least in triplicate.
30. The calculated IC_{50} is an “apparent IC_{50} ” value, which serves as a comparative measure of inhibitor efficacy within the same assay setup. However, it is important to recognize that this value is influenced by various assay parameters such as protein concentration, the affinity of the probe for the enzyme, incubation times, and temperature. This context dependency means the apparent IC_{50} may vary across different experimental conditions.

Acknowledgments

We thank The Royal Society (The Royal Society International Collaboration Awards for Research Professors 2016: IC160044) and UKRI Grand Challenges Research Fund (“A Global Network for Neglected Tropical Diseases” grant number MR/P027989/1) for financial support; EP was supported, in part, by grant from the COFUND (Marie Skłodowska-Curie Actions, European Union and Durham University)—Durham International Fellowships for Research and Enterprise, Junior Research Fellowships (JRF) scheme.

References

1. de la Fuente M, Lombardero L, Gómez-González A et al (2021) Enzyme therapy: current challenges and future perspectives. *Int J Mol Sci* 22(17):9181. <https://doi.org/10.3390/ijms22179181>
2. Holdgate GA, Meek TD, Grimley RL (2018) Mechanistic enzymology in drug discovery: a fresh perspective. *Nat Rev Drug Discov* 17(2): 115–132. <https://doi.org/10.1038/nrd.2017.219>
3. Wiltshi B, Cernava T, Dennig A et al (2020) Enzymes revolutionize the bioproduction of value-added compounds: from enzyme discovery to special applications. *Biotechnol Adv* 40: 107520. <https://doi.org/10.1016/j.biotechadv.2020.107520>
4. Strelow J, Dewe W, Iversen PW et al (2012) Mechanism of action assays for enzymes. In: Markossian S, Grossman A, Brimacombe K et al (eds) *Assay guidance manual*. Eli Lilly & Company and the National Center for Advancing Translational Sciences, Bethesda (MD). PMID: 22553872
5. Szymański P, Markowicz M, Mikiciuk-Olasik E (2012) Adaptation of high-throughput screening in drug discovery-toxicological screening tests. *Int J Mol Sci* 13(1):427–452. <https://doi.org/10.3390/ijms13010427>
6. Porta EOJ, Steel PG (2023) Activity-based protein profiling: a graphical review. *Curr Res Pharmacol Drug Discov* 5:100164. <https://doi.org/10.1016/j.crphar.2023.100164>

7. Fang H, Peng B, Ong SY et al (2021) Recent advances in activity-based probes (ABPs) and affinity-based probes (AFBPs) for profiling of enzymes. *Chem Sci* 12(24):8288–8310. <https://doi.org/10.1039/d1sc01359a>
8. Cravatt BF, Wright AT, Kozarich JW (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* 77:383–414. <https://doi.org/10.1146/annurev.biochem.75.101304.124125>
9. Deng H, Lei Q, Wu Y et al (2020) Activity-based protein profiling: recent advances in medicinal chemistry. *Eur J Med Chem* 191: 112151. <https://doi.org/10.1016/j.ejmech.2020.112151>
10. Porta EOJ (2023) Mapping the evolution of activity-based protein profiling: a bibliometric review. *Adv Pharm Bull* 13(4):639–645. <https://doi.org/10.34172/apb.2023.082>
11. Chandramouli K, Qian PY (2009) Proteomics: challenges, techniques and possibilities to overcome biological sample complexity. *Hum Genomics Proteomics* 2009:239204. <https://doi.org/10.4061/2009/239204>
12. Reinecke M, Brear P, Vornholz L et al (2023) Chemical proteomics reveals the target landscape of 1,000 kinase inhibitors. *Nat Chem Biol* 20:577. <https://doi.org/10.1038/s41589-023-01459-3>
13. Vinogradova EV, Zhang X, Remillard D et al (2020) An activity-guided map of electrophile-cysteine interactions in primary human T cells. *Cell* 182(4):1009–1026.e29. <https://doi.org/10.1016/j.cell.2020.07.001>
14. Roberts AM, Ward CC, Nomura DK (2017) Activity-based protein profiling for mapping and pharmacologically interrogating proteome-wide ligandable hotspots. *Curr Opin Biotechnol* 43:25–33. <https://doi.org/10.1016/j.copbio.2016.08.003>
15. Simon GM, Cravatt BF (2010) Activity-based proteomics of enzyme superfamilies: serine hydrolases as a case study. *J Biol Chem* 285(15):11051–11055. <https://doi.org/10.1074/jbc.R109.097600>
16. Acker MG, Auld DS (2014) Considerations for the design and reporting of enzyme assays in high-throughput screening applications. *Perspect Sci* 1(1–6):56–73. <https://doi.org/10.1016/j.pisc.2013.12.001>
17. Wright MH, Sieber SA (2016) Chemical proteomics approaches for identifying the cellular targets of natural products. *Nat Prod Rep* 33(5):681–708. <https://doi.org/10.1039/c6np00001k>
18. Niphakis MJ, Cravatt BF (2014) Enzyme inhibitor discovery by activity-based protein profiling. *Annu Rev Biochem* 83:341–377. <https://doi.org/10.1146/annurev-biochem-060713-035708>
19. Wang S, Tian Y, Wang M et al (2018) Advanced activity-based protein profiling application strategies for drug development. *Front Pharmacol* 9:353. <https://doi.org/10.3389/fphar.2018.00353>
20. Bachovchin DA, Cravatt BF (2012) The pharmacological landscape and therapeutic potential of serine hydrolases. *Nat Rev Drug Discov* 11(1):52–68. <https://doi.org/10.1038/nrd3620>
21. Long JZ, Cravatt BF (2011) The metabolic serine hydrolases and their functions in mammalian physiology and disease. *Chem Rev* 111(10):6022–6063. <https://doi.org/10.1021/cr200075y>
22. Lucas A, Yaron JR, Zhang L et al (2018) Serpins: development for therapeutic applications. *Methods Mol Biol* 1826:255–265. https://doi.org/10.1007/978-1-4939-8645-3_17
23. Ćwilichowska N, Świdarska KW, Dobrzyń A et al (2022) Diagnostic and therapeutic potential of protease inhibition. *Mol Asp Med* 88: 101144. <https://doi.org/10.1016/j.mam.2022.101144>
24. Porta EOJ, Isern JA, Kalesh K et al (2022) Discovery of leishmania druggable serine proteases by activity-based protein profiling. *Front Pharmacol* 13:929493. <https://doi.org/10.3389/fphar.2022.929493>
25. Faucher F, Bennett JM, Bogyo M et al (2020) Strategies for tuning the selectivity of chemical probes that target serine hydrolases. *Cell Chem Biol* 27(8):937–952. <https://doi.org/10.1016/j.chembiol.2020.07.008>
26. Gillet LC, Namoto K, Ruchti A et al (2008) In-cell selectivity profiling of serine protease inhibitors by activity-based proteomics. *Mol Cell Proteomics* 7(7):1241–1253. <https://doi.org/10.1074/mcp.M700505-MCP200>
27. Liu Y, Patricelli MP, Cravatt BF (1999) Activity-based protein profiling: the serine hydrolases. *Proc Natl Acad Sci USA* 96(26): 14694–14699. <https://doi.org/10.1073/pnas.96.26.14694>
28. Janssen APA, van der Vliet D, Bakker AT et al (2018) Development of a multiplexed activity-based protein profiling assay to evaluate activity of endocannabinoid hydrolase inhibitors. *ACS Chem Biol* 13(9):2406–2413. <https://doi.org/10.1021/acscchembio.8b00534>
29. Xu H, Sabit H, Amidon GL et al (2013) An improved synthesis of a fluorophosphonate-polyethylene glycol-biotin probe and its use against competitive substrates. *Beilstein J Org*

- Chem 9:89–96. <https://doi.org/10.3762/bjoc.9.12>
30. Navarrete M, Ho J, Krokhn O et al (2013) Proteomic characterization of serine hydrolase activity and composition in normal urine. *Clin Proteomics* 10(1):17. <https://doi.org/10.1186/1559-0275-10-17>
31. Baggelaar MP, Van der Stelt M (2017) Competitive ABPP of serine hydrolases: a case study on DAGL- α . *Methods Mol Biol* 1491: 161–169. https://doi.org/10.1007/978-1-4939-6439-0_12
32. Abdel-Daim A, Ohura K, Imai T (2018) A novel quantification method for serine hydrolases in cellular expression system using fluorophosphonate-biotin probe. *Eur J Pharm Sci* 114:267–274. <https://doi.org/10.1016/j.ejps.2017.12.016>
33. Chen X, Wang Y, Ma N et al (2020) Target identification of natural medicine with chemical proteomics approach: probe synthesis, target fishing and protein identification. *Signal Transduct Target Ther* 5(1):72. <https://doi.org/10.1038/s41392-020-0186-y>
34. Isern JA, Porta EOJ, Kalesh K et al (2025) Profiling serine hydrolases in the Leishmania host-pathogen interactome using cell-permeable activity-based fluorophosphonate probes. *ChemBioChem* e202500160. <https://doi.org/10.1002/cbic.202500160>



Identifying Inhibitor Targets in Mycobacteria by Activity-Based Probe Profiling

Neetika Jaisinghani, Isabel Sakarin, Hiren V. Patel, Michael Li, Mary L. Previti, and Jessica C. Seeliger

Abstract

Activity-based protein profiling (ABPP) enables the detection of protein reactivity across entire proteomes. Here, we describe the application of ABPP to the identification of serine hydrolase inhibitor targets in *Mycobacterium tuberculosis* (*Mtb*), the causative agent of tuberculosis in humans. We highlight the adaptation of stable isotope labeling of amino acids (SILAC) to *Mtb* using a modified growth medium with an isotopically labeled sole nitrogen source and detail the resulting special requirements for proteomics analysis.

Key words Activity-based protein profiling, Mycobacteria, Tuberculosis, Inhibitor, SILAC, Proteomics

1 Introduction

Activity-based protein profiling (ABPP) uses chemical tools known as activity-based probes to identify enzyme classes or reactive residues across entire proteomes. Activity-based probes contain two elements: (1) a reactive group that covalently labels active sites, and (2) a reporter tag for downstream visualization or purification of labeled proteins [1, 2]. Distinct from total proteomics, activity-based protein profiling is used to assess enzymatic activity and protein reactivity within proteomes.

ABPP has been applied to mycobacteria to elucidate various aspects of mycobacterial biology and facilitate inhibitor characterization. For a review of the applications of ABPP in mycobacteria up to 2019, we refer the reader to Patel, Li, and Seeliger [3]. The first

Neetika Jaisinghani and Isabel Sakarin contributed equally.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-1-0716-4502-4_7.

enzyme classes profiled by ABPP in *Mycobacterium tuberculosis* (*Mtb*) were serine hydrolases [4, 5] and ATP-binding proteins [6, 7]. More recently, Levine and Beatty profiled penicillin-binding proteins (PBPs) and L,D-transpeptidases (LDTs) using activity-based probes that are analogs of β -lactams [8].

Additional ABPP studies in mycobacteria have focused on identifying the direct targets of anti-mycobacterial agents. In general, the inhibitors have been covalent modifiers that were then functionalized with an alkyne. Following validation that the alkyne derivative retained activity, subsequent click-chemistry to an appropriate moiety enabled target enrichment, detection, and/or identification. Alkyne derivatives have been used to identify the mycobacterial targets of β -lactones, tetrahydrolipstatin [9], lalostat [10], and EZ120 [11]. Quezada et al. profiled the targets of cephalosporins using an alkyne analog of a cephalosporin active against nonreplicating *Mtb* [12]. Canaan and Cavlier have developed cytophostin and cyclopostin (CyC) lipase inhibitors as anti-mycobacterials and have similarly reported using an alkyne derivative to identify CyC targets in the nontuberculous mycobacterium *M. abscessus* [13].

An alternative approach for target identification using ABPP is competitive ABPP, in which an inhibitor competes with an activity-based probe (Fig. 1). In this scenario, the selectivity of the

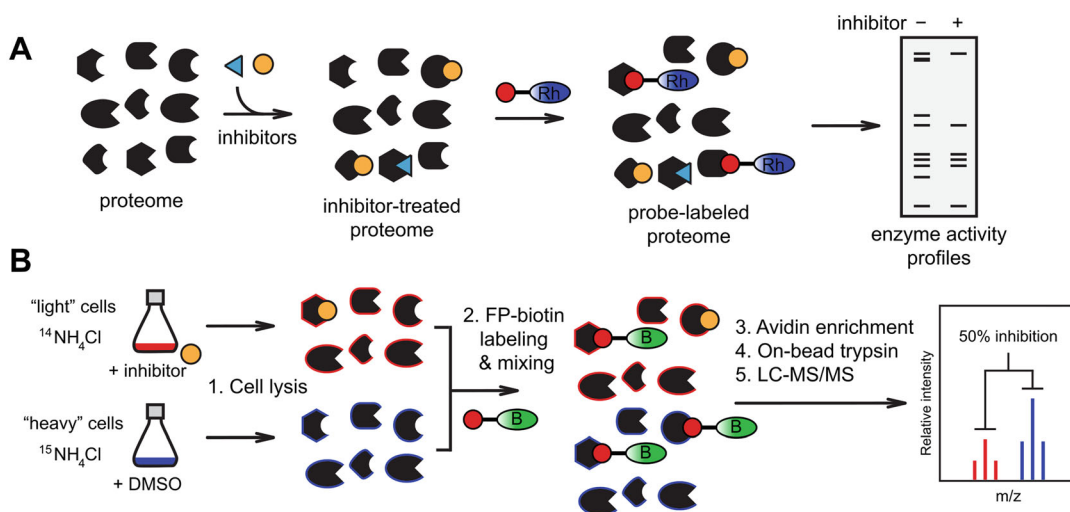


Fig. 1 Workflows for competitive activity-based protein profiling (ABPP) in mycobacteria. **(a)** Gel-based competitive ABPP: Live cells in culture are treated with inhibitor or vehicle control. Lysates of treated cells are labeled with an activity-based probe linked to an appropriate moiety for detection (e.g., Rh = rhodamine) and/or enrichment. The labeled proteome is then separated by SDS-PAGE and protein bands visualized (e.g., by fluorescence or western blot). **(b)** Competitive ABPP-SILAC. Samples are prepared as in (A) except cells are grown in medium with a ^{14}N - or ^{15}N -labeled source (for mycobacteria, a sole nitrogen source such as ammonium chloride) and then treated with inhibitor or vehicle control, respectively. Total lysate is then treated as an ABP linked to an affinity tag (e.g., B = biotin) and labeled proteins are enriched and analyzed by LC-MS/MS. In both **(a)** and **(b)**, inhibition of a given protein results in a decrease in detected ABP signal in the presence of inhibitor versus vehicle control

inhibitor—via protein interactions or modification chemistry—should overlap with that of the probe to ensure accurate and complete target identification. Examples in mycobacteria include competitive ABPP with an ATP-based activity probe to identify the targets of the noncovalent inhibitor staurosporine in *Mtb* [14]. In work from Canaan and Cavalier that preceded the development of alkyne derivatives of cyclophostin/cyclopostin compounds noted above, Nguyen et al. and Madani et al. used desthiobiotin-fluorophosphonate (FP) in competition with the inhibitor CyC₁₇ to identify serine hydrolase targets in *Mtb* and *M. abscessus* [15, 16]. These researchers also applied the same format to the oxadiazolone compound iBpPPOX in *M. abscessus* [17]. Most recently, our lab and the Boggy lab screened serine hydrolase inhibitor libraries against *Mtb* and then used competitive ABPP with biotin-fluorophosphonate to identify the targets of compounds that restricted *Mtb* growth and survival [18, 19]. Our study included comparing the targets of structurally related active vs. inactive compounds to distinguish enzymes potentially more relevant to the inhibitor mode of action [19].

This protocol details ABPP and competitive ABPP in *Mtb* using serine hydrolase-directed inhibitors and fluorophosphonate probes. In principle, the competitive ABPP protocol is generalizable to any matched (i.e., mechanistically similar, as described above) covalent inhibitor and activity-based probe pair, as has been demonstrated in the studies cited above. Since the subcellular localization of enzyme activity is experimentally underexplored in ABPP studies, we also include methods for differential sedimentation and provide exemplary data comparing fractions from several mycobacterial species (Fig. 2). Profiling enzymes in the pellet fraction, which is enriched in membrane proteins, and the supernatant, which is enriched in soluble cytosolic and periplasmic proteins, can provide additional guidance for exploring individual enzyme functions, especially in mycobacteria, which encode many hypothetical conserved proteins of unknown function.

Finally, another distinct aspect detailed here is the adaptation of stable isotope labeling by amino acids in cell culture (SILAC) for use in mycobacteria. In mammalian cells, SILAC entails culturing cells in medium supplemented with specific amino acids (e.g., lysine and arginine) containing either heavy or light stable isotopes under different sample conditions (e.g., with or without inhibitor). The resulting mass signatures associated with each condition enable samples to be combined and changes in protein abundance to be calculated within a single mass spectrometry run [20, 21]. SILAC presents a challenge in mycobacteria and other bacteria because they synthesize amino acids *de novo*, rather than obtain them only via salvage. We have reported a method that uses ammonium chloride as a sole nitrogen source and thus yields isotopic labeling

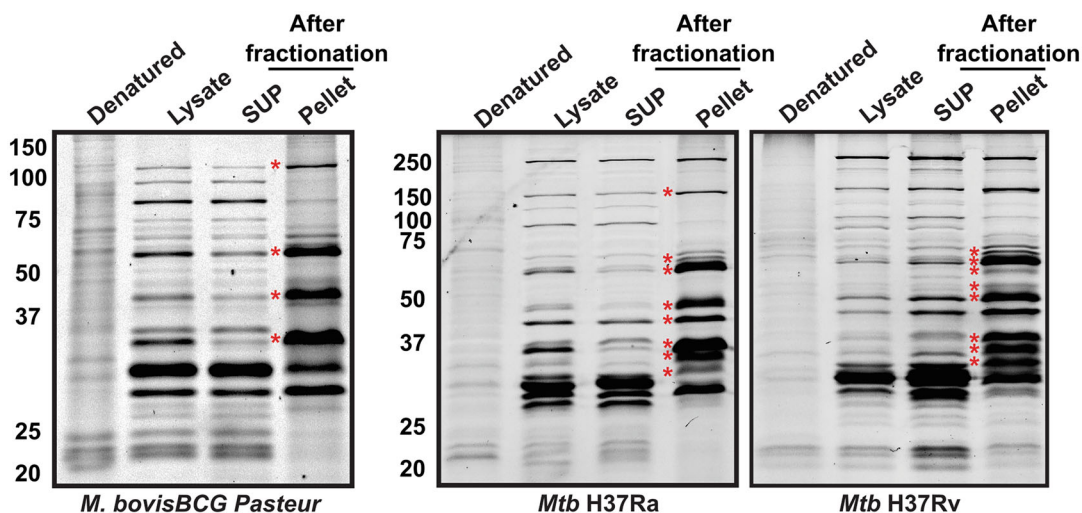


Fig. 2 Detection of differential serine hydrolase activity in different subcellular compartments of mycobacteria by ABPP. Total lysates from the indicated mycobacterial species were treated with ActivX TAMRA-FP and then fractionated by ultracentrifugation to yield the supernatant (SUP) comprising soluble proteins and the pellet fraction comprising proteins strongly associated with cell membrane and wall fragments. Red asterisks indicate serine hydrolases that were enriched in the pellet fraction. Heat-treated lysates (“Denatured”) serve as a negative control for activity-based probe labeling

of all protein nitrogens, not just specific amino acid side chains [19, 22]. Because this requires atypical data analysis, we provide the corresponding parameter file for use with the CIMAGE proteomics pipeline and provide comments about alternative approaches.

2 Materials

Prepare all solutions using ultrapure water (prepared by purifying deionized water, to attain a sensitivity of 18 MΩ-cm at 25 °C) and analytical grade reagents. All solutions are in ultrapure water unless otherwise specified. Prepare and store all reagents at room temperature unless indicated otherwise. Where filter sterilization is specified, this is achieved with a 0.2-μm polyethersulfone (PES) filter. Sterile technique is maintained until the cells are lysed. Follow all local regulations when disposing of waste materials, especially materials that come into contact with biohazardous organisms (*see Note 1*).

2.1 Culturing

1. *Mycobacterium tuberculosis* cell stock (stored at −80 °C) (*see Note 2*).
2. Culture flasks with screw caps appropriate for 100 mL of culture. For example, two 250 mL Erlenmeyer flasks with 50 mL culture each.

3. Shaking incubator.
4. Spectrophotometer appropriate for measuring optical density at 600 (OD₆₀₀).
5. Spectrophotometer cuvettes.
6. 50 mL and 15 mL conical tubes.
7. Centrifuge appropriate for 50 mL and 15 mL conical tubes up to at least 4000 $\times g$.
8. 10% (w/v) Tyloxapol (*see Note 3*).
9. Middlebrook 7H9-OADC-Tyloxapol medium (7H9 medium): Middlebrook 7H9 broth base supplemented with 10% (v/v) oleic acid-albumin-dextrose-catalase (OADC) supplement, 0.5% glycerol, and 0.05% Tyloxapol (*see Note 4*).
10. Phosphate-buffered saline (PBS): 1 mM monopotassium phosphate (KH₂PO₄), 154 mM sodium chloride (NaCl), and 5.6 mM disodium phosphate (Na₂HPO₄) in ultrapure water, adjusted to pH 7.4 and filter sterilized (*see Note 5*).
11. Phosphate-buffered saline with Tween (PBST): add 250 μ L of 20% v/v Tween 80 (0.05% v/v final concentration) to 100 mL PBS.

2.2 Lysis

1. Ice in insulated bucket.
2. Microcentrifuge suitable for at least 10,000 $\times g$.
3. BeadBug homogenizer (Benchmark Scientific) or similar bead beating homogenizer.
4. 2-mL screw-cap microfuge tubes (*see Note 6*).
5. 0.1 mm zirconia beads.

2.3 Labeling with Activity-Based Probe

1. 1.5 mL screw-cap microfuge tubes.
2. Fluorophosphonate conjugate depending on downstream application (e.g., detection, enrichment, identification) (*see Note 7*).
3. Dimethyl sulfoxide (DMSO), or other water-miscible organic solvent for solubilizing probe.
4. Bicinchoninic acid (BCA) protein quantitation kit (*see Note 8*).

2.4 Fractionation

1. Ultracentrifuge capable of at least 100,000 $\times g$.
2. Beckman T1–60 ultracentrifuge rotor or equivalent.
3. Ultracentrifuge tubes.

2.5 Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) and Detection

1. 5 $\times$ Laemmli sample buffer: 250 mM Tris-HCl, 10% SDS, 50% (v/v) glycerol, 5% (v/v) 2-mercaptoethanol, 0.04% (w/v) bromophenol blue (*see Note 9*).
2. Heat block or other heat source to heat samples up to 95 $^{\circ}$ C.
3. Protein ladder markers.

4. 10% SDS-PAGE gel.
5. SDS-PAGE running buffer: 0.025 M Tris-HCl, pH 8.3, 0.192 M glycine, 0.1% SDS.
6. Mini Protean 3 or similar gel electrophoresis and transfer system (*see Note 10*).
7. Western blot transfer buffer: 0.025 M Tris-HCl, 0.192 M glycine, 20% methanol.
8. Nitrocellulose membranes.
9. Blotting papers.
10. LI-COR Protein-Free Blocking Buffer (PBS) or equivalent blocking solution (*see Note 11*).
11. Fluorescence imager appropriate for Active-X-TAMRA FP (approximate excitation and emission maxima: 552 nm/575 nm) (*see Note 11*).
12. LI-COR Streptavidin IR 680 LT (*see Note 11*).
13. LI-COR Odyssey imager or similar image capable of near-infrared fluorescence detection (*see Note 11*).

2.6 Isotopic Labeling

1. ^{15}N - and ^{14}N -ammonium chloride.
2. Modified Roisin's medium: 1 g/L monobasic potassium phosphate (KH_2PO_4), 2.5 g/L sodium phosphate dibasic (Na_2HPO_4), 5.9 g/L ammonium chloride (NH_4Cl), 2.0 g/L potassium sulfate (K_2SO_4), 1.0 g/L citric acid, 0.08 mg/L zinc chloride (ZnCl_2), 0.4 mg/L iron chloride ($\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$), 0.02 mg/L copper sulfate (CuSO_4), 0.02 mg/L manganese (II) chloride ($\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$), 0.02 mg/L sodium borate ($\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$), 0.02 mg/L ammonium heptamolybdate ($(\text{NH}_4)_6\text{Mo}_7\text{O}_{24} \cdot 4\text{H}_2\text{O}$), 0.5 mM calcium chloride (CaCl_2), 0.5 mM magnesium chloride (MgCl_2), 0.5% (v/v) glycerol, 0.5 mg/L biotin, 0.05% (v/v) Tyloxapol, pH 6.6 (*see Note 12*).
3. Inhibitor compound.
4. DMSO or other water-miscible organic solvent for solubilizing inhibitor.
5. Refrigerated microcentrifuge.
6. PD Mditrap G-25 columns or other desalting size exclusion columns prepared and used according to the manufacturer's protocol.

2.7 Enrichment of Biotinylated Proteins

1. Streptavidin agarose beads.
2. 10% SDS (w/v): 10 g SDS in 100 mL of water.
3. 10 M urea solution in water (*see Note 13*).
4. 0.2- μm PES syringe filters.

2.8 Sample Preparation for Mass Spectrometry

1. 200 mM dithiothreitol (DTT) solution in water.
2. 400 mM iodoacetamide solution in water.
3. Sequence grade trypsin: 2 mg in 2 M urea and 2 mM CaCl₂ dissolved in PBS.
4. 5% formic acid (v/v) in water.
5. −20 °C freezer.

2.9 Mass Spectrometry Analysis

1. CIMAGE (<https://github.com/wangchulab/CIMAGE>).
2. FASTA file for Mtb H37Rv proteome—Supplementary File 1 (UniProt_rv_Mycobacterium_cravattlab_v1_05-10-2013) (*see Note 14*).
3. Custom parameter file for non-specific SILAC labeling—Supplementary File 2 (cimageN15.params).

3 Methods

3.1 Activity-Based Protein Profiling of Subcellular Fractions by SDS-PAGE

3.1.1 Culturing (See Note 1)

1. Thaw a frozen stock of *M. tuberculosis* and inoculate in 5 mL 7H9 medium containing 10% OADC, 0.5% glycerol and 0.05% Tyloxapol (*see Note 2*).
2. Incubate culture at 37 °C with orbital shaking at 110 rpm for 2–3 days until it reaches an OD₆₀₀ ~ 1.
3. Subculture to OD₆₀₀ 0.1 in 100 mL 7H9 medium (i.e., if the starter culture has OD₆₀₀ of 1, use 10 mL starter to inoculate 90 mL 7H9 medium).
4. Incubate culture for an additional 3–4 days until culture reaches an OD₆₀₀ ~ 1 (*see Note 15*).

3.1.2 Lysis

1. Centrifuge the culture at 4000 × g for 10 min at 22 °C.
2. Discard the supernatant and resuspend the pellet in 30 mL of PBS.
3. Centrifuge the suspension at 4000 × g for 10 minutes at 22 °C.
4. Repeat two additional washes as in b and c above.
5. Fill 2 × 2-mL tubes with ~0.5 mL 0.1 mm diameter zirconia beads.
6. Resuspend the cell pellet in 2 mL PBS. Transfer 1 mL of suspension to each of 2 mL tube.
7. Lyse the cells with a bead beater in four cycles of 45 s each at 4000 rpm with 5-min incubations on ice between each cycle (*see Note 16*).
8. Pellet cell debris by centrifugation at 10000 × g for 10 min at 4 °C.

9. Carefully transfer supernatant (without disturbing the cell pellet) to a fresh tube.
10. Filter sterilize twice using 0.2 μm polyethersulfone (PES) filters, each time into fresh tubes. Perform viability testing and remove lysates from the BSL3 per approved standard operating procedures.

3.1.3 Labeling with Activity-Based Probe

1. Determine protein concentration of the total lysates using a bicinchoninic acid (BCA) assay kit according to the manufacturer's instructions. Normalize lysates to 1 mg/mL with PBS in 2 mL total volume.
2. As a negative control, remove 1 mL from each sample into a fresh tube and heat denature by incubation at 95 °C for 10 min.
3. Add fluorophosphonate probe to each sample to a final concentration of 10 μM .
4. Incubate the suspension for 1 h at 22 °C.
5. Remove excess fluorophosphonate using an appropriate size-exclusion desalting column (such as PD Minitrapp G-25) according to the manufacturer's instructions.

3.1.4 Fractionation

1. For each sample, remove 50 μL for later SDS-PAGE analysis. Mix with 12.5 μL 5 $\times$ LSB and heat at 95 °C for 10 min.
2. Centrifuge the remaining lysate at 100,000 $\times$ g for 1 h to pellet membrane fragments.
3. Transfer supernatant (soluble fraction) to a fresh tube.
4. Repeat **steps 2 and 3**.
5. Resuspend pellets from both spins in 100 μL final volume of PBS with 0.2% SDS and incubate at 37 °C for 1 h with occasional gentle mixing to solubilize the membrane fragments (*see Note 17*).
6. Prepare SDS-PAGE samples from the soluble and pellet fractions as in **step 1**.

3.1.5 SDS-PAGE and Detection

1. Resolve 10 μg protein per sample on a 10% SDS-PAGE gel. Proceed to detection using one of the following methods depending on the probe used:
 - (a) *Fluorescent dye-FP*: To visualize labeled proteins, image SDS-PAGE gel using a fluorescence scanner with excitation and emissions settings appropriate for the fluorescent dye conjugate. In the case of ActivX TAMRA-FP, the approximate excitation and emission maxima are at 552 nm and 575 nm according to the manufacturer.
 - (b) *Biotin-FP*: To detect biotinylated proteins, transfer SDS-PAGE gel to a nitrocellulose membrane using a semi-dry transfer apparatus at 25 V for 60 min (*see Note 18*).

Incubate membrane with 10 mL LI-COR Blocking Buffer for 1 h. Incubate with 10 mL 1:10,000 Streptavidin IR-Dye 680 LT in LI-COR Odyssey Blocking Buffer containing 0.1% SDS and 0.1% Tween-20. Wash the membrane three times with 10 mL each PBST and once with PBS for 5 min per wash. Image using instrument settings appropriate for the excitation and emission wavelengths of the streptavidin conjugate.

3.2 Identification of Inhibitor Targets by Competitive Activity-Based Protein Profiling

3.2.1 Culturing

1. Thaw a frozen stock of *Mtb* and inoculate 5 mL 7H9 medium containing 10% OADC, 0.5% glycerol and 0.05% Tyloxapol (*see Note 19*).
2. Incubate at 37 °C with shaking at 110 rpm for 2–3 days until it reaches an OD₆₀₀ of ~1.

3.2.2 Isotopic Labeling and Inhibitor Treatment (See **Note 20)**

1. Centrifuge the culture at 4000 × g for 10 min at 22 °C.
2. Remove and discard the supernatant. Resuspend the pellet in 5 mL modified Roisin's medium (*see Note 21*).
3. Divide the cell suspension into two 2.5 mL aliquots. Centrifuge at 4000 × g for 10 min.
4. Remove and discard the supernatant. Resuspend the cell pellet in 5 mL modified Roisin's medium and centrifuge at 4000 × g at 22 °C for 10 min.
5. Remove and discard the supernatant. Resuspend each pellet in 5 mL modified Roisin's medium containing either heavy (¹⁵N) or light (¹⁴N) ammonium chloride as the sole nitrogen source. The final OD₆₀₀ will be 0.5.
6. Incubate cultures until OD₆₀₀ ~ 1 (*see Note 22*).
7. Add inhibitor or vehicle control (e.g., DMSO) to the desired final concentration to the ¹⁴N- and ¹⁵N-containing cultures, respectively (*see Note 23*). Incubate cultures for an additional 2 h at 37 °C with shaking at 110 rpm.

3.2.3 Lysis

1. Centrifuge the culture at 4000 × g for 10 min at 22 °C.
2. Discard supernatant and wash the pellet with 1 mL PBS containing 0.05% Tween 20.
3. Centrifuge the suspension at 4000 × g for 10 min at 22 °C.
4. Discard the supernatant and wash the pellet with 2 mL PBS.
5. Fill two 2 mL tubes with ~0.5 mL 0.1 mm diameter zirconia beads. Prewet beads with 0.5 mL PBS.
6. Resuspend the cell pellet in 2 mL PBS. Transfer 1 mL of suspension to each 2 mL tube containing beads.

7. Lyse the cells with a bead beater in 4 cycles of 45 s each at 4000 rpm with 5 min incubations on ice between each cycle.
8. Pellet cell debris by centrifugation at $10,000 \times g$ for 10 min at 4 °C.
9. Carefully transfer each supernatant (without disturbing the cell pellet) to a fresh tube.

3.2.4 Labeling with Activity-Based Probe

1. Determine protein concentration of the total lysates using a BCA assay kit according to the manufacturer's instructions. Normalize lysates to 1 mg/mL with PBS in a final volume of 2 mL.
2. Add biotin-FP to each sample to a final concentration of 10 μ M.
3. Incubate the suspension for 1 h at 22 °C.
 - (a) Remove excess biotin-FP using an appropriate size exclusion desalting column (such as PD Minitrapp G-25) according to the manufacturer's instructions. Combine the ^{14} N- and ^{15} N-labeled lysates in a 1:1 (v/v) ratio (500 μ L each) to a total volume of 1 mL.
 - (b) Add 200 μ L PBS, 400 μ L of 10% SDS solution (final concentration 2%) and 400 μ L of 10 M urea solution (final concentration 2 M) to each sample for a final volume of 2 mL.
 - (c) Incubate samples for 40 min at 22 °C with shaking at 110 rpm.
 - (d) Centrifuge samples at $4000 \times g$ for 20 min at 22 °C.
 - (e) Carefully transfer each supernatant (without disturbing the cell pellet) to a fresh 15 mL conical tube.
 - (f) Dilute each sample with 8 mL PBS and filter twice using 0.2 μ m PES filters, each time into fresh tubes. Perform viability testing and remove lysates from the BSL3 per approved standard operating procedures (*see* **Note 24**).

3.2.5 Enrichment of Biotinylated Proteins

1. For each sample, suspend 100 μ L streptavidin agarose beads in a 1.5 mL centrifuge tube in 1 mL PBS.
2. Centrifuge at $500 \times g$ for 5 min at 22 °C.
3. Repeat **steps 6** and **7** two more times for a total of three washes.
4. Add each lysate sample from Subheading 3.2.4 to an aliquot of washed beads. Incubate with gentle shaking at 22 °C for 2 h.
5. Centrifuge samples at $500 \times g$ for 5 min.
6. Discard the supernatant and resuspend beads in 1 mL PBS containing 0.25% SDS.

7. Repeat **steps 8** and **9** two additional times for a total of three washes.
8. Repeat **steps 8** and **9** except resuspend once in 1 mL PBS and once in 1 mL ultrapure water.
9. Centrifuge samples at $500 \times g$ for 5 min. Discard supernatant.

3.2.6 Sample Preparation for Mass Spectrometry

1. Resuspend each aliquot of beads in 500 μ L 10 M urea and 25 μ L 200 mM DTT. Incubate the mixture at 65 °C for 15 min.
2. Add 20 μ L 400 mM iodoacetamide to each bead suspension and incubate at 37 °C for 30 minutes.
3. Stop the reaction by adding 950 μ L PBS to each sample.
4. Centrifuge samples at $1400 \times g$ for 3 min. Discard the supernatant.
5. On-bead digestion: add 2 mg sequence-grade trypsin in PBS containing 2 M urea and 2 mM CaCl_2 for 12–14 h at 37 °C.
6. Centrifuge the bead suspension at $1400 \times g$ for 3 min.
7. Transfer each supernatant containing digested peptides to a fresh 1.5 mL centrifuge tube.
8. Add formic acid to the digested peptides to 5% v/v final concentration (*see* **Note 23**).

3.2.7 Mass Spectrometry Analysis

1. Submit samples for mass spectrometry analysis (*see* **Note 25**).
2. Extract the MS² spectra data from the RAW file using RAW-Xtract 1.9.9.2.
3. Use the ProLuCID algorithm to search spectra against a *Mtb* H37Rv reverse-concatenated nonredundant (gene-centric) FASTA database assembled from the UniProt database (*see* **Note 26**).
4. Quantify SILAC ratios with CIMAGE software using an appropriate parameter file to identify isotopically labeled peptides (*see* **Notes 27** and **28**).

4 Notes

1. *Mtb* is a Risk Level 3 organism that requires biosafety level 3 (BSL3) containment. Ensure that your use of biosafety containment and disposal of biohazardous waste conforms to local regulations and that all procedures are approved by the appropriate regulatory bodies at your institution. In particular, before following this protocol, ensure that approved procedures are in place for decontamination and removal of samples from the BSL3 laboratory for downstream analysis (e.g., SDS-PAGE, proteomics). For information regarding laboratory maintenance of *Mtb*, see Larsen, Biermann and Jacobs [23].

2. This protocol is designed for *Mtb*, but can be adapted for any (myco)bacterium by following the appropriate biosafety handling and culture conditions.
3. This protocol details gel-based ABPP. To perform gel-based *competitive* ABPP, culture 60 mL total, split into 2 30-mL aliquots and perform **step 7** of Subheading 3.2.2 and then return to Subheading 3.1.3.
4. Preparation of 10% (w/v) Tyloxapol: Tyloxapol is a highly viscous liquid. To prepare accurate solutions, do not aspirate, but dissolve aliquot of known mass in water and further dilute as needed to 10% (w/v). Filter sterilize.
5. Preparation of 7H9-OADC-Tyloxapol (7H9 medium): Dissolve 4.7 g Difco Middlebrook 7H9 Broth in 889.5 mL water and 10 mL 50% (v/v) glycerol. Autoclave. Cool to 50 °C and aseptically add 0.5 mL 10% (v/v) Tyloxapol and 100 mL BBL Middlebrook OADC Supplement. Store at 4 °C.
6. Preparation of PBS: Dissolve 0.14 g potassium phosphate (KH_2PO_4), 8.99 g sodium chloride (NaCl), 0.79 g disodium phosphate (Na_2HPO_4) in water to a final volume of 1 L. Adjust to pH to 7.4 with HCl or NaOH and filter sterilize. Avoid adding protease inhibitors containing PMSF or AEBSF or SDS and other denaturing detergents from lysis buffer as it will affect use of activity-based serine hydrolase probes.
7. Due to the mechanical stress associated with bead beating, use tubes recommended by the bead beater manufacturer. Any tube used for handling live *Mtb* should have a screw cap with a silicone O-ring to minimize possible contamination via leakage during centrifugation and other handling.
8. Several fluorophosphonate probes are commercially available, including conjugates to desthiobiotin (ActivX Desthiobiotin-FP), 5-carboxytetramethylrhodamine (ActivX TAMRA-FP), and azide (ActivX Azido-FP).
9. Any protein quantitation method may be used as long as it is compatible with the buffers used. Be sure to use protein standards made in the same buffer as your sample.
10. 5X Laemmli sample buffer: This solution is stable for only 1–2 weeks at room temperature and 1–2 months at –20 °C. The SDS will precipitate after freezing, so before use warm gently until SDS has fully dissolved.
11. Preparation of SDS-PAGE running buffer: Make a 10X stock solution. Dissolve 30 g Tris HCl, 144 g glycine, 10 g of SDS in water to final volume of 1 L. Store at room temperature. For 4 L of running buffer add 400 mL of 10× stock to 3600 mL of water. Mix with slow stirring to avoid excessive bubbles.

12. The protocol is reported for detection using Streptavidin IR 680 LT detection, but other streptavidin conjugates can be used with corresponding detection methods (e.g., visible fluorescence, luminescence). The blocking solution recommended for a given streptavidin conjugate should be used. We have found this blocking buffer to work best for LI-COR Odyssey detection of Streptavidin IR 680 LT.
13. Preparation of modified Roisin's medium: Dissolve 1 g KH_2PO_4 , 2.5 g Na_2HPO_4 , 5.9 g NH_4Cl *, 2.0 g K_2SO_4 , 0.08 mg ZnCl_2 , 0.4 mg $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$, 0.02 mg CuSO_4 , 0.02 mg $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$, 0.02 mg $\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$, 0.02 mg $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24} \cdot 4\text{H}_2\text{O}$ in 600 mL of water. Add 1 mL of 500 mM CaCl_2 , 1.0 mL of 500 mM MgCl_2 , 10 mL of 50% (v/v) glycerol, 0.5 mg biotin, 0.5 mL 10% (v/v) Tyloxapol to 600 mL water. Adjust the volume to 1 L with ultrapure water. Filter sterilize. *Substitute isotopically pure ^{15}N - and ^{14}N -ammonium chloride (NH_4Cl) as appropriate for the SILAC protocol.
14. Preparation of 10 M urea: Add 600.6 grams urea to stirring water (<500 mL warmed to >30 °C). While stirring adjust the volume to near 900 mL with gentle heat. Once urea dissolves, turn off heat and adjust to 1 L. Due to degradation and precipitation over time, solution should be made fresh each time.
15. This proteome is based on the genome of the *Mtb* H37Rv strain. A corresponding FASTA file should be generated for any other *Mtb* strain or bacterial species used with this protocol.
16. Incubations on ice reduce sample heating due to bead beating. Some bead beaters have an optional cooling jacket that obviates the need for this step.
17. Mixing can be performed by inversion every 15–20 min or by slow continuous rotation. Shaking that generates bubbles should be avoided to minimize protein aggregation and precipitation.
18. These parameters are appropriate for semi-dry transfer. Transfer can also be performed using wet transfer or other transfer methods such as the Bio-Rad Trans-Blot Turbo according to the manufacturer's recommendations.
19. In later steps, *Mtb* is cultured in modified Roisin's medium, but initial inoculation in rich 7H9 mediums helps recovery from the frozen stock and also shortens this early growth step. The doubling time of *Mtb* in 7H9 medium is ~24 h (see also **Note 22** regarding the doubling time of *Mtb* in modified Roisin's medium).

20. Other methods provide analogous opportunities for protein quantitation without the need for culturing in specialized medium, including label-free and peptide modification approaches using dimethylation reagents, iTRAQ, and tandem mass tags. For example, we have reported label-free affinity mass spectrometry in *Mtb* and others have reported multiplexed analysis using iTRAQ for *M. marinum*.
21. For wash steps, modified Roisin's medium with unlabeled ammonium chloride can be used. In general, this protocol uses modified Roisin's medium because ammonium chloride is the sole nitrogen source and is available in ^{14}N and ^{15}N isotopically pure forms. This enables isotopic labeling of all amino acid nitrogens. In principle, any medium can be used in which an isotopically pure sole nitrogen source can be provided.
22. The doubling time of *Mtb* in modified Roisin's medium is ~4 days. The final OD₆₀₀ can be modified, for example, to compare inhibitor targets across different phases of growth. Note *Mtb* reaches stationary phase at OD₆₀₀ ~ 1 in modified Roisin's medium. Thus, mid- to late exponential phase is OD₆₀₀ 0.5–0.8. If a different medium is used, growth kinetics should be determined before pursuing growth phase-dependent experiments. Here, the choice of culture volume and final cell density was informed by the need to balance the expense of isotopically pure reagents and availability of inhibitor compounds against the need for sufficient cell mass for downstream proteomics.
23. We strongly recommend performing at least one additional biological replicate in which the isotopic labeling is reversed (inhibitor added to ^{15}N -labeled and vehicle control added to ^{14}N -labelled cultures) to control for any isotope-related differences in labeling efficiency.
24. Samples can be stored at $-80\text{ }^{\circ}\text{C}$ until further processing.
25. Samples can be stored at $-20\text{ }^{\circ}\text{C}$ until further analysis.
26. Due to the broad availability of proteomics services, we do not provide an explicit protocol here. In general, any general MS/MS method for protein peptide identification can be used. Our reported ABPP results were obtained using Multi-Dimensional Protein Identification Technology (MuDPIT), which combines off-line and in-line liquid chromatography with mass spectrometric identification to increase the number of peptides that are identified.
27. A FASTA file with the *Mtb* H37Rv proteome is available as supplementary material.

28. As discussed in the Introduction, use of SILAC in mycobacteria results in nonspecific isotopic labeling of all amino acid nitrogens. As a result, analysis requires a custom CIAMGE parameter file to account for all possible substitutions. This file is provided as supplementary information. We note that other proteomics pipelines similarly allow for custom searches based on user-specific amino acid modifications and are alternatives to CIMAGE-based analysis. For example, the free tool FragPipe provides a user interface for SILAC, including an option to specify custom labels. A SILAC tutorial is available on the FragPipe website (https://fragpipe.nesvilab.org/docs/tutorial_silac.html). In the case of nonspecific isotope labeling, all modifications must be entered manually.

Acknowledgments

We thank Keriann Backus for providing the FASTA and parameter files for CIMAGE analysis. This work was supported by R35 GM128552 to J.C.S and I.S. was supported by T32 AI007539 and T32 GM008444.

References

1. Cravatt BF, Wright AT, Kozarich JW (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* 77(1):383–414
2. Fonović M, Bogoy M (2008) Activity-based probes as a tool for functional proteomic analysis of proteases. *Expert Rev Proteomics* 5(5):721–730
3. Patel HV, Li M, Seeliger JC (2019) Opportunities and challenges in activity-based protein profiling of mycobacteria. *Curr Top Microbiol Immunol* 420:49–72
4. Ortega C, Anderson LN, Frando A, Sadler NC, Brown RW, Smith RD et al (2016) Systematic survey of serine hydrolase activity in mycobacterium tuberculosis defines changes associated with persistence. *Cell Chem Biol* 23(2):290–298
5. Tallman KR, Levine SR, Beatty KE (2016) Small-molecule probes reveal Esterases with persistent activity in dormant and reactivating mycobacterium tuberculosis. *ACS Infect Dis* 2(12):936–944
6. Ansong C, Ortega C, Payne SH, Haft DH, Chauvigné-Hines LM, Lewis MP et al (2013) Identification of widespread adenosine nucleotide binding in Mycobacterium tuberculosis. *Chem Biol* 20(1):123–133
7. Wolfe LM, Veeraraghavan U, Idicula-Thomas S, Schurer S, Wennerberg K, Reynolds R et al (2013 Jun) A chemical proteomics approach to profiling the ATP-binding proteome of Mycobacterium tuberculosis. *Mol Cell Proteomics* 12(6):1644–1660
8. Levine SR, Beatty KE (2021) Investigating β -lactam drug targets in mycobacterium tuberculosis using chemical probes. *ACS Infect Dis* 7(2):461–470
9. Ravindran MS, Rao SP, Cheng X, Shukla A, Cazenave-Gassiot A, Yao SQ et al (2014 Feb) Targeting lipid Esterases in mycobacteria grown under different physiological conditions using activity-based profiling with Tetrahydrolipstatin (THL). *Mol Cell Proteom* 13(2):435–448
10. Lehmann J, Vomacka J, Esser K, Nodwell M, Kolbe K, Rämmer P et al (2016) Human lysosomal acid lipase inhibitor Ialstatat impairs mycobacterium tuberculosis growth by targeting bacterial hydrolases. *Med Chem Commun* 7(9):1797–1801
11. Lehmann J, Cheng TY, Aggarwal A, Park AS, Zeiler E, Raju RM et al (2018) An antibacterial beta-lactone kills mycobacterium tuberculosis by infiltrating mycolic acid biosynthesis. *Angew Chem Int Ed Engl* 7(1):348–353

12. Lopez Quezada L, Smith R, Lupoli TJ, Edoo Z, Li X, Gold B et al (2020) Activity-based protein profiling reveals that cephalosporins selectively active on non-replicating mycobacterium tuberculosis bind multiple protein families and spare peptidoglycan transpeptidases. *Front Microbiol* 11:1248
13. Barelrier S, Avellan R, Gnawali GR, Fourquet P, Roig-Zamboni V, Poncin I et al (2023) Direct capture, inhibition and crystal structure of HsaD (Rv3569c) from *M. tuberculosis*. *FEBS J* 290(6):1563–1582
14. Ortega C, Liao R, Anderson LN, Rustad T, Ollodart AR, Wright AT et al (2014) Mycobacterium tuberculosis Ser/Thr protein kinase B mediates an oxygen-dependent replication switch. *PLoS Biol* 12(1):e1001746
15. Madani A, Ridenour JN, Martin BP, Paudel RR, Abdul Basir A, Le Moigne V et al (2019) Cyclopostins and Cyclophostin analogues as multitarget inhibitors that impair growth of *Mycobacterium abscessus*. *ACS Infect Dis* 5(9): 1597–1608
16. Nguyen PC, Delorme V, Benarouche A, Martin BP, Paudel R, Gnawali GR et al (2017) Cyclopostins and Cyclophostin analogs as promising compounds in the fight against tuberculosis. *Sci Rep* 7(1):11751
17. Madani A, Mallick I, Guy A, Crauste C, Durand T, Fourquet P et al (2020) Dissecting the antibacterial activity of oxadiazolone-core derivatives against *Mycobacterium abscessus*. Neyrolles O, editor. *PLoS ONE* 15(9): e0238178
18. Babin BM, Keller LJ, Pinto Y, Li VL, Eneim A, Vance SE et al (2021) A screen of covalent inhibitors in mycobacterium tuberculosis identifies serine hydrolases involved in lipid metabolism as potential therapeutic targets. *bioRxiv*;2021.06.07.447460
19. Li M, Patel HV, Cognetta AB, Smith TC, Mallick I, Cavalier JF et al (2022) Identification of cell wall synthesis inhibitors active against mycobacterium tuberculosis by competitive activity-based protein profiling. *Cell Chem Biol* 29(5):883–896.e5
20. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A et al (2002 May) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 1(5):376–386
21. Ong SE, Mann M (2006) A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat Protoc* 1(6):2650–2660
22. Touchette MH, Van Vlack ER, Bai L, Kim J, Cognetta AB 3rd, Previti ML et al (2017) A screen for protein-protein interactions in live mycobacteria reveals a functional link between the virulence-associated lipid transporter LprG and the mycolyltransferase antigen 85A. *ACS Infect Dis* 3(5):336–348
23. Larsen MH, Biermann K, Jacobs WR Jr. (2007) Laboratory maintenance of *Mycobacterium tuberculosis*. In: *Current protocols in microbiology*. Chapter 10:Unit 10A 1



Application of Activity-Based Protein Profiling (ABPP) Strategy in the Identification of New Targets for Vascular Normalization and Binding Site Resolution

Jie Zhang and Yanchen Li

Abstract

Drug resistance stemming from vascular abnormalization poses a significant hurdle in the realm of tumor therapy. However, vascular normalization strategies offer promising avenues for addressing this challenge. One notable obstacle in this pursuit is the absence of well-defined targets for vascular normalization. The activity-based protein profiling strategy (ABPP) leverages highly reactive probes with remarkable capabilities for delving into uncharted proteins within living organisms, which allows us to explore transient protein–protein interactions, identify proteins within complex systems, and pinpoint the binding sites of small-molecule–protein complexes. In this chapter, we describe the construction of ABPP probes based on vascular normalization active molecules identified in previous studies, screening for activity, and their application to the identification and binding site resolution of the new target of vascular normalization, EphrinB2. This protocol serves as a successful case study, demonstrating its potential in validating new targets for vascular normalization. We anticipate its wide-ranging applicability not only for validating other novel vascular normalization targets but also for elucidating and resolving unidentified intracellular target proteins linked to active molecules.

Key words ABPP, Vascular normalization, Target identification, Binding site resolution, Subcellular localization, Target protein purification

1 Introduction

The processes of tumor development, metastasis, and immune escape are closely linked to abnormal angiogenesis, and structurally and functionally abnormal blood vessels lead to a near loss of drug and oxygen delivery at the tumor site, thereby profoundly influencing the formation of the tumor microenvironment and contributing to tumor therapy resistance [1]. Moreover, in anti-angiogenic treatment strategy, resistance to angiogenesis inhibitors triggers compensatory vascular proliferation, exacerbating the aberrations in tumor vasculature and perpetuating a vicious cycle [2]. There is

an urgent need for the development of strategies to normalize the structure and function of these blood vessels, enhancing tumor perfusion and thereby augmenting the effectiveness of chemotherapy and radiotherapy. Consequently, the concept of vascular normalization has been introduced and is gradually finding its way into clinical tumor therapy [3]. It is widely recognized that angiogenesis inhibitors combat angiogenesis by targeting pro-angiogenic factors, blocking angiogenesis-related pathways, and regulating pathological blood vessel formation. Interestingly, it has been observed that even low doses of angiogenesis inhibitors can induce a transient normalization of tumor vasculature [4]. However, it remains unclear whether the targets responsible for this effect are pro-angiogenic factors, which underscores the significant challenge of identifying clear therapeutic drug targets within the framework of vascular normalization theory.

The activity-based protein profiling (ABPP) strategy stands out as a powerful approach for target identification within the realm of chemical proteomics. This versatile approach aims to capture and label target proteins in complex biological samples using highly reactive specific probes [5]. Compared with other target identification and binding site analysis methods, such as genomics, structural biology, and bioinformatics, ABPP demonstrates distinct promise in exploring the detailed structural information of unknown proteins, transient weak protein–protein interactions (PPIs), and small-molecule–protein complex binding sites, identifying the functions of proteins in complex systems, and developing new “non-druggable” targets [6, 7].

The cornerstone of ABPP lies in probe design based on active molecules. Among the various techniques, multifunctional photoaffinity probes employing photoaffinity labeling technology currently represent the most efficacious method for pinpointing the direct targets of small bioactive molecules. An analysis of the structure–activity relationship (SAR) of the active molecule is conducted, facilitating the identification of the precise site for structure derivatization and ensuring the probe’s action target remains consistent. Based on the determination of SAR of the active molecule, a smaller diazirine photosensitizing group and an alkyne group as a bioorthogonal functional group were introduced to construct a minimized linker to minimize the effect on the structure and activity of the parent molecule. In addition, the bioorthogonal functional group, on the one hand, allows the introduction of the reporter group and the biotin tag to be accomplished in a stepwise manner, overcoming the effect of the conventional biotin label on the active molecule. On the other hand, the reporter motifs introduced by the bioorthogonal reaction can be imaged for subcellular-level localization of target proteins, which further ensures the accuracy of target protein identification (Fig. 1).

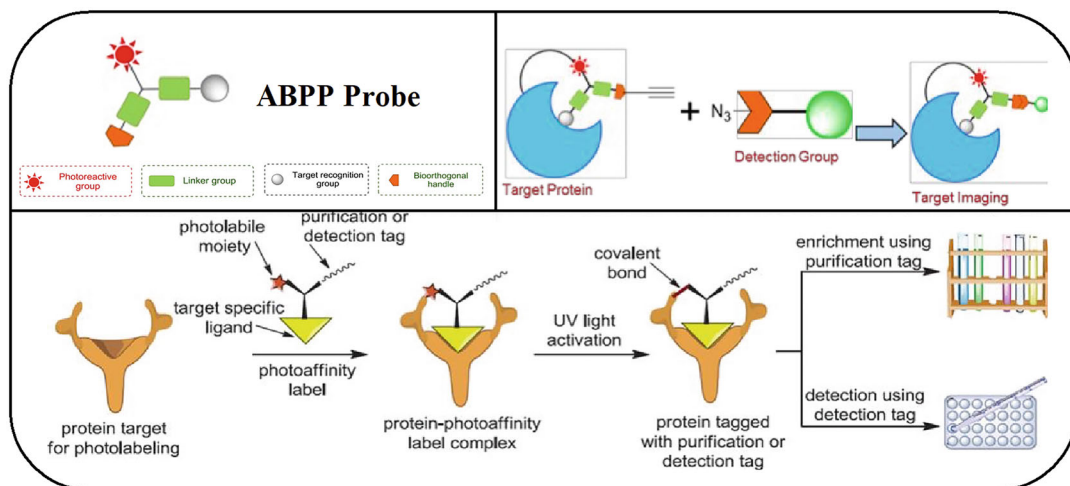


Fig. 1 Ideas for the application of the activity-based protein profiling (ABPP) strategy in the identification of new target for vascular normalization

Target proteins covalently cross-linked with biotinylated active molecules in complex protein systems are pulled down using streptavidin-modified magnetic beads. Magnetic bead-based pull-down offers easier handling, higher efficiency, and less protein loss than traditional affinity column chromatography separations. In addition, the target protein after pull-down can be separated from the magnetic beads by SDS-PAGE gel electrophoresis under denaturing conditions and can also be enriched by specific buffers. In contrast, the method of mixing reduced sample buffer with magnetic beads and incubating the beads at 100 °C for 5 min to detach the target proteins from the beads is more common, and the target proteins can be directly separated by SDS-PAGE and accurately localized by Coomassie brilliant blue staining. Subsequently, the target bands can be directly cleaved and in-gel digested with serial-grade trypsin, and the resulting peptide mixtures are analyzed by ultra-fast, high-resolution mass spectrometry for LC-MS, LC-MS/MS, and proteomic analysis, which allows for the identification of the target proteins after pull-down and the resolution of the binding sites of the active molecules.

Our previous study found that the quinazolinone-like active molecule QDAU5 has significant vascular normalizing activity with specific affinity for EphrinB2 and inhibits its phosphorylation [8]. In this chapter, we present a protocol for the identification and binding site resolution of a multifunctional photoaffinity probe-based ABPP for the new target of vascular normalization, EphrinB2. We describe here a vascular normalization target identification and binding site resolution procedure based on two types of multifunctional photoaffinity probes designed by QDAU5, including preparation and activity screening of multifunctional

photoaffinity probes, subcellular localization of target proteins, target protein confirmation, and binding site resolution. We anticipate that these guidelines can also be broadly applied to the validation of other novel targets for vascular normalization, as well as the validation and resolution of intracellular target proteins of active molecules of unknown targets.

2 Materials

2.1 Chemical Reagents for Molecular Preparation

1. 5-Amino-2-nitrobenzotrifluoride (Bidepharm, BD4412).
2. Triphosgene (BTC, Nanjing Xiezun Pharmaceutical Technology Co., 20220910).
3. 4-Aminophenylboronic acid pinacol ester (Bidepharm, BD14212).
4. Triethylamine (Tianjin Komeo Chemical Reagent Co.).
5. Dichloromethane (DCM, Tianjin Komeo Chemical Reagent Co.).
6. 2-Amino-4-bromobenzoic acid (Bidepharm, BD4525).
7. Methanamide (Tianjin Fuyu Fine Chemical Co.).
8. 1,1'-Bis(diphenylphosphino)ferrocene]dichloropalladium(II), PdCl₂(dppf) (damsa-beta, P2004838).
9. Cesium carbonate, Cs₂CO₃ (Macklin, B802456).
10. 1,4-Dioxane (Tianjin Fuyu Fine Chemical Co.).
11. Water.
12. Stannous chloride, SnCl₂ (damsa-beta, P1895197).
13. Concentrated hydrochloric acid.
14. Methanol (Tianjin Komeo Chemical Reagent Co.).
15. Malonic acid (damsa-beta, P1777749).
16. Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) (Pimio, BJS4266527).
17. N, N-Diisopropylethylamine (DIPEA) (damsa-beta, P2301015).
18. Tetrahydrofuran (THF) (Tianjin Komeo Chemical Reagent Co.).
19. N,N-Dimethylformamide (DMF, Tianjin Komeo Chemical Reagent Co.).
20. 2-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)ethan-1-amine (Nanjing Xiezun Pharmaceutical Technology Co., 20210425).
21. 2-((tert-butoxycarbonyl)amino)pent-4-ynoic acid (Acmechemical, B64890).

22. (4-(3-(trifluoromethyl)-3H-diazirin-3-yl)phenyl) methanamine (Nanjing Xiezun Pharmaceutical Technology Co., 20201210).
23. 1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDCI) (Macklin, N808856).
24. 1-Benzotriazolol (HOBt) (Leyan Reagent Platform, 1059815).
25. Trifluoroacetic acid (TFA) (Macklin, T818778).
26. Petroleum ether (Tianjin Komeo Chemical Reagent Co.).
27. Ethyl acetate (Tianjin Komeo Chemical Reagent Co.).

2.2 Activity Screening

1. U87 cell line.
2. HUVEC cell line.
3. Cell Count Kit-8 (Novonature, NK00801).
4. Cell Cycle and Apoptosis Detection Kit (DEEYEE, 1651104): contains propidium iodide stain (20 $\times$, DY80209), RNase A (50 $\times$, DY80209), staining buffer (25 mL, DY80209).
5. Annexin V-FITC/PI Apoptosis Detection Kit (DEEYEE, FA60426): contains Annexin V-FITC (DY20202-A), propidium iodide, PI (DY20202-B), 4 $\times$ binding buffer (DY20202-C).
6. Matrigel Matrix Basement Membrane (Corning, 356234).
7. 0.1% crystal violet staining solution (Solarbio, G1063).
8. Anti-EphrinB2 (phosphor Y316) antibody (Abcam, ab119323).

2.3 Subcellular Localization of Target Proteins

1. U87 cell line.
2. Copper sulfate pentahydrate ($\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, Guangdong Guanghua Sci-Tech Co., Ltd).
3. L-ascorbic acid sodium salt (Energy Chemical, E080152).
4. Sulfo Cy3-N₃ (Xi'an ruixi Biological Technology Co., Ltd., R-N3-001).
5. 4% Paraformaldehyde (4% PFA) (Coolaber, SL1830).
6. PBS (1 $\times$) (Corning, 21040077).
7. 4',6-diamidino-2-phenylindole (DAPI) (DEEYEE, 40930).

2.4 Target Protein Enrichment

1. U87 cell line.
2. HEK293 cell line.
3. 0.05% Trypsin, 0.53 mM EDTA, 1 $\times$ (Corning, 25-051-Cl).
4. RIPA strong lysis solution (Shaanxi Zhonghui Hecai Biomedical Technology Co., BA0615).

5. Phosphatase inhibitor(50X, Shaanxi Zhonghui Hecai Biomedical Technology Co., PL026).
6. PMSF (100×, Shaanxi Zhonghui Hecai Biomedical Technology Co., PL012-1).
7. Copper sulfate pentahydrate ($\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$).
8. L-ascorbic acid sodium salt.
9. N-(3-Azidopropyl)biotinamide.
10. Streptavidin magnetic beads (MCE, HY-K0208).
11. Reduced Loading buffer (5× denaturation, Shaanxi Zhonghui Hecai Biomedical Technology Co., PE012).
12. PBST (containing 0.05% Tween-20, MelunBio, 9005-64-5).
13. EphrinB2 Recombinant Protein (Abnova, H00001948-P01).
14. Acetone (Tianjin Komeo Chemical Reagent Co.).
15. Methanol (Tianjin Komeo Chemical Reagent Co.).

2.5 Gel Electrophoresis and Western Blot

1. Deionized water.
2. 30% Acr-Bis (29:1, Zhonghui Hecai, PE001C).
3. 1.5 M Tris-HCl, pH 8.8 (Zhonghui Hecai, PE001A).
4. 0.5 M Tris-HCl pH 6.8 (Zhonghui Hecai, PE001A).
5. 10% SDS (Zhonghui Hecai, PE001D).
6. 10% ammonium peroxydisulfate (Zhonghui Hecai, PE001E).
7. TEMED (Zhonghui Hecai, PE001F).
8. BCA protein quantification kit (DEEYEE, DY30205).
9. PVDF membrane (Millipore, 0000200915).
10. Electrophoresis buffer (10×): 0.25 mmol Tris, 1.92 mmol glycine, 34.68 mmol SDS, 1000 mL deionized water (Tris, MP, 819623; glycine, MP, 194825; SDS, MP, 190522).
11. Electrophoresis buffer (1×).
12. Transmembrane buffer (10×): 38.63 mmol glycine, 47.88 mmol Tris, 1.28 mmol SDS, 1000 mL deionized water.
13. Transmembrane buffer (1×).
14. TBST (10×): 20.06 mmol Tris-bace, 1.37 mmol NaCl, 1000 mL deionized water (NaCl, Tianjin Komeo Chemical Reagent Co.).
15. TBST (1×).
16. Rapid Sealing Solution (Beyotime Biotechnology, P0222).
17. Bovine serum albumin (BSA, Solarbio, A8020).
18. EphrinB2 antibody (Signalway Antibody, 49626-1).
19. GAPDH antibody (CST, 97166).

20. Goat anti-rabbit IgG (H + L) secondary antibody (DEEYEE, DY60202).
21. Goat anti-mouse IgG (H + L) secondary antibody (DEEYEE, DY60203).
22. ECL Chemiluminescent Substrate Kit (Extra Ultra Sensitive, Biosharp, BL520A).

2.6 Ultra-Fast, High-Resolution Mass Spectrometry

1. Acetonitrile (chromatographically pure, Merck).
2. Formic acid (LC-MS grade, Thermo Scientific, A117-50).
3. Ultrapure water (18.2 Ω , ultrapure water machine: mili-Q EQ. 7000).
4. Urea solution: 100 mM Tris-HCl + 8 M urea (2.4 mg urea + 500 μ L 1 M Tris-HCl + 4500 μ L ultrapure water).
5. 100 mM DTT: 5.642 mg DTT + 365 μ L ultrapure water. (DTT, MCE, HY-15917).
6. 500 mM IAA: 25.636 mg IAA + 277 μ L ultrapure water (2-iodoacetamide, MCE, HY-34477).

2.7 Main Consumables

1. Eppendorf tubes.
2. Magnetic force frame.
3. 6-well cell culture plate (Servicebio, CPP-6).
4. 6-well glass-bottom cell culture plates (Biosharp BS-CP-6B).
5. Transwell chambers with 24-well plates (PC membrane, 6.5 mm, pore size 8.0 μ m) (Corning, 3422).
6. Cell culture dish (100 $\times$ 20, Thermo Scientific, 150466).
7. Laser confocal dish (20 mm, NEST, 801001).
8. DMEM basal medium (Corning, 10013130).
9. Fetal bovine serum (Every Green, 11011-8611).
10. Penicillin-Streptomycin, sterile (100 $\times$, MCE, HY-K1006).
11. PBS buffer (Corning, 21-040-CVC).
12. 5% concentrate gel: 1.73 mL deionized water, 0.50 mL 30% Acr-Bis, 0.75 mL 0.5 M Tris-HCl, pH 6.8, 0.03 mL 10%SDS, 0.03 mL 10%AP, 0.003 mL TEMED.
13. 10% separation gel: 4.00 mL deionized water, 3.30 mL 30% Acr-Bis, 2.50 mL 1.5 M Tris-HCl, pH 8.8, 0.10 mL 10%SDS, 0.10 mL 10%AP, 0.008 mL TEMED.

2.8 Equipment

1. Oil bath (Greatwall, HWCL-1, 20030255).
2. Rotary evaporator (Greatwall, WB-2000, 16060127).
3. UV Crosslinker (SINSAGE, SGLinker III, SGL2022010).

4. Laboratory Microwave Synthesizer (XiangHu Technologies, XH-MC-1, HCM202109).
5. High-speed freezing centrifuge (Zonkia, HC-3018R, 05082100045).
6. Benchtop centrifuge (Cence, H1650, W16500021110474).
7. Cell Crusher (Ningbo Xinzhi Biotechnology Co., JY92-IIN, 10190637).
8. CNC Ultrasonic Cleaner (Kunshan Ultrasonic Instrument Co., KQ5200DE, 2019110080).
9. Low-temperature homogenizer (Hangzhou Mio Instrument Co., MTH-100, MU-C02-0340).
10. Electrophoresis (Beijing JUNYI Electrophoresis Co., Ltd., JY600C, Y07F22005018).
11. Orbital Shaker (Kylin-Bell Lab Instruments, TS-100, 2106399).
12. Thermostatic Incubation Shaker (Hangzhou Mio Instrument Co., ES-60, MU-C36-0143).
13. Rotary Homogenizer (Hangzhou Mio Instrument Co., YC-80, MU-C51-0337).
14. Mini-Chemiluminescent Imaging Analyzer (SINSAGE, Mini-Chemi 610, 306003).
15. Flow cytometer (Beckman Coulter Life Science, CytoFLEX, 21808271).
16. Fluorescent Inverted microscope (Nikon, Ts2-FL, 21903084).
17. Inverted biomicroscope (Nanjing Jiangnan Yongxin Optical Co., 1D-202, 10109306).
18. Laser confocal microscope (Nikon, Ti-E-C2, 21207771).
19. Fluorescent Enzyme Labeler (ALLSHENG, Feyond-F100, 244-14736-22120001).
20. Vacuum centrifugal concentrator (Beijing Jiam Technology Co., CV2000, YY/T0657-2008).
21. Ultra-Fast High-Resolution Mass Spectrometer (Thermo Scientific, Q Exactive Plus, 21910851).

2.9 Software

1. Image J (1.53, <https://imagej.nih.gov/ij/>).
2. GraphPad Prism 8 (v8.4.3).
3. NovoExpress (v1.5.0.2001).
4. Thermo Proteome Discoverer 2.3.

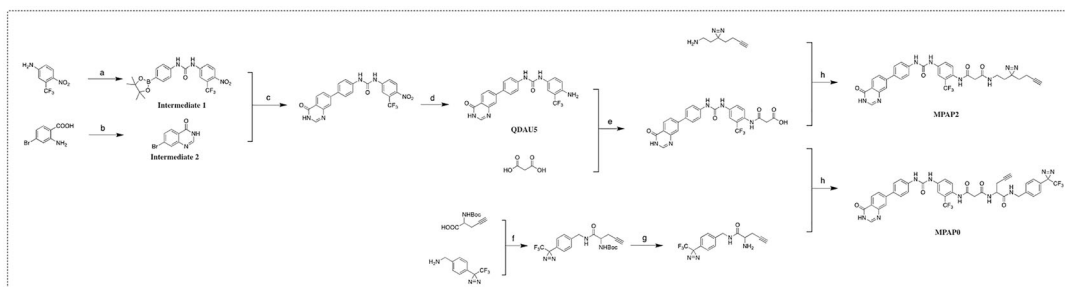
3 Methods

3.1 Preparation of the Active Molecule QDAU5 and a Multifunctional Photoaffinity Probe

To ensure that the probe is alignment with the target of the active molecule, SAR analysis identified the QDAU5-terminal aniline as the derivatization site. In terms of photosensitive groups, azides were deemed less photolabeling efficiency, benzophenone exhibits a prolonged photoactivation period, and trifluorotoluene diazene displayed significant site resistance. On the other hand, bioorthogonal handles, such as alkyne groups, offered the advantages of a more compact size and enhanced stability under UV light exposure. Consequently, a photosensitive moiety (diazirine) and a bioorthogonal handle (alkynyl group) were integrated to construct the minimal linker and minimize the effect on the structure and activity of the parent molecule. Leveraging above design strategy, two types of multifunctional photoaffinity probe molecules were designed (Fig. 2).

3.1.1 Preparation of QDAU5

1. Dissolve 0.8 g of BTC with 20 mL of anhydrous DCM in an ice bath and stir for 5 min.
2. Anhydrous DCM solution of 5-amino-2-nitrobenzotrifluoride was slowly added to the system and stirring was continued for 15 min.
3. Slowly add 10 mL of anhydrous DCM solution containing 1.1 mL of triethylamine and continue stirring for 15 min.
4. 10 mL of anhydrous DCM solution containing 1.5 g of 4-aminophenylboronic acid pinacol ester and 1.1 mL of triethylamine was added to the system and stirred for 40 min.
5. Separation by column chromatography (petroleum ether:ethyl acetate = 30:1) gave 1-(4-nitro-3-(trifluoromethyl)phenyl)-3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)urea (Intermediate 1).
6. 1.0 g 2-Amino-4-bromobenzoic acid was added to 16.3 mL formamide.



Reagents and reaction conditions: (a) 4-Aminophenyl boronicacid pinacolester, BTC, triethylamine, 0°C; (b) formamide, 150°C, 1h, N₂; (c) Pd(pddf)Cl₂, Cs₂CO₃, 100°C, 12h; (d) SnCl₄, HCl, MeOH, 80°C, reflux, 5h; (e) PyBOP, DIPEA, THF, r.t. 5h; (f) EDCI, HOBT, DIPEA, DCM, r.t., 5h; (g) TFA, DCM, r.t., 5h; (h) PyBOP, DIPEA, THF, r.t., 5h.

Fig. 2 Synthesis route of the active molecule QDAU5 and the multifunctional photoaffinity probes MPAP2/MPAP9

7. Protect the system with nitrogen.
8. Reacted at 150 °C, 300 W microwave for 1.45 h.
9. At the end of the reaction, the system was added to ice water and extracted three times with ethyl acetate.
10. The extracted organic phase was washed sequentially with NaHCO₃, saturated NaCl and dried with anhydrous Na₂SO₄.
11. The solvent was removed by vacuum evaporation to afford 7-bromoquinazolin-4(3H)-one (Intermediate 2).
12. 0.5 g of Intermediate 1, 0.25 g of Intermediate 2, 0.08 g of PdCl₂ (dppf) and 1.08 g of Cs₂CO₃ were dissolved in 30 mL of water and 10 mL of 1,4-dioxane
13. Reflux at 100 °C for 12 h.
14. Cooled to room temperature, extracted three times with ethyl acetate.
15. The extracted organic phase was washed sequentially with NaHCO₃, saturated NaCl and dried with anhydrous Na₂SO₄.
16. The solvent was removed in vacuum and the intermediate QDAU5-NO₂ was obtained by column chromatography (petroleum ether:ethyl acetate = 1:5).
17. 117.22 mg of SnCl₂ was dissolved in 10 mL of methanol and 3 mL of concentrated hydrochloric acid under ice bath conditions.
18. Reflux at 80 °C for 5 h.
19. Cool the system and adjust the pH of the system to 7–8 by adding saturated NaHCO₃ solution.
20. Extracted three times with ethyl acetate, the extracted organic phase was washed with NaHCO₃, saturated sodium chloride and dried with anhydrous Na₂SO₄.
21. The solvent was removed by vacuum evaporation to give the product QDAU5.

3.1.2 Preparation of the Multifunctional Photoaffinity Probe MPAP2

1. Dissolve 1 eq QDAU5 in anhydrous THF.
2. Add 3 eq malonic acid and 4 eq PyBOP.
3. Add 6 eq DIPEA under ice bath conditions.
4. Stir overnight at room temperature.
5. Column chromatographic separation (dichloromethane:methanol = 1:1) afforded the intermediate QDAU5-COOH.
6. 1.2 eq QDAU5-COOH was added to anhydrous THF containing 200 µL of DMF.
7. 2 eq PyBOP and 4 eq DIPEA were added under ice bath conditions.
8. Add 1 eq. 2-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)ethan-1-amine to the system, stir overnight.

9. Separation by column chromatography (petroleum ether: ethyl acetate = 1:5) afforded the product MPAP2.

3.1.3 Preparation of the Multifunctional Photoaffinity Probe MPAP0

1. Follow **steps 1–5** of Subheading **3.1.2**.
2. 1 eq. 2-((tert-butoxycarbonyl)amino)pent-4-ynoic acid, 1.8 eq. EDCI and 1.44 eq HOBt were dissolved in anhydrous DCM.
3. 10 eq DIPEA was added slowly in an ice bath and activated for 1 h.
4. To the system was added 1.2 eq. (4-(3-(trifluoromethyl)-3H-diazirin-3-yl)phenyl) methanamine.
5. Stir overnight at room temperature.
6. Column chromatographic separation (petroleum ether:ethyl acetate = 10:1) afforded 2-amino-N-(4-(3-(trifluoromethyl)-3H-diazirin-3-yl)benzyl)pent-4-ynamide (Intermediate 3).
7. Intermediate 3 was dissolved in 10 mL of anhydrous DCM.
8. Slowly add 2 mL of TFA in an ice bath and stir overnight at room temperature.
9. The pH was adjusted to 7 by adding saturated NaHCO_3 to the system.
10. The system was extracted three times with ethyl acetate, and the organic phase was washed with NaHCO_3 , saturated sodium chloride and dried with anhydrous Na_2SO_4 .
11. Separation by column chromatography (dichloromethane: methanol = 10:1) afforded 2-amino-N-(4-(3-(trifluoromethyl)-3H-diazirin-3-yl)benzyl)pent-4-ynamide (Intermediate 4).
12. Dissolve 1.5 eq QDAU5-COOH in 15 mL of anhydrous THF.
13. Add 4 eq PyBOP and 6 eq DIPEA to the system under ice bath conditions.
14. Add 1 eq of Intermediate 4 and stir overnight at room temperature. Column chromatographic separation (dichloromethane: methanol = 10:1) gave the product MPAP0.

3.2 Activity Screening of Multifunctional Photoaffinity Probes

3.2.1 In Vitro Antitumor Activity

Anti-Tumor Cell Proliferation

1. Cultured U87 cell line.
2. U87 cells were inoculated in 96-well plates (180 μL per well) at a density of 2000 cells/well, and six concentrations (100 μM , 10 μM , 1 μM , 0.1 μM , 0.01 μM , 0.001 μM) were set up for each compound, with six replicate wells set up for each concentration, as well as negative control wells and blank wells.
3. After inoculation, incubate at 37 $^{\circ}\text{C}$, 5% CO_2 for 24 h.
4. Prepare the compound solution according to the set concentration and administer 20 μL per well.

5. Continue incubation under the same conditions for 48 h.
6. Add 20 μL of CCK-8 to each well and continue incubation for 1–1.5 h (*see Note 1*).
7. Measure the absorbance at 450 nm on a multifunctional enzyme marker, calculate the cell survival rate in each well, and calculate the GI_{50} of the compounds on the cells with Prism 8.

Apoptosis

1. Cultivate U87 cell line.
2. Add 2 mL of medium to each well of a six-well plate and inoculate 1×10^6 cells.
3. After inoculation, the cells were incubated at 37 °C for 24 h under 5% CO_2 .
4. Administer each compound at four concentrations (Control, 1 μM , 5 μM , 10 μM) according to the GI_{50} value of the compound on U87 cells.
5. Continue incubation for 48 h.
6. Transfer the supernatant from each well to a new 4 mL Ep tube.
7. Add 200 μL of trypsin to each well for 1.5 min.
8. Terminate the digestion by adding 600 μL of medium.
9. Transfer the cells to the corresponding Ep tubes.
10. Centrifuge at 300 g for 5 min at 4 °C.
11. Discard the supernatant, add 800 μL of pre-cooled PBS, mix by blowing slowly, and transfer to a 1.5 mL Ep tube (*see Note 5*).
12. Centrifuge at 300 g for 5 min at 4 °C and discard the supernatant.
13. Mix according to the ratio of ultrapure water: binding buffer (4 $\times$) 3:1.
14. Add 100 μL of binding buffer (1 $\times$) to each sample (*see Note 2*).
15. Add 5 μL of Annexin V-FITC to each sample and incubate for 10 min at room temperature, protected from light.
16. After incubation, add 10 μL of PI to each sample and incubate for at least 5 min at room temperature (*see Note 6*).
17. Add 400 μL of pre-cooled PBS to each sample before running.
18. Flow cytometry test (*see Note 3*).

Cell cycle

1. Follow **steps 1–5** of subheading “**Apoptosis**”.
2. Discard the supernatant and wash the cells three times with pre-cooled PBS.
3. Add 200 μL of trypsin to each well for 1.5 min.

4. Terminate the digestion by adding 600 μ L of culture medium.
5. Transfer the cells to the corresponding Ep tubes.
6. Centrifuge at 800 g for 5 min at 4 °C and discard the supernatant.
7. For each sample, add 800 μ L of pre-cooled PBS to wash the cells once (*see Note 5*).
8. Centrifuge at 800 g for 5 min at 4 °C and discard the supernatant.
9. Add 600 μ L of pre-cooled 70% ethanol to each sample.
10. Fix the cells at 4 °C for 24 h.
11. Centrifuge at 800 g for 5 min at 4 °C and discard the supernatant.
12. For each sample, add 800 μ L of pre-cooled PBS to wash the cells once and discard the supernatant.
13. Prepare cell cycle staining buffer according to 500 μ L staining buffer +10 μ L Rnase A + 25 μ L PI per sample (*see Note 6*).
14. Add 500 μ L of cell cycle staining buffer to each sample.
15. Incubate at 37 °C for 30 min.
16. Flow cytometry test (*see Note 4*).

Anti-Tumor Cell Migration
(Scratch Assay)

1. Cultivate U87 cell line.
2. Add 2 mL of medium to each well of a six-well plate and inoculate 1×10^6 cells.
3. After inoculation, incubate at 37 °C, 5% CO₂ for 24 h.
4. Scratch with 200 μ L of sterile lance tip.
5. Discard the medium and wash three times with PBS.
6. Add 2 mL of medium to each well and administer each compound at six concentrations (Control, 0.1 μ M, 1 μ M, 5 μ M, 10 μ M, 20 μ M).
7. Scratch widths were taken at 0 h, 6 h, 12 h, 18 h, and 24 h with three different scratch fields per well.
8. Based on the scratch width, the ratio of different administered concentrations to the scratch width of control group was calculated.

Anti-Tumor Cell Migration
(Transwell Assay)

1. Cultivate U87 cell line.
2. Change the medium to starvation medium (containing 0.2% FBS) the day before digestion and continue to culture for 24 h.
3. Place 200 μ L of sterile lance tip, 96-well plate in advance in 4 °C overnight.
4. Digest the cells and count them.

5. Add 100–150 μL of cell suspension (cell density of 2×10^5 cells/mL) to the upper chamber, and add 800 μL of medium containing 20% FBS to the lower chamber and continue incubation for 24 h.
6. Discard the medium in the upper chamber and gently wipe off the cells on the upper surface with a cotton swab.
7. Soak the lower surface in 4% paraformaldehyde for 30–60 min for fixation.
8. Stain with 0.1% crystal violet staining solution for 20 min.
9. Place the chambers under an inverted microscope to take pictures, and take five fields of view for each chamber and count them.

Anti-Tumor Cell Invasion
(Transwell Assay)

1. Cultivate U87 cell line.
2. Change the medium to starvation medium (containing 0.2% FBS) the day before digestion and continue to culture for 24 h.
3. Place the matrix gel, 200 μL sterile lance tip, and 96-well plate into 4 °C overnight in advance (*see* **Note 7**).
4. Dilute the matrix gel to 2 mg/mL with pre-cooled serum-free medium.
5. Take 100 μL and spread it evenly in the upper chamber.
6. Incubate the small chamber at 37 °C for 3 h.
7. Remove and place in an ultra-clean bench for overnight drying.
8. Follow **steps 4–9** in subheading “[Anti-Tumor Cell Migration \(Transwell Assay\)](#)”.

3.2.2 Effect on the
Phosphorylation Level of
EphrinB2

1. Cultivate U87 cell line.
2. Add 2 mL of medium to each well of a six-well plate and inoculate 1×10^6 cells..
3. After inoculation, the cells were incubated at 37 °C with 5% CO_2 for 24 h.
4. Discard the medium and wash the cells three times with PBS.
5. Add 2 mL of medium to each well and administer each compound at six concentrations (Control, 0.1 μM , 1 μM , 5 μM , 10 μM , 20 μM).
6. Continue incubation under the same conditions for 48 h.
7. Discard the medium and wash the cells three times with PBS.
8. Add 200 μL of trypsin to each well for 1.5 min.
9. Terminate the digestion by adding 600 μL of medium.
10. Transfer the cells to the corresponding Ep tubes.
11. Add RIPA lysate containing phosphatase inhibitor and PMSF at 100 μL /sample.

12. Lysis on ice for 5 min.
13. Cells were broken by sonication, and each sample was sonicated six times on ice for 3 s each time.
14. Centrifuge at 4 °C, 12,000 rpm for 10 min.
15. Transfer the supernatant to a new Ep tube.
16. BCA protein quantification.
17. Configure separating gel and concentrating gel according to the procedure.
18. Upload samples according to the quantification results.
19. Electrophoresis at 90 V for 35 min, switch to 120 V and continue electrophoresis for 1.5 h.
20. Cut off the p-EphrinB2 band and transfer the membrane for 1.5 h (*see Note 8*).
21. Wash with TBST three times, 5 min each time.
22. Blocked with 5% skimmed milk for 2 h.
23. Rinse with TBST 3 times, 5 min each time.
24. Add p-EphrinB2 primary antibody and GAPDH primary antibody, incubate at 4 °C for 16 h.
25. Rinse the PVDF membrane with TBST three times, each time for 10 min.
26. Add the corresponding secondary antibody (1:20,000) and incubate for 1.5 h at room temperature.
27. Rinse the PVDF membrane with TBST three times, 10 min each time.
28. Develop protein bands by ECL method and take pictures.

3.2.3 Anti-Angiogenic Activity

Cytotoxicity

1. Cultured HUVEC cell line.
2. U87 cells were inoculated in 96-well plates (180 µL per well) at a density of 5000 cells/well, and eight concentrations (100 µM, 20 µM, 4 µM, 0.8 µM, 0.16 µM, 0.032 µM, 0.0064 µM, 0.00128 µM) were set up for each compound, with six replicate wells for each concentration, as well as a negative control well and a blank well.
3. Follow **steps 3–7** of subheading “[Anti-Tumor Cell Proliferation](#)”.

Anti-Migration (Scratch Test)

1. Cultured HUVEC cell line.
2. Follow **steps 2–8** of subheading “[Anti-Tumor Cell Migration \(Scratch Assay\)](#)”.

Anti-HUVEC Tubulogenesis

1. Cultured HUVEC cell line.
2. The complete medium was replaced with starvation medium (containing 0.2% FBS) the day before digestion and incubated for 24 h.
3. The matrix gel, 200 μ L sterile lance tip, and 96-well plate were placed at 4 °C overnight ahead of time (*see Note 7*).
4. Operated on ice, 50 μ L of matrix gel was added to each well.
5. Place the 96-well plate in a 37 °C incubator for 30 min.
6. 200 μ L of drug-containing HUVEC cell suspension was added to each well (5×10^4 cells at a drug concentration of 10 μ M).
7. Incubate at 37 °C, 5% CO₂ for 3–12 h.
8. Photographing network-like structure generation under an inverted microscope.
9. Analyzing correlation data with the Angiogenesis Analyzer plug-in for ImageJ.

3.3 Probes

Covalently Bind to Target Protein

3.3.1 Pure Protein

1. Cultivate EphrinB2 low-expressing cells HEK293 cells in six-well plates and incubate at 37 °C for 24 h with light protection.
2. Wash three times with pre-cooled PBS.
3. Add 200 μ L of trypsin to each well and digest for 1.5 min at room temperature.
4. Terminate the digestion by adding 600 μ L of medium to each well and transfer the cells to Ep tubes.
5. Wash the cells three times with pre-cooled PBS.
6. Add RIPA lysate containing phosphatase inhibitor and PMSF at 100 μ L/sample.
7. Lysed on ice for 5 min.
8. Cells were broken by sonication.
9. Centrifuge the cells at 12,000 rpm for 10 min at 4 °C.
10. Transfer the supernatant to a new Ep tube.
11. Add EphrinB2 recombinant protein according to the ratio of lysate: pure protein 10:1 (v:v), each sample contains 1–2 μ g of EphrinB2 recombinant protein.
12. Add different concentrations of multifunctional photoaffinity probes.
13. Co-incubate at 4 °C for 1 h.
14. Add 800 μ L of pre-cooled acetone per sample and incubate at –20 °C for more than 2 h or overnight.
15. Centrifuge at 4 °C, 12,000 rpm for 10 min.
16. Wash twice with pre-cooled methanol.

17. Place the Ep tube on ice and irradiate with 365 nm UV light for 20 min.
18. Repeat **steps 14–16**.
19. Add 100 μ L of PBS to each sample and set aside at -80°C (*see Note 10*).

3.3.2 Cell Lysate

1. EphrinB2-overexpressing U87 cells were cultured in six-well plates and incubated at 37°C for 24 h with light protection.
2. Follow **steps 2–10** in Subheading **3.3.1**.
3. Follow **steps 12–19** in Subheading **3.3.1**.

3.3.3 Live Cell

1. Cultivate EphrinB2 high-expressing cells U87 cells in six-well plates and incubate at 37°C for 24 h under light protection.
2. Add different concentrations of multifunctional photoaffinity probes to each well.
3. Incubate at 37°C for 4 h.
4. Wash the cells three times with PBS.
5. Add 2 mL of pre-cooled PBS to each well.
6. Place the six-well plate on an ice box and irradiate with 365 nm UV light for 20 min.
7. Wash the cells three times with pre-cooled PBS.
8. Add 200 μ L of trypsin to each well and digest for 1.5 min at room temperature.
9. Terminate the digestion by adding 600 μ L of medium per well and transfer the cells to Ep tubes.
10. Follow **steps 5–10** in Subheading **3.3.1**.
11. BCA protein quantification was performed for each sample.
12. Remaining samples were stored at -80°C for backup (*see Note 10*).

3.4 Subcellular Localization of Target Proteins Based on Cy3 as a Reporter Motif

1. Cultured U87 cell line.
2. 500 μ L of cell suspension containing 1×10^4 U87 cells was inoculated in a 35 mm laser confocal dish.
3. Incubated at 37°C , 5% CO_2 for 24 h.
4. The medium was discarded, and 500 μ L of 10 μM multifunctional photoaffinity probe medium was added and incubated for 4 h.
5. Cells were washed three times with pre-cooled PBS.
6. Add 500 μ L of pre-cooled PBS and irradiate with 365 nm UV light for 20 min.
7. Cells were washed three times with pre-cooled PBS.

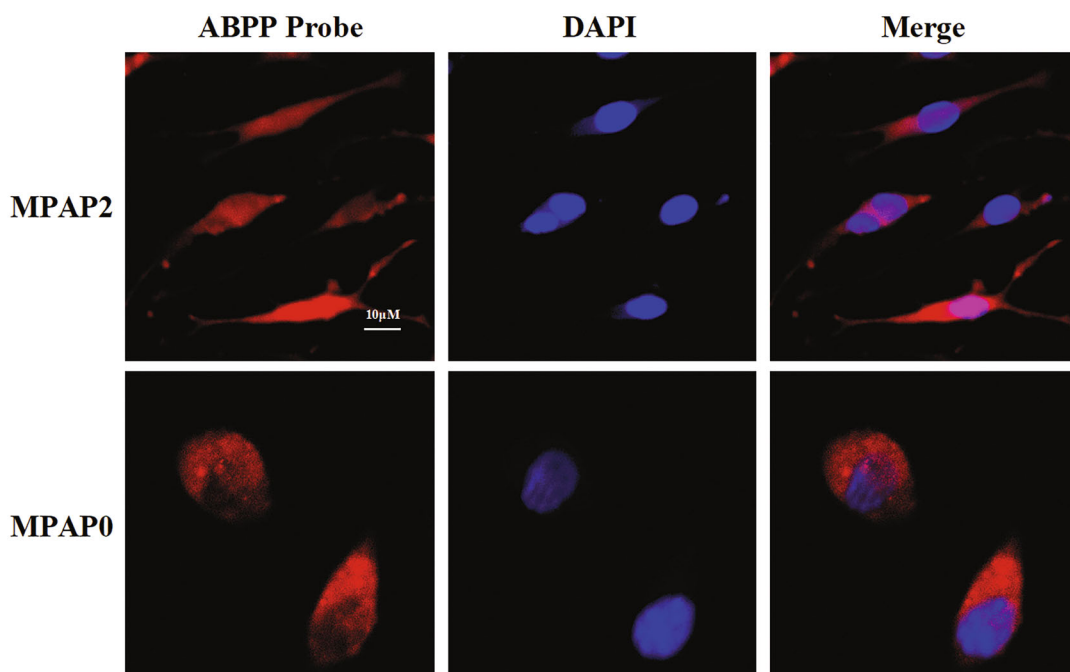


Fig. 3 Example of subcellular localization of the target protein EphrinB2

8. Add 500 μL of 4% paraformaldehyde solution and fix the cells for 30–50 min at room temperature.
9. Cells were washed three times with pre-cooled PBS.
10. 500 μL of bioorthogonal reaction solution (containing 20 μM L-ascorbic acid sodium, 7 μM Sulfo Cy3- N_3 and 2 μM CuSO_4) was placed.
11. Incubate in a constant temperature shaker at 37 $^\circ\text{C}$ for 1.5 h.
12. Wash the cells with pre-cooled PBS three times, each time for 5 min.
13. Add 200 μL of DAPI (1 $\mu\text{g}/\text{mL}$) solution and leave for 3 min at room temperature.
14. Wash the cells three times with pre-cooled PBS for 5 min each time.
15. The cells were excited by 405 nm and 561 nm light under an inverted fluorescence microscope or laser confocal microscope and photographed (Fig. 3).

3.5 Enrichment of Target Protein Based on Biotin as a Purification Moiety

3.5.1 Cell Lysate

1. Preparation of U87 cell lysates according to Subheading 3.3.1.
2. Protein quantification of cell lysate by BCA protein quantification kit.
3. Take an equal amount of lysate in a new Ep tube as needed and dilute to 100 μL with PBS.

4. Follow **steps 12–19** of Subheading **3.3.1**.
5. Add 100 μL of bioorthogonal reaction solution (containing 20 μM sodium L-ascorbate, 7 μM biotin- N_3 , and 2 μM CuSO_4).
6. Incubate at 37 $^\circ\text{C}$ in a constant temperature shaker for 1.5 h.
7. At the end of the incubation, 800 μL of pre-cooled acetone was added to each sample to terminate the reaction, and the reaction was incubated at -20°C for more than 2 h or overnight (*see Note 9*).
8. Centrifuge at 12,000 rpm for 10 min at 4 $^\circ\text{C}$.
9. Wash twice with pre-cooled methanol.
10. Add 200 μL of PBST to each sample (*see Note 10*).
11. Vortex the beads in a mixer for 20 s to mix well.
12. Take streptavidin magnetic beads according to 30 μL of each sample into Ep tube and put it on the magnetic rack for magnetic separation.
13. Discard the supernatant, add 1 mL of PBST, vortex and shake for 20 s, and magnetically separate the beads.
14. Repeat **step 25** once and discard the liquid.
15. Resuspend the beads in 200 μL of PBST per sample.
16. Add the beads (200 μL /sample) to the Ep tube in **step 22** and seal with sealing film.
17. Mix for 1.5 h in a rotary mixer.
18. Magnetic separation (*see Note 11*).
19. Add 800 μL of PBST to each sample, vortex for 15 s, separate magnetically, and discard the supernatant.
20. Repeat **step 31** three times.
21. For each sample, add 20 μL of PBST to resuspend the magnetic beads, and add 5 μL of reductive loading buffer (5 $\times$, denaturing).
22. Incubate at 100 $^\circ\text{C}$ for 5 min on a metal bath and separate magnetically.
23. Transfer the supernatant to a new Ep tube and store at -20°C for later use.

3.5.2 Live Cell

1. Cultivation of the U87 cell line.
2. Cells were inoculated in six-well plates at a cell density of 5×10^5 /well, ensuring that each well contained 2 mL of medium.
3. Incubated at 37 $^\circ\text{C}$, 5% CO_2 for 24 h.

4. The medium was aspirated, and 2 mL of multifunctional photoaffinity probe medium containing 10 μ M was added for 4 h of incubation.
5. Cells were washed three times with pre-cooled PBS.
6. Add 2 mL of pre-cooled PBS and irradiate with 365 nm UV light for 20 min.
7. Wash the cells three times with pre-cooled PBS.
8. Place 2 mL of bioorthogonal reaction solution (containing 20 μ M L-ascorbic acid sodium, 7 μ M biotin- N_3 , and 2 μ M $CuSO_4$).
9. Incubate in a constant temperature shaker at 37 °C for 1.5 h.
10. Wash the cells three times with pre-cooled PBS for 5 min each time.
11. Add 200 μ L of trypsin to each well and digest for 1.5 min at room temperature.
12. Terminate the digestion by adding 600 μ L of medium per well and transfer the cells to Ep tubes.
13. Wash the cells three times with pre-cooled PBS.
14. Add RIPA strong lysate containing phosphatase inhibitor and PMSF at 100 μ L/sample.
15. Lysis on ice for 5 min.
16. The cells were broken by sonication.
17. Centrifuge the cells at 4 °C, 12,000 rpm for 10 min.
18. Transfer the supernatant to a new Ep tube.
19. Add 800 μ L of acetone to each sample and leave overnight at –20 °C.
20. Centrifuge at 12,000 rpm for 10 min at 4 °C.
21. Wash the proteins twice with pre-cooled methanol and repeat **step 20**.
22. Follow **steps 10–23** of Subheading [3.5.1](#).

3.6 Purification of Target Protein

1. Prepare separator and concentrate according to the procedure.
2. Sample the enriched protein samples at 20 μ L per well.
3. Electrophoresis at 90 V for 35 min, switch to 120 V, and continue electrophoresis for 1.5 h.
4. Cut down the EphrinB2 and GAPDH bands, and flow the membrane at 200 mA for 1.5 h under the condition of 0–4 °C.
5. After membrane transfer, rinse the membrane with TBST three times, each time for 5 min.
6. Seal the membrane with 5% skimmed milk for 2 h.
7. Rinse with TBST 3 times, 5 min each time.

8. Add EphrinB2 primary antibody and GAPDH primary antibody, and incubate at 4 °C for 16 h.
9. Rinse the PVDF membrane three times with TBST for 10 min each time.
10. Add the corresponding secondary antibody (1:20,000) and incubate for 1.5 h at room temperature.
11. Rinse the PVDF membrane with TBST three times, each time for 10 min.
12. Develop protein bands by ECL method.

3.7 Target Protein Identification and Binding Site Resolution

1. 1 μ L of probe was added to 99 μ L of EphrinB2 recombinant protein (EphrinB2 recombinant protein quality assurance 10 μ g, probe concentration 100 μ M).
2. Incubate at 4 °C for 1 h.
3. 365 nm UV illumination for 20 min.
4. Add 800 μ L of pre-cooled acetone and incubate overnight at – 20 °C.
5. Centrifuge at 12,000 rpm for 10 min at 4 °C and discard the supernatant.
6. Wash the protein twice with methanol and repeat the centrifugation.
7. Add 200 μ L of PBS (*see Note 12*).
8. Add 100 mM Tris–HCl + 8 M urea to the sample.
9. Centrifuge at 12,000 rpm for 20 min at 4 °C, discard the supernatant.
10. Repeat **step 9** once.
11. After protein concentration, prepare 72/85/95/95 μ L samples.
12. Add 3.8/4.5/5/5 μ L of DTT solution at 100 mM (final concentration 5 mM) to the sample.
13. Incubate at 60 °C for 20 min.
14. After cooling to room temperature, add 2.34/2.77/3.1/3.1 μ L of IAA solution at a concentration of 500 mM (final concentration of 15 mM) and incubate for 15 min away from light.
15. Add 20 μ g of serial grade trypsin dry powder to 40 μ L of solution containing 1 M urea.
16. Add 1 μ L of enzyme suspension in the ratio of protein: enzyme = 1:20 and incubate at 35 °C for 2–18 h.
17. Dry by vacuum centrifugal concentrator.
18. Dissolve in 0.1% FA.

19. Activate Millipore C18 ZipTip by adding 10 μL of 100% acetonitrile neutral solution, repeat this step once.
20. Equilibrate Millipore C18 ZipTip by adding 10 μL of 0.1% FA solution, repeat this step once.
21. Add 10 μg of the sample to be desalted and repeat the combination 3–8 times.
22. Add 10 μL of 0.1% FA solution to wash the sample; this step of the experiment was repeated one time.
23. Add 10 μL of 70% acetonitrile and the effluent is the purified sample.
24. The sample was dried under vacuum with a vacuum centrifugal concentrator to a dry powder.
25. Add 20 μL of 2% acetonitrile/98% H_2O /0.1% FA to dissolve the sample.
26. LC-MS and LC-MS/MS analysis, U300-QE Plus, 2 μL sample volume.
27. Liquid phase conditions: phase A 0.1% formic acid aqueous solution, phase B acetonitrile (acetonitrile Merck mass spectrometry grade, 18.2 Ω ultrapure water) column: trap column, Thermo Fisher, 3 μm i.d, 75 $\mu\text{m} \times 2.0$ cm; the analytical column is a direct spray self-filling column with a particle size of 1.9 μm i.d 75 $\mu\text{m} \times 15$ cm (Fig. 4a).
28. Mass spectrometry conditions: spray voltage 2.1 kV, ion transfer tube temperature 300 $^\circ\text{C}$, lens voltage (S-lens RF level) 60, the rest of the detailed parameters are shown in Fig. 4b, c.
29. Data processing: after the samples were repeatedly collected three times, the data were imported into Thermo Proteome Discoverer 2.3 for processing (Fig. 5).
30. The FASTA database was downloaded from the UniProt database for library searching and analysis (Fig. 6).

4 Notes

1. Incubation time after CCK-8 addition was standardized according to the absorbance value of 0.2–0.8 at 450 nm, generally 1–1.5 h.
2. Apoptosis requires sample cell concentration to be adjusted to 1×10^6 cells/mL.
3. Annexin V-FITC and PI staining need to be detected on the machine within 1 h.
4. Cell cycle assays were performed as much as possible within 5 h after staining.

A

No	Time	Flow [μ l/min]	%B	Curve
1	0.000	Run		
2	0.000	0.600	2.0	5
3	5.000	0.600	2.0	5
4	10.000	0.600	8.0	5
5	43.000	0.600	28.0	5
6	53.000	0.600	40.0	5
7	57.000	0.600	100.0	5
8	62.000	0.600	100.0	5
9	65.000	0.600	2.0	5
10	70.000	0.600	2.0	5
11	New Row			
12	70.000	Stop Run		

B

Sheath gas flow rate	0
Aux gas flow rate	0
Sweep gas flow rate	0
Spray voltage ([kV])	2.10
Spray current (μ A)	
Capillary temp. ($^{\circ}$ C)	300
S-lens RF level	60.0

C

Properties of the method

Use lock masse: best
Chrom. peak wi 20 s

Time
Method duration 70.00 min

Method duration
Duration of the method

Properties of Full MS / dd-MS² (

General
Runtime 0 to 70 min
Polarity positive
Default charge : 2
Inclusion —
Exclusion —
Tags —

Full MS
Resolution 70,000
AGC target 1e6
Maximum IT 240 ms
Scan range 350 to 1500 m/z

dd-MS² / dd-SIM
Resolution 17,500
AGC target 1e5
Maximum IT 50 ms
Loop count 10
TopN 10
Isolation window 2.0 m/z
Fixed first mass —
(N)CE / stepped nce: 27

dd Settings
Minimum AGC t: 1.00e3
Intensity thresh 2.0e4
Apex trigger —
Charge exclusio 1, 7, 8, >8
Peptide match preferred
Exclude isotope: on
Dynamic exclus 20.0 s

Fig. 4 LC-MS analysis conditions: (a) U300 liquid chromatography conditions; (b, c) QE Plus mass spectrometry conditions

- When cells in Ep tubes are washed with PBS, try not to use vortex shaking mixing to avoid mechanical damage to cells.
- All fluorescent dyes are subject to quenching and should be stored and used away from light.
- When using matrix gel, it is necessary to operate strictly on ice, and the consumables used need to be pre-cooled in advance to avoid matrix adhesive solidification and failure.
- When performing western blot membrane transfer operation, it needs to be performed at low temperature, and the voltage should not be higher than 80 V, otherwise the protein will be easily burned.
- Bioorthogonal reactions can be configured with a bioorthogonal reaction solution or the components can be added to the system sequentially, taking care to add the CuSO₄ solution last to initiate the reaction. Cell lysate-level experiments can be terminated by adding pre-cooled acetone; live cell levels are

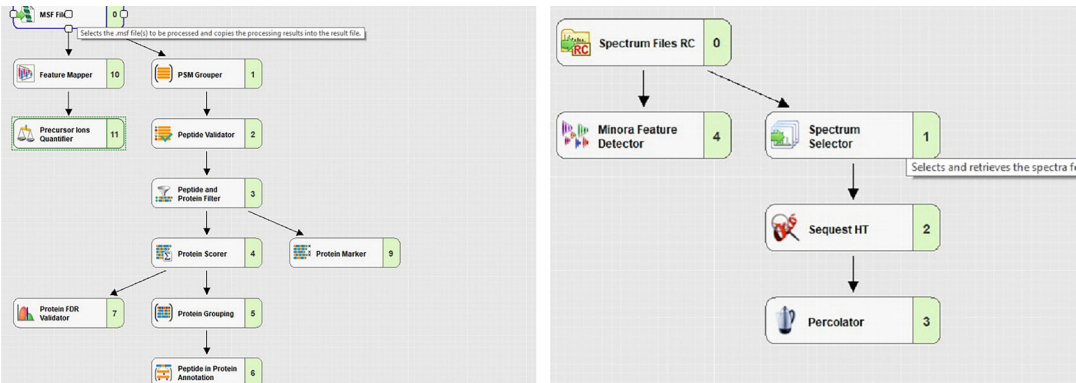


Fig. 5 Thermo Proteome Discoverer 2.3 data processing flow

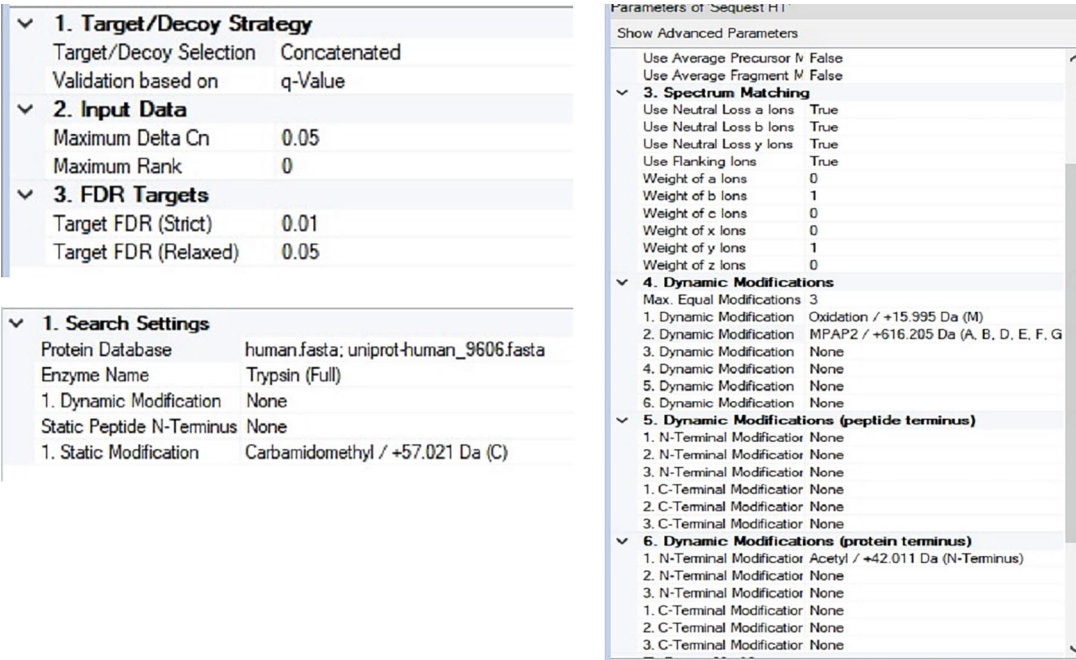


Fig. 6 FASTA database search parameter settings

terminated by discarding the reaction solution and washing the cells with PBS.

10. All protein solutions should be stored at -80°C .
11. Each magnetic separation operation lasted 1–3 min, and it was necessary to wait until all the magnetic beads were adsorbed and the supernatant was discarded when the solution was clarified.
12. Target protein identification can also be initially verified by competition experiments.

Acknowledgments

This work was supported by the Shaanxi Engineering Research Center for Vascular Drugs and the Instrumental Analysis Center of Xi'an Jiaotong University. This work was funded by the National Natural Science Foundation of China (NSFC, nos. 82373793 and 82173742 and 823B2096), the Science Fund for Distinguished Young Scholars of Shaanxi (no. 2022JC-54), and the Key Research and Development Program of Shaanxi (no. 2023-YBSF-131).

References

1. Rani V, Prabhu A (2020) Combining angiogenesis inhibitors with radiation: advances and challenges in cancer treatment. *Lancet Oncol* 21: 1395–1397
2. Zhang S, Cao M, Hou Z et al (2021) Angiotensin-converting enzyme inhibitors have adverse effects in anti-angiogenesis therapy for hepatocellular carcinoma. *Cancer Lett* 501:147–161
3. Jain RK (2014) Antiangiogenesis strategies revisited: from starving tumors to alleviating hypoxia. *Cancer Cell* 26:605–622
4. Mattheolabakis G, Mikelis CM (2019) Nanoparticle delivery and tumor vascular normalization: the chicken or the egg? *Front Oncol* 9: 1227
5. Chen X, Wong Y, Wang J et al (2017) Target identification with quantitative activity-based protein profiling (ABPP). *Proteomics* 17:3–4
6. Huth S, Oakley J, Seath C et al (2023) μ Map Photoproximity labeling enables small molecule binding site mapping. *J Am Chem Soc* 145: 16289–16296
7. Kawamata Y, Ryu K, Hermann G et al (2023) An electroaffinity labelling platform for chemoproteomic-based target identification. *Nat Chem* 15:1267–1275
8. Dai B, Shi X, Ma N et al (2018) HMQ-T-B10 induces human liver cell apoptosis by competitively targeting EphrinB2 and regulating its pathway. *J Cell Mol Med* 22:5231–5243



Activity-Based Protein Profiling to Study Virus–Host and Viral-Induced MicroRNA–Host Interactions

Parrish Evers and John Paul Pezacki

Abstract

Activity-based protein profiling (ABPP) is a robust technique enabling the user to explore the functional implications of disease and pathology on key protein effectors. As such, a variety of activity-based probes have been developed and applied broadly to study virus–host interactions and the consequences of essential noncoding RNAs, like microRNAs, which viruses exploit for their propagation. The following protocol describes the process of infecting and transfecting cells for ABPP and covers the main ways in which ABPP can be used to study infection: in-gel fluorescence, western blotting, and mass spectrometry. Taken together, the findings of these experiments can be used to identify new therapeutic targets and to test novel antiviral inhibitors.

Key words Activity-based protein profiling (ABPP), Chemoproteomics, Virus, microRNA, Hepatic lipase (LIPC), Mass spectrometry, Western blotting, In-gel fluorescence, Chemical Proteomics

1 Introduction

Since its conceptualization, activity-based protein profiling (ABPP) has filled a critical role in studying disease states by allowing researchers to expand beyond conventional, abundance-based experiments. ABPP reveals the functional proteome of a given enzyme class, which can be especially important when a post-translational modification (PTM) activates or inhibits a target protein [1] or when designing selective active site inhibitors [2]. Fluorophosphonate (FP) containing activity-based probes (ABPs) [3] targeting the serine hydrolases are particularly well-employed, given the physiological relevance of this enzyme family [4]. Unsurprisingly, many viral pathogens differentially regulate the serine hydrolases to promote pathogenesis, and as such, FP ABPP has previously been applied to study virus–host interactions [5, 6] and

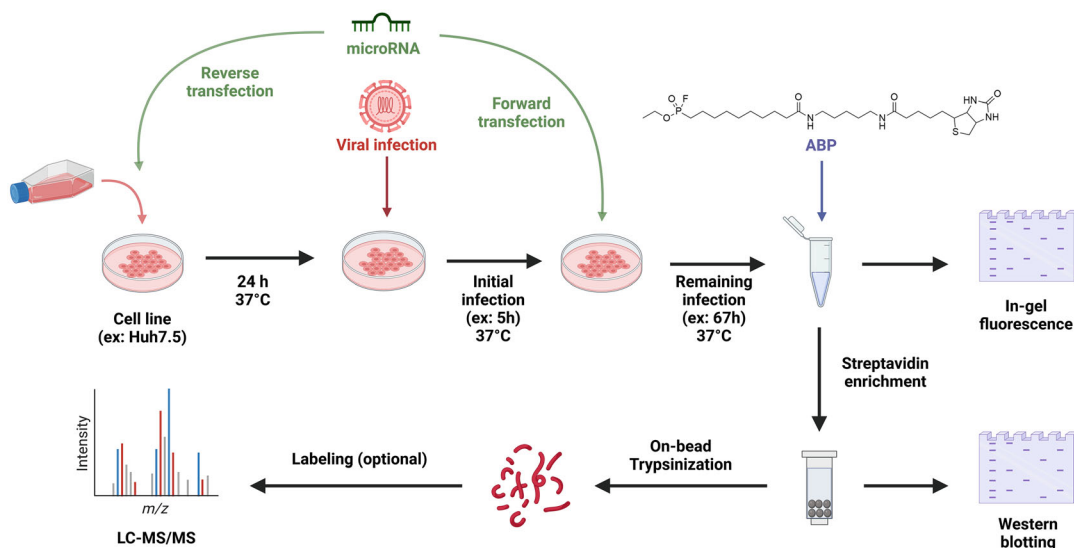


Fig. 1 Using ABPP to study viral infections and the consequences of microRNA regulation. Cells can (optionally) be reverse or forward transfected with a microRNA (or siRNA) of choice prior to infection or during infection to study the consequences. Following viral infection, the cells can be lysed and labeled with the activity-based probe (ABP), FP-biotin in this example. If necessary, a click reaction can be performed after ABP labeling to attach the appropriate reporter tag. These labeled lysates can then either be used directly to study in-gel fluorescence or enriched using streptavidin-agarose beads for either western blotting or mass spectrometry preparation. An optional labeling step on the digested peptides, such as dimethyl labeling, allows for improved comparison between experimental conditions before mass spectrometry

noncoding RNAs modulated by viruses [6, 7], where it can aid in identifying downstream effectors of viral pathogenesis (reviewed in ref. [8]).

ABPP can be applied *in vivo*, *in cellulo*, or *in vitro* following lysis depending on the nature of the probe [1]. Given our lab's success labeling *in vitro* post-lysis, we will focus on this for the current protocol where we describe ABPP using the FP-TAMRA fluorescent probe for in-gel fluorescence and FP-biotin probe for streptavidin enrichment (Fig. 1). Following streptavidin enrichment, FP-biotin-labeled samples can either be prepared for liquid chromatography mass spectrometry (LC-MS/MS) (labeled or label-free) or eluted for western blotting. Unlike many ABPP probes, FP-TAMRA and FP-biotin discussed in this work do not require an azide-alkyne click-chemistry reaction [9] to attach a reporter group; however, this process will still be described to more broadly suit virus-ABPP labeling purposes. It is also important to note that the ABPP techniques described build upon the pioneering work of Benjamin Cravatt's group [1, 10], which we have modified and expanded to include ABPP for western blotting and how to use ABPP in studying virus and microRNA (miRNA) regulation.

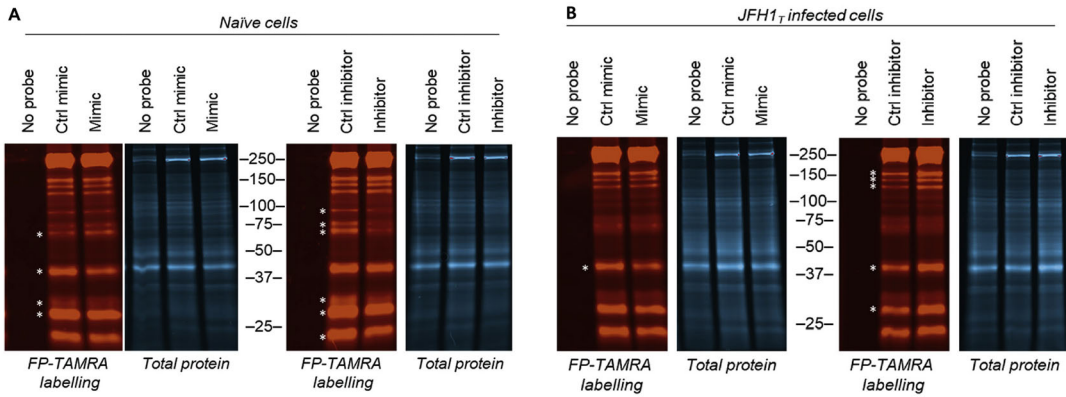


Fig. 2 FP-TAMRA labeling of miR-27b mimic and inhibitor-treated cells followed by HCV infection. Huh7.5 cells were mock infected (**a**) or infected (**b**) for 5 h with the JFH1_T strain of HCV. The cells were then transfected with fresh media containing either miR-27b (mimic), a negative control miRNA (ctrl mimic), an inhibitor of miR-27b (inhibitor), or an miRNA inhibitor-negative control (ctrl inhibitor) for 67 h. Cells were then lysed and labeled with FP-TAMRA before running on 10% acrylamide SDS-PAGE. Samples were imaged on a rhodamine channel (FP-TAMRA labeling) and a stain-free channel (total protein). Changes in protein labeling were quantified by densitometry and are denoted by an * beside the band. (This figure was adapted from ref. [6] with the permission of Elsevier Inc. on behalf of the American Society for Biochemistry and Molecular Biology)

As a workflow, our group will typically first perform FP-TAMRA labeling and screen for initial targets because it requires much less protein and is less time-intensive. In-gel fluorescence is particularly effective for visualizing activity changes in highly abundant and easily identifiable proteins like fatty acid synthase (FASN) [5]; however, in-gel fluorescence has lower sensitivity [1, 7] and is limited when trying to resolve proteins of similar molecular weight. Therefore, after FP-TAMRA labeling, our group generally performs FP-biotin labeling for mass spectrometry and then validates mass spectrometry hits with FP-biotin labeling for western blotting. The following method will be presented with selected figures from our group's recent work studying hepatitis C virus (HCV) and the role of miR-27b in promoting HCV through modulation of serine hydrolase activity [6]. This work led to the identification of human hepatic lipase (LIPC) as an antiviral effector downregulated by HCV through miR-27b expression. The examples of FP-TAMRA (Fig. 2) and FP-biotin for western blotting (Fig. 3) and mass spectrometry (Fig. 4) from this work will be presented to demonstrate the described protocol.

2 Materials

2.1 Cell Culture, miRNA Transfection, Viral Infection, and Lysis

1. Serological pipettes (25 mL, 10 mL, and 5 mL).
2. 1.5 mL microcentrifuge tubes.
3. Temperature and CO₂-controlled cell incubator (*see Note 1*).

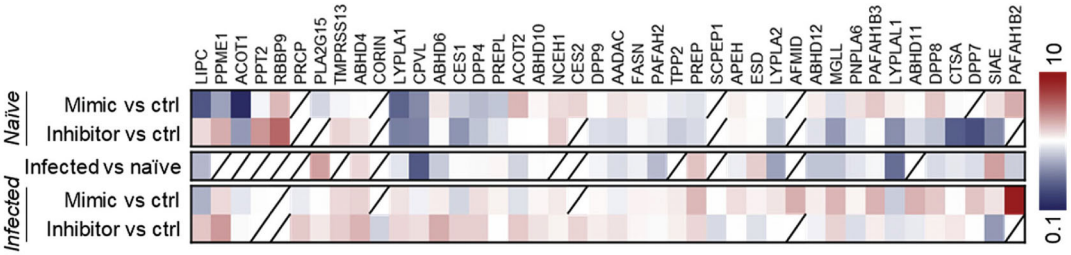


Fig. 3 LC-MS/MS following FP-biotin streptavidin enrichment of miR-27b mimic and inhibitor-treated cells followed by HCV infection. Huh7.5 cells were mock infected (a) or infected (b) for 5 h with the JFH1_T strain of HCV. The cells were then transfected with fresh media containing either miR-27b (mimic), a negative control miRNA (ctrl mimic), an inhibitor of miR-27b (inhibitor), or an miRNA inhibitor-negative control (ctrl inhibitor) for the rest of the infection (67 h). Cells were then lysed and labeled with FP-biotin before performing a streptavidin enrichment and on-bead trypsinization. The data shown reflects $n = 3$ biological replicates where each biological replicate was run through LC-MS/MS in technical triplicates ($n = 3$). (This figure was adapted from ref. [6] with the permission of Elsevier Inc. on behalf of the American Society for Biochemistry and Molecular Biology)

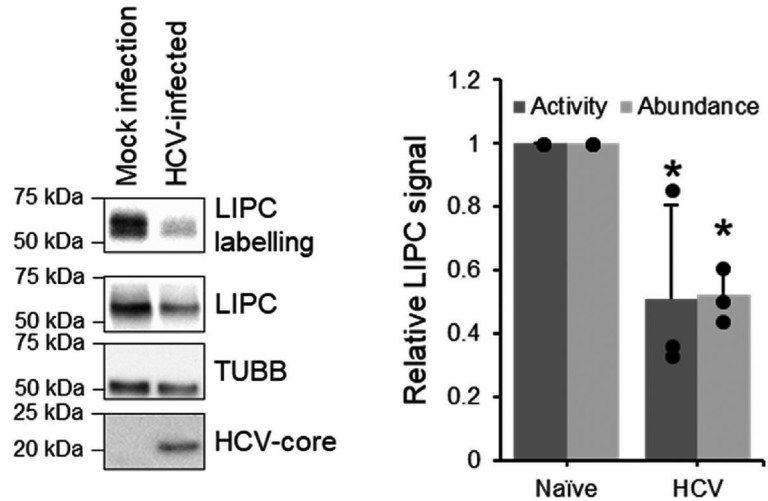


Fig. 4 Western blotting following FP-biotin streptavidin enrichment HCV-infected and null infected cells. Huh7.5 cells were infected the JFH1_T strain of HCV for 72 h. Cells were then lysed and labeled with FP-biotin before performing a streptavidin enrichment and elution using SDS-PAGE elution buffer. Samples were performed in biological replicates ($n = 3$). (This figure was adapted from ref. [6] with the permission of Elsevier Inc. on behalf of the American Society for Biochemistry and Molecular Biology)

4. Eukaryotic cell media: Eagle Minimum Essential Medium (EMEM) or Dulbecco's Minimum Essential Medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS) (store at 4 °C) (*see* **Note 2**).

5. Minimal eukaryotic cell media: Eukaryotic Cell media: EMEM or DMEM.
6. Opti-MEM™ I Reduced Serum Medium media (Gibco) (store at 4 °C).
7. Adherent culture plates or dishes (*see* **Note 3**).
8. mirVana miRNA mimic (Thermo Fisher Scientific) (store at –20 °C).
9. mirVana miRNA mimic negative control (Thermo Fisher Scientific) (store at –20 °C).
10. Lipofectamine™ RNAiMAX Transfection Reagent (Thermo Fisher Scientific) (store at 4 °C).
11. Virus stock (e.g., HCV) with known plaque-forming units (PFUs) or focus-forming units (FFUs) (store aliquots at –80 °C).
12. Phosphate-buffered saline (PBS): 137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, and 1.8 mM KH₂PO₄ (store at room temperature) (*see* **Note 4**).
13. Centrifuge capable of temperature control and high speeds ($\geq 14,000$ g) for 1.5 mL microcentrifuge tubes and (≥ 4000 g) 15 mL conical tubes.
14. Lysis buffer: 1% (v/v) Triton X-100 in PBS (prepare fresh).
15. Cell scrapers.

**2.2 ABP Labeling,
Washes, Streptavidin
Enrichment, and
Preparation for LC-MS/
MS Submission**

1. 100 μ M fluorophosphonate-tetramethylrhodamine (FP-TAMRA) (Thermo Fisher Scientific) in dimethyl sulfoxide (DMSO) (store at –20 °C).
2. 1 mM fluorophosphonate-PEG4-biotin (FP-biotin) (Santa Cruz Biotech.) in DMSO (store at –20 °C).
3. Rhodamine-azide (store at –20 °C).
4. 5 mM biotin-azide (biotin-PEG4-azide or similar) prepared in DMSO (store at –20 °C).
5. 2 $\times$ Laemmli buffer: 65.8 mM Tris–HCl (pH 6.8), 26.3% (w/v) glycerol, 2.1% SDS (store at room temperature); add 50–100 mM dithiothreitol (DTT) in deionized H₂O freshly before use.
6. Conical 15 mL tubes.
7. End-over-end rotator.
8. Acrylamide pre-cast, fast cast, or manually casted gels (capable of 40–50 μ L well volume).
9. Acetone (pre-chill at –20 °C).
10. Methanol (pre-chill at –20 °C).
11. 2.5% sodium dodecyl sulfate (SDS) (w/v) in PBS (store at room temperature).

12. 50 mM Tris(2-carboxylethyl)phosphine chloride, HCl (TCEP) prepared in deionized H₂O (prepare fresh).
13. 50 mM Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA) prepared in DMSO (store at -20 °C) and diluted in 4:1 ratio of *t*-butanol:DMSO prior to use.
14. 50 mM copper (II) sulfate prepared in deionized H₂O (store at room temperature).
15. 50 mM ammonium bicarbonate (prepare fresh).
16. 0.5 mg/mL sequencing grade modified trypsin prepared from 20 µg lyophilized powder using provided trypsin suspension buffer (store at -80 °C).
17. Micro Bio-Spin™ Chromatography Columns (Bio-Rad) (*see Note 5*).
18. Streptavidin agarose resin 50% aqueous slurry with 0.02% sodium azide (*see Note 6*) (store at 4 °C).
19. 6 M urea in deionized H₂O (prepare fresh).
20. 10 mM DTT in 50 mM ammonium bicarbonate (prepare fresh).
21. 25 mM iodoacetamide in deionized H₂O (prepare fresh).
22. 1% SDS (w/v) prepared in PBS or deionized H₂O (store at room temperature).
23. Vacuum centrifuge.
24. 4% (w/v) formaldehyde prepared from 36.5–38% (w/v) in H₂O (prepare fresh).
25. 4% (w/v) deuterated formaldehyde prepared from 20% (w/v) 98 atom % D in deionized H₂O (prepared fresh).
26. 4% (w/v) deuterated C13 formaldehyde prepared from 20% (w/v) 99 atom % D in deionized H₂O (prepared fresh).
27. 0.6 M sodium cyanoborohydride in deionized H₂O (prepare fresh).
28. 0.6 M sodium cyanoborodeuteride in deionized H₂O (prepare fresh).
29. 1% (v/v) ammonia hydroxide prepared in deionized H₂O from 28–30% (v/v) ammonia hydroxide in H₂O (prepare fresh).
30. 0.1% and 5% (v/v) formic acid prepared in deionized H₂O from 88% formic acid (prepare fresh).
31. Litmus pH paper.
32. 80% and 50% (v/v) acetonitrile prepared in deionized H₂O from 99% acetonitrile (LC-MS grade) (prepare fresh).
33. 80% acetonitrile with 0.1% formic acid (prepare fresh).

34. C18 pipette tips (Pierce) (*see* **Note 7**).
35. Micropipette tips (1000–1250 μ L, 200 μ L, and 10 μ L).

3 Methods

3.1 Cell Culture and Viral Infection

Perform all steps in accordance with the biosafety regulations specified by the appropriate government agency(s) regulating your laboratory practices.

1. Grow cells in appropriate culture medium and conditions specified by the manufacturer or determined through optimization [For example, we grow Huh7.5 s in DMEM supplemented with 10 mM non-essential amino acids (NEAAs) and 10% (v/v) FBS.]
2. Seed cells at appropriate density for viral infection. Cells should be seeded to achieve 1–2 mg/mL total protein after lysis in either 70 μ L or 1 mL depending on whether the cells will be prepared for in-gel fluorescence or streptavidin enrichment, respectively. Alternatively, reverse transfect cells with appropriate miRNA or siRNA using Subheading 3.2 (*see* **Note 8**).
3. Incubate at 37 °C overnight.
4. The next day (roughly 24 h later) use the cell growth factor to determine the number of cells in preparation for infection (*see* **Note 9**).
5. Prepare to infect the cells at the appropriate multiplicity of infection (MOI) for the virus being studied (for HCV JFH1_T 2A, our group uses an MOI = 0.1). To determine the total PFUs or FFUs required, multiply the number of cells by the number of wells/plates and by the MOI. Using the previously determined PFU/FFU of your viral stock, prepare a dilution in minimal growth media lacking supplements (e.g., no FBS) that will be used to infect the cells (*see* **Note 10**).
6. Aspirate the media off the wells/plate.
7. Add the minimal media containing the virus to cells being infected and add minimal media without virus to the no infection control. Incubate at 37 °C for appropriate infection time (e.g., for HCV this is 5 h) (*see* **Note 11**).
8. Aspirate and carefully discard virus-containing media and minimal media from cells. Replace with normal supplemented media and incubate for the remaining infection period (e.g., 67 h for HCV). Alternatively, transfect cells for remaining infection period with appropriate miRNA or siRNA using Subheading 3.2.

3.2 Reverse Transfecting miRNA or siRNA (Optional)

Prepare the following steps at room temperature. Allow RNAiMAX to warm briefly to room temperature before combining with miRNA/siRNA. The following subsection is closely adapted from the manufacturer's protocol.

1. For each experimental condition, prepare correct dilution of siRNA/miRNA in Opti-MEM media from 100 mM stock to achieve 10–100 nM final concentration in wells/plates. Prepare a separate RNAiMAX transfection reagent tube in opti-MEM where 2.5 μ L of RNAiMAX is added for every 1 μ L of 100 mM miRNA/siRNA in the first tube. The final opti-MEM concentration in the well/plate should be 1:5 (v/v) opti-MEM:media in the final well volume; therefore, this should be 1:10 (v/v) for the siRNA/miRNA tube and 1:10 (v/v) for the RNAiMAX tube (e.g., for a 15 cm dish with 15 mL final volume, there should be 1.25 mL opti-MEM each in the miRNA/siRNA tube and in the RNAiMAX tube; this results in 2.5 mL opti-MEM being added to each plate) (*see Note 12*).
2. Combine opti-MEM-siRNA/miRNA tube and opti-MEM-RNAiMAX tube for each sample. Mix by gently pipetting up and down. Wait at least 5 min but not more than an hour.
3. Add opti-MEM-RNAiMAX-miRNA/siRNA mixture to its corresponding plate.
4. Add cells in media at appropriate density to achieve either 70 μ L or 1 mL at 1–2 mg/mL total protein after lysis depending on if the cells are being prepared for in-gel fluorescence or streptavidin enrichment, respectively (*see Note 13*).

3.3 Collection, Lysis, and Preparation for Labeling with ABP

After collecting the cells in step 5, perform all additional steps on ice to minimize proteome degradation.

1. Remove cell media using a micropipette or serological pipette (*see Note 14*).
2. Wash cells with sterile 1 $\times$ PBS by slowly adding the half the well/plate volume (e.g., 5 mL 1 $\times$ PBS in a 10 cm dish) to the well wall, gently tilting the plate a few times, and then aspirating the PBS with a micropipette or serological pipette (*see Note 15*).
3. Add pre-chilled 1 $\times$ PBS with 1% Triton lysis buffer (e.g., 1 mL for a 15 cm dish). The lysis buffer should be added to cover as much plate surface area as possible while maintaining appropriate protein concentration (*see Note 16*).
4. If using a six-well plate or larger, scrape cells with a cell scraper by tilting plate at roughly 45° and scraping from the top to bottom of the plate. If using a 12-well plate or smaller, pipette up and down and gently agitate to lift cells.

5. Transfer lysis buffer with cells to a microcentrifuge tube and keep on ice (*see Note 17*).
6. Lyse the cells by sonicating for 15 pulses set to 30% duty and 1 s per pulse (*see Note 18*).
7. Centrifuge at 4 °C maximum speed ($\geq 20,000$ g) for 5 min. Transfer the supernatant to a new microcentrifuge tube on ice.
8. Take a 10–20 μ L aliquot of the lysate and quantify protein abundance using a common assay such as a DC assay or BCA assay (*see Note 19*).

3.3.1 FP-TAMRA Labeling for In-Gel Fluorescence

1. Prepare an aliquot of lysate that is 70 μ L and 1 mg/mL total protein concentration using lysis buffer.
2. Add the ABP of choice to the lysate at its ideal concentration (e.g., 0.7 μ L of 100 μ M FP-TAMRA). At this step, you will also add empty vehicle (e.g., 0.7 μ L DMSO) to the no probe control lysate (*see Note 20*).
3. Incubate on a rotating stand at 37 °C for 1 h (*see Note 21*).
4. If the ABP does not require click-chemistry to attach the fluorescent reporter, proceed directly to **step 7**.
5. Prepare a 1 mL mastermix adding reagents in the following order and vortexing after each addition: 20 μ L of freshly prepared 50 mM TCEP, 20 μ L of 50 mM copper (II) sulfate, 2 μ L of 50 mM TBTA, 956 μ L of PBS, and then 2 μ L of 50 mM rhodamine-azide (*see Note 22*).
6. Add 100 μ L of click mastermix to each labeled sample and incubate rotating end-over-end on a rotating stand for 45 min (*see Note 23*).
7. Terminate the labeling by adding 1 mL pre-chilled acetone and incubating at –20 °C or –80 °C for at least 20 min to precipitate proteome (*see Note 24*).
8. Stop incubating, remove supernatant, and allow the pellet to thaw at room temperature or on ice before centrifuging at maximum speed ($\geq 20,000$ g), 4 °C, for 15 min (*see Note 25*).
9. Let the pellet air dry for 5 min (*see Note 26*).
10. Terminate labeling or click reaction by adding 30 μ L of 2x Laemmli buffer with 30–100 mM DTT and heat for 5–10 min at 95 °C.
11. Run on acrylamide gel in the dark (to minimize TAMRA degradation).
12. To normalize rhodamine signal to total protein loaded, take a stain-free image (if gel contains trichloroethanol) or perform your preferred total protein stain (Coomassie, ponceau, etc.) on your gel.

3.3.2 FP-Biotin Labeling for Streptavidin Enrichment

1. Prepare an aliquot of lysate that is 1 mL and between 1 and 2 mg/mL total protein concentration using lysis buffer (*see Note 27*).
2. Add the ABP of choice to the lysate at its ideal concentration (e.g., 5 μ L of 1 mM FP-biotin). At this step, you will also add empty vehicle (e.g., 5 μ L DMSO) to the no probe control lysate (*see Note 28*).
3. Incubate on a rotating stand at 37 °C for 1 h (*see Note 21*).
4. If the ABP does not require click-chemistry to attach the biotin reporter, proceed directly to **step 12**.
5. Split the 1 mL of 1–2 mg/mL ABP-biotin labeled lysate into two 500 μ L fractions. It is important that the click reagents added in **steps 2–5** are added in order.
6. Add 10 μ L of 5 mM biotin-azide. Vortex.
7. Add 10 μ L of freshly prepared 50 mM TCEP. Vortex.
8. Add 10 μ L of 5 mM TBTA. Vortex.
9. Add 10 μ L of 50 mM copper (II) sulfate. Vortex.
10. Rotate end-over-end on a rotator for 45 min (*see Note 23*).
11. Recombine the two 500 μ L fractions.
12. Terminate the labeling or click reaction by transferring the sample to a 15 mL conical tube and adding 5 mL pre-chilled acetone. Incubate at –20 °C or – 80 °C for at least 20 min to precipitate proteome (*see Note 24*).
13. Stop incubating and allow the pellet to thaw at room temperature or on ice before centrifuging at maximum speed ($\geq 20,000$ g), 4 °C, for 15 min (*see Note 25*).
14. Remove supernatant and add 750 μ L of pre-chilled methanol. Resuspend pellet by sonicating for 5, 1 s pulses (30% duty) (*see Note 29*).
15. Centrifuge at maximum speed ($\geq 20,000$ g), 4 °C, for 5 min.
16. Repeat **steps 14** and **15** (methanol washes) two more times.
17. Air-dry pellet for 5 min (*see Note 26*).
18. Add 650 μ L of 2.5% SDS in PBS and resuspend pellet by sonicating for 15, 1 s pulses (30% duty). Transfer to a new 1.5 mL microcentrifuge tube (*see Note 30*).
19. Heat samples at 60 °C for 5 min and centrifuge for 4 min at 6500 g. Take a 20 μ L input/loading control from the supernatant and transfer the remaining supernatant to a 15 mL conical tube (*see Note 31*).
20. Bring the samples to 8 mL total volume by adding 7.370 mL 1 $\times$ PBS and proceed to Subheading 3.4. Alternatively, if storing overnight, bring the samples to 3.5 mL total volume and freeze.

3.4 Streptavidin Enrichment of FP- Biotin-Labeled Proteome

1. Place a Micro Bio-Spin™ centrifugal column in 2 mL collection tube and then aliquot 100 μ L of 50% slurry streptavidin-agarose resin/bead slurry into the column (*see Note 32*).
2. Wash the beads with 700 μ L PBS and then spin at 1000 g for 1 min. Discard flowthrough.
3. Repeat **step 3** two more times (*see Note 33*).
4. Transfer beads to the 8 mL containing the ABP-labeled proteome using two 500 μ L volumes of PBS. Ensure beads have equilibrated to room temperature before proceeding to the next step (*see Note 34*).
5. Incubate samples for 1.5 h at room temperature rotating end-over-end on a rotating stand (*see Note 35*).
6. Pellet the beads by centrifuging at 1400 g for 2 min at room temperature. Remove the supernatant leaving roughly 500 μ L. Resuspend the beads in the 500 μ L PBS and transfer back to columns. Use an additional 500 μ L of PBS to transfer any remaining beads to the columns. The same transfer technique described in **Note 34** may be used.
7. To remove nonspecifically bound proteins, wash the beads by adding 700 μ L 1% SDS in PBS, centrifuging at 1000 g for a minute, and discarding the flowthrough. Repeat two more times.
8. To further remove nonspecifically bound proteins, wash the beads by adding 700 μ L of 6 M urea, centrifuging at 1000 g for 1 min, and discarding the flowthrough. Repeat two more times.

3.4.1 Preparation for Western Blotting/Western blotting Following Streptavidin Enrichment

Perform all of the steps for this section at room temperature.

1. Wash the beads to remove excess urea by adding 700 μ L PBS, centrifuging at 1000 g for 1 min, and discarding the flowthrough. Repeat two more times (*see Note 36*).
2. Using two 500 μ L volumes of PBS, transfer the beads to a microcentrifuge tube. Centrifuge at 1400 g for 5 min and then remove PBS using a gel-loading tip (*see Note 37*).
3. Then, using a gel-loading tip elute the protein by adding 50 μ L of 2 $\times$ Laemmli buffer with 30–100 mM DTT and boil at 95 $^{\circ}$ C for 15 min (*see Note 38*).
4. Using a gel-loading tip, extract the 2 $\times$ Laemmli buffer and transfer to a new microcentrifuge tube. The samples are now ready to run on an acrylamide gel for a western blot.

3.4.2 Preparation for LC- MS/MS Following Streptavidin Enrichment

*Perform all of the steps for this section at room temperature. Following the wash steps (see Subheading 3.4.2, **step 3**), perform all remaining steps in a fume hood. Exercise caution to avoid introducing keratin (e.g., via dust or nongloved hands) to the samples as this will limit LC-MS/MS analysis.*

1. Wash the beads of excess urea by adding 700 μL PBS, centrifuging at 1000 g for 1 min, and discarding the flowthrough (*see Note 36*).
2. To remove residual detergent and urea wash, add 700 μL of freshly prepared 50 mM ammonium bicarbonate, centrifuge at 1000 g for 1 min, and discard the flowthrough. Repeat four times.
3. Using two 500 μL volumes of 50 mM ammonium bicarbonate, transfer the beads to a microcentrifuge tube. Using a gel loading tip, remove the ammonium bicarbonate.
4. In preparation for LC-MS/MS, reduce the disulfides for alkylation by adding 500 μL of 10 mM DTT in 50 mM ammonium bicarbonate and incubating at 65 °C for 15 min.
5. Alkylate the samples in 25 mM iodoacetamide by adding 25 μL of 500 mM iodoacetamide. Cover the microcentrifuge tubes in foil (iodoacetamide is photoactive) and rotate end-over-end at room temperature for 30 min.
6. Centrifuge the tubes at 1000 g for 1 min and remove the supernatant with a gel-loading tip.
7. Wash the beads by adding 500 μL of 50 mM triethylammonium bicarbonate, resuspending the beads, and centrifuging at 1000 g for 1 min. Following centrifugation, discard the supernatant. Repeat the wash two times (*see Note 39*).
8. Following the last wash, add 100 μL of 50 mM triethylammonium bicarbonate to resuspend the beads.
9. Add 2 μL of 0.5 mg/mL LC-MS/MS grade trypsin and rotate end-over-end at 37 °C overnight (*see Note 40*).
10. The following morning, centrifuge at 1000 g for 1 min and transfer supernatant (using a gel-loading tip) to a new biospin column in a microcentrifuge tube. Centrifuge at 1000 g for 1 min to remove any residual beads and discard column; the digested peptides will be in the flowthrough in the microcentrifuge tube.
11. Dry the samples through vacuum centrifugation until completely dry (generally 1–1.5 h). At this step, if performing label-free LC-MS/MS, the samples can now be desalted (proceed to Subheading 3.6); however, we most frequently dimethyl label samples first (proceed to Subheading 3.5) (*see Note 41*).

3.5 Dimethyl Labeling of Trypsinized Proteome (Optional)

Perform all of the steps for this section in a fume hood at room temperature or on ice.

1. Reconstitute each of the digested peptides in 100 μL of 100 mM triethylammonium bicarbonate and pipette up and down to mix briefly.

2. To the respective sample, add 4 μL of 4% (v/v) light (CH_2O) formaldehyde to the light sample, medium (CD_2O) formaldehyde to the medium sample, and heavy formaldehyde ($^{13}\text{CD}_2\text{O}$) to the heavy labeled sample.
3. Mix briefly by vortexing or flicking tubes and then centrifuge samples briefly with a low-speed bench-top centrifuge to ensure all the samples are at the bottom of the tube.
4. Add 4 μL of 0.6 M sodium cyanoborohydride (NaBH_3CN) to the light and medium formaldehyde samples. Add 4 μL of 0.6 M sodium cyanoborodeuteride (NaBD_3CN) to the heavy formaldehyde-labeled sample.
5. Mix briefly by flicking the bottom of the tube and then rotate end-over-end on a rotating stand for 1 h at room temperature.
6. Quench the reaction by adding 16 μL of 1% (v/v) ammonia hydroxide. Vortex the samples to mix and centrifuge with a bench-top centrifuge to ensure all the samples are at the bottom of the tube.
7. Begin reacidifying the solution by adding 8 μL of 5% formic acid.
8. For all samples being compared directly (light-, medium-, and heavy-labeled samples), combine into a single microcentrifuge tube (*see* **Note 42**).
9. Add 8 μL 5% formic acid to further quench the reaction. Apply 0.5–1 μL of sample to a litmus pH strip and check pH. If the pH is >4 , repeat by adding 8 μL 5% formic acid and verifying pH. Continue adding 5% formic acid in 8 μL increments until $\text{pH} \leq 4$. Proceed to Subheading 3.6 (*see* **Note 43**).

3.6 C18 Tip Desalting

The C18 tip cleanup protocol is adapted from the manufacturer's protocol outlined by Pierce (Thermo Fisher Scientific).

1. Prepare a new microcentrifuge tube with 100 μL 80% acetonitrile 0.1% formic acid that will be used for elution at the last step.
2. Solvate the C18 tip sorbent with 100 μL 50% acetonitrile and discard solution. Repeat for an additional wash (*see* **Note 44**).
3. Wash C18 tip with 100 μL 0.1% formic acid and discard solution. Repeat for an additional wash.
4. Load samples onto the C18 tip sorbent by pipetting up 100 μL of dimethyl-labeled samples and ejecting into a new microcentrifuge tube. Pipette up and down 5–10 times slowly to ensure binding. Keep flowthrough in case peptide binding to the sorbent is unsuccessful.
5. Repeat **step 4** until all the samples have been loaded onto sorbent.

6. Wash peptides bound to the C18 tip with 100 μ L 0.1% formic acid and discard solution. Repeat for an additional wash.
7. Elute in the tube from **step 1** by placing the C18 tip in the tube containing 100 μ L 80% acetonitrile 0.1% formic acid and pipetting up and down 5–10 times.
8. Dry samples by vacuum centrifuging the samples until completely dry (generally 1–1.5 h) (*see Note 45*).
9. The dimethyl-labeled peptides have now been desalted and are ready for submission to a mass spectrometry facility or for running on LC-MS/MS (*see Note 46*).

4 Notes

1. We typically incubate human-origin cell lines at 37 °C and 5% CO₂. These may be modified to suit the cell origin or specific cell line.
2. Other supplements such as NEAAs are also required for growing Huh7.5 s but are not required for most cell lines. It is important to ensure the correct media supplements are used for the corresponding cell line.
3. The size of dish will depend on the protein requirements and cell type. Generally, the smallest plate we will use for streptavidin enrichment is a 10 cm dish while a 12-well plate is the smallest plate we will use for in-gel fluorescence with FP-TAMRA.
4. We recommend preparing in a concentrated 10 $\times$ stock and filtering. 1 L 1 $\times$ PBS can then be prepared from this and autoclaved prior to use.
5. Any centrifugal column may be used with this so long as it is compatible with the beads, reagents, and centrifugal speeds in the present protocol.
6. Similar streptavidin or avidin resins/beads may be used so long as they are suitable for centrifugal pull downs, have comparable binding capacity (~1–3 mg biotinylated bovine serum albumin (BSA)/mL resin) to suit the concentrations used, and can withstand up to 1400 g.
7. Alternatively, C18 columns or spin-tips or any other desalting method may be used so long as it is compatible with LC-MS/MS.
8. The amount of protein and growth rate in a cell varies greatly from cell line to cell line. It is, therefore, important to have a rough idea of how many cells to seed prior to beginning. For MRC-5 s, to achieve a concentration of 1.5–2 mg/mL we will typically seed 5 million cells in a 15 cm dish while similar concentrations can be achieved with 1.5–2 million Huh7.5 or

A549 cells in a 10 or 15 cm dish. No matter the concentration, avoid letting the cells reach more than 80 or 90% confluency by the time for lysis. It is also advised to prepare a lysate that will later serve as a no probe control. This is to rule out off-target and nonspecific false-positive hits. The transfection of miRNA or siRNA can be performed prior to infection or during the infection period (after the initial incubation) depending on the application. In our examples, the effects of HCV-induced upregulation of miR-27b were assessed, and therefore, the transfection was performed during the infection at **step 8** of Subheading 3.1.

9. The growth factor can be calculated with the formula as follows: $\frac{N_t}{N_0} = e^{\left(\frac{\ln(2)}{t_{\text{doubling}}}\right) t_{\text{elapsed}}}$, whereby N_t is the number of cells at timepoint t , N_0 is the number of cells seeded, t_{doubling} is the doubling time per hour for the cell line (found from manufacturer/supplier or, preferably, established experimentally), and t_{elapsed} is the number of hours since the cells were seeded.
10. The PFU of a virus stock is determined by performing a viral plaque assay [11]. This is a common technique used to quantify any virus that produces viral plaques during a viral life cycle. However, the virus used in the current protocol, HCV is a nonlytic virus and, therefore, does not produce plaques that can be quantified. Instead, an FFU can still be calculated using immunofluorescence for the viral core or other structural protein [12]. Ensure the method used to calculate PFU/FFU is correct and validated for your specific virus. Minimal media indicated here is the normal media used for the cell line but lacking any nutritional supplements. For our purposes with HCV infecting Huh7.5 s, prior to infection we use DMEM media supplemented with 10% FBS and 10 nM NEAAs; therefore, minimal media is DMEM without FBS or NEAAs. It is not essential to use minimal media.
11. The volume of virus-minimal media to add will also be the minimum volume for the well or plate being used. Our group uses the following volumes: 350 μL /well for a 12-well plate, 500 μL /well for a 6-well plate, 2.5 mL per 10 cm dish, and 5 mL per 15 cm dish.
12. The final concentration of miRNA or siRNA is specified by the manufacturer. Check this and optimize if needed; for miR-27b, 100 nM was ideal. It is important to also transfect a control (scramble) miRNA/siRNA in addition to the miRNA/siRNA. In addition, inhibitors (antisense oligos) against the miRNA and siRNA being studied are often useful experimental conditions to include.
13. It is important to add the cells at the correct volume achieving 1/6 opti-MEM or transfection may be impacted. For example,

for a 15 cm dish with 15 mL and 2.5 mL opti-MEM, 2 million cells should be prepared in 12.5 mL.

14. If performing FP-TAMRA labeling with a low-volume well (e.g., 12-well plate), a micropipette can be used. If FP-biotin labeling is being performed, a serological pipette is most sensible given the larger well volume (10 or 15 cm dish).
15. For cells that lift easily such as HEK293Ts, this wash can be skipped to avoid losing too many cells; however, it is possible that the probe will react with extracellular proteases.
16. It is essential to *not* add protease inhibitor to the lysis buffer as this will inhibit several of the serine hydrolases or other enzymes attempting to be labeled. The buffer is chilled on ice prior to use. The specific volume of lysis buffer can vary depending on the size of plate/dish used to ensure that most cells can be collected while ensuring the protein concentration will be greater than the 1–2 mg/mL required for the labeling. To achieve this, we generally use 150 μ L for a 12-well plate, 300 μ L for a 6-well plate, 600 μ L for a 10 cm dish, and 1–2 mL for a 15 cm dish. Keep in mind that for FP-biotin labeling, 1 mL of 1–2 mg/mL lysate is necessary but if a 10 cm dish is used, only 600 μ L will be collected; therefore, it is ideal to have at least 1.7 mg/mL in this 600 μ L.
17. From this point until labeling is completed, it is advised to keep the samples on ice in between steps to avoid protein degradation. This is especially important given that the lysis buffer does not contain protease inhibitor. It is, however, advised to avoid freezing the samples at this stage to avoid impacting enzyme activity with freeze–thaw cycles.
18. With larger plates, the cells should be visible appearing like clumps sinking to the bottom of the microcentrifuge tube. During the sonication, they should be disrupted and mostly disappear. If some remains, after 15 pulse cycles, this is fine, move to the next step; however, most of the original pellet should be gone. Avoid oversonicating as this may harm the proteome.
19. We generally use a 10 μ L aliquot diluted twofold in lysis buffer. We typically perform a DC assay using the kit provided by Bio-Rad and for standards we use lyophilized BSA serially diluted in a range of 0.125 mg/mL to 4 mg/mL.
20. In the current example, we added FP-TAMRA, which should be used at 1 μ M final concentration; therefore, from a 100 μ M FP-TAMRA stock, 0.5 μ L is added.
21. These conditions are for FP ABPs and will likely work with most ABPs; however, some ABPs may be incubated at room temperature or for a different length of time. It is advised to

confirm ideal incubation temperature or time for appropriate labeling.

22. 1 mL of mastermix will be sufficient for nine samples but a larger volume can be made if there are more than nine samples.
23. For maximal click reaction, the samples can be incubated for 2 h; however, longer incubation times will result in higher background and false-positive labeling. In our experience, sufficient biotin-azide and rhodamine-azide labeling occurs after 45 minutes with limited background labeling. Generally, we also cover the samples in foil, especially when working with rhodamine-azide.
24. Samples can be stored at this step (at -80°C) and are stable for days to weeks.
25. It is important to let the sample thaw prior to centrifugation to prevent sample loss. If the sample is not completely thawed, protein may pellet over the ice and then, as the ice thaws following centrifugation, it will disrupt the pellet.
26. Avoid overdrying to ensure the pellet will still resuspend.
27. For mass spectrometry analysis, at least 1.5 mg/mL but ideally 2 mg/mL has led to the best results for our group. For western blotting, 1 mg/mL can be used but we generally opt for 2 mg/mL if the samples are concentrated enough. It is critical, however, that all samples are at the same concentration prior to labeling.
28. In the current example, we added FP-biotin, which should be used at 1–5 μM final concentration; therefore, from a 1 mM FP-biotin stock, 5 μL is added to achieve 5 μM .
29. The pellet may not fully disperse during the first sonication in methanol; however, during the second and third washes, it should disperse well.
30. Samples can also be stored here, although we generally store the samples during the acetone precipitation.
31. The heating and centrifugation are to resolubilize any protein aggregates prior to beginning the FP-biotin streptavidin-agarose pulldown. At this stage, there is generally either a very small pellet or no pellet. However, if the pellet is more than 20% of the original pellet size from the last methanol wash, the samples can be heated briefly (2–5 min) at 90°C and the centrifugation repeated. Though the protein concentrations are normalized before labeling, it is still critical to take the input control here to account for any sample losses prior to the streptavidin-agarose pulldown.
32. The beads will precipitate over time. It is, therefore, important to swirl the bottle before each time an aliquot is taken. Do not

vortex the beads or centrifuge at speeds beyond 1400 g or the manufacturer-specified limit to avoid sheering.

33. It is important to limit the time that the beads are dry throughout the pulldown. We recommend leaving them suspended in PBS for the last step and centrifuging only when ready to proceed to the pulldown.
34. It is not necessary to use two 500 μ L volumes of PBS to transfer the beads; however, the final volume should be 9 mL for appropriate bead concentration. We often use two 500 μ L volumes to minimize sample/bead lost during transfer steps. The following method/steps can be done for extra caution: (i) use pipette #1 to add 500 μ L PBS to the column; (ii) pipette up and down and then transfer the beads slowly to the 15 mL conical tube; (iii) set aside pipette #1 with the tip still attached; (iv) using pipette #2, add 500 μ L PBS to the column; and (v) using pipette #1 again, aspirate the column PBS and add to 15 mL conical tube. This achieves a transfer and a wash of the column and pipette tip to transfer any residual sample/beads.
35. During the second half of the incubation is when we would advise preparing any reagents needed “fresh” for the subsequent washing steps.
36. We will generally leave the beads suspended in the PBS when preparing any further reagents to limit the amount of time they are dry; then, when it is time to proceed we will centrifuge and remove the flowthrough to avoid leaving the beads dry for very long.
37. Theoretically, a transfer step could be skipped here and the elution of proteins done in the column; however, in our experience it tends to be better to transfer and elute in a microcentrifuge tube. Therefore, we use a gel loading tip to remove all PBS as the tip is narrow enough to extract PBS while losing minimal beads.
38. It is not essential to use a gel-loading tip here, but in our experience, this tends to reduce bead losses in the pipette tip when resuspending the beads in Laemmli buffer. The DTT should be prepared fresh in a 1 M aliquot and then added to the 2 $\times$ Laemmli buffer purchased commercially or prepared in-house. We opt for the inexpensive commercial 2 $\times$ Laemmli buffer from Bio-Rad for convenience. After boiling for 15 min at 95 $^{\circ}$ C, a bench-top mini centrifuge spin for 3–5 s can be useful to collect any sample that has condensed on the microcentrifuge cap and sides.
39. When resuspending the beads in the ammonium bicarbonate, pipetting up and down can lead to some beads being lost in the

pipette tip. When resuspending, do so with care to minimize bead loss.

40. For best results, it is best to start the overnight trypsin digestion in the late afternoon or evening.
41. Dimethyl labeling is convenient if comparing two experimental conditions or a maximum of three. If two experimental samples are being compared, it is best to medium (+28 Da) label a no probe or negative control condition, allowing improved resolution between light and heavy conditions. If more than three samples are being compared, alternative labeling methods such as TMT (Thermo Fisher Scientific) or iTRAQ (Millipore-Sigma) are an option.
42. At this stage, the samples will have been dimethyl labeled with light, medium, and heavy chains.
43. Generally, this takes between 24 μ L and 48 μ L of 5% formic acid to quench three 100 μ L dimethyl-labeled samples. The exact amount can vary with quality of 1% ammonia and the initial quenching. For the desalting, there are different techniques that can be applied but we find the C18 tips to be the most effective.
44. It is not necessary to pipette up and down at this step; simply aspirate and eject the solution. Per the manufacturer, it is critical that the sorbent remains solvated/wet; therefore, avoid ejecting completely when discarding the solution and leave a very small amount of solution to maintain hydration in the sorbent. We also advise completing the desalting steps in relatively quick succession to limit any time the tip/sorbent may be dry.
45. After drying, it is normal for the tube to appear empty. At most, there may be a small amount of pellet at the bottom of the tube.
46. Following desalting, the samples can be vacuum eluted and run either by the user or a local mass spectrometry facility. These steps can vary depending on available machinery and are not unique to the application of ABPP for virus-host interactions; therefore, we suggest the methodology employed by our mass spectrometry facility [6, 7], but note this is by no means comprehensive for all applications. To analyze the data following LC-MS/MS, our mass spectrometry facility generally processes the raw data using Proteome Discover (Thermo Fisher Scientific) and filters the MS2 spectra using SEQUEST HT engine against the UniProt/SwissProt *Homo sapiens* database.

References

1. Speers AE, Cravatt BF (2009) Activity-based protein profiling (ABPP) and click chemistry (CC)-ABPP by MudPIT Mass Spectrometry. *Curr Protoc Chem Biol* 1:29–41. <https://doi.org/10.1002/9780470559277.ch090138>
2. Niphakis MJ, Cravatt BF (2014) Enzyme inhibitor discovery by activity-based protein profiling. *Annu Rev Biochem* 83:341–377. <https://doi.org/10.1146/annurev-biochem-060713-035708>
3. Liu Y, Patricelli MP, Cravatt BF (1999) Activity-based protein profiling: The serine hydrolases. *Proc Natl Acad Sci USA* 96:14694–14699. <https://doi.org/10.1073/pnas.96.26.14694>
4. Long JZ, Cravatt BF (2011) The metabolic serine hydrolases and their functions in mammalian physiology and disease. *Chem Rev* 111:6022–6063. <https://doi.org/10.1021/cr200075y>
5. Nasheri N, Joyce M, Rouleau Y et al (2013) Modulation of fatty acid synthase enzyme activity and expression during hepatitis C virus replication. *Chem Biol* 20:570–582. <https://doi.org/10.1016/j.chembiol.2013.03.014>
6. Desrochers GF, Filip R, Bastianelli M et al (2022) microRNA-27b regulates hepatic lipase enzyme LIPC and reduces triglyceride degradation during hepatitis C virus infection. *J Biol Chem* 298:101983. <https://doi.org/10.1016/j.jbc.2022.101983>
7. Filip R, Desrochers GF, Lefebvre DM et al (2021) Profiling of microRNA targets using activity-based protein profiling: linking enzyme activity to microRNA-185 function. *Cell Chem Biol* 28:202–212. <https://doi.org/10.1016/j.chembiol.2020.12.009>
8. Evers P, Pezacki JP (2023) Unraveling complex microRNA signaling pathways with activity-based protein profiling to guide therapeutic discovery. *Isr J Chem:e202200088*. <https://doi.org/10.1002/ijch.202200088>
9. Desrochers GF, Sherratt AR, Blais DR et al (2015) Profiling kinase activity during hepatitis C virus replication using wortmannin probe. *Acs Infect Dis* 1:443–452. <https://doi.org/10.1021/acsinfecdis.5b00083>
10. Speers AE, Cravatt BF (2004) Profiling enzyme activities in vivo using click chemistry methods. *Chem Biol* 11:535–546. <https://doi.org/10.1016/j.chembiol.2004.03.012>
11. Juarez D, Long KC, Aguilar P et al (2013) Assessment of plaque assay methods for alphaviruses. *J Virol Methods* 187:185–189. <https://doi.org/10.1016/j.jviromet.2012.09.026>
12. Zhong J, Gastaminza P, Cheng GC et al (2005) Robust hepatitis C virus infection in vitro. *Proc Natl Acad Sci USA* 102:9294–9299. <https://doi.org/10.1073/pnas.0503596102>



Activity-Based Protein Profiling of Active Host Cell Serine Hydrolases for Bioprocess Development of Therapeutic Proteins and Vaccines

Xuanwen Li, Taku Tsukidate, Alyssa Q. Stiving, and Hillary A. Schuessler

Abstract

High-risk host cell serine hydrolases and other process residues have been shown to have a significant impact on the quality and safety of biologics and vaccines. In order to monitor and control process-related enzymes, particularly those present at trace levels during bioprocess development, an optimized activity-based protein profiling (ABPP) method has been developed. This method enables the identification and quantification of their activities within a single assay. This chapter provides a detailed protocol for ABPP and presents typical results obtained from the analysis of process intermediates derived from a live viral vaccine candidate.

Key words ABPP, Bioprocess development, Excipient stability, Product stability, LC-MS

1 Introduction

Biologics, such as therapeutic proteins, and vaccines, such as recombinant subunit vaccines or live viral vector vaccines, are biopharmaceuticals that are often produced through cell culture processes. The upstream cell culture process involves using living cells, such as mammalian cells or microbial cells, to produce the desired therapeutic proteins or antigens. The proteins or antigens produced by the cultured cells are then purified in downstream process to remove impurities and ensure product safety and quality [1]. The purification process typically involves multiple steps, such as filtration, chromatography, and ultrafiltration. Even with multiple purification processes employed during the production of biologics and vaccines, it is still possible for process impurities, such as host cell proteins (HCPs), to be present in the final drug substance [1]. It is challenging to completely eliminate all HCPs due to their similarities in physicochemical properties or direct interactions to the target proteins or purification resins [2].

HCPs are proteins derived from the cells used as the production system, and their presence in the final product can have potential safety and efficacy implications. Beyond the potential direct biological activity and impact on immunogenicity, HCP can have significant adverse effects on the stability of both product and excipients in the drug formulation [3–5]. For example, a residual host cell protein hexosaminidase B was identified to be associated with N-glycan degradation of a therapeutic recombinant monoclonal antibody product in a stability study [4]. Moreover, even when present in very small amounts, a group of enzymes called serine hydrolases, which includes lipases and esterases, have been found to degrade formulation excipient polysorbate 80 or 20 [3, 5]. Another class of serine protease from the complement cascade complement Component 1's (C1s) was reported to be responsible for the proteolysis of HIV envelope protein gp120 used for HIV vaccine development [6]. To address the potential risks associated with residual HCPs, regulatory authorities, such as the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA), provide guidelines for HCPs measurement and control in biopharmaceuticals. These guidelines require the biopharmaceutical industry to establish sensitive analytical methods for the detection and quantification of HCPs to ensure product quality and safety. The readout from analytical tests can help drive the process efforts, such as optimizing upstream cell culture conditions to reduce HCP expression, and developing more selective purification techniques, to refine purification strategies and improve the removal of HCPs from biologics and vaccines.

Analytical techniques used for HCP analysis include enzyme-linked immunosorbent assays (ELISAs) and liquid chromatography-mass spectrometry (LC-MS)-based proteomics approach. ELISA is typically only able to detect certain subsets of HCPs based on the availability of specific antibodies, which may not cover all potential HCPs present. LC-MS offers a broader coverage of HCPs as it can detect and quantify a wide range of proteins without prior knowledge of their identity. In addition, LC-MS can identify and quantify specific HCPs, allowing manufacturers to monitor and control their levels during the manufacturing process [1]. The challenges associated with LC-MS-based proteomics approach include (1) the identification of low-abundant HCPs due to the high dynamic range of trace-level HCPs compared to high-abundant drug substance, and (2) traditional abundance-based proteomics does not measure the activity of identified host cell enzymes. To fill both gaps, we have established an activity-based protein profiling (ABPP) approach using LC-MS for active serine hydrolases identification to improve the understanding of polysorbate degradation in biological formulations (Fig. 1) [5]. ABPP is a powerful technique used to study the function and activity of enzymes in complex biological systems. ABPP workflow

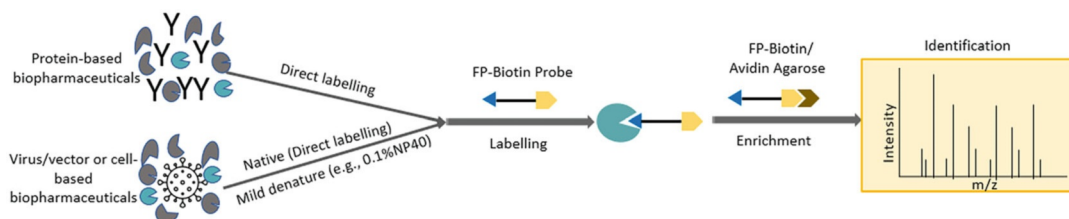


Fig. 1 The scheme diagram of the ABPP workflow for serine hydrolases from protein, virus/vector, or cell-based biopharmaceuticals. Host cell proteins in native status, or proteins optionally extracted from virus/vector or cell-based biopharmaceuticals are incubated with synthesized chemical probes. The active enzymes captured by the chemical probes are specifically enriched by an affinity tag. The enriched proteins can be identified by LC-MS/MS. Note: The specific details of the workflow may vary depending on the experimental design, the nature of the probes used, and the goals of the study.

involves the design and utilization of active site-directed probes that covalently bind to the active site of enzymes or enzyme families, which has been adapted for activity detection of more than a dozen enzyme classes, including various proteases, kinases, phosphatases, glycosidases, and oxidoreductases [7].

This chapter focuses on the detailed description of the ABPP workflow for the identification of active host cell serine hydrolases and other process-related residues with serine protease activity, such as trypsin (a commonly used reagent for certain types of vaccine production), from bioprocess intermediates derived from a live viral vector vaccine candidate (Fig. 1). This project aims to shed light on the potential role of these hydrolases in the degradation of viral structure proteins observed during stability studies (data not shown). It is important to note that the described ABPP workflow has been successfully applied not only in the investigation of viral vector vaccine candidates but also in supporting bioprocess development for various therapeutic proteins and vaccines (Fig. 1).

2 Materials

The specific experimental protocol may have additional safety instructions or requirements that should be followed according to local institutional guidelines specific to your facility for potent compounds and biosafety materials being handled to ensure compliance with the appropriate safety practices.

2.1 Sample (See Note 1)

1. For demonstration purposes, this protocol utilized a harvested virus fluid (HVF) sample obtained from a live recombinant viral vector vaccine candidate called LVV1. The LVV1 vaccine candidate was produced by harvesting it from transfected Vero cells. The Vero cell lineage was isolated from kidney epithelial cells extracted from an African green monkey.

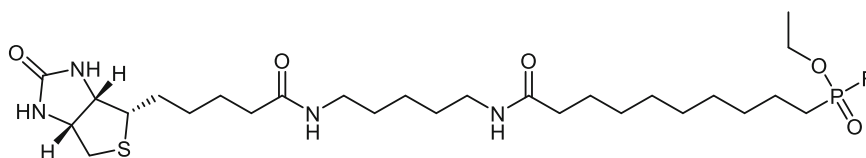


Fig. 2 The chemical structure of the fluorophosphonate (FP)-biotin probe

2.2 Chemical Probes (See Note 2)

1. Dimethyl sulfoxide (DMSO, sterile-filtered, purity $\geq 99.7\%$).
2. Fluorophosphonate (FP)-biotin probe (Fig. 2) was synthesized by Wuxi AppTec (*see* Note 3).
3. Dissolve FP-biotin probe in DMSO to make a 0.1 mM stock solution (*see* Note 4).

2.3 Sample Labelling and Purification

2. 1 M Tris-HCl (pH 8.0).
3. NP-40, Protein Grade® detergent, 10% solution, sterile-filtered.
4. Methanol: acetone (50:50), kept at 4 °C.
5. Pierce™ high-capacity streptavidin agarose.
6. 8 M urea: 1 viral of pre-weighted urea after reconstitution with 16 mL high-purity water (*see* Note 5).
7. SDS solution (20%).
8. Sample suspension buffer: 2 M urea and 1%SDS, 50 mM Tris-HCl (pH 8).
9. Sample washing buffer: 2 M urea/50 mM Tris-HCl (pH 8).
10. Tube revolver rotator.

2.4 Protein Digestion

1. Trypsin: 0.5 $\mu\text{g}/\mu\text{L}$ in 50 mM Tris-HCl (pH 8).
2. Pierce™ DTT (Dithiothreitol), No-Weigh™ Format.

2.5 LC-MS

1. Nano LC system, such as EASY-nLC™ 1200 System (Thermo Fisher) or Evosep One System (Evosep).
2. Mass spectrometry, such as Q Exactive HF-X (Thermo Fisher) or Orbitrap Exploris™ 480 (Thermo Fisher) or TimsTOF (Bruker).

2.6 Database Search and Analysis

1. Proteome Discoverer 2.5 (Thermo Fisher).

3 Methods

Carry out all procedures at room temperature unless otherwise specified.

3.1 Sample Labelling and Purification

1. Prepare four Eppendorf tubes (1.5 mL), proceed with **steps 2–7** in a biosafety cabinet as LVV1 was classified as Biosafety Level 2 (BSL2) material.
2. Transfer 500 μ L of HVF from LVV1 to each Eppendorf tube (*see Note 1*).
3. Add 25 μ L of 1 M Tris (pH 8) to each tube to adjust the pH (*see Note 6*).
4. Add NP-40 into two tubes to a final concentration of 0.1% and incubate for 30 min with vortex for every 10 min (*see Note 7*).
5. Add 20 μ L of FP-biotin probe (100 μ M stocking concentration) for a final concentration of 4 μ M in two of the tubes (*see Note 8*), or 20 μ L of control DMSO to each of the rest two tubes. There were four samples in this experiment: sample 1 (HVF + 0.1%NP40 + FP-biotin), sample 2 (HVF + FP-biotin), sample 3 (HVF + 0.1%NP40 + DMSO), and sample 4 (HVF + DMSO) (*see Note 9*).
6. The samples were incubated at room temperature for 2 h with constant mixing on a rotator (*see Note 10*).
7. Add 1 mL of ice-cold methanol: acetone (50:50) to each sample on ice to remove the excess FP-biotin probes. Mix the sample three times using a pipette. The samples were incubated on ice for approximately 30 min (*see Note 11*).
8. The samples were centrifuged at 16,000 g for 10 min at 4 °C to pellet the precipitated protein (*see Note 12*).
9. The samples were washed with 1 mL of methanol: acetone (50:50) on ice. The methanol:acetone mixture was added to each pelleted sample, vortexed to mix and wash, and then the samples were centrifuged at 16,000 g for 10 min at 4 °C to re-pellet the precipitated protein (*see Note 13*).
10. Most of the supernatant was removed and discarded. The pellet was dried using N₂ gas (*see Note 14*).
11. A 1 mL of sample suspension buffer (2 M urea, 1% SDS, 50 mM Tris, pH 8) was added to each sample to re-solubilize the pellets. The suspension buffer was pipetted up and down several times, ensuring the entire pellet was resolubilized (*see Note 15*).
12. A total 60 μ L of agarose beads were added to each sample and incubated for 1 h at room temperature with constant mixing on a rotator (*see Note 16*).
13. The samples were centrifuged at 1000 g for 1 min. The supernatant was removed and discarded (*see Note 17*).
14. A total of 500 μ L of sample washing buffer (2 M urea/50 mM Tris, pH 8) was added to each sample and vortexed briefly to

mix. The samples were centrifuged at 1000 g for 1 min, and the supernatant was removed and discarded. This washing step was repeated for a total of three times.

15. The beads were washed with 500 μL of 50 mM Tris. This was repeated two times.

3.2 Protein Digestion

1. The beads were resuspended in 50 μL of 50 mM Tris-HCl (pH 8) (*see Note 18*).
2. A total of 2 μL (1 μg) of trypsin enzyme was added to each sample (*see Note 19*).
3. The samples were incubated overnight (16–18 h) at 37 °C with gentle mixing (500 RPM) (*see Note 20*).
4. The beads were centrifuged at 1000 g for 1 min. The supernatant was removed and transferred to a fresh collection tube for each sample.
5. A total of 50 μL of water was added to each tube containing the beads and flicked gently to wash the beads. The beads were centrifuged at 1000 g for 1 min. The supernatant was removed and combined with the respective tube from the previous step.
6. A total of 1 μL of 1 M DTT was added to each sample of the combined supernatants. The samples were incubated at 80 °C for 10 min (*see Note 21*).
7. The samples were cooled down to room temperature and then placed at –70 °C for 10 min to cool the samples.
8. The samples were dried using a SpeedVac until approximately 10 μL remained (*see Note 22*).

3.3 Evotip Preparation and LC Separation

1. The Evotip preparation was done according to the manufacturer's protocol for sample loading as shown below.
2. Washing Evotips: the Evotips were first washed with 20 μL of buffer B (acetonitrile with 0.1% formic acid) through centrifugation at 700 g for 1 min.
3. Conditioning Evotips: the Evotips were immersed in isopropanol for 20 seconds and subsequently equilibrated with 20 μL of buffer A (water with 0.1% formic acid) through centrifugation (*see Note 23*).
4. Loading samples: a volume of 20 μL of sample was loaded onto the Evotips through centrifugation.
5. Washing the tips: the Evotips were then spun and washed with 20 μL of buffer A.
6. Preservation: the Evotips was covered with 100 μL of buffer A, and a brief 10-second spin was performed.

7. LC separation: the peptides were separated on an Endurance analytical C18 column (EV-1106, 1.9 μm beads, 115 cm $\times$ 150 μm ID) using the pre-set gradient of 30 samples per day on the Evosep One system (*see* **Note 24**).

3.4 MS Setup

1. Mass spectrometry was run with an Orbitrap Exploris™ 480 (Thermo Fisher). The data-dependent acquisition (DDA) was performed with the following parameters.
2. Ion source:
 - (a) Positive ion (V): 1800
 - (b) Sweep gas (Arb): 2
 - (c) Ion transfer tube temp ($^{\circ}\text{C}$): 300
3. Full scan:
 - (a) Orbitrap resolution: set at 60,000
 - (b) Scan range: 350–1800 m/z
 - (c) RF Lens (%): 30
 - (d) Normalized AGC target (%): 300
 - (e) Maximum inject time: 100
 - (f) Microscan: 1
 - (g) Polarity: positive
4. dd MS2 parameters:
 - (a) Precursor fragmentation mode: high-energy collision dissociation (HCD)
 - (b) Isolation window (m/z): 1.2
 - (c) Normalized collision energy: 27%
 - (d) Orbitrap resolution: 15,000
 - (e) Maximum injection time (ms): 200
5. Data-dependent scan:
 - (a) Number of dependent scans: 10
 - (b) Precursor charge state: 2–7
 - (c) Intensity threshold: 9.0e3
6. Dynamic exclusion:
 - (a) Exclude after n times: 2
 - (b) If occurs with (s): 15
 - (c) Exclusion duration (s): 20
 - (d) Mass tolerance: 5 ppm

3.5 Database Search and Analysis

1. DDA data were analyzed using Proteome Discoverer 2.4 (*see Note 25*). SEQUEST HT was used to search against the UniProt *Chlorocebus sabaenus* proteome.
2. Search parameters included a precursor mass tolerance of 10 ppm, a fragment mass tolerance of 0.02 Da, and trypsin digestion with up to two missed cleavages allowed. Oxidation (M) was set as a dynamic modification. Percolator was used to filter the results to a 1% false discovery rate at the peptide level using the target-decoy strategy. Only protein identified with two or more unique peptides were considered for further data analysis (*see Note 26*). Label-free quantification was performed using “Minora Feature Detector,” “Feature Mapper,” and “Precursor Ions Quantifier” nodes as recommended. Minora Feature Detector detects chromatographic peaks and associate them with peptide spectral matches. Feature Mapper performs retention time alignment, which helps link features across all raw files. For label-free quantification of precursor ions, the Precursor Ions Quantifier tool was utilized.
3. Protein relative abundance was estimated using the summed MS1 peak area of all identified peptides. In order to identify potential active serine hydrolases, only proteins that exhibited at least a fourfold enrichment in FP-biotin-labeled samples compared to the corresponding controls with DMSO incubation were considered (*see Note 27*).
4. As shown in Table 1 and Fig. 3, a total of 31 potential serine hydrolases were identified from an LVV1 HVF sample. Interestingly, only eight (26%) of these proteins were detected using the abundance-based traditional proteomics workflow [a total of 1004 proteins were identified with at least two unique peptides (data not shown)]. This data highlights the distinct advantage of the ABPP workflow in identifying low-abundance HCPs and providing insights into their enzymatic activities. Moreover, relatively higher levels of serine hydrolases were observed in the sample pretreated with NP-40 (Fig. 4), indicating some serine hydrolases may be potentially incorporated/encapsulated in this live viral vector. Furthermore, it is noteworthy that several of these HCPs were also identified in the final purified drug substances of LVV1 (data not shown).
5. These findings are significant as they offer a potential explanation for the stability issues observed in the structure proteins of LVV1. The identification of HCPs, especially those high-risk serine hydrolases, suggests their potential involvement in protein degradation and highlights the importance of clearing these HCPs in order to improve the stability and quality of the LVV1 vaccine.

Table 1
List of potential serine hydrolases identified by ABPP from Vero cells used to product LW1

Accession ^a	Gene ^a	Protein description ^a	Unique peptides ^b	Sample 1 ^c	Sample 2 ^c	Sample 3 ^c	Sample 4 ^c	Sample 1/3 ^d	Sample 2/4 ^d	Sample 1/2 ^d	Abundance ranking ^e
A0A0D9RQG8	PREP	Post-proline cleaving enzyme	29	7.95E +09	6.02E +09	2.07E +08	3.51E +08	38.35	17.15	1.32	947
A0A0D9RYT5	ESD	S-formylglutathione hydrolase	9	5.36E +09	3.06E +09	1.11E +08	3.67E +07	48.37	83.38	1.75	ND
A0A0D9RSY5	DPP4	Dipeptidyl peptidase-4	21	5.10E +09	3.50E +09	ND	ND	ND	ND	1.46	ND
A0A0D9RLJ3	PREPL	Prolyl endopeptidase-like	22	4.57E +09	2.92E +09	ND	ND	ND	ND	1.56	ND
A0A0D9SL50	PAFAH1B3	SGNH_hydro domain-containing protein	9	4.31E +09	3.58E +09	ND	1.39E +07	ND	258.29	1.20	667
A0A0D9S009	ABHD11	AB hydrolase-1 domain-containing protein	8	4.21E +09	2.62E +09	6.61E +06	ND	636.80	ND	1.61	ND
A0A0D9RQ15	CTSA	Carboxypeptidase	5	4.01E +09	3.09E +09	ND	ND	ND	ND	1.30	838
A0A0D9R2L6	ABHD10	AB hydrolase-1 domain-containing protein	13	2.93E +09	2.75E +09	ND	ND	ND	ND	1.07	ND
A0A0D9RK62	NCEH1	Neutral cholesterol ester hydrolase 1	9	2.64E +09	2.66E +09	ND	ND	ND	ND	0.99	ND
A0A0D9S470	SIAE	SASA domain-containing protein	10	2.30E +09	1.48E +09	ND	ND	ND	ND	1.55	ND
A0A0D9RR27	APEH	Acyl-peptide hydrolase	10	2.29E +09	1.97E +09	4.77E +08	1.26E +07	4.81	155.87	1.16	ND
A0A0D9RDL3	DPP9	Dipeptidyl peptidase-9	13	1.32E +09	1.30E +09	ND	ND	ND	ND	1.01	ND

(continued)

Table 1
(continued)

Accession ^a	Gene ^a	Protein description ^a	Unique peptides ^b	Sample 1 ^c	Sample 2 ^c	Sample 3 ^c	Sample 4 ^c	Sample 1/3 ^d	Sample 2/4 ^d	Sample 1/2 ^d	Abundance ranking ^e
A0A0D9RGE5	RBBP9	RB Binding Protein 9, Serine Hydrolase	6	9.64E+08	5.88E+08	ND	ND	ND	ND	1.64	ND
A0A0D9RMX3	OVCA2	Esterase	6	9.05E+08	7.56E+08	ND	ND	ND	ND	1.20	ND
A0A0D9S887	LYPLA2	Abhydrolase_2 domain-containing protein	5	8.43E+08	5.96E+08	2.69E+07	1.99E+07	31.29	29.98	1.41	ND
A0A0D9S859	PAFAH2	Platelet-activating factor acetylhydrolase	9	6.90E+08	4.53E+08	6.01E+06	4.63E+06	114.82	97.86	1.52	ND
A0A0D9QX48	HTRAI	HtrA Serine Peptidase 1	7	6.43E+08	2.17E+08	1.04E+06	ND	619.74	ND	2.96	701
A0A0D9RN32	AKRIC3	Aldo-keto reductase family 1 member C3	7	4.18E+08	3.60E+08	8.90E+06	ND	46.98	ND	1.16	380
A0A0D9QVP0	PLA2G15	Uncharacterized protein	4	2.90E+08	1.50E+08	7.53E+06	ND	38.59	ND	1.94	ND
A0A0D9QUP5	PRCP	Prolylcarboxypeptidase	3	2.79E+08	3.89E+08	ND	ND	ND	ND	0.72	ND
A0A0D9S3N9	DDX6	RNA helicase	2	2.74E+08	3.43E+08	1.44E+07	8.36E+06	19.06	41.07	0.80	712
A0A0D9QX46	CEHs	Carboxylic ester hydrolase	3	2.00E+08	1.59E+08	ND	ND	ND	ND	1.26	ND
A0A0D9RT88	LYPLAL1	Lysophospholipase-like protein 1	3	1.96E+08	6.54E+07	ND	ND	ND	ND	3.00	ND
A0A0D9S3B0	PAFAH1B2	SGNH_hydro domain-containing protein	2	1.82E+08	3.07E+08	ND	ND	ND	ND	0.59	922

A0A0D9S173	LIPE	Hormone-sensitive lipase	3	1.57E +08	1.19E +08	4.10E +06	3.70E +07	38.27	3.22	1.32	ND
A0A0D9RU97	PSMB5	Proteasome subunit beta	2	1.50E +08	9.68E +07	ND	ND	ND	ND	1.55	285
A0A0D9RA68	PPT2	Palmitoyl-Protein Thioesterase 2	2	1.44E +08	3.27E +07	ND	ND	ND	ND	4.42	ND
A0A0D9RL12	GSTA4	Glutathione transferase	2	1.36E +08	6.50E +07	ND	ND	ND	ND	2.09	ND
A0A0D9RIY2	BCHE	Carboxylic ester hydrolase	3	1.13E +08	3.56E +07	ND	ND	ND	ND	3.18	ND
A0A0D9S448	AFM1D	Kynurenine formamidase	2	4.96E +07	2.99E +07	ND	ND	ND	ND	1.66	ND
A0A0D9S4V3	SLC16A3	MFS domain-containing protein	4	1.43E +07	4.25E +08	ND	ND	ND	ND	0.03	ND

ND not detected

^aIdentified potential serine hydrolases with accession ID, gene name, and protein description from UniProt *Chlorocibus sabaeus* proteome

^bUnique peptides used to identify each protein after database search

^cThe total peak area for each identified protein after database search used for label-free quantification; sample 1: HVF+0.1%NP40+FP-biotin; sample 2: HVF+FP-biotin; sample 3: HVF+0.1%NP40+DMSO; and sample 4 HVF +DMSO

^dThe ratio of peak areas between two different samples

^eThe abundance ranking of potential serine hydrolases identified from a traditional abundance-based proteomics where the same HVF sample was denatured, reduced, and alkylated before trypsin digestion and LC-MS. A total of 1,004 unique HCPs (>=2 unique peptides) were identified

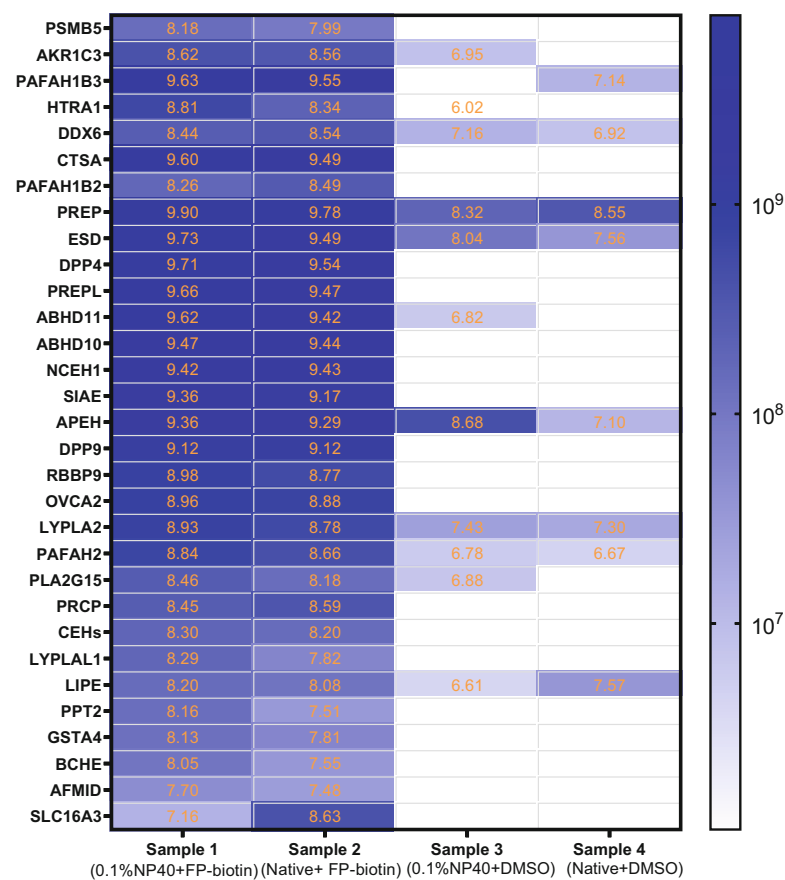


Fig. 3 The typical results obtained from the ABPP assay are visualized using a heatmap. The peak area of the identified proteins was transformed with Log10

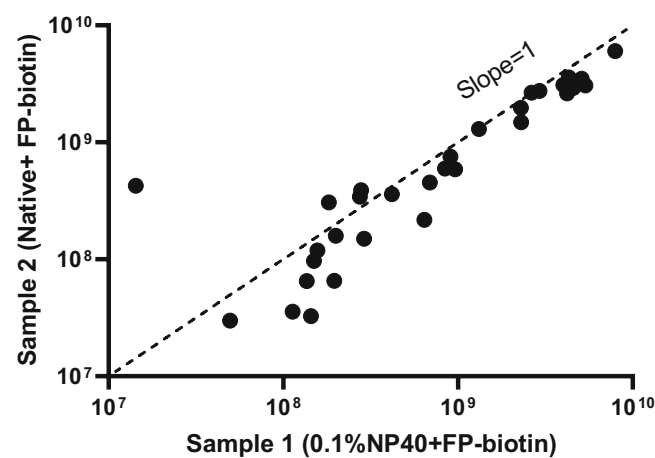


Fig. 4 The comparison of the identified serine hydrolases with or without mild denature using 0.1% NP40. The dotted line shows the line of identity (slope = 1)

4 Notes

1. Sample selection: Various process intermediates used to produce biologics or vaccines, ranging from harvested cell culture fluid (HCCF) to final drug substance (DS), have been successfully tested by this protocol (Fig. 1). Depending on the estimated serine hydrolases concentrations in a particular sample, the initial sample starting materials can be from 500 μg for a less purified sample (e.g., HCCF) to 5 mg or even more for a most purified sample (e.g., DS). Considering the protein precipitation step using a 1:2 ratio (V/V) of methanol/acetone in a 1.5-mL Eppendorf tube, the recommended starting volume is 500 μL . If the initial sample volume is less than 500 μL , it can be diluted with 50 mM Tris-HCl (pH 8) to reach a final volume of 500 μL . On the other hand, if the initial sample volume exceeds 500 μL , it can be concentrated to 500 μL using a SpeedVac machine or a sample desalting/concentration column.
2. Chemical probes: In this protocol, a range of chemical probes can be utilized to label various enzyme classes, each offering different specificity and coverage [5, 8]. A chemical probe without enrichment tag or with different enrichment tag other than biotin as used in this protocol can be employed to design a competitive ABPP experiment to confirm the specificity of the identified protein as active enzymes [5]. In an ideal scenario, the protein abundances of the identified active enzymes would ideally decrease if the samples were pre-incubated with a chemical probe targeting the same class of enzymes, but without an enrichment tag or with a different enrichment tag other than biotin.
3. FP-biotin can be readily obtained commercially from various sources.
4. FP-biotin can be prepared as a stock solution with a concentration of 0.1 mM or higher. It is recommended to aliquot the solution and store it in a $-70\text{ }^{\circ}\text{C}$ freezer. When stored under these conditions, the FP-biotin solution should maintain good quality for several years.
5. Always use freshly prepared urea-based reagents in this protocol.
6. Taking into consideration the diverse pH optima of various serine hydrolases and wide range of buffer pHs for in-process biopharmaceutical samples (usually between pH 3.0 and 8.0) [9], it is conceivable that the enzymatic activity of certain serine hydrolases might be increased to improve the limit of detection of the ABPP assay by adjusting the buffer pH to the respective pH optimum, but this may compromise the enzymatic activity of other hydrolases [10].

7. NP-40 was added to break down the potential interaction of HCPs with LVV1 viral vector or extract HCPs incorporated/encapsulated into such vectors [11].
8. The optimal final concentration of FP-biotin may vary and should be optimized for specific experimental conditions. Typically, a concentration ranging from 1 to 10 μM is effective. However, it is important to note that excessive probe concentrations can potentially reduce the purification efficiency of streptavidin beads.
9. In order to identify proteins that specifically interact with the chemical probes, it is crucial to include a control sample. In this protocol, a control sample treated with DMSO was used.
10. The incubation time for the protocol can be optimized. Typically, an incubation period of 0.5–2 h is sufficient.
11. To remove excess chemical probes, there are multiple methods available, including the use of desalting or buffer exchange columns, as well as protein precipitation, as suggested in this protocol. When using organic solutions for protein precipitation, it is crucial to keep the samples chilled on ice. Failure to do so may make it difficult to dissolve the protein pellet in subsequent steps.
12. Before placing the samples in the centrifuge, it is crucial to ensure that the centrifuge has been cooled down to 4 °C.
13. When removing the solution from the tube, it is essential to avoid touching the pellet. Furthermore, it is critical to fully dissolve the pellet to ensure the complete removal of any excess chemical probes.
14. Before subjecting the samples to N_2 gas or using a SpeedVac to dry any remaining organic solution, it is crucial to remove most of the residual solution.
15. To ensure that all labeled proteins can be effectively captured and reduce nonspecific adsorption of proteins on the agarose resin during subsequent purification steps, it is crucial to fully dissolve the pellet.
16. The optimal amount of agarose beads and the incubation time may be subject to optimization.
17. When removing the supernatant, it is critical to avoid touching and disrupting the agarose beads.
18. **Steps 13 and 14** of the protocol were specifically designed to remove any residual SDS or urea as these compounds can potentially impact protein digestion or MS performance.
19. For protein digestion of the enriched proteins, there are multiple methods available. In this specific protocol, an on-beads

digestion method was employed. It is important to note that the amount of trypsin enzyme used for digestion may be excessive for the limited enriched proteins.

20. To prevent sample loss, it is recommended to collect all the supernatants during the experimental procedure.
21. DTT was added to the experimental procedure to facilitate the breakdown of disulfide bond-linked peptides.
22. To prevent potential sample loss, it is advisable to avoid completely drying the digested peptide during the experimental procedure. Alternatively, this step can be avoided; instead up to 300 uL of samples can be loaded directly on Evotip.
23. Make sure all tips are pale white after soaking. This step might take up 30–60 seconds.
24. Considering that the enriched sample is relatively simple, it may not necessarily increase the number of protein identifications by changing the LC method from 30 SPD to 15 SPD.
25. This protocol can be effectively executed using an early version of Proteome Discoverer or any other protein identification search engine. These tools are designed to facilitate the protein identification and relative quantification of proteomic data.
26. In order to enhance the confidence of protein identification at the proteome level, we employed a strategy of using two or more unique peptides. However, it is worth noting that this approach may inadvertently overlook certain promising serine hydrolases that are identified based on only one unique peptide. Therefore, particularly in the case of highly purified process intermediates, it becomes crucial to thoroughly examine the proteins that are identified with only one unique peptide. By considering proteins with only one unique peptide, we can mitigate the risk of disregarding potentially important serine hydrolases that could play significant roles in the process under investigation.
27. This stringent criterion was employed to ensure robust identification of active serine hydrolases.

Acknowledgments

We thank our colleagues in Process R&D for providing us with material and scientific support. We also gratefully thank our colleagues Dr. Gregory C Adam and Dr. Smaranda Bodea for the technical input during ABPP method development, and Douglas Richardson for the support.

References

1. Li X (2023) Recent applications of quantitative mass spectrometry in biopharmaceutical process development and manufacturing. *J Pharm Biomed Anal* 234:115581
2. Oh YH, Becker ML, Mendola KM, Choe LH, Min L, Lee KH, Yigzaw Y, Seay A, Bill J, Li X, Roush DJ, Cramer SM, Menegatti S, Lenhoff AM (2024) Factors affecting product association as a mechanism of host-cell protein persistence in bioprocessing. *Biotechnol Bioeng* 121:1283
3. Li X, Wang F, Li H, Richardson DD, Roush DJ (2022) The measurement and control of high-risk host cell proteins for polysorbate degradation in biologics formulation. *Antib Ther* 5(1): 42–54
4. Li X, An Y, Liao J, Xiao L, Swanson M, Martinez-Fonts K, Pavon JA, Sherer EC, Jawa V, Wang F, Gao X, Letarte S, Richardson DD (2021) Identification and characterization of a residual host cell protein hexosaminidase B associated with N-glycan degradation during the stability study of a therapeutic recombinant monoclonal antibody product. *Biotechnol Prog* e3128
5. Li X, Chandra D, Letarte S, Adam GC, Welch J, Yang RS, Rivera S, Bodea S, Dow A, Chi A, Strulson CA, Richardson DD (2021) Profiling active enzymes for polysorbate degradation in biotherapeutics by activity-based protein profiling. *Anal Chem* 93(23):8161–8169
6. Li SW, Yu B, Byrne G, Wright M, O'Rourke S, Mesa K, Berman PW (2019) Identification and CRISPR/Cas9 inactivation of the CLs protease responsible for proteolysis of recombinant proteins produced in CHO Cells. *Biotechnol Bioeng* 116(9):2130–2145
7. Cravatt BF, Wright AT, Kozarich JW (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* 77:383–414
8. Liu GY, Nie S, Zheng X, Li N (2022) Activity-based protein profiling probe for the detection of enzymes catalyzing polysorbate degradation. *Anal Chem* 94(24):8625–8632
9. Gupta SK, Graf T, Edelmann FT, Seelmann H, Reintinger M, Hillringhaus L, Bergmann F, Wiedmann M, Falkenstein R, Wegele H, Yuk IH, Leiss M (2023) A fast and sensitive high-throughput assay to assess polysorbate-degrading hydrolytic activity in biopharmaceuticals. *Eur J Pharm Biopharm* 187:120–129
10. Dehghani A, Binder F, Zorn M, Feigler A, Fischer KI, Felix MN, Happersberger P, Reisinger B (2024) Investigating pH effects on enzymes catalyzing polysorbate degradation by activity-based protein profiling. *J Pharm Sci* 113(3):744–753
11. Bracewell DG, Smith V, Delahaye M, Smales CM (2021) Analytics of host cell proteins (HCPs): lessons from biopharmaceutical mAb analysis for Gene therapy products. *Curr Opin Biotechnol* 71:98–104



Synthesis and Application of Activity-Based Probes for Serine Proteases

Alba Ramos-Llorca, Valeria Parravicini, Koen Augustyns, and Michaela Prothiwa

Abstract

Activity-based probes (ABPs) are chemical molecules that covalently react with the active form of the targeted enzyme. These probes are useful for target and biomarker identification, as well as drug discovery. This chapter describes the creation of an ABP library tailored to serine proteases, featuring a bioorthogonal tag and a range of diphenyl phosphonate warheads to improve selectivity and reactivity. We detail a protocol for synthesizing substituted warheads, enhancing probe affinity. The effectiveness of these ABPs is validated biochemically through analyses of their inhibitory potencies and kinetic behaviors. Furthermore, we illustrate their application in labeling serine proteases, using fluorescence scanning for detection.

Key words Activity-based probes, Diphenyl phosphonate, Serine proteases, Chemical proteomics, Click chemistry

1 Introduction

Serine proteases comprise over 40% of all known proteases in humans and are defined by the presence of a serine as a catalytic element in their active site [1]. These enzymes play a vital role in numerous physiological processes and are implicated in various human diseases, including cancer, cardiovascular, lung, and intestinal disorders. As such, they have been recognized as promising biomarkers and drug targets [2].

During the last decade, there has been a significant effort to investigate the expression and function of proteases in living organisms. One technique that has emerged as particularly promising is activity-based protein profiling (ABPP), which employs activity-based probes (ABPs) to facilitate the identification of active proteins in complex proteomes. ABPs are designed to react covalently with the active form of an enzyme, providing information about its activity rather than expression [3]. ABPP has a wide range of

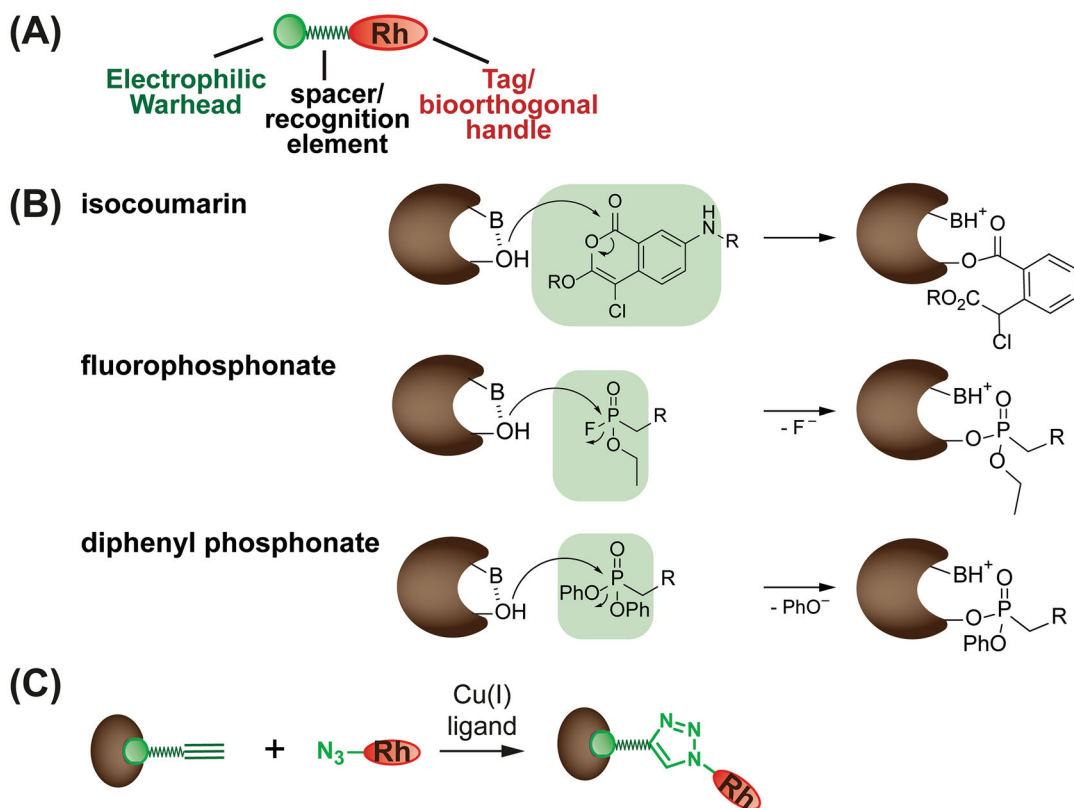


Fig. 1 Activity-based probes for serine proteases. **(a)** Design of a chemical probe containing an electrophilic warhead, a spacer or recognition element, and a terminal fluorescent reporter tag. **(b)** Examples of electrophilic warheads that covalently modify nucleophilic serine residues. **(c)** Copper (I) catalyzed alkyne-azide click chemistry allows for the attachment of a fluorescent tag to visualize the probe labeled serine protease. Rh = rhodamine

potential applications, including the identification of enzyme targets, in vivo visualization, or the discovery of new biomarkers and new inhibitors [4].

ABPs are structured into three main components: (1) a warhead or reactive group; (2) a recognition element, often peptidic, to improve selectivity; and (3) a reporter tag for detection or isolation (Fig. 1a) [5]. Warheads typically exhibit electrophilic properties to facilitate covalent bonds with the nucleophilic serine in the active site [6]. A wide range of electrophilic warheads are well-documented as protein-reactive components in chemical probes, including isocoumarin, oxolactams, chloromethyl ketone, sulfonyl fluorides, sulfonyloxypthalimides, and phosphorylated warheads [7–10]. The precise tuning of their electrophilicity, combined with the probe's ligand characteristics, ensures selectivity and reactivity essential for targeting only the nucleophilic serine residues (Fig. 1b). The α -amino diphenyl phosphonate (DPP) warhead

stands out as the most commonly used for serine-protease-specific ABPs. Its advantages include exclusive binding to serine proteases, synthetic accessibility, and high stability in plasma and buffers, making it highly effective for biological applications [11, 12]. The recognition element in a probe not only separates the warhead from the tagging moiety but is also key to enhance the probe's selectivity. Utilizing knowledge about protease substrates or their specificities enables the design of probes that are selectively targeted toward a particular protease or protease family. Numerous methods for optimizing the recognition element, including positional scanning libraries, phage and mRNA display methods, and mass spectrometry proteomic techniques, have yielded highly specific ABPs for use in complex biological samples [13]. The chemical probe's reporter tag can be a fluorescence dye, an affinity tag, or a functional group designed for bioorthogonal conjugation. To avoid the potential issue of bulky fluorophores hindering the probe's access to the active site, designs incorporating a terminal alkyne tag or an azide group are often favored (Fig. 1c). After binding to a protein, these tags can be conjugated to a corresponding detection tag through Huisgen cycloaddition under denaturing conditions.

This chapter delves into the synthesis and application of ABPs featuring a DPP warhead for targeting serine proteases. We showcase a strategy for fine-tuning the DPP warhead with different substituents to improve its reactivity. This approach enables the creation of an ABP library with warheads tailored for enhanced effectiveness and equipped with a “clickable” reporter tag, broadening their application spectrum. To assess the efficacy of these probes, we describe how to evaluate their inhibitory activity and kinetic profile using enzymatic assays. Additionally, we explain how to employ these probes for the labeling of recombinant proteases, using thrombin and chymotrypsin as illustrative examples.

2 Materials

2.1 General

Synthesis of α -Amino Diphenyl Phosphonate Activity-Based Probes

2.1.1 Birum-Oleksyzyn

*Reaction: Synthesis of Alkyl
Activity-Based Probe*

1. Reagents: selected aldehyde (with the desired selectivity-enhancing group), pent-4-yn-1-yl carbamate, a triphenyl phosphite (commercial or synthesized following Subheading 3.1.2), and a Lewis acid ($\text{Cu}(\text{OTf})_2$, $\text{Bi}(\text{OTf})_3$ or BF_3OEt_2).
2. Anhydrous solvent (DCM, THF or ACN).
3. Purification: celite or silica, DCM, EtOAc heptane, and a silica column.
4. Equipment: safety hood, round-bottom flask, magnetic stir bar, magnetic stir plate, LCMS, rotary evaporator, and (optional) automated flash chromatography system.

**2.1.2 Synthesis of
Substituted
Triphenylphosphite**

1. Reagents: substituted phenol, triethylamine, phosphorus trichloride.
2. Solvent: anhydrous THF.
3. Workup: celite pad.
4. Equipment: safety hood, vacuum pump, oven, round-bottom flask, magnetic stir bar, magnetic stir plate, ice bath, rotary evaporator, glass filter funnel.

**2.2 Biochemical
Validation of Probes**

1. Recombinant enzyme: recombinant thrombin (10 μg , 20,000 pmol/min/ μg).
2. Substrate stock solution of 100 mM Boc-Val-Pro-Arg-AMC in DMSO. Store in aliquots at $-20\text{ }^{\circ}\text{C}$.
3. 10 mM stock solution of the desired synthesized ABP in DMSO. Store in aliquots at $-20\text{ }^{\circ}\text{C}$.
4. Reconstitution buffer (HEPES buffer), 50 mM HEPES, 580 mM NaCl at pH 7.5 ($25\text{ }^{\circ}\text{C}$).
5. Tris buffer, 0.5 M Tris base, 0.5 M Tris HCl, 100 mM CaCl_2 , 1.5 M NaCl at pH 8.3 ($25\text{ }^{\circ}\text{C}$).
6. Reconstitute the enzyme by adding 100 μL of reconstitution buffer to a vial containing 10 μg of recombinant thrombin (enzyme stock solution final concentration 100 $\mu\text{g}/\text{mL}$). Store in aliquots at $-20\text{ }^{\circ}\text{C}$.
7. Solvents: Milli-Q water, DMSO.
8. Equipment: pipettes, black 96-well plate, spectrophotometer, graphic software such as GraphPad Prism 9 or GraFit7.

**2.3 Biochemical
Application of Probes**

1. 1.25 mM 5-carboxytetramethylrhodamine azide (5-TAMRA-azide) working stock in DMSO. Store at $-20\text{ }^{\circ}\text{C}$.
2. 2.5 mM tris-hydroxypropyltriazolylmethylamine (THPTA) ligand working stock in DMSO. Store at $-20\text{ }^{\circ}\text{C}$.
3. 50 mM sodium ascorbate solution in water (freshly prepared).
4. 50 mM CuSO_4 stock solution in water (freshly prepared).
5. HEPES reaction buffer, 100 mM HEPES, 0.5 M NaCl at pH 7.5 ($25\text{ }^{\circ}\text{C}$).
6. $2\times$ SDS-loading buffer, 63 mM Tris-HCl, 10% (v/v) glycerol, 2% (w/v) SDS, 0.25% (w/v) bromophenol blue and 5% (v/v) β -mercaptoethanol in H_2O deion.
7. Coomassie staining solution, 25% (w/v) Coomassie Brilliant Blue R250, 9.2% (v/v) acetic acid (conc.), and 45.5% ethanol (abs.) in H_2O deion.
8. Coomassie destainer, 10% acetic acid (conc.), and 20% ethanol (abs.) in H_2O deion.

9. 500 μM stock solution of the desired synthesized ABP in DMSO. Store in aliquots at $-20\text{ }^{\circ}\text{C}$.
10. Solvents: H_2O deion, DMSO.
11. Equipment: disposable 1.5 mL microcentrifuge tubes, incubator (temperature set to $37\text{ }^{\circ}\text{C}$), PerfectBlueTM Twin S system for small gels ($10 \times 10\text{ cm}$) or other SDS-PAGE unit with power supply, instrument for in-gel fluorescence scanning (e.g., Bio-Rad Chemidoc MP).

3 Methods

3.1 General

Synthesis of α -Amino Diphenyl Phosphonate Activity-Based Probes

ABPs with a DPP warhead are synthesized through a Birum-Oleksyszyn reaction. This involves a three-component reaction catalyzed by a Lewis acid, using an aldehyde, a carbamate, and a triphenyl phosphite [11, 14]. By decorating the phenyl groups in the triphenyl phosphite with different substituents, different reactivity can be added to the warhead. The selectivity-enhancing group (R_1) is carried by the aldehyde, and it is chosen based on the specificity of the targeted serine protease. The carbamate carries a terminal alkyne to allow the introduction of the desired reporter tag by click chemistry.

The following methodology describes a common and accessible synthetic strategy for an ABP library targeting serine proteases. First, Subheading 3.1.1 involves the formation of the DPP warhead. Additionally, if the reader wants to synthesize a substituted DPP warhead, Subheading 3.1.2 describes the synthesis of a substituted triphenyl phosphite.

3.1.1 Birum-Oleksyszyn Reaction: Synthesis of Alkyl Carbamate Probe (Fig. 2)

1. The selected aldehyde (1 eq.), pentyl-4-yn-1-yl carbamate (1 eq.), and triphenyl phosphite (1.1 eq., commercial or synthesized following Subheading 3.1.2) are added to a round-bottom flask, equipped with a magnetic stir bar, and anhydrous solvent (DCM, THF or ACN, to a final concentration of aldehyde 0.3 M, *see Note 1*) is added. In cases where the

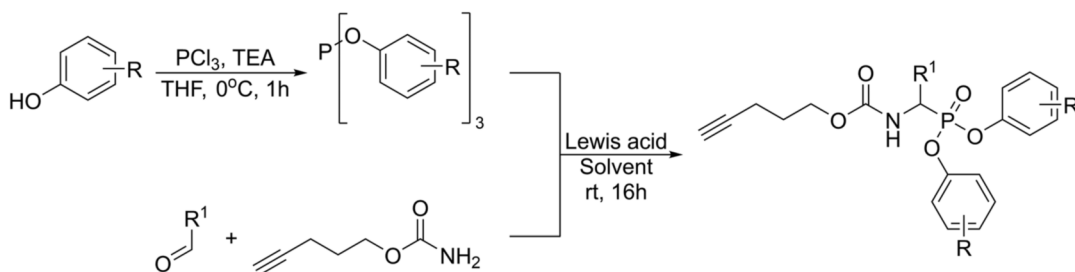


Fig. 2 General synthesis of an α -amino diphenyl phosphonate alkyne probe by the Birum-Oleksyszyn reaction. Lewis acid is $\text{Cu}(\text{OTf})_2$, $\text{Bi}(\text{OTf})_3$, or BF_3OEt_2 ; Anhydrous solvent is DCM, THF, or ACN

aldehyde contains a basic amine, it is necessary to protect it with a protecting group. Tert-butyloxycarbonyl (Boc) is the most convenient option, though alternative protection strategies may be required to achieve higher yields [15, 16].

5. The Lewis acid ($\text{Cu}(\text{OTf})_2$, $\text{Bi}(\text{OTf})_3$ or BF_3OEt_2 , 0.1 eq., *see Note 1*) is then added to the same round-bottom flask with the help of a syringe, spatula, or weighting paper (depending on the nature of the Lewis acid).
6. The solution is stirred at room temperature for 16 h under anhydrous conditions. Monitoring the reaction by Liquid chromatography-mass spectrometry (LCMS) shows the formation of the desired phosphonate and consumption of the starting aldehyde (*see Note 2*).
7. The solvent is evaporated using a rotary evaporator. It is crucial to maintain the temperature of the rotary evaporator below 30 °C to ensure the stability of the compound.
8. The residue is loaded in celite/silica by dissolving it in the minimum amount of DCM in a round-bottom flask and adding celite/silica. The solvent is evaporated using a rotary evaporator.
9. The remaining celite/silica is used as dry loading for flash column chromatography in a silica column. The mobile eluents are EtOAc and heptane. The gradient used goes from 0% to 100% EtOAc. The column chromatography is done using a flash chromatography automated system or by manual column chromatography.
10. The fractions containing the pure product are collected in a round-bottom flask and the solvent is evaporated using a rotary evaporator.
11. The characterization and purity determination can be done by ^1H NMR and/or LCMS.
12. Yields for this reaction go from 5 to 62%.

3.1.2 Synthesis of Substituted Diphenyl Phosphite (Fig. 2)

1. A substituted phenol (3 eq.), anhydrous THF (0.1 M), and dry triethylamine (3 eq.) (*see Note 3*) are added under argon to a round-bottom flask equipped with a magnetic stir bar (*see Note 4*). The flask is placed in an ice bath.
2. After 10 min, phosphorus trichloride (1 eq.) is added dropwise to the phenol solution.
3. The reaction mixture is stirred for 2 h at 0 °C.
4. The mixture is filtered through a celite pad and is washed with anhydrous THF (*see Note 5*).

5. The filtrate is concentrated under reduced pressure using a rotary evaporator. It is crucial to maintain the temperature of the rotary evaporator below 30 °C to ensure the stability of the compound.
6. The product is left in the vacuum pump for 6 h to eliminate all the solvent (*see Note 6*).
7. The characterization and purity determination can be done by ¹H NMR and/or LCMS.
8. Yields for this reaction go from 70 to 98%.

3.2 Biochemical Validation of Probes

3.2.1 IC₅₀ Determination

The IC₅₀ value is a measure of a molecule's inhibitory potency. It refers to the concentration required to reduce the activity of the target by 50%. This value can be used as an initial validation and screening cut-off to determine if the synthesized ABP is effectively targeting the desired enzyme. In this chapter, we outline the procedure for using thrombin as a model for trypsin-like serine proteases. If testing a different enzyme, the concentrations, selected substrate, and buffer need to be adjusted accordingly.

1. Prepare the working buffer solution by diluting 0.5 M Tris base, 0.5 M Tris HCl, 100 mM CaCl₂, and 1.5 M NaCl buffer with Milli-Q water to make a solution of 50 mM Tris base, 50 mM Tris HCl, 10 mM CaCl₂, 1 and 50 mM NaCl, and store it at 4 °C.
2. Prepare the enzyme working solution by diluting 2 µL of enzyme stock solution in 7 mL of working buffer solution.
3. Dilute 20 µL of ABP stock solution in 480 µL of working buffer solution. Prepare seven additional serial dilutions (1:4; 200 µL of previous dilution and 300 µL of working buffer solution) (with 4% DMSO) (*see Note 7*). Compounds can be initially screened at three concentrations (10, 1, and 0.1 µM) to estimate the range of the IC₅₀ value. If the compound presents an IC₅₀ < 10 µL, an exact IC₅₀ determination can be performed with more serial dilutions.
4. Prepare the substrate working solution by diluting 3 µL of substrate stock solution in 5 mL of working buffer solution.
5. In a black 96-well plate; dispense 5 µL of compound dilution and 145 µL of enzyme working solution using a pipette. It is necessary to have a blank and a control when conducting experiments. Instead of adding 5 µL of ABP dilution, dispense 5 µL of DMSO to blank and control wells. Each dilution of the ABP should be added in a separate well (*see Note 8*). Duplicates are also essential; thus, place each compound in two adjacent wells.
6. Incubate the plate for 15 min at 37 °C.

7. Dispense 50 μL of substrate working solution using a multi-channel pipette.
8. Measure enzymatic activity for 30 min at 37 °C using a spectrophotometer. Fluorescence is monitored at an excitation of 380 nm and an emission of 440 nm. The readout consists of evaluating the protease-mediated release of the fluorophore 7-amino-4-methylcoumarin (AMC) from the substrate.
9. Repeat the experiment in triplicate.
10. IC_{50} values can be obtained by fitting the data using graphics software such as GraphPad Prism 9.

3.2.2 Second-Order Rate Constant (k_{app}) Determination

Many irreversible compounds follow a two-step mechanism, where the formation of a reversible complex is followed by the formation of an irreversible complex. This process is defined by the pseudo-first-order rate constant (k_{obs}). To measure the potency of an irreversible compound, the apparent second-order rate constant (k_{app}) is calculated from the correlation between k_{obs} and the compound concentration. Kinetic progress curves are recorded in pseudo-first-order conditions to determine k_{obs} [17].

1. In a black 96-well plate, dispense 5 μL of compound dilution and 50 μL of substrate working solution using a pipette. It is necessary to have a blank and a control when conducting experiments. Instead of adding 5 μL of ABP dilution, dispense 5 μL of DMSO to blank and control wells. Each dilution of the ABP should be added in a separate well. Duplicates are also essential; thus, place each compound in two adjacent wells.
2. Incubate the plate for 15 min at 37 °C.
3. Dispense 145 μL of enzyme working solution (preincubated at 37 °C) with a multichannel pipette (*see Note 9*).
4. Measure enzymatic activity for 1 h at 37 °C using a spectrophotometer (*see Note 10*). Fluorescence is monitored at an excitation of 380 nm and an emission of 440 nm. The readout consists of evaluating the protease-mediated release of the fluorophore AMC from the substrate.
13. Repeat the experiment in triplicate.
14. Fit the data for each compound concentration, using a graphing software such as GraFit7 or GraphPad Prism 9, to the following equation to obtain k_{obs} (*see Note 11*).

$$A = A_0 + v_s \times t + (v_0 - v_s) \times (1 - \exp(-k_{\text{obs}} \times t)) / k_{\text{obs}}$$

15. Fit the k_{obs} values against compound concentration to calculate k_{app} , which corresponds to the slope of the plot's linear part.

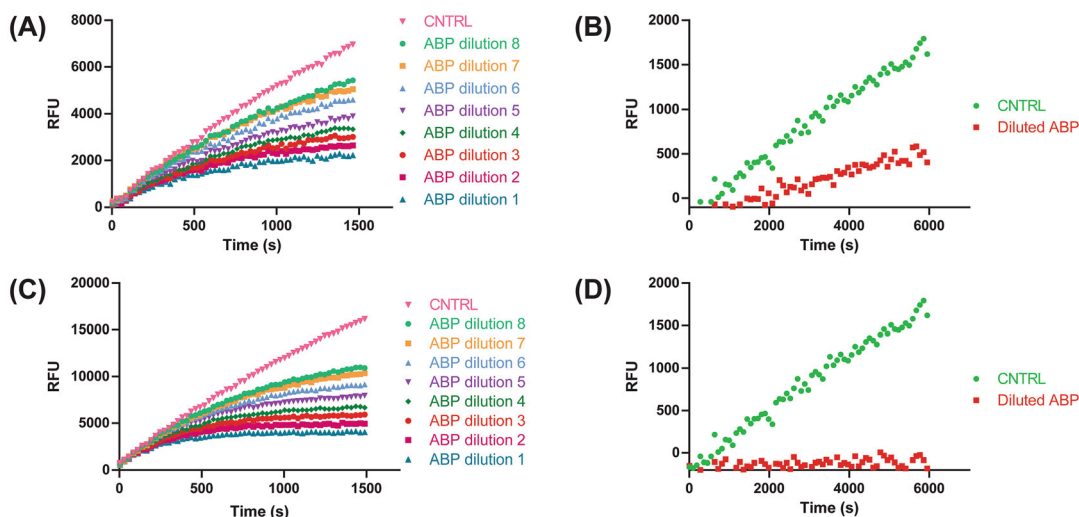


Fig. 3 Kinetic progress curves (a, c) and jump dilutions (b, d) in thrombin. Reversible ABPs (a, b) and irreversible ABPs (c, d)

16. Typical results are illustrated in Fig. 3. If the ABP is reversible, the progress curve will not lead to a plateau and the progress curve will stabilize linearly (Fig. 3a), while, with an irreversible ABP, a plateau will appear in the progress curve when all the enzyme is inhibited (Fig. 3c).

3.2.3 Determination of the Mechanism of Action: Jump Dilutions

The kinetic progress curve can provide insights into the type of mechanism of inhibition. However, to confirm whether a compound has an irreversible mechanism of action, additional assays such as jump dilutions are necessary. These assays involve using a compound concentration of 10 times the IC_{50} for the incubation and then diluting it 100 times before adding the substrate [18].

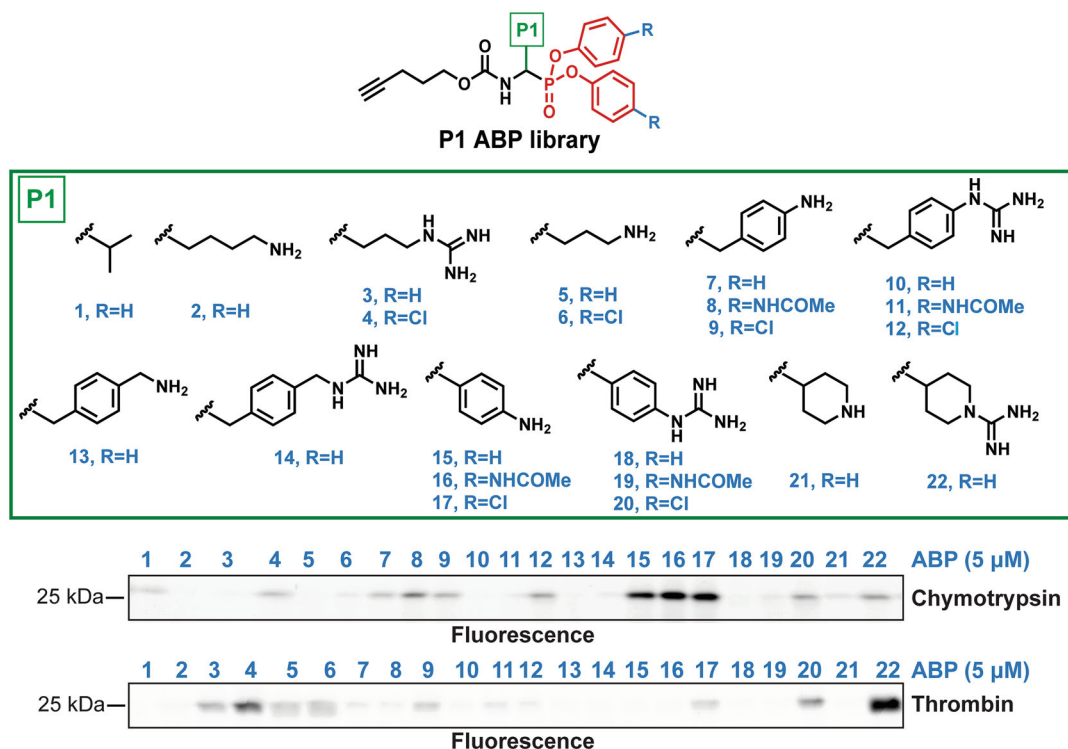
1. Prepare the enzyme working solution by diluting 2 μ L of enzyme stock solution in 5 mL of working buffer solution.
2. Dilute the ABP stock solution in working buffer solution to prepare a working solution with a concentration $400 \times$ the IC_{50} value (final concentration in the well $10 \times IC_{50}$).
3. Prepare the substrate working solution by diluting 3 μ L of substrate stock solution in 5 mL of working buffer solution.
4. In a black 96-well plate, dispense 5 μ L of ABP working solution and 145 μ L of enzyme working solution. It is necessary to have a blank and a control when conducting experiments. Instead of adding 5 μ L of ABP dilution, dispense 5 μ L of DMSO to blank and control wells. Each dilution of the ABP should be added in a separate well. Duplicates are also essential; thus, place each compound in two adjacent wells.
5. Incubate the plate for 15 min at 37 $^{\circ}$ C.

6. Transfer 2 μL of the previous mixture to an empty well, and dilute with 148 μL of working buffer solution.
7. Incubate the plate for 30 min at 37 °C.
8. Dispense 50 μL of working substrate solution.
9. Measure enzymatic activity for 16 h at 37 °C using a spectrophotometer. Fluorescence is monitored at an excitation of 380 nm and an emission of 440 nm. The readout consists of evaluating the protease-mediated release of the fluorophore AMC from the substrate.
10. Repeat the experiment in duplicate.
11. Typical results are illustrated in Fig. 3. Inhibition of enzyme activity by reversible and irreversible ABPs results in different outcomes. Reversible ABPs cause a gradual increase in activity due to the slow release of the ABP (Fig. 3b), while irreversible ABPs lead to permanent loss of enzyme activity (Fig. 3d).

3.3 Biochemical Application of Probes

This protocol outlines the use of ABPs with a bioorthogonal alkyne tag for the activity-dependent detection of serine proteases through gel electrophoresis (SDS-PAGE) and in-gel fluorescence imaging. We detail the attachment of a fluorescent tag to the probe-labeled protease using click chemistry. This method is adaptable for labeling any recombinant protease, demonstrated here with recombinant chymotrypsin and thrombin.

1. Prepare 50 μL of a solution containing 15 $\mu\text{g/mL}$ chymotrypsin or thrombin in HEPES buffer or any other buffer compatible with the protease of interest and click reagents (*see Note 12*).
2. Add ABP stock solution (5 μM final concentration, 1% DMSO) to each sample, and mix by gentle vortexing or pipetting. A control reaction should be conducted either without the probe or after pretreatment with suitable protease inhibitors.
3. Incubate for 60 min at 37 °C without shaking (*see Note 13*).
4. Add click reagents in the following order to start the click reaction: (1) 1 μL rhodamine azide stock, (2) 1 μL ligand stock, (3) 1 μL ascorbic acid stock solution, (4) 1 μL CuSO_4 stock solution.
5. Mix immediately strongly by vortexing.
6. Incubate for 60 min at room temperature without shaking. Avoid direct sunlight exposure to prevent photobleaching.
7. Add 2 $\times$ SDS loading buffer to stop the click reaction and mix by vortexing.
8. Perform SDS-PAGE according to standard methods. Avoid direct sunlight exposure to prevent photobleaching.



25 kDa—

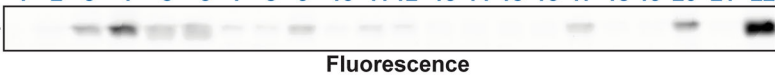


Fluorescence

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 ABP (5 μM)

Chymotrypsin

25 kDa—



Fluorescence

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 ABP (5 μM)

Thrombin

Fig. 4 In-gel fluorescence detection of recombinant chymotrypsin and thrombin labeled with probes from an ABP library featuring diverse natural and nonnatural amino acids in the P1 position and different substitutions on the DPP warhead [15]

9. After running the gel, visualize the labeled serine proteases using an in-gel fluorescence readout. This can be performed directly within the glass plates of the gel using a suitable fluorescence reader like the Bio-Rad Chemidoc MP utilized in this procedure.
10. Typical results are illustrated in Fig. 4. This figure shows the in-gel fluorescence detection of recombinant chymotrypsin and thrombin. These proteases were labeled using probes from an ABP library, which includes a variety of both natural and nonnatural amino acids at the P1 position and different substitutions on the DPP warhead.
11. Perform Coomassie staining to confirm equal protein content.

In conclusion, this chapter outlines a method for synthesizing ABPs targeting serine proteases, featuring a DPP warhead with various substitutions to enhance its reactivity. This synthesis process is practical for many biochemistry, bioorganic laboratories, or any facility equipped for organic synthesis. We demonstrate the evaluation of ABPs' inhibitory activity and kinetic profiles. Additionally,

we highlight the use of ABPs in gel-based detection methods, facilitating straightforward visualization of active proteases with standard biochemistry equipment.

4 Notes

1. The pair of Lewis acid and solvent will be determined by the nature of the R1 moiety. The most common pair used is Cu (OTf)₂/DCM. However, if the yield of the reaction was lower than 15–20%, it is recommended to try other solvents, such as THF, or other Lewis acid/solvent pairs, such as Bi(OTf)₃ with THF or ACN or BF₃OEt₂ with ACN.
2. Avoid contact with water and methanol, which can hydrolyze the phosphonate.
3. Ensure that all reagents are completely dry before use. The substituted phenol must be vacuum-dried for 24 h before use. Molecular sieves can be used to dry triethyl amine. To do so, place the triethylamine in an Erlenmeyer flask, add the molecular sieves to cover the bottom of the flask, and flush the flask with argon before closing it with a septum. To measure the desired amount of triethylamine, use a syringe.
4. The round-bottom flask should be dried in the oven (~130 °C) for a minimum of 24 h.
5. To filter a solution using a celite pad, add celite to cover two-thirds of a glass filter funnel, and then pour the solution into the funnel.
6. Please be cautious while using the product, as it tends to form foam when vacuum-dried.
7. DMSO can affect the assay. Therefore, it is important to maintain the same % of DMSO in all serial dilutions.
8. The IC₅₀ value determined is a good start as the smallest concentration for the progress curves.
9. The enzyme should be added last to avoid any inhibition before the start of the measurement.
10. The first data points play a crucial role in determining the initial velocity of the reaction. It is essential to start taking measurements immediately after dispensing the enzyme. The time interval between readings should be set to the smallest possible value by the spectrophotometer.
11. In the equation: A_t refers to the fluorescence at time t , A_0 is the fluorescence at time zero, v_i is the initial rate, and v_s is the velocity at steady-state.

12. Avoid using Tris buffer, CHAPS detergent, and high concentrations of SDS, as they can inhibit the click reaction.
13. For labeling other proteases, employ a buffer that maintains the activity of your target proteases. Ensure the sample volume is adequate to meet the detection threshold of your readout method. When using different probes, it may be necessary to optimize labeling conditions, such as probe concentration, labeling duration, and pH.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 765608 and from the Research Fund Flanders grant agreement FWO-SBO S 001017 N and by the University of Antwerp SEP-BOF grant number 44875.

References

1. Di Cera E (2009) Serine proteases. *IUBMB Life* 61:510–515. <https://doi.org/10.1002/iub.186>
2. Heutink KM, ten Berge IJM, Hack CE et al (2010) Serine proteases of the human immune system in health and disease. *Mol Immunol* 47: 1943–1955. <https://doi.org/10.1016/j.molimm.2010.04.020>
3. Cravatt BF, Wright AT, Kozarich JW (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* 77:383–414. <https://doi.org/10.1146/annurev.biochem.75.101304.124125>
4. Niphakis MJ, Cravatt BF (2014) Enzyme inhibitor discovery by activity-based protein profiling. *Annu Rev Biochem* 83:341–377. <https://doi.org/10.1146/ANNUREV-BIOCHEM-060713-035708>
5. Jessani N, Cravatt BF (2004) The development and application of methods for activity-based protein profiling. *Curr Opin Chem Biol* 8:54–59. <https://doi.org/10.1016/j.cbpa.2003.11.004>
6. Powers JC, Asgian JL, Ekici ÖD, James KE (2002) Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem Rev* 102: 4639–4750. <https://doi.org/10.1021/cr010182v>
7. Haedke U, Götz M, Baer P, Verhelst SHL (2012) Alkyne derivatives of isocoumarins as clickable activity-based probes for serine proteases. *Bioorg Med Chem* 20:633–640. <https://doi.org/10.1016/J.BMC.2011.03.014>
8. Shannon DA, Gu C, McLaughlin CJ et al (2012) Sulfonyl fluoride analogues as activity-based probes for serine proteases. *ChemBiochem* 13:2327–2330. <https://doi.org/10.1002/cbic.201200531>
9. Kam CM, Abuelyaman AS, Li Z et al (1993) Biotinylated isocoumarins, new inhibitors and reagents for detection, localization, and isolation of serine proteases. *Bioconjug Chem* 4: 560–567. <https://doi.org/10.1021/bc00024a021>
10. Ruivo EFP, Gonçalves LM, Carvalho LAR et al (2016) Clickable 4-Oxo- β -lactam-based selective probing for human neutrophil elastase related proteomes. *ChemMedChem* 11: 2037–2042. <https://doi.org/10.1002/CMDC.201600258>
11. Oleksyszyn J, Powers JC (1991) Irreversible inhibition of serine proteases by peptide derivatives of (alpha-aminoalkyl)phosphonate diphenyl esters. *Biochemistry* 30:485–493. <https://doi.org/10.1021/bi00216a026>
12. Ides J, Thomae D, Wyffels L et al (2014) Synthesis and in vivo preclinical evaluation of an 18F labeled uPA inhibitor as a potential PET imaging agent. *Nucl Med Biol* 41:477–487. <https://doi.org/10.1016/j.nucmedbio.2014.03.019>
13. Faucher F, Bennett JM, Bogoy M, Lovell S (2020) Strategies for tuning the selectivity of chemical probes that target serine hydrolases.

- Cell Chem Biol 27:937–952. <https://doi.org/10.1016/J.CHEMBIOL.2020.07.008>
14. Van der Veken P, El Sayed I, Joossens J et al (2004) Lewis acid catalyzed synthesis of N-protected diphenyl 1-Aminoalkylphosphonates. *Synthesis (Stuttg)* 2005:634–638. <https://doi.org/10.1055/s-2004-837307>
 15. Ramos-Llorca A, Decraecker L, Cacheux VMY et al (2023) Chemically diverse activity-based probes with unexpected inhibitory mechanisms targeting trypsin-like serine proteases. *Front Chem* 10:1089959. <https://doi.org/10.3389/FCHEM.2022.1089959/BIBTEX>
 16. Pan Z, Jeffery DA, Chehade K et al (2006) Development of activity-based probes for trypsin-family serine proteases. *Bioorg Med Chem Lett* 16:2882–2885. <https://doi.org/10.1016/j.bmcl.2006.03.012>
 17. Lambeir A (1996) Dipeptide-derived diphenyl phosphonate esters: mechanism-based inhibitors of dipeptidyl peptidase IV. *Biochim Biophys Acta Gen Subj* 1290:76–82. [https://doi.org/10.1016/0304-4165\(96\)00012-8](https://doi.org/10.1016/0304-4165(96)00012-8)
 18. Copeland RA (2013) Evaluation of enzyme inhibitors in drug discovery a guide for medicinal chemists and pharmacologists. Wiley



Chapter 12

CellEKT: A Chemical Proteomics Workflow to Profile Cellular Target Engagement of Kinase Inhibitors - A Step-by-Step Guide

Berend Gagestein, Joel Rüegger, Antonius P. A. Janssen, Stephan Kirchner, Uwe Grether, Arne C. Rufer, and Mario van der Stelt

Abstract

The human genome encodes for over 600 kinases, which are important targets for drug discovery in many diseases, such as cancer and autoimmune disorders. The majority of kinase inhibitors target the conserved ATP-binding pocket, which contributes to the difficulty of developing selective inhibitors. A lack of selectivity may contribute to high drug candidate attrition rates due to an unfavorable off-target profile. To prevent unforeseen side effects, as well as to accurately link inhibitor function to pharmacological effects and phenotypic readouts, efficient methods for determining target engagement across the cellular kinome are necessary. One way to address this problem is to profile the interaction landscape of kinase inhibitors in living cells using the broad-spectrum kinase probe XO44. Here, we describe a detailed protocol of the CellEKT workflow using XO44 to profile kinase inhibitors in a full dose-response manner across 200+ kinases. The protocol allows to determine the EC₅₀ of up to three kinase inhibitors in a cell line of interest in 4 days.

Key words In situ target engagement, Drug selectivity, ABPP, Kinome profiling, Pull-down, XO44, CellEKT, EC₅₀

1 Introduction

The human genome encodes for over 600 catalytically active kinases [1], including protein kinases as well as metabolite kinases, such as lipid and carbohydrate kinases. These enzymes mediate the phosphorylation of substrate proteins by transferring the γ -phosphate groups from ATP to serine, threonine, and tyrosine residues. Protein phosphorylation is a delicate process of regulating signaling pathways throughout the cell, enabling the downstream cellular function. Disturbance of the kinase and phosphatase

Authors Berend Gagestein and Joel Rüegger have equally contributed to this chapter.

Zhipeng Tao (ed.), *Activity-Based Proteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2921, https://doi.org/10.1007/978-1-0716-4502-4_12,
© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2025

interplay by aberrant kinase activity is implicated in various diseases, including cancer [7], neurodegenerative disorders [2], and inflammatory conditions [3], highlighting the significance of kinases as therapeutic targets. Inhibiting kinase activity has proven successful in the treatment of many diseases and is reflected by the 80 FDA-approved kinase inhibitors as of writing [4].

The majority of kinase inhibitors target the highly conserved ATP-binding pocket, making it challenging to design an inhibitor which is active and selective for a specific kinase or kinase subset without interacting with other potentially unfavorable kinases [5]. Many kinase inhibitors in development eventually fail due to an unfavorable safety profile [6]. Profiling kinase inhibitors at an early stage is crucial to comprehend potential toxicity concerns, but also to fully understand the mode of action of already FDA-approved drugs to allow drug repurposing. This need is even stronger for kinase inhibitors advancing into fields beyond oncology [7].

Biochemical activity assays based on purified kinases (or their catalytic domains) are the classical way to determine the selectivity of kinase inhibitors across the kinome. A more recent technique is the Kinobeads assay [8], which are beads immobilized with a set of promiscuous kinase inhibitors. To determine the selectivity profile of a kinase inhibitor the inhibitor is incubated with Kinobeads to compete for available binding sites in an affinity-based readout in cell lysates. Kinobeads account for the kinase expression profile and activation states of kinases in a certain cell line but are limited to the application on lysate level. Therefore, the full complexity of a living cell can be underestimated due to various factors, for example, cellular penetration of the drug, post-translational modifications, and cellular ATP concentration (2–8 mM in mammalian cells) [9]. Techniques such as cellular thermal shift assay (CETSA) [10] and energy transfer assays (NanoBRET) [11, 12] allow for the selectivity readout of kinase inhibitors in living cells. Both CETSA and NanoBRET come with limitations regarding coverage or requiring a recombinant construct, respectively. A more versatile technique is the use of chemical probes such as XO44 [13], which is not limited to a specific cell line. XO44 is a molecule with a promiscuous hinge binder moiety, which functions as a recognition moiety for kinases (*see* Fig. 1a). Furthermore, XO44 is equipped with a sulfonyl fluoride, which forms a covalent bond with the conserved lysine in the active site of the kinase (*see* Fig. 1b). Enrichment of the bound kinases is achieved through click chemistry of the alkyne handle and biotin-azide, which enables isolation with streptavidin beads.

We previously described the CelleKT (Cellular Endogenous Kinase Targeting) workflow [14] to assess the cellular selectivity of kinase inhibitors using XO44 and two new probes ALX005 and ALX011. Engaged kinases identified by CelleKT were thoroughly validated using complementary platforms such as NanoBRET and

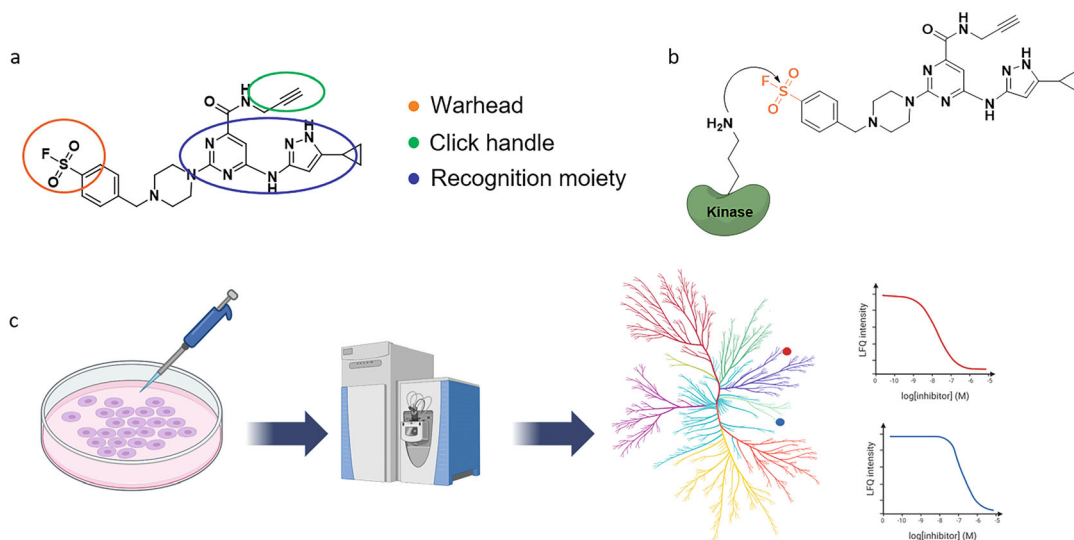


Fig. 1 X044 and chemical proteomics workflow used in this protocol. (a) Chemical structure of X044. (b) Nucleophilic attack of conserved lysine on sulfonamide warhead of X044 (c) Chemical proteomics workflow to determine the full dose response of a kinase inhibitor in live cells using X044. Two kinases are inhibited. In red, kinase is inhibited with higher activity, in blue kinase is inhibited with lower activity [15]

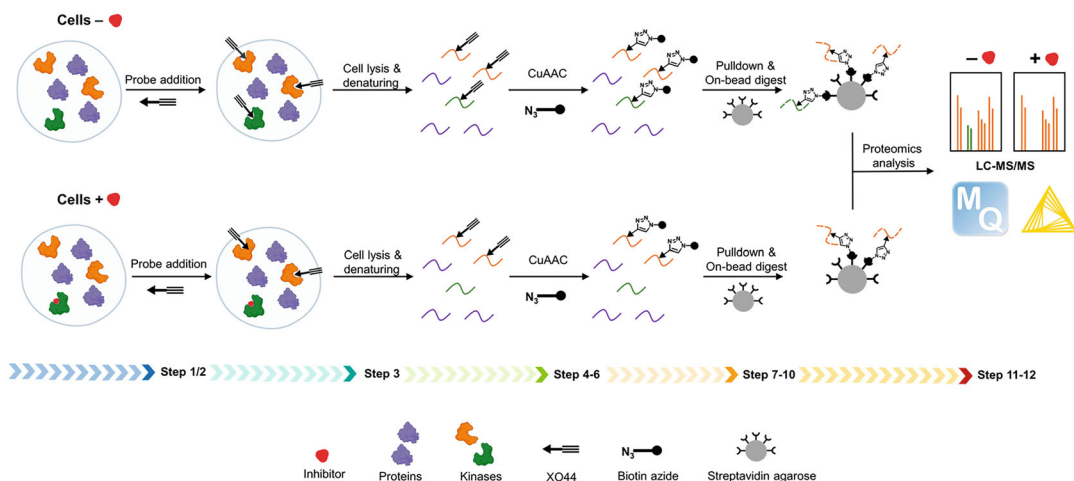


Fig. 2 Overview of the CelleKT workflow [15]

phosphoproteomics. This chapter presents a detailed, step-by-step protocol for the CelleKT workflow (*see* Fig. 1c), using X044 as a representative example. Combined with an automated data processing pipeline, the protocol enables concentration-dependent target engagement profiling of up to three kinase inhibitors across over 200 kinases per cell line within just four working days.

Figure 2 illustrates the CelleKT workflow: cells are treated with inhibitor or vehicle followed by X044 or vehicle treatment for any

kind of cell type (Subheadings 3.1 and 3.2, **steps 1, 2**). After cell lysis (Subheading 3.3, **step 3**), the probes are biotinylated through click chemistry followed by protein precipitation, reduction, and alkylation (Subheadings 3.4, 3.5 and 3.6, **steps 4–6**). The probe-bound proteins are enriched with streptavidin beads, digested and desalted using StageTips (Subheadings 3.7, 3.8, 3.9 and 3.10, **steps 7–10**). The tryptic peptides are measured by LC-MS in a data-dependent acquisition (DDA) fashion (Subheading 3.11, **step 11**). We employ MaxQuant to process the raw data, which is then refined through our KNIME pipeline to calculate and plot full dose-response curves of inhibited kinases (Subheading 3.12, **step 12**). Recently, XO44 derivatives with increased metabolic stability [16] and coverage [17] have been developed; however, XO44 itself remains the only commercially available broad-spectrum kinase probe, which can be used *in situ*.

2 Materials

The used solvents need to be ULC-MS grade and reagents need to be of the highest purity in order to prevent contamination in LC-MS. Store all prepared reagents at room temperature (RT) unless otherwise indicated. The amounts indicated for 48 samples have a small overhead to account for minor losses using, for example, a dispenser.

2.1 Cell Treatment

1. Culture medium (HEK293T cells): Culture cells at 37 °C under 7% CO₂ in DMEM with GlutaMAX, 10% NBS, and penicillin and streptomycin (200 µg/mL each). Store at 4 °C (*see Note 1*).
2. Culture medium (MV4–11 cells): Culture cells at 37 °C under 5% CO₂ in IMDM with GlutaMAX, 10% FCS, and penicillin and streptomycin (200 µg/mL each). Store at 4 °C (*see Note 1*).
3. 1 M HEPES: Dissolve 2.38 g HEPES in Milli-Q, adjust pH to 7.5 with aq. NaOH, and fill to 10 mL.
4. Treatment medium: Prepare treatment medium by combining corresponding serum-free culture medium with 1 M HEPES pH 7.5 to a final concentration of 20 mM HEPES (*see Note 2*). Add lipid-free bovine serum albumin 0.1% (w/v), and perform sterile filtering over a 0.45 µm filter. Store at 4 °C (*see Note 3*).
5. 10 mM XO44 stock: Dissolve 5 mg PF-6808472 (Merck) in 928 µL DMSO. Store at –20 °C.
6. 10 mM inhibitor stock: Dissolve inhibitor accordingly in DMSO. Store at –20 °C.
7. DPBS: Dulbecco's Phosphate Buffered Saline.

2.2 Harvesting and Lysis

1. Milli-Q: Generated by a Millipore Advantage A10 device equipped with Millipak[®] Express 40 filter (*see Note 4*).
2. 1 M MgCl₂ stock: Dissolve 2.0 g of magnesium chloride hexahydrate (MgCl₂·6 H₂O) in Milli-Q to a final volume of 10 mL.
3. 1 M NaCl stock: Dissolve 0.58 g NaCl in 10 mL Milli-Q.
4. 10% (w/v) Triton X-100 stock: Dissolve 1.0 g Triton X-100 in 9 mL Milli-Q.
5. 1 M Tris stock: Dissolve tris(hydroxymethyl)aminomethane (Tris) in Milli-Q, adjust pH to 7.8 with aq. HCl, and add Milli-Q.
6. Benzonase storage buffer: Prepare storage buffer by combining 200 μL 1 M Tris HCl (pH 8.0), 20 μL 1 mM MgCl₂, 200 μL 1 mM NaCl, 5.88 mL 85% glycerol (v/v), and 3.70 mL Milli-Q.
7. 10 U/μL Benzonase stock: Dilute benzonase nuclease according to the specific activity in the storage buffer. Aliquots can be stored at −20 °C for at least 6 months.
8. 25X PIC stock: Dissolve one tablet of cComplete[™] EDTA-free protease inhibitor cocktail in 2 mL Milli-Q. Aliquots can be stored at −20 °C for at least 6 months.
9. Triton Lysis buffer: For 48 samples, prepare 20 mL of lysis buffer by combining 1 mL of 1 M HEPES stock, 3 mL of 1 M NaCl, 20 μL of 1 M MgCl₂, 200 μL of 10% Triton X-100, and 14.93 mL Milli-Q. Aliquots can be stored at −20 °C for at least 1 year. Add 800 μL of PIC stock and 50 μL of Benzonase stock fresh before cell lysis.
10. 10% SDS stock (w/v): Combine 1 g sodium dodecyl sulfate (SDS) with 10 mL Milli-Q (*see Note 5*).

2.3 Click Chemistry

1. 10 mM biotin-azide: Dissolve 3.8 mg biotin-azide in 1.0 mL DMSO. Store at −20 °C.
2. 25 mM CuSO₄: Dissolve 399 mg anhydrous CuSO₄ in 100 mL Milli-Q. Store at −20 °C.
3. 50 mM TCEP-HCl: Dissolve 14.4 mg tris(2-carboxyethyl) phosphine hydrochloride (TCEP-HCl) in 1.0 mL DPBS. Prepare directly before use.
4. 25 mM THPTA: Dissolve 10.9 mg tris-hydroxypropyltriazolylmethylamine (THPTA) in 1.0 mL DMSO. Store at −20 °C.
5. Pull-down click mix: For 48 samples, combine 800 μL 25 mM CuSO₄, 160 μL 25 mM THPTA, 200 μL 10 mM biotin-N₃, and 400 μL 50 mM TCEP-HCl (*see Note 6 and 7*).

6. 2 mM AF647-N₃: Dissolve 0.51 mg of AF647-N₃ in 300 μ L DMSO. Store at -20°C .
7. 1 M EDTA stock: Dissolve 2.92 g EDTA in 10 mL Milli-Q, adjust pH to 8.0 with NaOH.
8. Gel click mix: For each sample combine 1.6 μ L 25 mM CuSO₄, 0.32 μ L 25 mM THPTA, 0.4 μ L 2 mM AF647-N₃, and 0.8 μ L 50 mM TCEP-HCl (*see* **Note 6** and **7**).
9. 50 mM HEPES/50 mM EDTA: Combine 250 μ L 50 mM HEPES, 250 μ L 50 mM EDTA, and 4.50 mL MilliQ.
10. 4X Laemmli Buffer: 240 mM Tris-HCl pH 6.8, 8% w/v SDS, 40% glycerol, 5% v/v β -mercaptoethanol, 0.04% v/v bromophenol blue.

2.4 Reduction and Alkylation

1. 0.5 M IAA stock: For 48 samples, dissolve 46.2 mg iodoacetamide (IAA) in 500 μ L Milli-Q. Light-sensitive, prepare directly before use.
2. 1 M DTT stock: Dissolve 1.54 g dithiothreitol (DTT) in Milli-Q to a final volume of 10 mL. Aliquots can be stored at -20°C for up to 3 months.
3. 10X PBS stock: Dissolve 1.43 g KH₂PO₄, 87.7 g NaCl, and 3.60 g NaH₂PO₄ in Milli-Q, adjust pH to 7.5 with NaOH, and bring up to 1 L (*see* **Note 8**).
4. PBS: Dilute 10X PBS stock ten times in Milli-Q (pH should be 7.5).
5. PBS + 0.5% SDS: Combine 50 mL of 10% (w/v) SDS stock and 100 mL 10X PBS with 850 mL Milli-Q to obtain PBS with 0.5% (w/v) SDS.
6. PBS + 0.5% SDS + 5 mM DTT: For 48 samples, combine 125 μ L of 1 M DTT, 1.25 mL 10% (w/v) SDS stock, and 2.5 mL 10X PBS stock with 21.1 mL Milli-Q to obtain 0.5% SDS in PBS containing 5 mM DTT.

2.5 Streptavidin Enrichment

1. High Capacity Streptavidin Agarose: Thermo Scientific (+ Pierce™ Control Agarose: Thermo Scientific in a 2:3 mixture, *see* **Note 9**).
2. PBS/10 mM DTT: For 48 samples, add 250 μ L of 1 M DTT stock to 25 mL PBS to obtain PBS containing 10 mM DTT.

2.6 On-Bead Digestion

1. 1 M CaCl₂ stock: Dissolve 147 mg calcium chloride dihydrate (CaCl₂·2 H₂O) in 1 mL Milli-Q.
2. OB-digestion buffer: For 48 samples, prepare 10 mL of lysis buffer by combining 1 mL 1 M Tris pH 7.8, 1 mL 1 M NaCl, 10 μ L 1 M CaCl₂, 200 μ L acetonitrile. Fill up to 10 mL using Milli-Q. Stocks can be stored for at least 1 year at -20°C .

3. Trypsin stock (0.5 $\mu\text{g}/\mu\text{L}$): Dissolve 20 μg of trypsin (sequencing grade; Promega) in 40 μL 1 mM HCl solution (diluting 37% HCl into Milli-Q, pH 3). Snap freeze with liquid nitrogen and store at -20°C . Avoid freeze-thaw cycles.
4. OB-digestion buffer + trypsin: For 48 samples, combine 10 mL OB-digestion buffer with 50 μL trypsin stock. Addition of 200 μL of resulting OB-digestion buffer per sample = 0.5 μg trypsin per sample.

2.7 Sample Preparation and Desalting

1. 20% Formic acid: Prepare a solution of 20% (v/v) formic acid (FA) in Milli-Q. Discard after 3 months.
2. StageTip solution A: Prepare a solution of 0.5% (v/v) FA in Milli-Q. Discard after 3 months.
3. StageTip solution B: Prepare a solution of 0.5% (v/v) FA in 80% (v/v) acetonitrile/Milli-Q. Discard after 3 months.
4. LC-MS sample solution: Combine Milli-Q, acetonitrile, and formic acid in 97:3:0.1 (v/v/v) ratio. Discard after 3 months.
5. 2 mL StageTip tubes: Remove the cap of 2 mL tubes with scissors, and add cap with a hole to place the StageTip. The caps can be reused, and the holes can be conveniently prepared with a leather hole puncher (*see* Fig. 3).
6. C18 StageTips: Commercially available (CDS Analytical™ Empore™ C18 StageTips) or prepare according to the *Nature Protocols* by Rappsilber et al. [18]. Two punched disks of Empore™ C18 extraction disks are used in a 200 μL pipette tip (*see* Note 10).

2.8 LC-MS Analysis

1. LC setup: UltiMate 3,000 RSLCnano system set in a trap-elute configuration with a nanoEase M/Z Symmetry C18 100 $\text{\AA}$, 5 μm , 180 $\mu\text{m} \times 20\text{ mm}$ (Waters) trap column for peptide

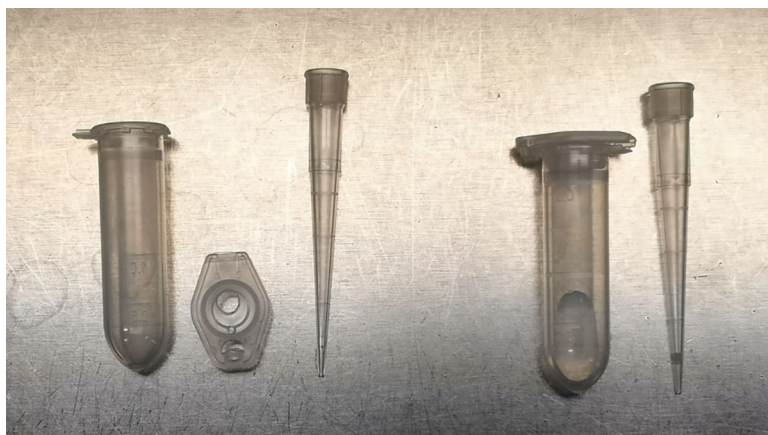


Fig. 3 Setup for StageTip (Left) and a conditioned StageTip (Right)

loading/retention and nanoEase M/Z HSS C18 T3 100 Å, 1.8 µm, 75 µm × 250 mm (Waters) analytical column for peptide separation. Column oven set to 40 °C.

2. Mobile phase A: 0.1% (v/v) FA in ULC-MS grade water.
3. Mobile phase B: 0.1% (v/v) FA in ULC-MS grade acetonitrile.

2.9 Data Processing and Analysis

1. FASTA file: Go to <http://www.uniprot.org> and download the according proteome. In the case of the human proteome, search with UPID: UP000005640.
2. Kinase list: In order to filter probe-enriched kinases from the human proteome, a list of kinases needs to be downloaded from <http://www.uniprot.org>. Use the search parameters: (keyword:KW-0418) AND (organism_id:9606) AND (reviewed:true).
3. MaxQuant: Version 2.4.13.0 is run with label-free quantification (LFQ) enabled and an LFQ min. ratio count of 1. iBAQ and match between runs is enabled.
4. KNIME workflow: To process the MaxQuant output in an automated fashion, we wrote a data analysis workflow in KNIME Analytics Platform 4.5.2 with KNIME Python Integration and the Vernalis KNIME nodes as additional modules. The workflow is fully annotated to help the user understand the data analysis steps (*see Note 11*). The workflow can be downloaded here: https://github.com/APAJanssen/KNIME_MaxQuant_processing. Installation instructions and template sample lists are provided there.

3 Methods

Example experiment in Table 1 (*see Notes 12 and 13*):

Table 1
Overview of full dose-response selectivity profiling for three kinase inhibitors

Sample nr.	Sample type	Concentration	Purpose
1–2	DMSO	–	Negative control
3–6	XO44	1 µM	Positive control
7–10	XO44	10 µM	Peptide library
11–22	Inhibitor 1	0.1 nM–10,000 nM	Test samples
23–34	Inhibitor 2	0.1 nM –10,000 nM	Test samples
35–46	Inhibitor 3	0.1 nM –10,000 nM	Test samples

3.1 Option 1:
Adherent Cells (e.g.
HEK293T Cells) ~6
h Total + Overnight

1. Seed 1.5×10^6 cells per well (6 well plate). Let cells adhere overnight (*see Note 14*).
2. Aspirate culture medium and add treatment medium (1.8 mL) with 1.11X inhibitor or DMSO (*see Note 15*).
3. Incubate for 60 min at 37 °C. (*see Note 16*).
4. Add treatment medium (0.2 mL) with 10X XO44 or DMSO (*see Note 17*).
5. Incubate for 30 min at 37 °C.
6. Aspirate probe solution and add 0.9 mL ice-cold DPBS.
7. Collect cells by scraping, and transfer to 1.5 mL tubes on ice.
8. Spin down (1,000 g, 5 min, 4 °C), and aspirate supernatant.
9. Snap-freeze pellets using liquid nitrogen and store at -80 °C

3.2 Option 2:
Suspension Cells (e.g.
MV4-11 Cells) ~4 h

1. Count cells, spin down (300 g, 5 min, rt), and resuspend to 1 M/mL in treatment medium (*see Note 14*).
2. Divide cells over pre-warmed 15 mL falcon tubes (7 mL to receive a total of 7×10^6 cells per condition) already containing 1.11X inhibitor or DMSO (*see Note 18*).
3. Incubate for 60 min at 37 °C (*see Note 16*).
4. Add treatment medium (777 µL) with 10X XO44 or DMSO. (*see Note 19*).
5. Incubate for 25 min at 37 °C.
6. Spin down (300 g, 5 min, 37 °C), aspirate supernatant, and place the tubes on ice.
7. Add 0.9 mL ice-cold DPBS and transfer the cells to 1.5 mL tubes on ice.
8. Spin down (1,000 g, 5 min, 4 °C), and aspirate supernatant.
9. Snap-freeze pellets using liquid nitrogen and store at -80 °C.

3.3 Cell Lysis ~4 h

1. Thaw pellets in the fridge.
2. Add 380 µL ice-cold lysis buffer and lyse cells by pipetting up and down + vortexing (*see Note 20*).
3. Determine protein concentration with Qubit/BCA assay, and adjust concentration to the lowest value (*see Note 21 and 22*).
4. Transfer 360 µL to a 2 mL tube containing 40 µL 10% SDS, giving 1% final, vortex, and leave at RT.
5. Proceed with click chemistry for gel analysis or snap-freeze samples using liquid nitrogen, and continue with the pull-down the following day (*see Note 23 and 24*).

3.4 Click Chemistry (for Pull-Down) ~0.5 h + 1 h Incubation

1. Thaw samples at RT (*see Note 25*).
2. Add 31.2 μL pull-down click mix to 400 μL lysate and vortex.
3. Incubate for 60 min at RT.
4. Add 80 μL 50 mM HEPES/50 mM EDTA (*see Note 26*).

3.5 Protein Precipitation ~1 h

1. Add 666 μL MeOH, and vortex.
2. Add 166 μL CHCl_3 , and vortex.
3. Add 150 μL Milli-Q, and vortex thoroughly.
4. Spin down (1,500 g, 5 min).
5. Remove the upper and lower layers without disturbing the protein pellet (*see Note 27*).
6. Add 600 μL MeOH to the pellet.
7. Resuspend by sonication with a microplate sonicator (10% amplitude, 2×10 s).
8. Spin down (20,000 g, 5 min).
9. Remove MeOH, and air-dry pellet for ± 5 min (*see Note 28*).

3.6 Reduction and Alkylation ~1 h

1. Redissolve protein pellet in 500 μL PBS + 0.5% SDS + 5 mM DTT.
2. Incubate for 15 min at 65 °C with 800 RPM shaking.
3. Spin down briefly and inspect for undissolved material (*see Note 29*).
4. Allow the samples to reach RT for ± 5 min (*see Note 30*).
5. Add 10 μL of 0.5 M IAA stock, vortex, and incubate at RT in the dark for 30 min (*see Note 31*).
6. Spin down samples briefly (20,000 g, 2 min).

3.7 Streptavidin Enrichment (*see Note 32*) ~1 h + 2 h Incubation

1. Add 20 μL of 50% PierceTM high-capacity streptavidin agarose slurry per sample to a 15 mL falcon tube (*see Note 9*).
2. Wash the beads with PBS + 0.5% SDS (2×6 mL) by vortexing and spinning down (3,000 g, 2 min), then discard supernatant.
3. Wash the beads with PBS (6 mL) by vortexing and spinning down (3,000 g, 2 min), then discard supernatant.
4. Add 500 μL PBS + 10 mM DTT per sample to the falcon tube containing the beads (*see Note 33*).
5. Divide the beads over 1.5 mL low-binding tubes (*see Notes 34 and 35*).
6. Add the pull-down sample to the 1.5 mL tubes containing the beads, resulting in a total volume of 1 mL with 0.25% SDS (*see Note 36*).
7. Agitate the beads by shaking (1,000 RPM, 2 h) or tumbling.

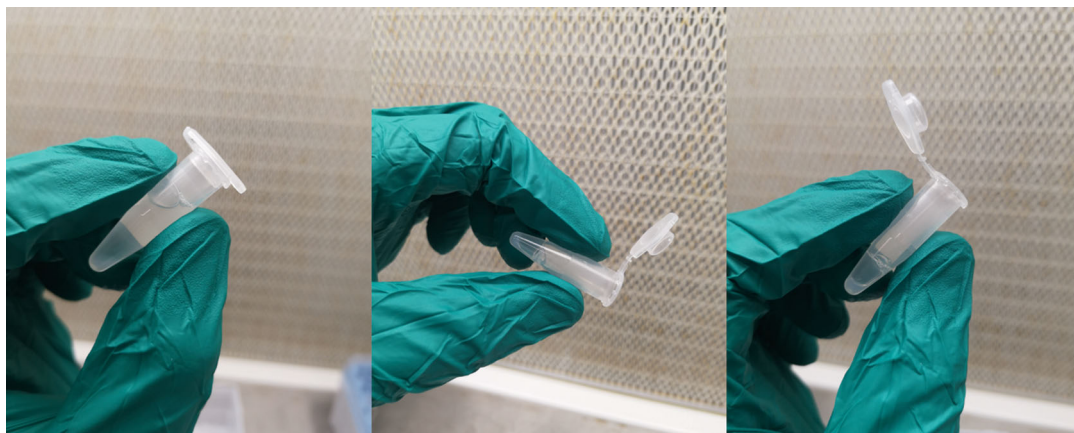


Fig. 4 Washing beads with PBS + SDS. Tip over the 1.5 mL tube without excess shaking to avoid losing beads



Fig. 5 Washing beads with PBS. Due to the high surface tension more liquid will remain in the 1.5 mL tube. Tip over the tube without excess shaking to avoid losing beads

3.8 Washing Beads **~2 h**

1. Spin down sample (3,000 g, 2 min) and discard supernatant by pouring out.
2. Wash the beads with 1 mL of 0.5% SDS in PBS. Spin down sample (3,000 g, 2 min) and discard supernatant by pouring out. Perform this washing step three times (*see* Fig. 4 and **Note 37**).
3. Wash the beads with 1 mL PBS. Spin down sample (3,000 g, 2 min) and discard the supernatant by pouring out. Perform this washing step three times (*see* Fig. 5).
4. Resuspend the beads in MS-grade Milli-Q (0.8 mL) and transfer to 1.5 mL low-binding tubes.



Fig. 6 Taking off the supernatant by pipetting. Make sure to leave a volume equal to the pellet volume to avoid aspirating beads

5. Spin down sample (3,000 g, 2 min) and remove most of the supernatant by pipetting, being careful not to aspirate any beads. (*see* Fig. 6).

3.9 On-Bead Digestion 5 min + Overnight Incubation

1. Add 200 μ L on-bead digestion buffer + trypsin per sample (*see* Note 38).
2. Digest at 37 °C with vigorous shaking (1,000 RPM) overnight.

3.10 Sample Preparation and Desalting ~4 h + 2–3 h Speedvac

1. Place StageTips (**steps 2–4**) in 2 mL StageTip tubes and condition by passing through the following mixtures by spinning them down in a centrifuge (*see* Figs. 3 and 7).
2. 50 μ L MeOH, 1–2 min, 450 g.
3. 50 μ L StageTip solution B, 1–2 min, 450 g (*see* Note 39).
4. 50 μ L StageTip solution A, 1–2 min, 450 g (*see* Note 40).
5. To the digested samples, add 50 μ L 20% FA in Milli-Q and vortex.
6. Spin down the beads (3,000 g, 2 min).
7. Transfer 90% of supernatant to a prepared, conditioned StageTip from **step 4** (*see* Note 41).
8. Place StageTip back into the centrifuge and spin down for 2 min at 300 g, then for 5 min at 400 g, checking the progression every 1 min.
9. Add 100 μ L StageTip solution A. Spin down for 2.5 min at 800 g.
10. Transfer the cap containing the StageTip to a new 1.5 mL low protein binding tube. Elute the peptides using 100 μ L StageTip solution B and spinning down for 5 min at 800 g.
11. Evaporate solvents completely using a speedvac (45 °C, 2–3 h), and store samples at –80 °C.

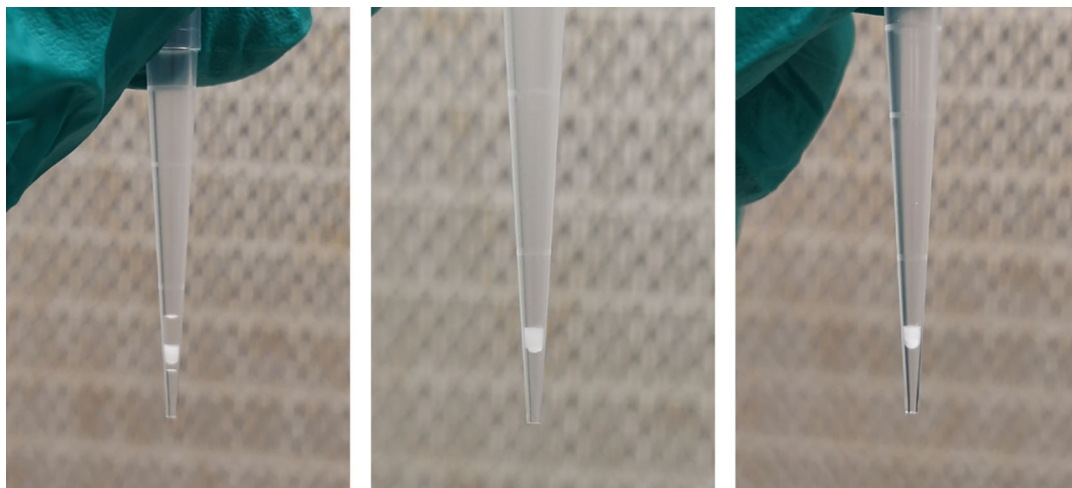


Fig. 7 Reference StageTips. The StageTip on the left is conditioned and has some remaining liquid left, which prevents it from drying. The StageTip in the middle is conditioned but is drying out and should be used immediately. The StageTip on the right has dried out and should be remade

3.11 LC-MS Analysis
~1 h Practical + 85
min Gradient per
Sample

1. Reconstitute peptides in 25 μL LC-MS solution by vortexing + shaking (1,300 RPM, 10 min), and spin down briefly.
2. Pool 1–2 μL from each sample as quality control (QC) sample.
3. Inject 3–5 μL per sample depending on QC results.
4. Inject samples on trap column at a flow rate of 15 $\mu\text{L}/\text{min}$ for 2 min with 99% mobile phase A, 1% mobile phase B eluent. Program an 85 min gradient with linear increment to 1% B from 0 to 2 min, 5% B at 5 min, 22% B at 55 min, 40% B at 64 min, 90% B at 65 to 74 min, and 1% B at 75 to 85 min using mobile phase A and mobile phase B controlled by a flow sensor at 0.3 $\mu\text{L}/\text{min}$ with an average pressure of 400–500 bar (5,500–7,000 psi). Introduce eluent by electro-spray ionization (ESI) via the nanoESI source (Thermo) using stainless steel nano-bore emitters (40 mm, OD 1/32", ES542, Thermo Scientific).
5. MS setting: Operate Q Exactive HF in positive mode with data-dependent acquisition without the use of lock mass, and default charge of 2+ and external calibration with LTQ Velos ESI positive ion calibration solution every 5 days to less than 2 ppm. Set the tune file for the survey scan to a scan range of 350–1,400 m/z , 120,000 resolution (m/z 200), 1 microscan, automatic gain control (AGC) of 3e6, max injection time of 100 ms, no sheath, aux or sweep gas, spray voltage ranging from 1.7 to 3.0 kV, capillary temp of 250 $^{\circ}\text{C}$. and an S-lens value of 80. Set the loop count to 10 for the 10 data-dependent MS/MS events. Use a resolution of 15,000, AGC target 1e5, max IT time 50 ms, isolation window of 1.6 m/z , fixed first

mass of 120 m/z, and normalized collision energy (NCE) of 28 eV for the general settings. For individual peaks, use data-dependent settings of 1.00×10^3 for the minimum AGC target resulting in only triggering an MS/MS event if an intensity threshold of 2.0×10^4 is reached. Disable apex trigger. Exclude unassigned +1 charges, and charges greater than +8 with peptide match mode are preferred. Enable isotope exclusion, and apply dynamic exclusion of 10 s.

3.12 Data Processing and Analysis

1. Import .raw folders of each LC-MS run within an experiment into MaxQuant, highlight all files in MaxQuant and set to “No fractions.”
2. Import FASTA file.
3. Adjust the number of threads according to the number of cores available (*see Note 42*).
4. MaxQuant will output a proteingroups.txt file, which can, along with a sample list and the kinase list, be read into our KNIME workflow and further processed to receive dose-response fitted data with graphical visualization (*see Note 11*).

4 Notes

1. Cells should be passaged twice a week at 80–90% confluency by resuspension in a fresh medium after trypsinization if necessary. Cell lines should be tested regularly for mycoplasma contamination and discarded after 2–3 months of use.
2. Probe and inhibitor treatment can be conducted outside of the CO₂ incubator for convenience. In that case, use carbonate-buffered media with 20 mM HEPES. The addition of HEPES to the treatment medium is not required if incubation is performed in a CO₂ incubator.
3. Fatty acid-free BSA (0.1%) is added to the serum-free medium to prevent probe or inhibitor adsorption to the plastic.
4. Be aware that filters such as Biopak Polisher filters (cat. No. CDUFBI001) could introduce contaminants for mass spectrometry.
5. Weigh SDS in a fumehood to prevent inhalation.
6. Mix CuSO₄, THPTA, AF647-N₃ (for gel), or biotin-N₃ (for pull-down), vortex, add TCEP-HCl, vortex. Pipetting instead of vortexing reduces the mixture’s exposure to oxygen, minimizing unwanted reactions.
7. Reagents, such as Cy5-N₃ or TAMRA-biotin-N₃, suffer from high SDS concentrations.

8. Different formulations with higher potassium content will result in the precipitation of potassium dodecyl sulfate.
9. To improve bead handling during the washing steps add 30 μL Pierce™ control agarose resin slurry per sample in addition to 20 μL of 50% Pierce™ high-capacity streptavidin agarose slurry. Vortex the bead bottles thoroughly before pipetting out the required amount of slurry. The beads do not fit through most 200 μL pipette tips, therefore use 1,000 μL tips.
10. If StageTips tips are prepared by hand, it is recommended to prepare all of them in one batch for consistency. As an alternative to StageTips, especially for a large amount of samples, desalting can also be performed with a 96-well C18 Oasis sample extraction plate. For the Oasis sample extraction plate, use 100 μL MeOH, 200 μL solution B, and 200 μL solution A for the conditioning steps. For washing and elution, use 200 μL solution A and 200 μL solution B.

11.

- (a) Prepare sample list template: Create a sample list in Microsoft Excel using the provided template to specify .raw file names, treatment conditions, and proteins of interest (POI), which are cross-correlated downstream in the workflow.
- (b) Data input: Load the sample list template and the MaxQuant output into KNIME.

The following steps are performed in an automated fashion by the KNIME data analysis workflow.

- (c) Sanitization of MaxQuant output:
 - Removal of reverse peptides and contaminants from the MaxQuant output.
 - Grouped identifications are split, retaining only the first (predominant) identification.
 - Elimination of proteins with fewer than two unique peptides (configurable).
- (d) Normalization and imputation of missing values:
 - Gaussian distribution is fit to all log₂-transformed LFQ values using Python's SciPy norm.fit.
 - Transformation of the distribution according to Perseus [20] default settings by shifting μ left 1.8σ and reducing σ by 70%.
 - Imputation of missing values for enriched proteins based on the transformed distribution.

- (e) Calculation protein enrichment:
 - Enrichment computation of proteins in probe-treated positive controls versus DMSO-treated negative controls based on raw intensity values.
 - Identification of proteins with an enrichment ratio ≥ 2 as probe targets.
- (f) Quality filtering:
 - Positive controls with at most 1 missing LFQ intensity value are retained.
 - Imputation of missing values for enriched proteins.
 - Normalization of LFQ intensities with respect to mean unimputed positive control values.
- (g) pEC₅₀ fitting:
 - Identification of proteins where the highest concentration of inhibitor shows $\leq 60\%$ probe labeling.
 - Preparation for pEC₅₀ fitting by calculating the EC₅₀ and Hill slope using Python's SciPy optimize.curve_fit algorithm using the equation:

$$\%Effect = \frac{100}{1 + e^{Hill * (\log([inhibitor]) - \log(EC_{50}))}}$$

- *Hill* is the Hill slope, while [*inhibitor*] is the concentration of inhibitor used (typically in M). Initial values for Hill slope and EC₅₀ are provided as 1 and 1,000, respectively. The Hill slope is bound to be fit between 0 and 5. Fits are classified as “Good” if $0.3 \leq \text{Hill slope} \leq 3.0$, $\sigma(\text{pEC}_{50}) \leq 1.0$ and $\text{pEC}_{50} \geq \text{Max}(\text{p}[\text{inhibitor}]) - 0.5$, and as “Poor” otherwise.
- (h) Data visualization:
 - Generation of LFQ intensity histogram plots to visualize the relative abundance of proteins.
 - MS/MS counts and unique peptide counts per concentration of inhibitor are plotted to show available MS evidence.
 - The phylogenetic tree is created based on the seminal paper of Manning et al. [21] and data visualization tools developed in the Phanstiel Lab [22].
- (i) Export:
 - All plots are exported as scaled vector graphic files combined with the curve fits written to a Microsoft Excel file.

12. Two vehicle-treated samples (DMSO) are used as a negative control as comparison to the four positive control 1 μM XO44 samples to identify specifically probe-enriched proteins. In

addition, four 10 μM XO44 samples are used as a “peptide library” to increase protein identification coverage using the “Match between runs” function in the data processing software. We recommend measuring a full-dose response curve of kinase inhibitors in duplicates from 0.1 nM to 10 μM ($\log_{10}$ steps, 6 concentrations in total). Inhibitors can also be measured at only one or two concentrations, but then, we recommend measuring quadruplicates of each concentration point. If the expression of kinases is investigated, 10 μM of XO44 instead of 1 μM can be used since competition of an inhibitor is not an issue. The workflow can be used with other probes than XO44, for example, probes enriching serine hydrolases [19]. However, incubation time and probe concentration will have to be adjusted for optimal performance.

13. If the workflow is done with >10 samples, it is recommended to use an electronic multi-dispenser pipette (e.g., Eppendorf Multipette) throughout the whole pull-down workflow to achieve consistent results.
14. For optimal performance, 0.5–1 mg of protein per sample is needed. For HEK293T cells, this corresponds to around 1.5×10^6 cells seeded in six-well plates and for MV4–11 around 7×10^6 cells. For every cell line, the amount of cells required, and the corresponding well size for adherent cells needs to be determined beforehand. The amount of cells (and protein) needs to be consistent over all the conditions within an experiment. This ensures that inhibitor and probe ratio to cells is the same. Moreover, label-free quantification also requires consistent sample preparation.
15. Inhibitor treatment medium (adherent cells):
 - (a) Dilute 10 mM stock and resulting dilutions tenfold by diluting 2 μL inhibitor stock in 18 μL DMSO down to 0.0001 mM.
 - (b) From each inhibitor concentration, 4.44 μL is added to a falcon tube (*see* Table 2).
 - (c) 4 mL serum-free medium is added, resulting in an inhibitor concentration of 1.11X (*see* Table 2).
 - (d) To each well, 1.8 mL of the corresponding treatment medium with an inhibitor concentration of 1.11X is added for the first incubation step (60 min) according to Table 2:
16. For experiments in which only expression levels of kinases are investigated (no competition), the incubation step with inhibitor can be skipped (steps 2 and 3 of Subheadings 3.1 or 3.2).

Table 2
Dilution scheme for adherent cells (e.g., HEK293T cells) in a six-well plate

Sample	Final conc. [μM]	Medium per tube [mL]	DMSO [μL]	0.0001 mM [μL]	0.001 mM [μL]	0.01 mM [μL]	0.1 mM [μL]	1 mM [μL]	10 mM [μL]
DMSO	–	20 mL	20						
Inhibitor	0.0001	4 mL		4.44					
Inhibitor	0.001	4 mL			4.44				
Inhibitor	0.01	4 mL				4.44			
Inhibitor	0.1	4 mL					4.44		
Inhibitor	1.0	4 mL						4.44	
Inhibitor	10	4 mL							4.44

Table 3
Dilution scheme of XO44 for treating adherent cells (e.g., HEK293T cells) cultured in a six-well plate

Sample	Final conc. [μM]	samples	Medium per tube [μL]	DMSO [μL]	1 mM [μL]	10 mM [μL]
DMSO	–	2	500	0.5		
XO44	1	4 (positive control per dataset)+ 12 (per inhibitor)	6,000		6	
XO44	10	4	1,000			10

17. XO44 treatment medium (adherent cells):
- (a) Dilute 10 mM XO44 stock in serum-free medium according to Table 3:
 - (b) Add 0.2 mL 10X XO44 treatment medium to the corresponding well to receive a final concentration of 1X inhibitor and 1X probe for the second incubation step (30 min).
18. Inhibitor treatment medium (suspension cells):
- (a) Dilute 10 mM stock and resulting dilutions tenfold by diluting 2 μL inhibitor stock in 18 μL DMSO down to 0.0001 mM.
 - (b) From each inhibitor concentration, 7.77 μL is added to a falcon tube (see Table 4).
 - (c) 7 mL medium containing 1×10^6 cells/mL is added resulting in an inhibitor concentration of 1.11X for the first incubation step (60 min) according to Table 4:

Table 4
Dilution scheme for suspension cells (e.g., MV4–11 cells) in a falcon tube

Sample	Final conc. [μM]	samples	Medium per tube [mL]	DMSO [μL]	0.0001 mM [μL]	0.001 mM [μL]	0.01 mM [μL]	0.1 mM [μL]	1 mM [μL]	10 mM [μL]
DMSO	–	2 + 8	7	7.77						
Inhibitor	0.0001	2	7	7.77						
Inhibitor	0.001	2	7		7.77					
Inhibitor	0.01	2	7			7.77				
Inhibitor	0.1	2	7				7.77			
Inhibitor	1.0	2	7					7.77		
Inhibitor	10	2	7						7.77	

Table 5
Dilution scheme of XO44 for treating suspension cells (e.g., MV4–11 cells) in a falcon tube

Sample	Final conc. [μ M]	Replicates (N falcon tubes)	Medium per tube [μ L]	DMSO [μ L]	1 mM [μ L]	10 mM [μ L]
DMSO	–	2	777	7.77		
XO44	1	4 (positive control per dataset) + 12 (per inhibitor)	777		7.77	
XO44	10	4	777			7.77

19. XO44 treatment medium (suspension cells):
 - (a) Dilute 10 mM XO44 stock tenfold by diluting 2 μ L in 18 μ L DMSO to receive a 1 mM stock.
 - (b) From 1 mM and 10 mM probe stock 7.77 μ L is dissolved in 777 μ L to receive a 10X probe dilution (*see* Table 5).
 - (c) 777 μ L with 10X probe dilution is added to the corresponding falcon tube to receive a final concentration of 1X inhibitor and 1X probe for the second incubation step (30 min) according to Table 5:
20. If a probe is used, which forms a reversible bond, for example, the imine formation between a salicylaldehyde warhead and the lysine [16], a reductive amination with NaBH₄ can be done on lysate level to form a covalent bond between the probe and the according kinases.
 - (a) Weigh 9.45 mg NaBH₄ and dissolve in 2.5 mL cold lysis buffer (50 mL Falcon tube because of H₂ formation) = 100 mM stock.
 - (b) To 9.5 mL lysis buffer, add 0.5 mL of 100 mM NaBH₄ stock.
 - (c) Add 380 μ L lysis buffer with NaBH₄ to the cell pellet, and incubate for 30 min on ice (vortex or pipet up and down upon addition).
 - (d) Transfer 360 μ L to a 2 mL tube containing 40 μ L 10% SDS giving 1% final, vortex, and leave at RT.
 - (e) Proceed with click chemistry for pull-down or gel analysis (*see* Note 23 and 24), or snap freeze with liquid nitrogen and store at –80 °C.
21. Bradford assay is not compatible with SDS concentrations used here.
22. If there are a few samples with low protein yield (< 0.5 mg) compared to the other samples (e.g. 0.75 mg), it is better to find an average concentration (e.g. 0.6 mg) to prevent diluting the samples too much and losing overall performance.

23. Samples can be stored at -80°C for at least 3 months after snap freezing.
24. Clicking (an aliquot of) samples to a fluorophore-azide, and running on SDS-PAGE gel before doing a pull-down is highly recommended if novel probes are used. This way, incubation times and probe/competitor concentrations can be optimized. Altered lysis conditions should also first be validated on gel.

Sample preparation for gel:

 - (a) Thaw samples at RT (*see* **Note 25**).
 - (b) Add 3.12 μL gel click mix to 40 μL lysate.
 - (c) Click for 30 min at RT.
 - (d) Quench with 4x Laemmli buffer (5 min, 95°C).
 - (e) Load 20 μL per lane on 1.5 mm gel or 10 μL per lane on 0.75 mm gel.
25. Freezing samples will lead to SDS precipitation which can be observed while thawing up the samples. SDS will go in solution once the samples reach RT. Continue with the procedure once SDS is fully dissolved.
26. EDTA will chelate copper and halt oxidative damage, beneficial for experiments involving western blots but optional for proteomics.
27. If the pellet is a disk, remove both layers without disturbing the floating pellet. This works best by holding the tube at a 45° angle to stick the protein pellet against the side. If the disk is fragile or does not get stuck to the tube, use a 1,000 μL pipet tip to remove most of the liquid. Put the 1,000 μL pipet tip in a 200 μL pipet tip, and remove all the remaining liquid by holding the tube at a 45° angle to stick the protein pellet against the side.
28. Air-dry pellets for at most 5 min as completely dried-out protein pellets are very hard to redissolve.
29. In case of undissolved protein, sonicate the samples with microplate sonicator (10% amplitude, 2×10 s), and/or incubate for an additional 15 min at 65°C with 800 RPM shaking.
30. Exposing IAA to heat or light can result in unwanted peptide modifications.
31. Use the 30-min incubation step with IAA to wash the beads (Subheading 3.7, steps 1–4).
32. From this point on, work in a flow cabinet to prevent keratin contamination. Always wear gloves. A hairnet is also recommended.
33. DTT is added to quench unreacted IAA.

34. Amount of low binding tubes = amount of samples. The beads need to be transferred to a 50 mL falcon tube if the pull-down is conducted with more than 24 samples. Use the required amount of PBS + 10 mM DTT to transfer the beads, and wash the 15 mL falcon tube 1–2 times. Mix the bead suspension thoroughly before dividing the beads over the 1.5 mL tubes. After pipetting 3–4 aliquots, mix the beads again.
35. We recommend using low protein binding tubes from Eppendorf as we have found that other brands might lead to increased background signal in the LC-MS.
36. Do not transfer any undissolved material to the beads.
37. It is vital to prevent downstream SDS contamination. Clean the working surface, and change gloves after the PBS + SDS washing steps.
38. 200 μ L on-bead digestion buffer containing 0.5 μ g of trypsin. 0.5 μ g trypsin per sample is an indication for pull-down experiments. Check the number of missed cleavages in downstream analysis, and adjust the amount of trypsin if necessary.
39. To be as efficient as possible, leave the 2 mL StageTip tubes equipped with a StageTip in the centrifuge, and StageTip solutions B and A can directly be added without removing them from the centrifuge. This is not recommended for sample loading, however, to prevent potential sample loss.
40. Add StageTip solution A to the tips, spin briefly, and wait to fully elute the solution until the sample is ready to load. This prevents the StageTip from drying out.
41. Make sure to not transfer beads. Alternatively, beads can be removed by filtering over a Bio-Spin column (Bio-Rad) into a new low-bind tube. Bio-Spin columns can be reused. Clean them by rinsing with UPLC-grade Milli-Q, MeOH, and acetonitrile.
42. On a desktop pc (Intel core i7 tenth generation, 32 GB RAM, 8 cores), it takes around 10 h to process 46 samples.

Acknowledgments

Research reported in this publication was supported by Oncode Accelerator, a Dutch National Growth Fund project under grant number NGFOP2201, Oncode Institute, and by F. Hoffmann-La Roche through the Access to Distinguished Scientist (ROADS) Program (ROADS-032).

References

1. Consortium TU, Bateman A, Martin M-J et al (2022) UniProt: the universal protein knowledgebase in 2023. *Nucleic Acids Res* 51: D523–D531. <https://doi.org/10.1093/nar/gkac1052>
2. Jennings D, Huntwork-Rodriguez S, Henry AG et al (2022) Preclinical and clinical evaluation of the LRRK2 inhibitor DNL201 for Parkinson's disease. *Sci Transl Med* 14:eabj2658. <https://doi.org/10.1126/scitranslmed.abj2658>
3. Sato S, Sanjo H, Takeda K et al (2005) Essential function for the kinase TAK1 in innate and adaptive immune responses. *Nat Immunol* 6: 1087–1095. <https://doi.org/10.1038/ni1255>
4. Roskoski R (2024) Properties of FDA-approved small molecule protein kinase inhibitors: a 2024 update. *Pharmacol Res* 200:107059. <https://doi.org/10.1016/j.phrs.2024.107059>
5. Cohen P, Alessi DR (2013) Kinase drug discovery – what's next in the field? *ACS Chem Biol* 8:96–104. <https://doi.org/10.1021/cb300610s>
6. Attwood MM, Fabbro D, Sokolov AV et al (2021) Trends in kinase drug discovery: targets, indications and inhibitor design. *Nat Rev Drug Discov* 1–23. <https://doi.org/10.1038/s41573-021-00252-y>
7. Ferguson FM, Gray NS (2018) Kinase inhibitors: the road ahead. *Nat Rev Drug Discov* 17: 353–377. <https://doi.org/10.1038/nrd.2018.21>
8. Bantscheff M, Eberhard D, Abraham Y et al (2007) Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat Biotechnol* 25:1035–1044. <https://doi.org/10.1038/nbt1328>
9. Greiner JV, Glonek T (2021) Intracellular ATP concentration and implication for cellular evolution. *Biology* 10:1166. <https://doi.org/10.3390/biology10111166>
10. Savitski MM, Reinhard FBM, Franken H et al (2014) Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* 346:1255784. <https://doi.org/10.1126/science.1255784>
11. Vasta JD, Corona CR, Wilkinson J et al (2018) Quantitative, wide-Spectrum kinase profiling in live cells for assessing the effect of cellular ATP on target engagement. *Cell Chem Biol* 25:206–214.e11. <https://doi.org/10.1016/j.chembiol.2017.10.010>
12. Robers MB, Wilkinson JM, Vasta JD et al (2021) Single tracer-based protocol for broad-spectrum kinase profiling in live cells with NanoBRET. *Star Protoc* 2:100822. <https://doi.org/10.1016/j.xpro.2021.100822>
13. Zhao Q, Ouyang X, Wan X et al (2017) Broad-Spectrum kinase profiling in live cells with lysine-targeted sulfonyl fluoride probes. *J Am Chem Soc* 139:680–685. <https://doi.org/10.1021/jacs.6b08536>
14. Rüegger J (2025) CellEKT: a robust chemical proteomics workflow to profile cellular target engagement of kinase inhibitors. *Mol. Cell. Proteom.* 100961 <https://doi.org/10.1016/j.mcpro.2025.100961>
15. Scientific Image and Illustration Software | BioRender. <https://www.biorender.com/>
16. Yang T, Cuesta A, Wan X et al (2022) Reversible lysine-targeted probes reveal residence time-based kinase selectivity. *Nat Chem Biol* 1–8. <https://doi.org/10.1038/s41589-022-01019-1>
17. Tang G, Wang W, Zhu C et al (2024) Global reactivity profiling of the catalytic lysine in human kinome for covalent inhibitor development. *Angew Chem Int Ed.* <https://doi.org/10.1002/anie.202316394>
18. Rappsilber J, Mann M, Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2: 1896–1906. <https://doi.org/10.1038/nprot.2007.261>
19. van Rooden EJ, Florea BI, Deng H et al (2018) Mapping in vivo target interaction profiles of covalent inhibitors using chemical proteomics with label-free quantification. *Nat Protoc* 13: 752–767. <https://doi.org/10.1038/nprot.2017.159>
20. Tyanova S, Temu T, Sinitcyn P et al (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* 13:731–740. <https://doi.org/10.1038/nmeth.3901>
21. Manning G, Whyte DB, Martinez R et al (2002) The protein kinase complement of the human genome. *Science* 298:1912–1934. <https://doi.org/10.1126/science.1075762>
22. Metz KS, Deoudes EM, Berginski ME et al (2018) Coral: clear and customizable visualization of human Kinome data. *Cell Syst* 7:347–350.e1. <https://doi.org/10.1016/j.cels.2018.07.001>



Isolation of Exosome from Senescent Preadipocytes for Mass Spectrometric Analysis

Yao Wang

Abstract

Senescent cell-released exosomes play a pivotal role in regulating intercellular communication by transferring bioactive molecules and signaling components. These extracellular vesicles, derived from senescent cells, carry a cargo that includes proteins, lipids, and nucleic acids, influencing the functional status of recipient cells. To facilitate a comprehensive understanding of the intricate molecular mechanisms underlying senescence-associated communication, a detailed and robust protocol is introduced herein. This protocol is specifically designed for the isolation of exosomes derived from senescent cells, laying the groundwork for downstream analyses aimed at unraveling the functional and molecular characteristics of these critical mediators of intercellular signaling.

Key words Senescence, Exosome, C12FDG, Preadipocytes

1 Introduction

Cellular senescence is a hallmark of aging and aging-related diseases, characterized by irreversible cell cycle arrest [1]. Unlike apoptotic cells, senescent cells persist in the body during aging and aging-related diseases, and a recent revelation highlights their role in secreting proteins known as the senescence-associated secretory phenotype (SASP) [2]. This phenotype involves the release of inflammatory cytokines, chemokines, and matrix metalloproteinases into the extracellular space, influencing organismal homeostasis [3]. Senescent cell-released exosomes are part of the SASP, have been reported to affect neighboring and distant cells in different perspectives [4, 5]. The heterogeneity of exosomes arises from the diverse types and functional states of the releasing cells, as well as the various biogenetic pathways involved in their formation. In this context, understanding the physiological and pathological roles

of exosome and its relationship between exosome and senescent-rich microenvironment is crucial for unraveling the mechanisms underlying aging and age-associated diseases [6].

In healthy tissues, endogenous immune surveillance mechanisms naturally restrict the accumulation of senescent cells by specifically eliminating them [7]. If immune surveillance fails to effectively identify and target the removal of senescent cells, it may lead to the accumulation of these cells. Recently, many studies have demonstrated senescent cells escape immune surveillance by regulation of immunoregulatory molecules or ligand to suppress the response of certain immune cells. Indeed, Pereira et al. found that the upregulation of the nonclassical major histocompatibility complex (MHC) molecule HLA-E in the senescent fibroblasts plays an inhibitory role in suppressing cytotoxicity of natural killer (NK) cells and CD8⁺ T cells [8]. The interaction between HLA-E and NK and CD8⁺ T cells is mediated by the inhibitory receptor NKG2A [8]. In addition, senescence-rich microenvironment exhibits an altered immune regulatory response. For example, senescent preadipocytes accumulate in the adipose tissue from the mice fed high-fat-diet (HFD), which might result in the decreasing invariant natural killer T (iNKT) cells [9, 10]. While senescent preadipocytes might directly regulate immunoregulation [11, 12], it is interesting to investigate whether exosome from senescent preadipocytes affect immune cells function.

In this protocol, senescent preadipocytes are isolated from mice subjected to a high-fat diet. Viable senescent cells are identified using a fluorescence-based flow cytometric assay [13], followed by flow cytometry and cell sorting (FACS). Exosomes derived from senescent cells are collected and isolated using the ExoQuick-TC kit [14], employing an ultracentrifugation-free method. Subsequently, the protein content of the isolated exosomes is validated by western blot before being prepared for mass spectrometric analysis. By comparing the protein profiles of senescent and non-senescent cells, specific protein molecules associated with senescence will be identified. To elucidate the immunoregulatory characteristics of senescent cell-derived exosomes, molecules from major histocompatibility complexes (MHC) class I and II will be investigated.

2 Materials

2.1 Primary Cell Isolation

1. Age-matched C57BL6 mice with a chow or high-fat-diet.
2. Collagenase D (11088882001, Roche), prepared into 100 mg/mL with digestion buffer.
3. Dulbecco's phosphate-buffered saline (PBS, 14190144, Gibco).

4. Hanks' Balanced Salt Solution (HBSS, 14025092, Gibco).
5. Bovine serum albumin (BSA, BP1605-100, Fisher BioReagents).
6. Digestion buffer (2% BSA in HBSS).
7. Cell culture medium (DMEM high glucose, 10%FBS, 1% Antibiotic-Antimycotic).
8. Fetal bovine serum (FBS, 100-106-500, Gemini Bio).
9. DMEM high glucose (11965092, Gibco).
10. Antibiotic-Antimycotic (15240096, Gibco).
11. ACK lysis buffer (A1049201, Gibco).

2.2 Senescent Cell Sorting

1. Cell staining buffer (0.5% BSA, 2 mM EDTA in PBS).
2. 5-Dodecanoylaminofluorescein di- β -D-galactopyranoside (C12FDG, ab273642, Abcam).
3. C12FDG stock (16.5 mM in DMSO).
4. Antibodies for staining (CD16/32 blocking antibody; CD45-APC, clone 30-F11; CD31-PE/CY7, clone 390).
5. Viability dye (DAPI, 1 mg/mL, 62248, Thermo Scientific).

2.3 Senescent Cell-Derived Exosome Collection and Isolation

1. Exosome collection medium (DMEM high glucose containing 10% Exo-FBS and 1% Antibiotic-Antimycotic).
2. Exo-FBS (EXO-FBS-50A-1, System Biosciences).
3. ExoQuick-TC (EXOTC50A-1, System Biosciences).

2.4 Western Blot Analysis

1. RIPA Buffer (ab156034, Abcam).
2. Protease and Phosphatase Inhibitor Cocktail (PPC1010, Sigma-Aldrich).
3. Pierce™ BCA Protein Assay Kits (23227, Thermo Fisher).
4. 4 × Bolt™ LDS Sample Buffer (B0008, Thermo Fisher).
5. 1 M DTT (R0861, Thermo Fisher).
6. Precast gel (Bolt™ Bis-Tris Plus Mini Protein Gels, 4–12%, Thermo Fisher).
7. 20X Bolt™ MES SDS Running Buffer (B000202, Thermo Fisher); prepare 1× MES running buffer using deionized water.
8. iBlot™ Transfer Stack, mini-PVDF blot (IB401002, Thermo Fisher).
9. iBlot 2 Gel Transfer Device (IB21001, Thermo Fisher).

2.5 Sample Preparation for Mass Spectrometric Analysis

1. HPLC water (270733, Sigma-Aldrich).
2. Protein lysis buffer (1% sodium deoxycholate (SDC), 100 mM Tris-HCl (PH 8.5), 10 mM Tris (2-chloroethyl) phosphate, and 40 mM chloroacetamide).
3. Trypsin (T7575-1KT, Sigma Millipore).
4. Pierce C18 Spin Columns (89873, Pierce).
5. Desalting activation buffer (50% acetonitrile, ACN, in HPLC water).
6. Desalting equilibration solution (0.5% trifluoroacetic acid, TFA, in 5% ACN).
7. Desalting sample buffer (2% TFA in 20% ACN).
8. Desalting wash solution (0.5% TFA in 5% ACN).
9. Desalting elution buffer (70% ACN in HPLC water).

3 Methods

3.1 Primary Senescent Cell Isolation

1. Euthanize the mouse and harvest the epididymal white adipose tissue (eWAT).
2. Wash eWAT with PBS, and transfer eWAT to HBSS.
3. Chop eWAT into pieces (1 mm² to 2 mm² pieces) and transfer it to a 50 mL tube, and make up the volume to 5 mL with HBSS+1%BSA solution. Add 60 μ L Collagenase D (made in HBSS with 1% BSA) to each tube. *The working concentration of Collagenase D is 1.2 mg/mL.
4. Incubate at 37 °C for 30 min with shaking at 50 rpm.
5. Prewet a 100 μ m cell strainer with cell culture medium and add digested eWAT. Rinse cell strainer with 5 mL cell culture medium.
6. Spin at 400 g for 5 mins at room temperature.
7. Aspirate the supernatant (*see Note 1*), and resuspend the cell pellet with 1 mL of ACK lysis buffer and incubate at room temperature for 3 min.
8. Make up the volume to 10 mL with cell culture medium, and spin at 400 g for 5 mins at room temperature.
9. Remove the supernatant and suspend the cells in 1 mL of cell culture medium and perform a cell count.

3.2 Senescent Cell Sorting

1. For each eWAT from mouse fed an HFD for 16 weeks, approximately 10⁷ cells could obtain. However, C₁₂FDG assay is sensitive to cell concentration. To get a reproducible result, adjust cells into 3 $\times$ 10⁶ cells/mL using cell culture medium.

2. For each $C_{12}FDG$ assay, add 2 μL of 16.5 mM (prepared in DMSO) stock of $C_{12}FDG$ (final conc. 33 μM) to 1 mL cell suspension (*see Note 2*).
3. Mix well by pipetting and incubate at 37 °C for 1 h.
4. After incubation, spin down the cells at 400 g for 5 min at room temperature.
5. Resuspend cells with cell staining buffer and spin down the cells at 400 g for 5 min.
6. Prepare cell staining cocktail: 50 μL cell staining buffer for each sample, which contains CD16/CD32 antibody, 1 $\mu g/mL$ CD45-PE, and 1 $\mu g/mL$ CD31-PECy7. Prepare samples with single-color stain and with no staining for flow cytometry calibration purpose.
7. Resuspend the pellet cells with cell staining cocktail and incubate at 4 °C for 30 mins in the dark.
8. After incubation, add 1 mL cell staining buffer to each sample, and spin down the cells at 400 g for 5 min at room temperature.
9. Remove supernatant and resuspend in 1 mL cell staining buffer, and spin down at 400 g for 5 min at room temperature.
10. Remove supernatant and add cell staining buffer containing DAPI, and transfer to flow tubes.
11. Samples are immediately processed for FACS Aria II Sorter (BD Biosciences). To set flow cytometer by using unstained sample and single-color stained samples. Preadipocytes are identified as DAPI⁻ CD45⁻ CD31⁻ population. To set a threshold for $C_{12}FDG$ staining, samples from mice fed a chow diet are recorded. The sample from mice fed a chow diet is calibrated as 5% of $C_{12}FDG^+$ (*see Note 3*).
12. Senescent preadipocytes derived from mice fed an HFD that identified as $C_{12}FDG^+$ (18%) and sorted into cell culture medium (*see Note 4*).
13. Sorted cells are resuspended in exosome collection medium at a concentration of 10^6 cells/mL and plated into 10 cm Petri dish.
14. Based on median fluorescence intensity of $C_{12}FDG$, we identified a subpopulation of CD45⁻ CD31⁻ preadipocytes from HFD mice that had elevated SA- β gal activity when compared to those from control diet-fed counterparts (Fig. 1).

3.3 EV Collection and Isolation

1. Three days after cell culture, collect supernatants. Centrifuge at 3000 g for 15 mins to separate cells and cell debris from the supernatant.

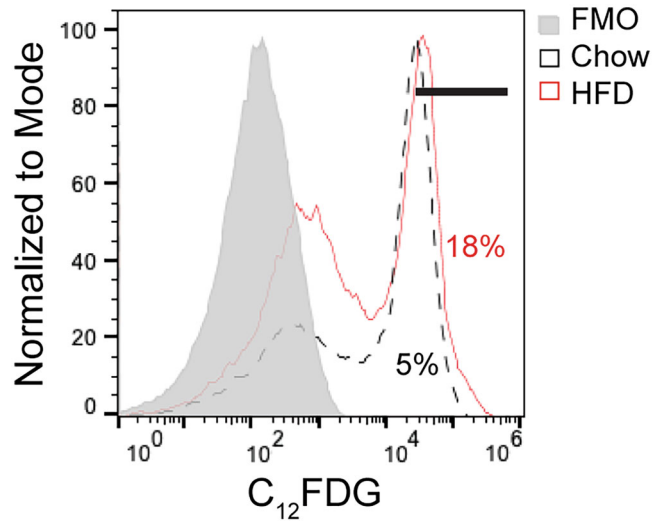


Fig. 1 The representative histogram shows C12FDG staining and gating on the subset with the fluorescence ($C_{12}FDG^+$) in the same $DAPI^-$, $CD45^-$, $CD31^-$ subpopulation. $C_{12}FDG$ fluorescence of the sample from mice fed a chow diet (Chow, black line) or high-fat diet (HFD, red line) was 5% and 18%, respectively

2. Transfer the cleared supernatant to a sterile tube, and add the appropriate volume of ExoQuick-TC directly. Mix thoroughly by inverting or flicking the tube. Incubate the tube at 4 °C overnight (*see Note 5*).
3. After the overnight incubation, centrifuge the ExoQuick-TC/supernatant mixture at 1500 g for 30 min. Following centrifugation, observe the exosomes as a beige or white pellet at the bottom of the tube.
4. Carefully aspirate the supernatant. Spin down any residual ExoQuick-TC solution by centrifugation at 1500 g for 5 min.
5. Remove all remaining fluid by aspiration, ensuring not to disturb the precipitated exosomes in the pellet.
6. Resuspend pellet in RIPA lysis buffer for western blot analysis or in PBS for mass spectrometric analysis.

3.4 Western Blot Analysis

1. Exosome lysates are processed for quantification of protein concentration by BCA assay.
2. Equal amount of protein for each sample are mixed with 4 × Bolt™ LDS Sample Buffer and 1 M DTT and heated at 70 °C for 10 min.
3. Prepare 4–12% Bis-tris precast gel by rinsing wells with 1× MES running buffer, and refill gel tank with 1× MES running buffer.
4. Load sample for each well. Load 10 μL pretained protein marker to first and last lane.

5. Set the voltage at 220 V for 30 min or until blue dye reaches the bottom of the gel.
6. After gel running, open the mini-gel cassette using the gel knife, and discard the plate without the gel, and slowly remove the gel adhered to the other plate.
7. Gently transfer the pre-run gel into deionized water.
8. Unseal one iBlot™ 2 PVDF mini-transfer stack, and separate the top stack and place the bottom stack with the plastic tray directly on the blotting surface.
9. Transfer the pre-run gel on the PVDF membrane of bottom stack and remove air bubbles by blotting roller.
10. Soak the iBlot™ filter paper with deionized water. Place the presoaked iBlot™ filter paper on the pre-run gel. Use the blotting roller to remove any air bubbles between the filter paper and gel.
11. Remove and discard the white plastic separator from the top stack.
12. Take the top stack from the bench and place it on top of the presoaked filter paper with the copper electrode facing up (and transfer gel layer facing down). Remove any air-bubbles using the blotting roller.
13. Place the iBlot™ 2 absorbent pad on top of the iBlot™ 2 Transfer Stack such that the electrical contacts are aligned with the corresponding electrical contacts on the blotting surface of the iBlot™ 2 gel transfer device.
14. Use the blotting roller to flatten any protrusions in the transfer stack.
15. Choose the preset program “P0: 20 V for 1 minute 23 V for 4 mins 25 V for 2 mins.”
16. After gel transfer, wash blot with 1 × TBST for 5 min, and block the blot with 5% milk in TBST for 1 h.
17. After blocking blots are incubated with either 1 µg/mL anti-mouse TSG101 or CD81 antibody in 5% milk in TBST for overnight at 4 °C with shaking.
18. After overnight incubation, blots are washed with 1 × TBST for 5 min for four times.
19. After wash, blots are incubated with HRP-conjugated secondary antibody at a concentration of 1 µg/mL in 5% milk in TBST for 1 h at room temperature.
20. After incubation, blots are washed for four times and incubate with 1 mL Pierce™ ECL Plus Western Blotting Substrate for 1 min before ChemiDoc Imaging.
21. Good quality of exosome protein sample should show a clear band for both antibodies. The band of CD81 and TSG101 should be around 25 kD and 40 kD, respectively (*see Note 6*).

3.5 Mass Spectrometric Analysis Preparation

1. Freshly isolated exosomes are resuspended in 100 μ L PBS, and add 200 μ L lysis buffer containing 1% sodium deoxycholate (SDC), 100 mM Tris-HCl (PH 8.5), 10 mM Tris (2-chloroethyl) phosphate, and 40 mM chloroacetamide, and incubate at 4 °C for 5 min.
2. After lysis, samples are heated at 100 °C for 10 min and centrifuged at 12,000 g, 4 °C for 15 min.
3. Supernatants are collected for protein concentration measurement using BCA assay.
4. Same protein amount for each sample is diluted with double distill water to reduce the SDC concentration to <0.5%.
5. Trypsin is added at a ratio of 1:50 (enzyme: protein) and incubated at 37 °C overnight.
6. After trypsin digestion, an equal volume of 1% formic acid in ethyl acetate is added to quench the trypsin. The mixture is centrifuges at 2500 g for 10 min at room temperature.

During the centrifuge, start to prepare sample desalting using Pierce C18 Spin Columns

7. Activation: Place Pierce C18 Spin Column in a new tube. Each column is activated by 200 μ L of desalting activation buffer and spined down at 1500 g for 1 min. Discard the flow-through. Repeat once using the same tube.
8. Equilibration: Place Pierce C18 Spin Column in a tube. Add 200 μ L of desalting equilibration solution for each column. Centrifuge at 1500 g for 1 min. Discard the flow-through. Repeat once.
9. Sample preparation: After centrifuge, remove supernatant to a new tube, and mix 3 parts sample to 1 part desalting sample buffer. The final sample will contain 0.5% TFA in 5% can.
10. Sample Load: Place Pierce C18 Spin Column in a new tube. Load sample on the top of column and centrifuge at 1500 g for 1 min. Since each Pierce™ C18 Spin Column can process 10–150 μ L of sample, it is necessary to repeat loading sample to ensure all the sample can be bind to the column (*see Note 7*).
11. Wash sample: Place the column into a new tube. Add 200 μ L of desalting wash solution to the column and centrifuge at 1500 g for 1 min. Discard the flow-through. Repeat this once.
12. Elute sample: Place the column in a new tube. Add 20 μ L of desalting elution buffer to the top of the resin bed. Centrifuge at 1500 g for 1 min. Repeat this step in the same collection tube.
13. The eluted peptides are dried in a vacuum centrifuge for spectral acquisition.

4 Notes

1. Carefully discard the supernatant, avoiding cell loss during aspiration. The supernatant contains mature adipocytes and triglycerides.
2. The C₁₂FDG assay for detecting cellular senescence relies on the presence of β -galactosidase, which is enriched in senescent cells, to hydrolyze C₁₂FDG. Following hydrolysis and laser excitation, C₁₂FDG emits green fluorescence, allowing for its detection using flow cytometry.
3. Visualization of C₁₂FDG signal using histogram would help identify the C₁₂FDG-positive population.
4. To enhance downstream analysis, it is crucial to use FBS-free medium when isolating exosomes, as fetal bovine serum inherently contains bovine exosomes.
5. To get a better isolation, tubes should not be rotated or mixed during the incubation period and should remain upright.
6. There are many exosomes specific markers to use for validation. In addition to TSG101 and CD81, other markers, like CD9, Flotillin-1, HSP90, etc., are widely used [15].
7. To ensure complete binding, it is recommended to recover the flow-through and repeat the centrifuge.

Acknowledgments

This work was supported by Diabetes Research Connection and Bakar Aging Research Institute Postdoctoral Fellowship to YW.

References

1. Di Micco R, Krizhanovsky V, Baker D, d'Adda di Fagagna F (2021) Cellular senescence in ageing: from mechanisms to therapeutic opportunities. *Nat Rev Mol Cell Biol* 22:75–95. <https://doi.org/10.1038/s41580-020-00314-w>
2. Muñoz-Espín D, Serrano M (2014) Cellular senescence: from physiology to pathology. *Nat Rev Mol Cell Biol* 15:482–496. <https://doi.org/10.1038/nrm3823>
3. Yousefzadeh MJ, Zhao J, Bukata C et al (2020) Tissue specificity of senescent cell accumulation during physiologic and accelerated aging of mice. *Aging Cell* 19:e13094. <https://doi.org/10.1111/acer.13094>
4. Takasugi M, Okada R, Takahashi A et al (2017) Small extracellular vesicles secreted from senescent cells promote cancer cell proliferation through EphA2. *Nat Commun* 8:15729. <https://doi.org/10.1038/ncomms15728>
5. Wallis R, Mizen H, Bishop CL (2020) The bright and dark side of extracellular vesicles in the senescence-associated secretory phenotype. *Mech Ageing Dev* 189:111263. <https://doi.org/10.1016/j.mad.2020.111263>
6. Palmer AK, Tchkonja T, Kirkland JL (2022) Targeting cellular senescence in metabolic disease. *Mol Metab* 66:101601. <https://doi.org/10.1016/j.molmet.2022.101601>
7. Kale A, Sharma A, Stolzing A et al (2020) Role of immune cells in the removal of deleterious senescent cells. *Immun Ageing* 17:16. <https://doi.org/10.1186/s12979-020-00187-9>

8. Pereira BI, Devine OP, Vukmanovic-Stejjic M et al (2019) Senescent cells evade immune clearance via HLA-E-mediated NK and CD8+ T cell inhibition. *Nat Commun* 10:2387. <https://doi.org/10.1038/s41467-019-10335-5>
9. Arora S, Thompson PJ, Wang Y et al (2021) Invariant natural killer T cells coordinate removal of senescent cells. *Med* 2:938–950. <https://doi.org/10.1016/j.medj.2021.04.014>
10. Lynch L, Nowak M, Varghese B et al (2012) Adipose tissue invariant NKT cells protect against diet-induced obesity and metabolic disorder through regulatory cytokine production. *Immunity* 37:574–587. <https://doi.org/10.1016/j.immuni.2012.06.016>
11. Xiao L, Yang X, Lin Y et al (2016) Large adipocytes function as antigen-presenting cells to activate CD4(+) T cells via upregulating MHCII in obesity. *Int J Obes* 40:112–120. <https://doi.org/10.1038/ijo.2015.145>
12. Park J, Huh JY, Oh J et al (2019) Activation of invariant natural killer T cells stimulates adipose tissue remodeling via adipocyte death and birth in obesity. *Genes Dev* 33:1657–1672. <https://doi.org/10.1101/gad.329557.119>
13. Cahu J, Sola B (2013) A sensitive method to quantify senescent cancer cells. *J Vis Exp* 50494. <https://doi.org/10.3791/50494>
14. Kim M, Son IT, Noh GT et al (2023) Exosomes derived from colon cancer cells promote tumor progression and affect the tumor micro-environment. *J Clin Med* 12:3905. <https://doi.org/10.3390/jcm12123905>
15. Deng F, Miller J (2019) A review on protein markers of exosome from different bio-resources and the antibodies used for characterization. *J Histotechnol* 42:226–239. <https://doi.org/10.1080/01478885.2019.1646984>



Identification of Anticancer ROS Targets by Cysteine Reactivity Protein Profiling

Junbing Zhang and Liron Bar-Peled

Abstract

Cysteine reactivity profiling is recognized as a powerful and versatile chemoproteomic technology to study the cysteinome in native biological systems. It involves the use of chemical probes that covalently bind to reactive cysteines, enabling the enrichment and quantitative analysis of cysteine abundance through mass spectrometry. Cysteines play critical roles in maintaining cellular redox homeostasis by dynamically switching between their reduced and oxidized forms during redox perturbations. Here, we described a detailed workflow of sample preparation for cysteine-based chemical proteomics, to systematically identify cysteines that are modified by anticancer drug induced ROS.

Key words Anticancer drugs, ROS, Cysteine, Chemical proteomics, Posttranslational modifications

1 Introduction

Reactive oxygen species (ROS) are reactive small molecules derived from O₂, which can be generated during normal metabolism [1–3]. Maintaining redox homeostasis is a fundamental necessity for all living organisms. In humans, deviations in ROS generating and detoxification pathways can result in major human pathologies, like cancer and neurodegeneration [4]. Compared to normal cells, cancer cells usually have high proliferation rates and exist in a more stressful environment. As a result, cancer cells have rewired core metabolic pathways to maintain unnatural levels of biosynthetic molecules and energy [5]. This rewiring creates toxic metabolic byproducts, some of which are ROS. Due to their highly reactive nature, ROS can go on to react with major macromolecules (to differing degrees), like DNA, lipids, and proteins, which provokes lethal cellular outcomes [6]. Accordingly, cancers have heightened needs for redox control. Therefore, approaches that overcome redox control and elevate ROS levels in malignant cells hold considerable therapeutic potential [4, 7]. Moreover, multiple

anticancer drugs have been suggested to kill cancer cells partly by increase the steady-state ROS levels [1, 8, 9]. However, a critical hurdle in the development or improvement of ROS-based therapies is the knowledge gap in the molecular mechanisms of ROS targets and subsequent ROS-mediated cell death [10].

Although much of the ROS cytotoxicity in cancer cells has been attributed to oxidative DNA damage, there is growing appreciation that ROS mediated post-translation modifications (PTMs) of key signaling proteins also contribute to their cytotoxicity [7, 11]. Cysteine is one of the most nucleophilic amino acids and, as a result, can be modified by ROS into different redox forms, including sulfenylation (SOH), sulfinylation (SO₂H), and disulfide formation. Many of these modifications are reversible endowing cysteines as redox sensors, which play key roles in response to increased ROS levels and maintenance of redox homeostasis [12]. However, compared with other kinds of protein modifications, knowledge of ROS-modified protein targets is still at its infancy. The global dissection of ROS modified proteins has been advanced using chemical proteomic technologies [13–17], including isoTOP-ABPP (isotopic tandem orthogonal proteolysis–activity-based protein profiling). The isoTOP-ABPP platform was pioneered by the Cravatt laboratory [18, 19], which uses reactive probes designed to specifically bind protein active sites in complex biological systems [20]. In general, the reactive probes include three parts: (1) a warhead covalently labeling reactive cysteine of target proteins; (2) a tag for affinity purification, like biotin; (3) a linker to minimize the steric hindrance between warhead and affinity tag [21, 22]. The reactive cysteine-based isoTOP-ABPP platform has been used for studies in covalent drug discovery, reactive metabolite modifications, and small-molecule screening, which accelerate the discovery of new lead compounds and druggable hotspots for human diseases [21, 23, 24, 25–28].

This chapter describes the workflow for using multiplexed isoTOP-ABPP to systematically map cysteine reactivity following treatment with anticancer agents that alter steady-state ROS levels [8].

2 Materials

2.1 Cell Culture and Drug Treatment

1. Cell line: K562.
2. Cell medium: RPMI-160 supplemented with 10% fetal bovine serum, Glutamax I, and penicillin/streptomycin.
3. DMSO.
4. β -lapachone (LAP).
5. Auranofin (AUR).

6. Arsenic trioxide (ATO).
7. Doxorubicin (DOXO).
8. Elesclomol (ELC).
9. Cisplatin (DDP).
10. Bleomycin (BLE).
11. 5-Fluorouracil (5FU).
12. Sulfasalazine (SUL).
13. 2-Methoxyestradiol (2-ME).
14. NOV-002 (NOV).
15. CellTiter-Glo.
16. Dulbecco's phosphate buffered saline (DPBS).
17. 0.9% NaCl.
18. T75 tissue culture flask.
19. T175 flask.
20. 96-well plate.
21. 50 mL centrifuge tubes.
22. 1.5 mL microcentrifuge tubes.
23. 37 °C incubator.
24. Megafuge 8R.
25. Microcentrifuge.

2.2 Cell Lysis and Probe Labeling

1. Iodoacetamide polyethyleneoxide desthiobiotin (IA-DTB).
2. 2-Iodacetamid.
3. IGEPAL[®] CA-630.
4. 15 mL centrifuge tubes.
5. Bradford solution.
6. Triton-X100.
7. Protease Inhibitor Cocktail.
8. Benzonase.
9. DTT.

2.3 Protein Extraction

1. HPLC-grade water.
2. HPLC-grade methanol.
3. HPLC-grade chloroform.

2.4 Trypsinization, Peptide Enrichment, and TMT-Labeling

1. Trypsin protease mix, MS grade.
2. Hepes.
3. HPLC-grade acetonitrile.
4. Urea.

5. Triethylammonium bicarbonate (TEAB) buffer.
6. Pierce™ Streptavidin Agarose.
7. 10% formic acid.
8. C18 Micro-Spin Columns.
9. 1.5 mL microcentrifuge tubes.
10. Isobaric tandem mass-tag (TMT) reagents.

3 Methods

3.1 Cell Treatment with Anticancer Drugs

1. Two thousand five hundred K562 cells were seeded into 96-well plate and treated with different concentrations of anti-cancer drugs (ATO, AUR, DDP, LAP, NOV, 2-ME, SUL, 5FU, DOXO, ELC and BLE).
2. Four days after drug treatment, IC₅₀ values were determined using CellTiter-Glo reagent.
3. K562 cells were maintained in T175 flask at a cell density of 0.25×10^6 to 1×10^6 per mL. All the K562 cells were pooled together (*see Note 1*), and 30 million cells were treated with vehicle control or anticancer drugs at ~5X their IC₅₀ concentrations for 3 h (*see Note 2*).
4. Cells were harvested in 50 mL centrifuge tubes and spin down at 1500 rpm for 2 min.
5. Cells were washed with cold DPBS once and transfer to 1.5 mL EP tube.
6. Spin down at 1200 g for 2 min and then snap-freeze the cell pellet in liquid nitrogen; store at -80°C .

3.2 Cell Lysis

1. Prepare lysis buffer (PBS + 1:1000 Benzonase) on ice (*see Note 3*).
2. Thaw K562 cells, add 250 μL lysis buffer, and leave cell pellet on ice to defrost; pipette up and down to resuspend the cell pellet.
3. Sonicate samples at 4°C for a total of 10 minutes (pulse on 30 seconds, pulse off 30 seconds for a total of 20 minutes).
4. Rotate samples end-over-end for 30 minutes at 4°C .
5. 300 g for 3 min at 4°C (soft centrifugation to get rid of membrane proteins).
6. Transfer supernatant into a new 1.5 mL EP tube on ice.
7. Measure total protein concentration with Bradford assay (2 μL of sample in 1 mL 1:5 Bradford solution); normalize protein concentration of each sample to 2 mg/mL.

3.3 Probe Labeling

1. Take 500 μ L protein lysate per sample, add 5 μ L IA-DTB (10 mM stock in DMSO, final concentration 100 μ M) for each sample, and rotate for 1 h at ambient temperature (*see Note 4*).

3.4 Protein Extraction

1. Prepare labeled 15 mL falcon tubes during the probe incubation time (use the polypropylene ones).
2. Add 4 mL MeOH, 3.5 mL autoclave water and 1 mL chloroform into falcon tubes above.
3. Stop probe labeling by adding the samples (spin samples down briefly before adding) into the mixture from **step 2**.
4. Vortex each sample for 5 s each (make sure whole sample is cloudy).
5. Spin down in centrifuge at 4200 rpm for 10 min, there will be a white disc at interface between aqueous and chloroform layer.
6. Aspirate aqueous and chloroform layers, and be careful to not aspirate the white disc (this contains the proteins) (*see Note 5*).
7. Add 1 mL of MeOH to each disc, and make sure disc is fully submerged in MeOH.
8. Sonicate 15 pulses until protein disc is fully broken up for each sample.
9. Transfer MeOH containing proteins to proteomics Eppendorf tubes; spin down samples at max speed (13,000 rpm) at 4 °C for 10 min.
10. Aspirate MeOH, leave enough MeOH to cover top of pellet, and store in -80°C (*see Note 6*).

3.5 Reduction, Alkylation and Trypsinization

1. Make a fresh stock of buffer A (9 M urea, 10 mM DTT, and 50 mM Hepes) (*see Note 7*).
2. Spin down samples from -80°C for 10 min at 4 °C with max speed, and aspirate methanol.
3. Add 180 μ L of buffer A solution to each sample.
4. Resuspend the pellets by thorough pipetting up and down or sonication, briefly spin down before heating.
5. Heat samples at 65 °C Celsius for 20 min.
6. Briefly spin down samples, let them cool down to room temperature, add 20 μ L of 500 mM Iodoacetamide, and shake in the dark at 37 °C for 30 min.
7. After 30 min incubation, transfer samples into new low affinity Eppendorf tubes.
8. Add 700 μ L 50 mM Hepes (1:20 dilution of 1 M Hepes) to old tubes, vortex, briefly spin down, and then transfer into new Eppendorf tubes, along with the rest of the sample (*see Note 8*).

9. Add trypsin solution (10 uL of 0.5 µg/µL trypsin in trypsin buffer, containing 25 mM CaCl₂) made above to each sample; incubate in shaker at 37 °C Celsius for 3 h or overnight.

3.6 Peptide Enrichment with Streptavidin Beads

1. Prepare the streptavidin beads before the trypsin digestion ends. Use 40 uL of 50:50 beads:slurry per sample (20 uL beads:20 uL slurry). Wash beads/slurry tubes 3X times with 1 mL wash buffer (1xTBS, 1:1000 IGEPAL).
2. After beads have been washed, add wash buffer to make 50:50 beads/wash buffer, mix the beads and wash buffer very well.
3. Cut enough tips (200 uL non-filter tips) for all samples.
4. After trypsinization, add 40 µL of washed streptavidin beads to each sample, invert streptavidin beads in between each sample to ensure beads are evenly distributed throughout wash buffer (*see Note 9*).
5. Let samples rotate end-over-end with streptavidin beads at room temperature for 2 h.

3.7 Peptide Elution

1. Prepare labeled Bio-Spin column accordingly.
2. Washed 2X with 1 mL of wash buffer (1XTBS, 1:1000 IGEPAL), 2X with 1 mL PBS, 2X with 1 mL HPLC-grade H₂O (*see Note 10*).
3. Samples were transferred to Bio-Spin columns, and peptides were eluted following the addition of 500 uL of elution buffer (50% ACN, 0.1% FA). Sample was allowed to gravity drip into new tube, followed by applied pressure from the top to dry the beads using bulb.
4. The eluate was then evaporated to dryness in SpeedVac overnight.

3.8 TMT Labeling

1. Resuspend the peptide of each sample with 30 uL of buffer B (200 mM EPPS (pH 8.4), 30% ACN), vortex, quick spin down, and then sonicate for 5 min in Branson sonicator bath.
2. Add 4.0 µL (stock concentration is 2.5 µg/ul) corresponding TMT tag (126, 127 N, 127C, 128 N, 128C, 129 N, 129C, 130 N, 130C, 131 N, 131C, 132 N, 132C, 133 N, 133C, 134 N) to each peptide solution (*see Note 11*).
3. Short vortex, brief spin down, and incubate the reaction at room temperature for 1 h.
4. Quench the reaction by adding 4.5 µL 5% hydroxylamine to each sample. Incubate at room temperature for 15 min.
5. Add 1% of formic acid to acidify the mixture. Spin down and combine each 16-plex into a low bind EP tube.
6. Desalt the mixture over a C18 Sep-Pak column.

Table 1
TMT tags used (Only ATO, LAP, and AUR were showed here)

Sample #	Sample name	TMT label
1	DMSO	TMTpro-126
2	DMSO	TMTpro-127 N
3	DMSO	TMTpro-127C
4	Saline	TMTpro-128 N
5	Saline	TMTpro-128C
6	Saline	TMTpro-129 N
7	ATO	TMTpro-129C
8	ATO	TMTpro-130 N
9	ATO	TMTpro-130C
10	LAP	TMTpro-131 N
11	LAP	TMTpro-131C
12	LAP	TMTpro-132 N
13	AUR	TMTpro-132C
14	AUR	TMTpro-133 N
15	AUR	TMTpro-133C
16	AUR	TMTpro-134 N

7. Place in SpeedVac and speed vac to dryness.
8. Fractionate the sample using high-pH reversed-phase high-pressure liquid chromatography (HpRP) kit as instructed and analyzed on an Orbitrap Fusion Lumos or Orbitrap Eclipse mass spectrometer (Thermo Fisher Scientific) coupled with a Proxeon EASY-nLC 1200 liquid chromatography (LC) pump (Thermo Fisher Scientific) (Table 1).

4 Notes

1. Pool all the cells together and then divide into different groups for the drug treatment, which will decrease the variation between different groups.
2. Concentration for drug treatment here depends on the specific purpose.
3. Benzonase helps destroy the nucleus of cells.
4. IA-DTB is light sensitive and makes a stock solution in DMSO (10 mM), and store aliquots at -80°C in the absence of light, avoiding frequent freeze and defrost.

5. If the disk does not form properly, aspirate just the top layer.
6. The protein pellets can be frozen here at -80°C for future use.
7. Urea and DTT should be made fresh on the day of the experiment.
8. Dilution of urea to lower than 2.0 M concentration is critical to ensure proper enzymatic activity of trypsin.
9. Make sure an equal amount of streptavidin beads was added to each sample.
10. Solution doesn't really go through the column. Vacuum can be applied from the bottom with a cut tip on the in-house line vacuum line.
11. TMT tags are in 100% dry ACN; do not leave tubes open for extended periods of time, and quickly pipette tags into peptide solution.

References

1. Chio IIC, Tuveson DA (2017) ROS in cancer: the burning question. *Trends Mol Med* 23(5): 411–429. <https://doi.org/10.1016/j.molmed.2017.03.004>
2. Sies H, Jones DP (2020) Reactive oxygen species (ROS) as pleiotropic physiological signaling agents. *Nat Rev Mol Cell Biol* 21(7): 363–383. <https://doi.org/10.1038/s41580-020-0230-3>
3. Murphy MP (2009) How mitochondria produce reactive oxygen species. *Biochem J* 417(1):1–13. <https://doi.org/10.1042/BJ20081386>
4. Tretter V, Hochreiter B, Zach ML, Krenn K, Klein KU (2021) Understanding cellular redox homeostasis: a challenge for precision medicine. *Int J Mol Sci* 23(1). <https://doi.org/10.3390/ijms23010106>
5. Chandel NS, Tuveson DA (2014) The promise and perils of antioxidants for cancer patients. *N Engl J Med* 371(2):177–178. <https://doi.org/10.1056/NEJMcibr1405701>
6. Wu K, El Zowalaty AE, Sayin VI, Papagiannakopoulos T (2024) The pleiotropic functions of reactive oxygen species in cancer. *Nat Cancer* 5(3):384–399. <https://doi.org/10.1038/s43018-024-00738-9>
7. Reczek CR, Chandel NS (2015) ROS-dependent signal transduction. *Curr Opin Cell Biol* 33:8–13. <https://doi.org/10.1016/j.ccb.2014.09.010>
8. Zhang J, Simpson CM, Berner J, Chong HB, Fang J, Ordulu Z, Weiss-Sadan T, Possemato AP, Harry S, Takahashi M, Yang TY, Richter M, Patel H, Smith AE, Carlin AD, Hubertus de Groot AF, Wolf K, Shi L, Wei TY, Durr BR, Chen NJ, Vornbaumen T, Wichmann NO, Mahamdeh MS, Pooladanda V, Matoba Y, Kumar S, Kim E, Bouberhan S, Oliva E, Rueda BR, Soberman RJ, Bardeesy N, Liao BB, Lawrence M, Stokes MP, Beausoleil SA, Bar-Peled L (2023) Systematic identification of anticancer drug targets reveals a nucleus-to-mitochondria ROS-sensing pathway. *Cell* 186(11): 2361–2379 e2325. <https://doi.org/10.1016/j.cell.2023.04.026>
9. Yang H, Villani RM, Wang H, Simpson MJ, Roberts MS, Tang M, Liang X (2018) The role of cellular reactive oxygen species in cancer chemotherapy. *J Exp Clin Cancer Res* 37(1): 266. <https://doi.org/10.1186/s13046-018-0909-x>
10. Ranasinghe R, Mathai ML, Zulli A (2022) Cisplatin for cancer therapy and overcoming chemoresistance. *Heliyon* 8(9):e10608. <https://doi.org/10.1016/j.heliyon.2022.e10608>
11. Collins Y, Chouchani ET, James AM, Menger KE, Cocheme HM, Murphy MP (2012) Mitochondrial redox signalling at a glance. *J Cell Sci* 125(Pt 4):801–806. <https://doi.org/10.1242/jcs.098475>
12. Zhang J, Bar-Peled L (2023) Chemical biology approaches to uncovering nuclear ROS control. *Curr Opin Chem Biol* 76:102352. <https://doi.org/10.1016/j.cbpa.2023.102352>
13. Zhang N, Wu J, Zheng Q (2024) Chemical proteomics approaches for protein post-

- translational modification studies. *Biochim Biophys Acta Proteins Proteom* 1872(4): 141017. <https://doi.org/10.1016/j.bbapap.2024.141017>
14. Kisty EA, Falco JA, Weerapana E (2023) Redox proteomics combined with proximity labeling enables monitoring of localized cysteine oxidation in cells. *Cell Chem Biol* 30(3):321–336 e326. <https://doi.org/10.1016/j.chembiol.2023.02.006>
 15. Ali MY, Bar-Peled L (2024) Chemical proteomics to study metabolism, a reductionist approach applied at the systems level. *Cell Chem Biol* 31(3):446–451. <https://doi.org/10.1016/j.chembiol.2024.02.015>
 16. Burger N, Chouchani ET (2024) A new era of cysteine proteomics – technological advances in thiol biology. *Curr Opin Chem Biol* 79: 102435. <https://doi.org/10.1016/j.cbpa.2024.102435>
 17. Shi Y, Carroll KS (2020) Activity-based sensing for site-specific proteomic analysis of cysteine oxidation. *Acc Chem Res* 53(1):20–31. <https://doi.org/10.1021/acs.accounts.9b00562>
 18. Liu Y, Patricelli MP, Cravatt BF (1999) Activity-based protein profiling: the serine hydrolases. *Proc Natl Acad Sci USA* 96(26): 14694–14699. <https://doi.org/10.1073/pnas.96.26.14694>
 19. Nomura DK, Dix MM, Cravatt BF (2010) Activity-based protein profiling for biochemical pathway discovery in cancer. *Nat Rev Cancer* 10(9):630–638. <https://doi.org/10.1038/nrc2901nrc2901>
 20. Weerapana E, Wang C, Simon GM, Richter F, Khare S, Dillon MB, Bachovchin DA, Mowen K, Baker D, Cravatt BF (2010) Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* 468(7325): 790–795. <https://doi.org/10.1038/nature09472nature09472>
 21. Maurais AJ, Weerapana E (2019) Reactive-cysteine profiling for drug discovery. *Curr Opin Chem Biol* 50:29–36. <https://doi.org/10.1016/j.cbpa.2019.02.010>
 22. Niphakis MJ, Cravatt BF (2014) Enzyme inhibitor discovery by activity-based protein profiling. *Annu Rev Biochem* 83:341–377. <https://doi.org/10.1146/annurev-biochem-060713-035708>
 23. Bar-Peled L, Kemper EK, Suciú RM, Vinogradova EV, Backus KM, Horning BD, Paul TA, Ichu TA, Svensson RU, Olucha J, Chang MW, Kok BP, Zhu Z, Ihle NT, Dix MM, Jiang P, Hayward MM, Saez E, Shaw RJ, Cravatt BF (2017) Chemical proteomics identifies Drug-gable vulnerabilities in a genetically defined cancer. *Cell* 171(3):696–709 e623. <https://doi.org/10.1016/j.cell.2017.08.051>
 24. Vinogradova EV, Zhang X, Remillard D, Lazar DC, Suciú RM, Wang Y, Bianco G, Yamashita Y, Crowley VM, Schafroth MA, Yokoyama M, Konrad DB, Lum KM, Simon GM, Kemper EK, Lazear MR, Yin S, Blewett MM, Dix MM, Nguyen N, Shokhirev MN, Chin EN, Lairson LL, Melillo B, Schreiber SL, Forli S, Teijaro JR, Cravatt BF (2020) An activity-guided map of electrophile-cysteine interactions in primary human T cells. *Cell* 182(4):1009–1026. e1029. <https://doi.org/10.1016/j.cell.2020.07.001>
 25. Bachovchin DA, Cravatt BF (2012) The pharmacological landscape and therapeutic potential of serine hydrolases. *Nat Rev Drug Discov* 11(1):52–68. <https://doi.org/10.1038/nrd3620>
 26. Backus KM (2019) Applications of reactive cysteine profiling. *Curr Top Microbiol Immunol* 420:375–417. https://doi.org/10.1007/82_2018_120
 27. Takahashi M, Chong HB, Zhang S, Yang TY, Lazarov MJ, Harry S, Maynard M, Hilbert B, White RD, Murrey HE, Tsou CC, Vordermark K, Assaad J, Gohar M, Durr BR, Richter M, Patel H, Kryukov G, Brooijmans N, Alghali ASO, Rubio K, Villanueva A, Zhang J, Ge M, Makram F, Griesshaber H, Harrison D, Koglin AS, Ojeda S, Karakyriakou B, Healy A, Popoola G, Rachmin I, Khandelwal N, Neil JR, Tien PC, Chen N, Hosp T, van den Ouweland S, Hara T, Bussema L, Dong R, Shi L, Rasmussen MQ, Domingues AC, Lawless A, Fang J, Yoda S, Nguyen LP, Reeves SM, Wakefield FN, Acker A, Clark SE, Dubash T, Kastanos J, Oh E, Fisher DE, Maheswaran S, Haber DA, Boland GM, Sade-Feldman M, Jenkins RW, Hata AN, Bardeesy NM, Suva ML, Martin BR, Liao BB, Ott CJ, Rivera MN, Lawrence MS, Bar-Peled L (2024) DrugMap: a quantitative pan-cancer analysis of cysteine ligandability. *Cell* 187(10): 2536–2556. e2530. <https://doi.org/10.1016/j.cell.2024.03.027>
 28. Esther K, Kemper Yuanjin, Zhang Melissa M, Dix Benjamin F Cravatt (2022) Global profiling of phosphorylation-dependent changes in cysteine reactivity *Nature Methods* 19(3):341–352. <https://doi.org/10.1038/s41592-022-01398-2>



A Comprehensive Data Processing Pipeline for Phosphoproteomics in Viral Infection Studies

Sini Huuskonen, Xiaonan Liu, Kari Salokas, Antti Tuhkala, Salla Keskitalo, and Markku Varjosalo

Abstract

This protocol outlines a comprehensive methodology for investigating cellular responses to viral infections through quantitative global and phosphoproteome analysis. Phosphorylation, a crucial posttranslational modification, regulates various cellular processes and plays a key role in virus-host interactions. To study the phosphorylation events, we have developed an analysis pipeline, and the subsequent steps detail sample preparation techniques for obtaining high-quality protein samples suitable for phosphoproteomic analysis. These include lysing virus-infected cells, processing protein samples through reduction, alkylation, digestion, and desalting, and enriching phosphopeptides using CAE-Ti-IMAC microspheres. High-throughput DIA-PASEF MS-analysis using TimsTOF Pro 2 facilitates sensitive detection of phosphorylation events, enabling a comprehensive understanding of viral infection, replication, and cytotoxicity in the host cell. The protocol also integrates the DIA-NN program for processing and analyzing MS data. Overall, this protocol provides a robust framework for identifying potential therapeutic targets within these altered pathways for elucidating the molecular mechanisms underlying viral pathogenesis and host immune evasion.

Key words Phosphoproteomics, Viral infections, Posttranslational modifications (PTMs), Protein kinases, Virus-host interactions, SARS-CoV-2, Phosphorylation, Phosphopeptide enrichment, DIA-PASEF MS-analysis

1 Introduction

Phosphopeptide enrichment coupled with LC-MS represents a methodology for investigating the cellular signaling, such as the host cell responses to viral infection, through the analysis of phosphorylated peptides and proteins [1–3]. This approach often utilizes the enrichment of phosphopeptides and parallel global proteomic analyses, thereby allowing for the detection of modified amino acids, proteins, and pathways affected by viral agents

The order of the authors corresponds to their academic contribution to this paper.

[1, 4]. Phosphorylation events, typically mediated by protein kinases, are persistent during viral infections [5, 6]. Furthermore, viral entities have evolved strategies to exploit host phosphorylation machinery, often by mimicking cellular proteins or harboring motifs resembling phosphorylation sites [5–8]. This facilitates their interaction with host kinases, culminating in the hijacking of cellular signaling pathways and modulation of host cell functions to facilitate viral replication and dissemination [9, 10].

The protocol outlined here describes a systematic approach for phosphoproteomic analysis to study host cell responses to viral infections, laying the foundation for subsequent drug discovery processes (Fig. 1). One of the many benefits of this protocol is the enrichment of phosphorylated peptides that are of low abundance under atypical sample conditions, thus enhancing the detection of crucial phosphorylation events during viral infection. To demonstrate the usability and functionality of the pipeline, we use SARS-CoV-2 infected and uninfected mock VeroE6-TMPRSS2-H10 cells. Cells are harvested at various post-infection time points to observe the spatial progression of viral infection within the cell. The samples undergo a series of steps including deactivation of the virus via Urea lysis, protein concentration determination, reduction, alkylation, digestion, and peptide desalting [11], ultimately resulting in the peptide samples for both global proteome and subsequently after phosphopeptide enrichment and phosphoproteome analysis.

Particularly for phosphopeptide enrichment, employing CAE-Ti-IMAC (Carboxyamidomethyl-Modified Titanium Dioxide Immobilized Metal Affinity Chromatography) microspheres ensures the efficient capture of phosphorylation events (i.e., phosphorylated peptides) from complex cellular mixtures. CAE-Ti-IMAC microspheres work based on immobilized metal affinity chromatography (IMAC), which exploits the affinity of metal ions, particularly titanium (Ti), for phosphorylated peptides. These titanium ions have a high affinity for phosphate groups, allowing them to selectively bind phosphorylated peptides from a mixture of peptides. The CAE modification enhances the binding capacity of the microspheres by introducing carboxyamidomethyl (CAE) groups, which provide additional binding sites for the negatively charged phosphate groups on phosphorylated peptides. Non-phosphorylated peptides are washed away, leaving behind only the enriched phosphopeptides bound to the microspheres [12].

Subsequently, DIA-PASEF MS-analysis is conducted on high-end mass spectrometry instruments, such as the TimsTOF Pro 2 from Bruker coupled with the Evosep ONE LC system. This setup facilitates a comprehensive and efficient quantitative exploration of phosphorylation events, significantly enhancing our comprehension of complex host cell responses to viral infection at the

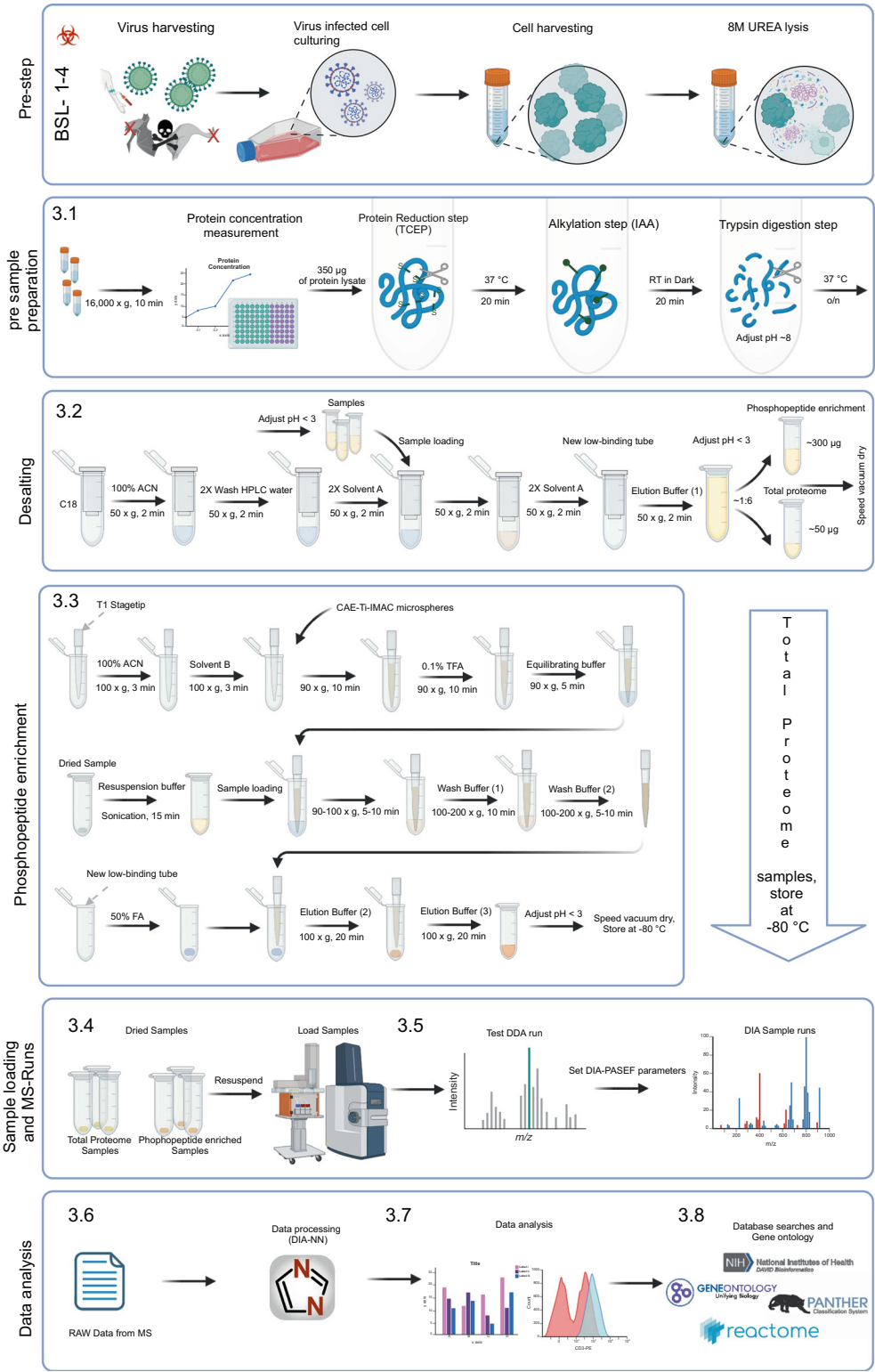


Fig. 1 Protocol for comprehensive quantitative proteomic and phosphoproteomic analysis of viral infection. The pipeline begins with a pre-step, where various sources of viruses are detected, identified, and extracted for further experimentation. These are then used for infecting cell models, such as Vero-E6-H10 cells, and cell

molecular level. In DIA, instead of isolating and analyzing individual ions one by one (as in data-dependent acquisition, or DDA), the mass spectrometer fragments all ions within a predefined mass range simultaneously. This generates a comprehensive set of fragment ions. In turn, PASEF is a technique where ions are accumulated in the trap, then released in packets and sequentially fragmented as they traverse through the collision cell. This allows for the fragmentation of multiple ions in parallel while maintaining their separation based on their ion mobility. This approach allows detection of alterations in phosphorylation patterns across a wide range of proteins [13].

The MS results are processed using the DIA-NN program, a computational tool specifically designed for analyzing data from DIA mass spectrometry experiments. DIA-NN facilitates the identification and quantification, with minimal missing values, of peptides and proteins in complex biological samples. While DIA-NN is the primary tool utilized in the protocol, other computational tools may also be employed. The accrued data is subsequently refined through analytical methods and further analyzed using data analysis techniques [14].

Through this approach, we can decipher the complex network of phosphorylation events modulated by viral infection(s), unraveling key amino acid residues modified on proteins and subsequently the signaling pathways involved in viral replication and host response. In conclusion, the protocol described here provides a comprehensive framework for the molecular-level study of viral infections, utilizing phosphoproteomic techniques to study the interplay between viruses and host cells. A more in-depth demonstration of the usability and results can be found in ‘The comprehensive SARS-CoV-2 “hijackome” knowledge base’ article, where we map intricate interplay of the SARS-CoV-2 variants’

Fig. 1 (continued) samples are collected at various time points to monitor viral infection progression. Subsequently, the collected samples are lysed in 8 M Urea to deactivate the viruses, allowing for safe processing under normal BSL-1 conditions. (3.1) Pre-sample preparation involves measuring the sample's protein concentration. A total of 350 µg of protein is then used for reduction, followed by alkylation and trypsin digestion. (3.2) Samples are desalted using BioPureSNP Macro C18 Columns. After desalting, the samples are split into two portions: total proteome sample (50 µg) and 300 µg for samples undergoing phosphopeptide enrichment. The samples are then speed-vacuum dried and stored at –80 °C until use. (3.3) Phosphopeptide enrichment is performed using T1 stage tips containing C8 material. CAE-Ti-IMAC microspheres are loaded onto the stage tips and conditioned. Resuspended samples are then loaded and washed, and phospho-enriched peptides are eluted and dried before running them in MS. (3.4) Dried total proteome and phosphopeptide enriched samples are resuspended and loaded onto MS using the manufacturer's loading protocol. (3.5) First, a test DDA run is performed to set final DIA-PASEF parameters for the samples. Subsequently, the samples are run using parameters mentioned in this paper. (3.6) Extracted RAW data is then processed using the DIA-NN program. (3.7) Data is normalized and imputed. (3.8) Finally, the data can be compared to preexisting information and databases for further validation and analysis

evolutionary changes on virus behavior in host cells [15]. This developed pipeline and protocol lays the groundwork for identifying potential therapeutic targets and developing novel antiviral strategies.

2 Materials

1. 0.1% trifluoroacetic acid (TFA).
2. 100 mM ammonium bicarbonate (ABC; Sigma-Aldrich).
3. 50% formic acid (FA).
4. Acetonitrile (ACN).
5. BCA protein assay kit (Thermo Fisher).
6. C18 Macro SpinColumns (The Nest Group, Inc.).
7. CAE-Ti-IMAC microsphere slurry.
8. Centrifuge capable of reaching $16,000\times g$, set to 4 °C.
9. Elution buffer 1: 80% ACN, 0.1% TFA in HPLC water.
10. Elution buffer 2: 10% ammonia.
11. Elution buffer 3: 80% ACN + 2% FA (40 mL ACN + 1 mL FA + 9 mL H₂O).
12. Equilibration buffer: 80% ACN + 6% TFA (40 mL ACN + 3 mL TFA + 7 mL H₂O).
13. Iodoacetamide (IAA).
14. Low-binding Eppendorf tubes.
15. Lysis buffer consisting of 8 M Urea, 50 mM NH₄HCO₃, and 1X phosphatase and protease inhibitors cocktail (e.g., Thermo Fisher Halt Protease and Phosphatase Inhibitor Cocktail).
16. Resuspension buffer: 90% ACN, 10% TFA (45 mL ACN + 5 mL TFA) (*see Note 1*).
17. Sequencing Grade Modified Trypsin (Promega).
18. Solvent A: 0.1% FA in H₂O.
19. Solvent B: 0.1% FA in 80% ACN.
20. Tris(2-carboxyethyl)phosphine (TCEP; Sigma Aldrich).
21. Wash buffer 1: 50% ACN, 6% TFA, 200 mM NaCl (25 mL ACN + 3 mL TFA + 2 mL 5 M NaCl + 20 mL H₂O).
22. Wash buffer 2: 30% ACN, 0.1% TFA (15 mL ACN + 50 μ L TFA + 34.95 mL H₂O).
23. Water bath for sonication.

3 Methods

3.1 Total and Phosphopeptide Sample Pre-preparation

To demonstrate the pipeline, we use SARS-CoV-2 variant infected and noninfected mock VeroE6-H10 or VeroE6-TMPRSS2-H10 cells. The virus-infected cell samples should be deactivated in lysis buffer consisting of 8 M Urea and 50 mM NH_4HCO_3 , supplemented with 1X phosphatase and protease inhibitors cocktail, before handling samples safely in BSL1 conditions.

1. Centrifuge the lysate at $16,000\times g$ for 10 min at 4 °C to clear it (*see Note 2*).
2. Determine the protein concentration in the cleared lysate using a BCA protein assay kit (Thermo Fisher) (*see Note 3*).
3. Obtain equal amounts of protein ($\sim 350\text{ }\mu\text{g}$) for all samples (*see Note 4* and *Note 5*).
4. Reduce proteins using TCEP to achieve a final concentration of 5 mM. Incubate the samples at 37 °C with gentle shaking for 20 min (*see Note 6*).
5. Cool the samples to room temperature (RT), and add freshly prepared IAA (iodoacetamide) to a final concentration of 10 mM for alkylation. Incubate the samples for 20 min at RT in the dark (*see Note 7*).
6. Dilute samples fourfold with 100 mM ammonium bicarbonate (ABC; Sigma-Aldrich) to reduce the Urea concentration to less than 2 M (*see Note 8*).
7. Add 5 μg Sequencing Grade Modified Trypsin (1:100–1:20 (wt/wt)) to the diluted samples at an approximate trypsin: protein ratio of 1:50 to 1:20 (w/w), and incubate at 37 °C for 16 h for digestion (*see Note 9*).

3.2 Desalting Peptides Using BioPureSPN Macro C18 Columns (Nest Group)

Preparing the samples for mass spectrometry analysis involves desalting peptides to remove interfering substances and enhance the sensitivity and accuracy of downstream analysis. Desalting is particularly crucial for ensuring optimal performance of mass spectrometry instruments and improving the quality of obtained data. Desalting of peptides is achieved through a series of steps using BioPureSPN Macro C18 (35–350 μg maximum peptide binding capacity) columns manufactured by Nest Group. These columns utilize C18 resin, a hydrophobic material that selectively retains hydrophobic peptides while allowing hydrophilic and interfering substances to pass through.

1. After digestion, ensure samples are acidic ($\text{pH} < 3$), adjust using 10% TFA or FA.
2. Add ACN to achieve a final concentration of 1% (vol/vol).
3. Set up C18 column in a 2 mL Eppendorf tube.

4. Condition C18 column with 200 μL of 100% ACN.
5. Spin column at $50\times g$ for 2 min at RT.
6. Wash column with 200 μL Mass spectrometry grade HPLC water (*see Note 10*).
7. Condition C18 column with 200 μL of Solvent A two times.
8. Load samples (up to 400 μL at a time) and empty waste flow through after every loading of $\sim 800\ \mu\text{L}$.
9. Spin column at $50\times g$ for 2 min at RT until entire sample is loaded.
10. Confirm all sample has passed through the column (*see Note 11*).
11. Wash column with 100 μL Solvent A; repeat wash step (*see Note 12*).
12. Transfer column to new 2 mL low-binding Eppendorf tube.
13. Elute samples with 100 μL Elution Buffer (1); repeat elution step.
14. Spin column at $50\times g$ for 4 min at RT; confirm all has flowed through.
15. Split the eluted samples two low binding Eppendorf tubes in ratio of $\sim 1:6$ ($\sim 30\ \mu\text{L}$: $170\ \mu\text{L}$) so approx. $\sim 50\ \mu\text{g}$ of sample is used for global proteomics and $\sim 300\ \mu\text{g}$ is used for phosphopeptide enrichment.
16. Speed vacuum dry samples using a SpeedVac for $\sim 1\text{--}2$ h. Store your samples in -80 until use (*see Note 13*).

3.3 Sample Preparation for Phosphopeptide Samples Using CAE-Ti- IMAC Microspheres

Phosphopeptides are enriched by using CAE-Ti-IMAC microspheres, facilitated by T1 StageTips. In this stage, phosphorylated peptides in the sample solution selectively bind to the immobilized titanium ions on the microspheres through coordination complexes formed between the phosphate groups and the metal ions. The elution process involves lowering the pH by adding elution buffer, which disrupts the coordination complexes between the phosphorylated peptides and the titanium ions, resulting in the release of the enriched peptides into solution.

1. Activate the T1 StageTips with 200 μL 100% acetonitrile, centrifuge at $100\ g$ for 3 min.
2. Equilibrate with 200 μL Solvent B to T1 StageTip centrifuge at $100\times g$ for 3 min.
3. Equilibrate with 200 μL Solvent A to T1 StageTip centrifuge at $100\times g$ for 3 min.
4. To load the microspheres, transfer 100 μL slurry of CAE-Ti-IMAC microspheres to T1 StageTip and centrifuge at $90\times g$ for 10 min.

5. Continue by washing CAE-Ti-IMAC microspheres with 150 μL 0.1% TFA; centrifuge at $90\times g$ for 10 min (*see Note 14*).
6. Wash microspheres with 50 μL Equilibration buffer, centrifuge at $90\times g$ for 5 min (*see Note 12*).
7. Dilute the sample (200–300 μg) with 200–300 μL resuspension buffer (sample concentration should be around 0.5–1 $\mu\text{g}/\mu\text{L}$). Sonicate for 15 min in a water bath. (*see Note 15*).
8. Load the sample to the column (150 μL –200 $\mu\text{L} \times 2$) and spin 90 – $100\times g$ for 5–10 min, repeating until entire sample is loaded (*see Note 11*).
9. Wash two times CAE-Ti-IMAC microspheres with 200 μL of wash buffer (1); centrifuge at 100 – $200\times g$ for 10 min.
10. Continue to wash once CAE-Ti-IMAC microspheres with 200 μL wash buffer (2); centrifuge at 100 – $200\times g$ for 5–10 min.
11. Before elution step, add 30 μL of 50% FA to the final low-binding elution collection tube (*see Note 16*).
12. To elute samples, pipet 100 μL of Elution buffer (2) into the stage tip and assemble the tip into an elution collection tube, set up in step 11, and centrifuge at $100\times g$ for 20 min (*see Note 17*).
13. Elute second time using 50 μL Elution buffer (3) into the same tube.
14. Check the pH (pH < 2–3); Adjust by adding 50–100% FA to the samples.
15. Dry samples using a SpeedVac for ~1–2 h. Store your samples at -80°C until analysis (*see Note 13*).

3.4 Sample Loading to Mass Spectrometer

In this step, the global proteomics samples and the enriched phosphopeptides, prepared using the aforementioned methods, are loaded onto Evotips for mass spectrometric analysis with Bruker timsTOF Pro 2.

1. Resuspend sample in 40 or 60 μL Solvent A (*see Note 18*).
2. To set up Evotips for loading, transfer 20 μL solvent B to all Evotips (*see Note 19*).
3. Place the Evotips in the centrifuge, $800\times g$ for 60 s.
4. To condition the Evotips, place the Evotip adaptor rack on top of the 96-well microtiter plate with 100 μL 1-propanol (*see Note 20*).
5. Soak for a minimum of 10 s and then visually inspect that the tips are pale white (*see Note 21*).

6. To equilibrate the Evotips still in the adapter rack, transfer 20 μL Solvent A to each Evotip.
7. Place the Evotips in the centrifuge, $800\times g$ for 60 s.
8. Place the Evotips in the original tray and transfer the sample (20 μL) to each Evotip (*see* **Note 22**).
9. Place the Evotips in the centrifuge, $800\times g$ for 60 s.
10. Wash by transferring 20 μL Solvent A to each Evotip.
11. Place the Evotips in the centrifuge, $800\times g$ for 60 s.
12. Transfer 100 μL Solvent A to each Evotip. Place the Evotips in the centrifuge, $800\times g$ for 5–10 s to keep the tips wet.
13. Add enough Solvent A to the bottom of the tray so the bottom of the Evotips is submerged, and store cold with lid on until analysis (*see* **Note 23**).

3.5 DIA-PASEF MS— Run Setup

Acquire the samples using the DIA-PASEF method Meier et al. [12], employing sample-optimized data-independent analysis (DIA) scan windows. The adjustment of DIA-PASEF windows is conducted based on a sample-specific DDA run. Additional adjustments are made for phosphorylation enrichment and total proteome analysis. Please refer to the instrument provider for more detailed instructions.

To adjust DIA-PASEF scan windows to each sample type, perform a test data-dependent acquisition (DDA) in PASEF mode from a pooled sample for global proteomics and phosphoproteomics (*see* **Notes 24** and **25**).

1. Adjust default DIA-short-gradient acquisition methods based on the sample-specific DDA-PASEF run using the “timsControl” software (Bruker Daltonics) for sample-specific DIA-PASEF parameter optimization.
2. Modify the following parameters for each sample type (*see* Table 1).
3. Set ion mobility windows to best match the ion cloud density observed in the sample type-specific DDA runs [13].

3.6 DIA-NN

DIA-NN [14] is a computational tool used for processing and analyzing mass spectrometry (MS) data obtained through data-independent acquisition (DIA) experiments. This program employs advanced algorithms to identify and quantify peptides and proteins from complex biological samples. In the described protocol, the DIA-NN program is utilized after MS data acquisition to process the raw data, including peak picking, spectral library generation, and peptide identification. By integrating the DIA-NN program into the workflow, we can efficiently analyze large datasets generated from phosphoproteomic experiments, ultimately

Table 1
Mass spectrometry parameters

Parameter	DDA-PASEF	DIA-PASEF
Ionization mode	Positive	Positive
Acquisition method	Data-dependent acquisition (DDA)	Data-independent acquisition (DIA)
Ion mobility method	PASEF	PASEF
Mass range	Default, 0.5 cycle time method	Adjusted based on DDA-PASEF run
Mobility range	0.85–1.3 1/K ₀	0.85–1.3 1/K ₀
Cycle time estimate	0.5 s	Adjusted based on DDA-PASEF run
Chromatography system	Evosep One	Evosep One
Chromatography column	8 cm × 150 µm, 1.5 µm C18 beads	8 cm × 150 µm, 1.5 µm C18 beads
Mobile phase A	0.1% formic acid in water	0.1% formic acid in water
Mobile phase B	0.1% formic acid in acetonitrile	0.1% formic acid in acetonitrile
Gradient time	21 min	21 min
Ion source	CaptiveSpray nano-electrospray	CaptiveSpray nano-electrospray
Mass spectrometer	Bruker timsTOF Pro 2	Bruker timsTOF Pro 2
Software	TimsControl (Bruker Daltonics)	TimsControl (Bruker Daltonics)
Ion mobility windows	Default/none	Adjusted based on DDA-PASEF run

facilitating the interpretation of the complex network of phosphorylation events modulated by viral infections¹⁶. Library generation and sample searches should be done in separate steps as a pipeline. Download the DIA-NN program and search the results using DIA-NN. Set the following parameters (*see* Table 2 and **Note 26**):

3.7 Data Processing, Normalization, and Imputation

This step involves a series of data preprocessing and analysis procedures across samples such as handling missing values through QRILC imputation and standardizing expression levels with median normalization. Additionally, we include the generation of volcano plots to visually represent significant alterations in protein expression between sample and control groups (*see* Fig. 2).

1. Ensure that Python and R are installed, and download the code provided at (<https://github.com/varjola/processing-paper>).
2. Set up a python virtual environment, or use system python to install the required packages (requirements.txt):
3. "pip3 install -r requirements.txt"
4. Install the required R packages (R_requirements.R).
5. "Rscript R_requirements.R"
6. Run the one-step analysis script:

Table 2
DIA-NN parameters (see Note 27)

Parameter	Total proteome DIA-NN	Phosphoproteomics DIA-NN
Missed cleavages	1	1
Maximum number of variable modifications	3	3
N-term M excision	Yes	Yes
C carbamidomethylation	Yes	Yes
Oxidation (M)	Yes	No (<i>see Note 28</i>)
Acetylation (N-term)	Yes	Yes
Phosphorylation (S/T/Y)	No	Yes
Peptide length range	7–30	7–30
Precursor charge range	2–4	2–4
Precursor m/z range	300–1600	300–1600 (<i>see Note 29</i>)
Fragment ion m/z range	100–1700	100–1700 (<i>see Note 29</i>)
Mass accuracy (ppm)	15	15
MS1 accuracy (ppm)	15	15

7. "python3 analysis_onestep.py"

Alternatively, use a Jupyter Notebook server to run the analysis step by step (analysis.ipynb).

The script will assume data is in the "example_report.pg_matrix.tsv", with sample table in "example_sample_table.tsv", and group comparisons in "example_comparisons.tsv". The script will then perform the analysis:

8. Filter the data based on missing values (proteins with missing values in over 40% of the samples of all sample groups will be dropped).
9. Normalize each column (median normalization).
10. Impute missing values using QRILC.
11. Compare the sample groups specified in the comparison table using independent t-test, correcting the p-values afterwards using the Benjamini-Hochberg procedure.
12. Differences with a q-value of <0.01 and a log₂ fold change greater than 1 or less than −1 will be marked as significant.

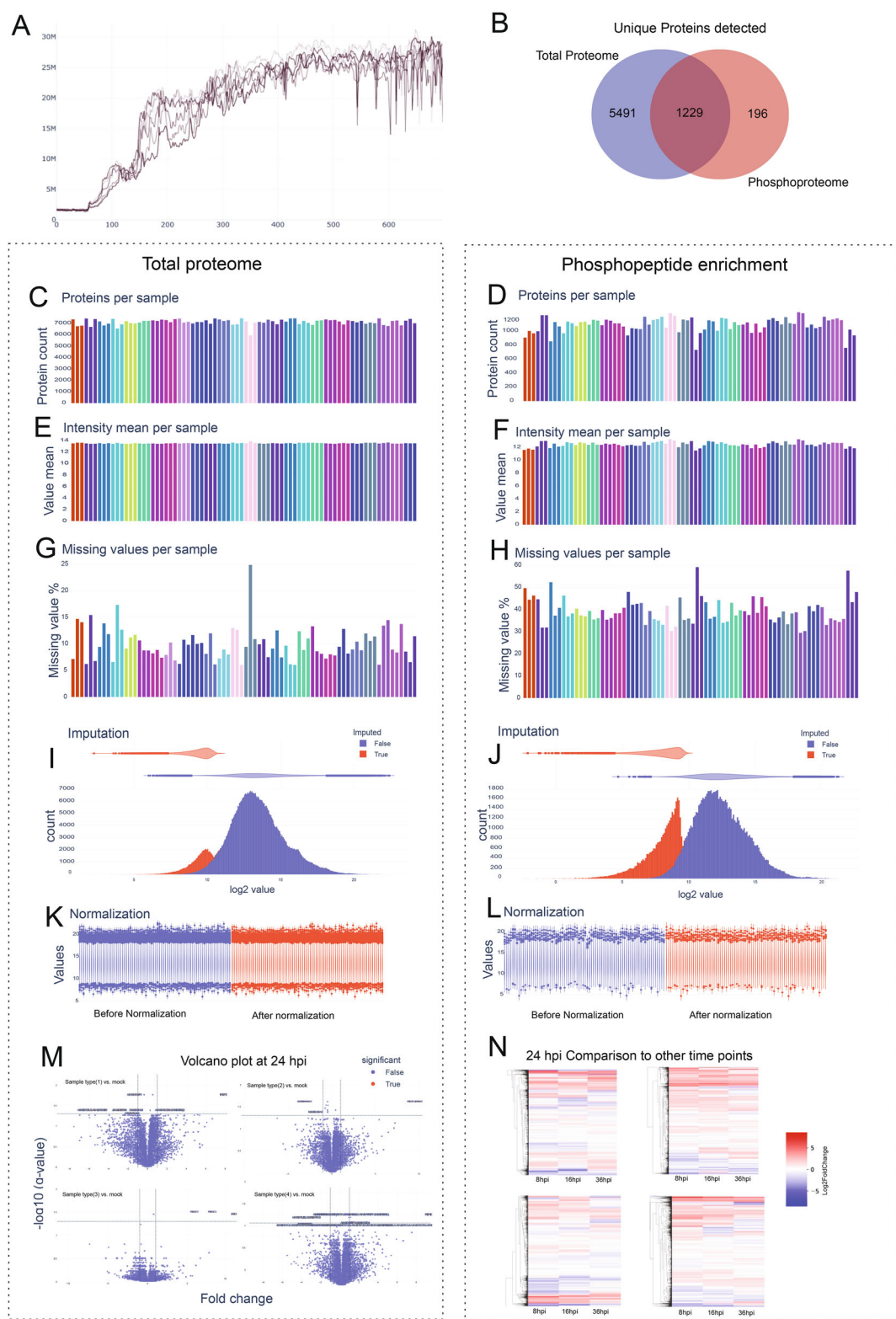


Fig. 2 Output of quality control data and compressive comparison analyses. (a) Example mass spectrometry graph of typical phosphopeptide analysis runs. (b) A Venn diagram of unique protein detected within total

3.8 Additional Advisable Analytical Framework for Phosphorylation in Viral Infections (Optional)

The following outlines a comprehensive analytical framework for exploring the dynamics of protein phosphorylation and functional enrichment in cellular systems, particularly during viral infections.

1. In scenarios involving non-human cells, employ the Gene ID Conversion Tool or an equivalent alternative to systematically convert protein IDs into human genes/proteins. This ensures alignment with established human-centric databases and facilitates coherent cross-species comparisons.
2. Perform functional enrichment analysis using DAVID (<https://david.ncifcrf.gov/>), Panther (<https://www.pantherdb.org/>), or Reactome (<https://reactome.org/>) databases. These analyses can provide valuable insights into enriched pathways, molecular functions, and cellular components associated with the identified proteins. Additionally, consider utilizing Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways for interpretation of the enriched pathways (*see* Fig. 3a).
3. Evaluate the raw phosphorylated DIA-NN run file for viral proteins to identify unique phosphosites. Subsequently, compare this dataset with normalized and imputed quantitative data. Cross-reference literature databases, UniProt, and Protein Data Bank (PDB) structural information to elucidate the effects of viral protein phosphorylation sites (*see* Fig. 3b).
4. Analyze phosphosites in comparison to controls. Identify common kinases present in the samples and discern the pathways influenced by these kinases. This sheds light on the intricate interplay between phosphorylation events and cellular signaling cascades.
5. Employ clustering analysis on normalized, imputed, and significantly expressed data files to uncover potential patterns and correlations within the dataset. This provides insights into the underlying biological mechanisms and facilitates the identification of novel regulatory networks.

Fig. 2 (continued) proteome samples and phospho-enriched samples. (c, d) Graph of protein detected per sample for total proteome samples and phosphopeptide enriched samples. (e, f) Graph of intensity values detected per sample's value mean for both sample types. (g, h) Graph of percentile of missing values detected per sample for both sample types. (i, j) Graph of imputation for both sample types. (k, l) Normalization graph for values before normalization and after normalization, for both sample types. (m) Example of output analysis of normal volcano plot, where four samples at 24 hpi (hours post infection) are compared to mock noninfected cell sample ($p < 0.05$, log2 Fold Change). Data is shown for total proteome data, but equally is performed for both data sets. (n) Alternative comparison analysis where four samples, mock background removed, at 24 hpi are compared to all other hpi points within sample type

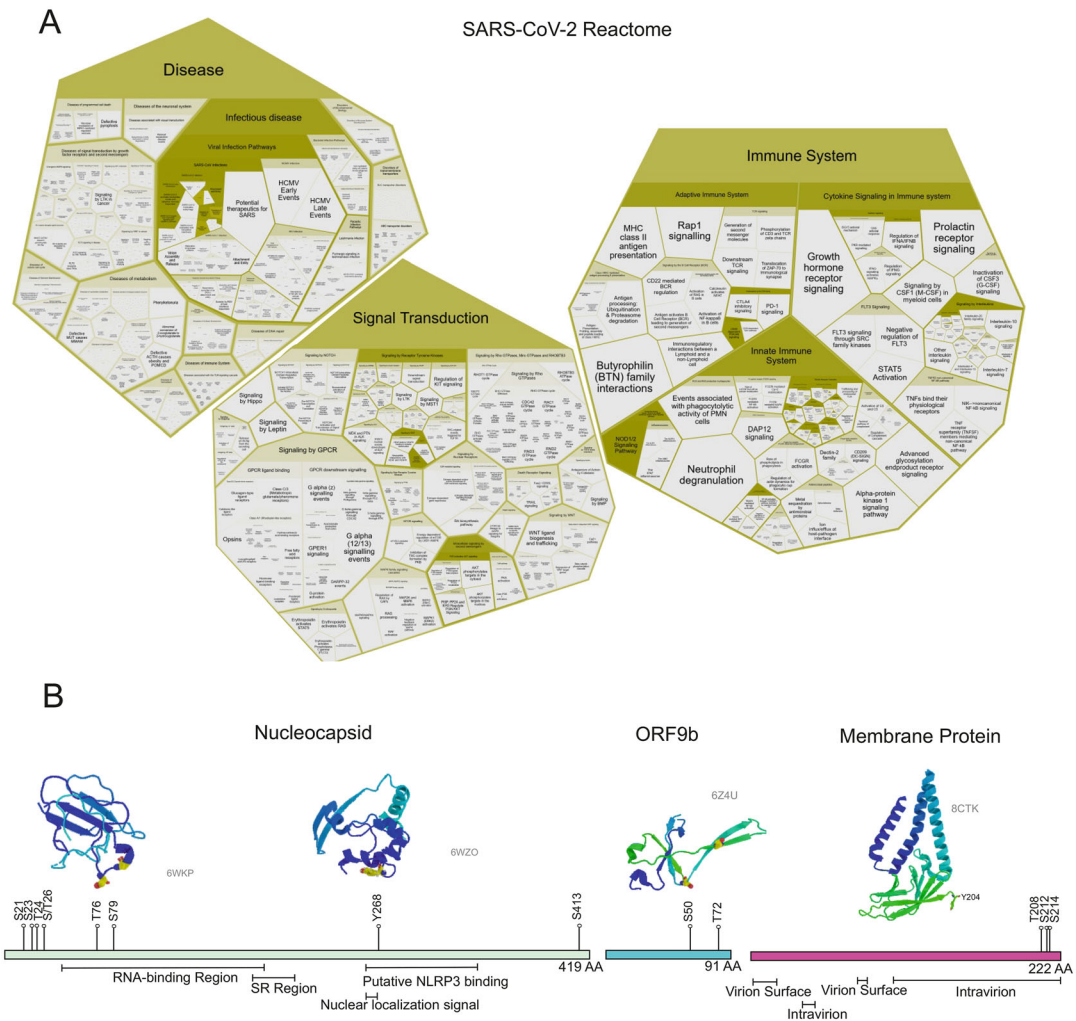


Fig. 3 Example of output results. (a) Reactome enrichment (<https://reactome.org/>) of phosphopeptide enriched proteins from SARS-CoV-2 variants and timepoints combined, showing coverage of Reactome pathways. Namely: disease, signal transduction, and immune system. (b) Overview of the SARS-CoV-2 viral protein's phospho-sites detected via phosphopeptide enrichment coupled with mass spectrometry. Phosphorylation sites were detected in nucleocapsid, ORF9b and membrane proteins Structural images of protein domains or proteins are delivered from RCSB Protein Data Bank and PDB code can be seen in light gray next to structure. Yellow color demonstrates detected phosphosites, and right below, exact localization can be seen in protein sequence. For instance, membrane proteins phosphosites are outside 3D-structures sequence and therefore not shown

4 Notes

1. In case the sample was resuspended in another buffer, use Equilibration Buffer for resuspension.
2. For safety concerns, make sure samples are deactivated using 8 M Urea; samples can be handled in normal laboratory conditions after proper deactivation of the virus.
3. Any protein concentration analysis is valid in this step.
4. Use the supernatant only; avoid pipetting the cell debris pellet at the bottom of the sample tube.
5. Use low-binding 2 mL Eppendorf tubes throughout the pipeline to avoid material loss.
6. TCEP or, alternatively, dithiothreitol (DTT) is a reducing agent that can cleave disulfide bridges and aid in protein denaturation.
7. Add freshly prepared IAA (iodoacetamide) for alkylation, as IAA is light sensitive. Alkylation prevents disulfide bond reformation and ensures protein stability during subsequent analysis.
8. Adjust the pH to >7 using 1 M ABC before adding trypsin to optimize enzymatic activity.
9. Trypsin incubation time may differ between manufacturers; to ensure complete digestion of the samples, we recommend overnight incubation to optimize enzymatic efficiency.
10. To avoid undesirable contamination in the mass spectrometer, please use HPLC-grade water specifically designed for use in MS runs.
11. After the final loading, ensure that the entire sample passes through the column. If there is still sample in the column, please respin until the sample has completely passed through the column before moving to the wash step.
12. Wash buffer is used to remove any detergents or unwanted substances before elution. To improve the sample purity, the wash step can be repeated multiple times.
13. Stop point. Store samples at -80°C until further analysis.
14. Minimize excessive spinning of the microspheres, as it may lead to damage or compaction of the material, impeding solvent flow.
15. This step can also be done at an earlier point, before microsphere setup. It is important that the sample stays at $+4^{\circ}\text{C}$ throughout handling. Sonication may heat up the sample; perform this step with care.

16. Add FA to elution tube (low binding) before the elution step. FA in the tube adjusts the pH back to acidic.
17. The first elution changes the pH to basic due to the presence of ammonium compounds in the eluate solution, thus eluting negatively charged phosphopeptides.
18. For samples where protein concentration is $>200 \mu\text{g}$, load $10\text{--}20 \mu\text{L}$ to Evotips. For global proteome lysis samples, the sample dilution rate should be $\sim 1:50$, and for phosphopeptide-enriched samples, $\sim 1:5$.
19. Follow the manufacturer's instructions.
20. Alternatively, 2-propanol can be used.
21. Pale white color indicates that the tips are soaked.
22. Transfer Evotips to Solvent A tray, if loading time exceeds a few minutes to keep Evotips wet.
23. Tips can be stored in $+4^\circ\text{C}$ for at least 2 weeks.
24. During a test run, check if any adjustments are needed.
25. Mass spectrometry parameters depend on equipment and samples. Parameters need to be modified accordingly for each sample type. Use a DDA run to set DIA-PASEF windows and parameters.
26. Results/output can be requantified using the DIA-NN R package if one wants to perform analysis only using the phospho-precursors.
27. Leave other DIA-NN settings at default.
28. Set oxidation (M) modification to "No" to decrease the search space.
29. For phosphoproteomics, the DIA-NN m/z range should be adjusted based on the created parameters.

References

1. Bouhaddou M, Memon D, Meyer B, White KM, Rezeli VV, Correa Marrero M, Polacco BJ, Melnyk JE, Ulferts S, Kaake RM, Batra J, Richards AL, Stevenson E, Gordon DE, Rojc A, Obernier K, Fabius JM, Soucheray M, Miorin L, Moreno E, Koh C, Tran QD, Hardy A, Robinot R, Vallet T, Nilsson-Payant BE, Hernandez-Armenta C, Dunham A, Weigang S, Knerr J, Modak M, Quintero D, Zhou Y, Dugourd A, Valdeolivas A, Patil T, Li Q, Hüttenhain R, Cakir M, Muralidharan M, Kim M, Jang G, Tutuncuoglu B, Hiatt J, Guo JZ, Xu J, Bouhaddou S, Mathy CJP, Gaulton A, Manners EJ, Félix E, Shi Y, Goff M, Lim JK, McBride T, O'Neal MC, Cai Y, Chang JCJ, Broadhurst DJ, Klippsten S, De Wit E, Leach AR, Kortemme T, Shoichet B, Ott M, Saez-Rodriguez J, tenOever BR, Mullins RD, Fischer ER, Kochs G, Grosse R, García-Sastre A, Vignuzzi M, Johnson JR, Shokat KM, Swaney DL, Beltrao P, Krogan NJ (2020) The global phosphorylation landscape of SARS-CoV-2 infection. *Cell* 182(3): 685–712.e19. <https://doi.org/10.1016/j.cell.2020.06.034>
2. Stukalov A, Girault V, Grass V, Karayel O, Bergant V, Urban C, Haas DA, Huang Y, Oubraham L, Wang A, Hamad MS, Piras A, Hansen FM, Tanzer MC, Paron I, Zinzula L, Engleitner T, Reinecke M, Lavacca TM, Ehmann R, Wölfel R, Jores J, Kuster B,

- Protzer U, Rad R, Ziebuhr J, Thiel V, Scaturro P, Mann M, Pichlmair A (2021) Multilevel proteomics reveals host perturbations by SARS-CoV-2 and SARS-CoV. *Nature* 594(7862):246–252. <https://doi.org/10.1038/s41586-021-03493-4>
3. Weber A, Dam S, Saul VV, Kuznetsova I, Müller C, Fritz-Wolf K, Becker K, Linne U, Gu H, Stokes MP, Pleschka S, Kracht M, Schmitz ML (2019) Phosphoproteome analysis of cells infected with adapted and nonadapted Influenza A virus reveals novel Pro- and antiviral signaling networks. *J Virol* 93(13):e00528–e00519. <https://doi.org/10.1128/JVI.00528-19>
 4. Gerritsen JS, White FM (2021) Phosphoproteomics: a valuable tool for uncovering molecular signaling in cancer cells. *Expert Rev Proteomics* 18(8):661–674. <https://doi.org/10.1080/14789450.2021.1976152>
 5. Keck F, Ataey P, Amaya M, Bailey C, Narayanan A (2015) Phosphorylation of single stranded RNA virus proteins and potential for novel therapeutic strategies. *Viruses* 7(10):5257–5273. <https://doi.org/10.3390/v7102872>
 6. Jaafar ZA, Kieft JS (2019) Viral RNA structure-based strategies to manipulate translation. *Nat Rev Microbiol* 17(2):110–123. <https://doi.org/10.1038/s41579-018-0117-x>
 7. Chamontin C, Bossis G, Nisole S, Arhel NJ, Maarifi G (2021) Regulation of viral restriction by post-translational modifications. *Viruses* 13(11):2197. <https://doi.org/10.3390/v13112197>
 8. Tavares LA, Januário YC, daSilva LLP (2021) HIV-1 hijacking of host ATPases and GTPases that control protein trafficking. *Front Cell Dev Biol* 9:622610. <https://doi.org/10.3389/fcell.2021.622610>
 9. Crow MS, Lum KK, Sheng X, Song B, Cristea IM (2016) Diverse mechanisms evolved by DNA viruses to inhibit early host defenses. *Crit Rev Biochem Mol Biol* 51(6):452–481. <https://doi.org/10.1080/10409238.2016.1226250>
 10. Liu X, Salokas K, Weldatsadik RG, Gawrysiak L, Varjosalo M (2020) Combined proximity labeling and affinity purification—mass spectrometry workflow for mapping and visualizing protein interaction networks. *Nat Protoc* 15(10):3182–3211. <https://doi.org/10.1038/s41596-020-0365-x>
 11. Capriotti AL, Antonelli M, Antonioli D, Cavaliere C, Chiarcos R, Gianotti V, Piovesana S, Sparnacci K, Laus M, Laganà A (2019) Effect of shell structure of Ti-immobilized metal ion affinity chromatography core-shell magnetic particles for phosphopeptide enrichment. *Sci Rep* 9(1):15782. <https://doi.org/10.1038/s41598-019-51995-z>
 12. Meier F, Brunner A-D, Frank M, Ha A, Bludau I, Voytik E, Kaspar-Schoenefeld S, Lubeck M, Raether O, Bache N, Aebersold R, Collins BC, Röst HL, Mann M (2020) diaPASEF: parallel accumulation-serial fragmentation combined with data-independent acquisition. *Nat Methods* 17(12):1229–1236. <https://doi.org/10.1038/s41592-020-00998-0>
 13. Demichev V, Szyrwił L, Yu F, Teo GC, Rosenberger G, Niewianda A, Ludwig D, Decker J, Kaspar-Schoenefeld S, Lilley KS, Müllender M, Nesvizhskii AI, Ralser M (2022) Dia-PASEF data analysis using FragPipe and DIA-NN for deep proteomics of low sample amounts. *Nat Commun* 13(1):3944. <https://doi.org/10.1038/s41467-022-31492-0>
 14. Huuskonen S, Liu X, Redchuk T, Salokas K, Lundberg R, Maljanen S, Belik M, Reinholm A, Kolehmainen P, Tripathi G, Laine P, Belanov S, Auvinen P, Vartiainen M, Laine L, Österlund P, Julkkunen I, Kakkola L, Varjosalo M (2024) The comprehensive SARS-CoV-2 ‘Hijackome’ knowledge base—reveals significant changes in host cell protein expression and activation by multiple SARS-CoV-2 variants. *Res Square*. <https://doi.org/10.21203/rs.3.rs-4322477/v1>
 15. Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M (2020) DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods* 17(1):41–44. <https://doi.org/10.1038/s41592-019-0638-x>



Activity-Based Protein Profiling of Bacterial Monooxygenases

John B. Joyce  and Michael R. Hyman 

Abstract

Activity-based protein profiling (ABPP) is a proteomics approach that enables the selective molecular characterization of specific proteins within complex mixtures. A variety of activity-based approaches have been used to characterize some important bacterial monooxygenases such as ammonia monooxygenase (AMO) that have resisted purification-based characterization. In this report, we describe a 2-step ABPP method for AMO and several other types of bacterial monooxygenases that play key roles in biogeochemical cycles and processes. The focus of this report is on SDS-PAGE-based methods to detect fluorescently labeled monooxygenase components. This report also draws attention to various features of these types of analyses that can impact the use of ABPP approaches for the quantification of monooxygenase components and their detection in different compartments of bacterial cells.

Key words Activity-based protein profiling, Monooxygenases, SDS-PAGE, Near-infrared (NIR) fluorescent imaging, Diyne probe

1 Introduction

Bacterial monooxygenases are a diverse group of enzymes involved in a wide variety of microbial metabolic processes ranging from degradation of xenobiotic compounds, modification of secondary metabolites, to catabolism of primary carbon and/or energy sources [1–3]. These reductant-requiring enzymes typically catalyze the incorporation of one atom of oxygen from O₂ into a substrate, while the other oxygen atom is reduced to water. These enzymes can catalyze chemo-, regio-, and/or enantioselective reactions that can generate hydroxylated, epoxidized, dehalogenated, or desaturated products [4].

Some of the best-characterized monooxygenases are the heme-containing cytochrome P450's, which are widely distributed in both eukaryotes and prokaryotes. Diverse types of nonheme monooxygenases are often only encountered in prokaryotes and some of

these have received considerable attention as they catalyze key reactions in biogeochemical cycles. Prominent examples include soluble di-iron-containing monooxygenases (SDIMOs) such as soluble methane monooxygenase (sMMO) [5], membrane-bound, mono-iron AlkB-containing alkane hydroxylases [6], and copper-containing monooxygenases, such as ammonia monooxygenase (AMO) and particulate methane monooxygenase (pMMO) [7].

Activity-based protein profiling (ABPP) is primarily a proteomics technique that uses selective probes to target proteins of interest in complex mixtures such as intact cells [8]. After interaction with the probe, simple or advanced protein analysis approaches are used to further characterize the targeted polypeptides. As the name implies, the interaction between the probe and target protein often capitalizes on an enzymatic activity. In the case of bacterial monooxygenases, the probes are frequently alkyne-containing molecules that act as irreversible mechanism-based inactivators of these enzymes [9].

The first bacterial monooxygenase we characterized using a form of ABPP was AMO in the aerobic nitrifying bacterium, *Nitrosomonas europaea*. This bacterium oxidizes ammonia (NH_3) to nitrite (NO_2^-), and AMO catalyzes the initial step in ammonia catabolism. Despite many prior efforts, AMO has not been purified in an active form, and a sequence of progressively more complex ABPP-based approaches has played an important role in refining our understanding of this integral membrane-bound enzyme (Fig. 1). For example, in intact cells, AMO activity is exquisitely sensitive to irreversible inactivation by acetylene, the simplest

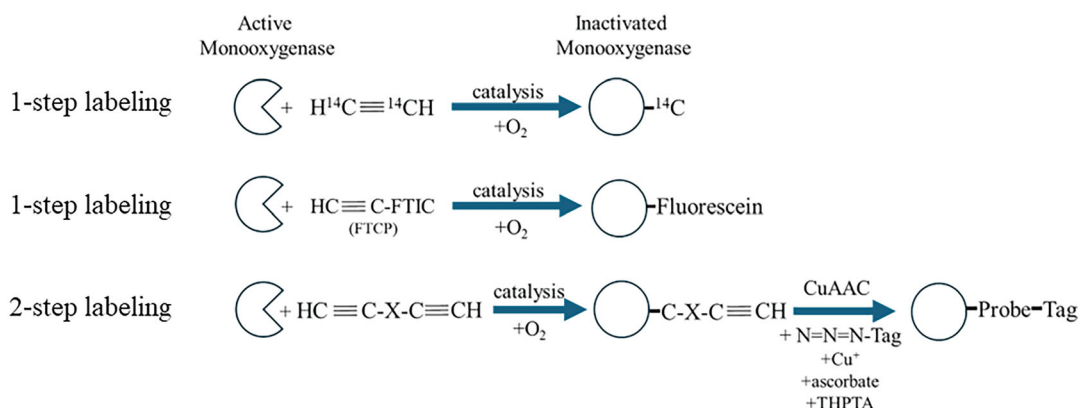


Fig. 1 Different types of ABPP approaches used to characterize AMO from *N. europaea*. The Figure shows three different types of ABPP approaches that have been used to characterize AMO in *N. europaea*. The two 1-step approaches include the activity based labeling of AMO components by carbon-14 (^{14}C) from $[\text{U-}^{14}\text{C}]\text{-C}_2\text{H}_2$ [7], and by FTCP [12]. The 2-step approach used to detect AMO components involves the use of diyne probes and subsequent detection using highly fluorescent fluorophores or ligands suitable for affinity purification [14]

alkyne [10]. The kinetics of inactivation of AMO activity by acetylene fit many of the criteria for mechanism-based inactivation [7]. Covalent binding of ^{14}C from $[\text{U-}^{14}\text{C}]\text{-C}_2\text{H}_2$ to a single membrane-associated polypeptide, as detected by SDS-PAGE and fluorography, provided the first molecular evidence for a component polypeptide of this enzyme. A similar radiolabeling approach has also been used to identify the functionally equivalent polypeptide in the membrane-bound form of methane monooxygenase (pMMO) in *Methylococcus capsulatus* (Bath) [11]. Subsequently, it was demonstrated that the ^{14}C -binding component of AMO could also be fluorescently labeled when cells were treated with the acetylenic fluorescein derivative, fluorescein thiocarbamoyl propargylamine (FTCP) [12]. After final purification by preparative SDS-PAGE, the N-terminal amino acid sequence of the FTCP-labeled polypeptide was used to identify the gene encoding AmoA, the first structural component of AMO identified at the molecular level. This one-step approach using FTCP-dependent labeling has also been used to demonstrate expression of AMO in the recently discovered comammox bacteria that are defined by their ability to completely oxidize ammonia to nitrate [13].

We have more recently developed a two-step ABPP approach for AMO [14]. Compared to the one-step approaches described above, this method provides greater flexibility in terms of the probes that can be used and the types of downstream proteomic analyses that can be conducted. The two-step approach involves the initial inactivation of AMO by terminal diyne probes such as 1,7-octadiyne (17OD). In this initial step, one terminal alkyne group on the probe is catalytically activated by AMO to a reactive intermediate (most likely to a ketene), which then reacts within, or near, the active site of AMO. This leads to an irreversible loss of activity through the formation of a covalent probe:enzyme adduct. The adduct can then be detected *in vivo* or captured *in vitro* using a simple copper-catalyzed azide-alkyne cycloaddition (CuAAC) reaction with azide-containing tags. The use of fluorescent tags enables the detection of labeled proteins by techniques such as SDS-PAGE coupled with near-infrared (NIR) scanning or in intact cells using techniques such as fluorescence microscopy or flow cytometry [15]. Conversely, the use of tags such as biotin enables the affinity-based purification of the covalent probe:enzyme adduct as well as identification and quantification of the captured polypeptides at the amino acid level using genome-enabled mass spectrometry [14].

The methods described here are based on our initial ABPP study with AMO in *N. europaea* [14] and emphasize the detection of labeled monooxygenase components using SDS-PAGE and NIR gel scanning. The methods described also include observations made as we have extended the original approach to other bacterial monooxygenases. Some of these more recent efforts have focused

on identifying selective probes for different types of monooxygenases or identifying key issues impacting the use of fluorescent labeling to quantify monooxygenase expression. We have also examined the impacts of the accessibility of common ABPP reagents for detecting monooxygenases in intact cells when these enzymes have different cellular locations and orientations.

2 Materials

Prepare all reagents with sterile deionized (DI) water or 99.9% dimethyl sulfoxide (DMSO), whichever is indicated. Store all reagents at the temperatures specified. Some of the instruments and materials listed may be changed or substituted based on the user's needs or equipment/supply access. Please follow all waste disposal procedures when handling and disposing of waste materials.

2.1 Bacterial Harvest Buffer

1. Phosphate Buffer (PB): Add 5.999 g of monobasic sodium phosphate, anhydrous, (NaH_2PO_4) to 0.95 L of sterile DI water. Adjust pH to 7.0 with 10 M NaOH or 10 M HCl, and then bring to a final volume of 1 L with sterile DI water for a final concentration of 50 mM. Store at room temperature.

2.2 Probe Stocks and Supplies

1. Diyne probes: Make a 1 M stock by mixing the diyne probe with an appropriate volume of DMSO (these volumes are typically 50–100 μL). Probes: 1,5-hexadiyne (15HD) (*see Note 1*), 1,6-heptadiyne (16HD), 1,7-octadiyne (17OD), 1,8-nonadiyne (18ND), 1,9-decadiyne (19DD), 1,4-diethynylbenzene (DEB), and dipropargylamine (DIP) (*see Note 2*). All stock solutions can be stored at 4 °C for 6 months without significant loss of function (*see Note 3*).
2. 10 mL glass serum bottles (larger or smaller glass or plastic bottles/tubes can be used).
3. Butyl rubber stoppers (or appropriate caps for the container).

2.3 Cell Lysis Reagents

1. Lysing Matrix B, Bulk: MP Biomedicals™, 0.1 mm silica spheres.
2. Bead Mill 4 Mini Homogenizer: Fisherbrand™.

2.4 Protein Assay Reagents

1. BSA stock solution: Prepare bovine serum albumin at 2 mg/mL in sterile DI water.
2. Reagent A: Add 1 g of sodium bicinchoninate (BCA), 2 g of sodium carbonate (Na_2CO_3), 0.185 g of sodium tartrate dihydrate ($\text{C}_4\text{H}_4\text{Na}_2\text{O}_6 \cdot 2\text{H}_2\text{O}$), 0.4 g sodium hydroxide (NaOH), and 0.95 g of sodium bicarbonate (NaHCO_3) to a

bottle with ~95 mL of sterile DI water. Adjust the pH to 11.25 with 10 M NaOH and bring it to a final volume of 100 mL with sterile DI water.

3. Reagent B: Add 0.4 g of cupric sulfate pentahydrate ($\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$) to 10 mL sterile DI water for a final concentration of 160.2 mM (4% w/v solution).
4. NanoDrop™ Eight Spectrophotometer with appropriate software: Thermo Scientific.

2.5 Click (CuAAC) Reaction Reagents

1. Alexa Fluor™ 647 Azide: Add 0.5 mg of Alexa Fluor™ 647 Azide, triethylammonium salt to 1 mL of DMSO for a final concentration of 0.588 mM. Aliquot into 25 μL volumes and immediately store in the dark at -20°C until ready to use (*see Note 4*).
2. THPTA: Add 4.3 mg of THPTA (tris-hydroxypropyltriazolylmethylamine) to 1 mL sterile DI water to make a 10 mM aqueous stock and aliquot into 50 μL volumes. Freeze immediately at -20°C until ready to use (*see Note 5*).
3. Copper solution: Add 0.0624 g of cupric sulfate and pentahydrate ($\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$) to 50 mL of sterile DI water for a final concentration of 5 mM. Can be stored at room temperature.
4. Sodium ascorbate: Add 0.0225 g of sodium ascorbate to 1 mL of sterile DI water and vortex to mix (*see Note 6*). Final concentration is 113.6 mM. Store at room temperature until ready to use. *Must be prepared the day of use.*
5. Aminoguanidine: Add 0.1106 g of aminoguanidine hydrochloride to 1 mL DMSO for a final concentration of 1 M. Can be stored at -80°C for extended use. The day of use, make an aqueous solution from the aminoguanidine stock by adding 25 μL of the 1 M stock solution to 75 μL of sterile DI water for a final concentration of 250 mM. Additionally, an aqueous stock solution can be prepared from solid aminoguanidine hydrochloride on the day of the experiment.
6. Diamide: Add 0.0172 g of diamide (1,1'-azobis(N, N-dimethylformamide; CAS 10465-78-8) to 1 mL of sterile DI water for a final concentration of 100 mM. *Must be prepared on the day of use.*

2.6 SDS Polyacrylamide Gel

1. Resolving gel buffer, 1.5 M Tris-HCl, pH 8.8: Add 90.85 g of Tris base to 400 mL of sterile DI water in a 1 L graduated cylinder. Mix with a magnetic stir bar until completely dissolved. Adjust to pH 8.8 with 12 M HCl. Remove the stir bar and adjust the volume to 500 mL with sterile DI water. Store at 4°C .

2. Stacking gel buffer, 0.5 M Tris-HCl, pH 6.8: Add 30.28 g of Tris base to 400 mL of sterile DI water in a 1 L graduated cylinder. Mix with a magnetic stir bar until completely dissolved. Adjust to pH 6.8 with 12 M HCl. Remove the stir bar and adjust the volume to 500 mL with sterile DI water. Store at 4 °C.
3. 40% acrylamide solution: Acrylamide/bis-Acrylamide 29:1 (40% solution) (*Fisher BioReagents™*).
4. 10% ammonium persulfate (APS): Add 0.1 g of ammonium persulfate to 1 mL of sterile DI water. Vortex to mix. Can be stored at room temperature but *must be prepared the day of use*.
5. 10% SDS: Add 10 g of sodium dodecyl sulfate (SDS) to about 80 mL of sterile DI water in a small beaker. Add a magnetic stir bar and mix until the SDS is completely dissolved. Adjust volume to 100 mL with sterile DI water. Store at room temperature.
6. *N,N,N',N'*-tetramethylethylenediamine (TEMED): Store at 4 °C.
7. 10X SDS-PAGE running buffer: Add ~800 mL of sterile DI water to a suitable container. Then, add 30.3 g of Tris base, 144.4 g of glycine, and 10 g of SDS. Ensure all the reagents have gone into solution and bring to a final volume of 1 L. Store at 4 °C. Make 1 L of 1X SDS-PAGE running buffer by diluting 100 mL of 10X SDS-PAGE running buffer in 900 mL of sterile DI water on the day of the experiment. Ensure the 1X buffer has been chilled to 4 °C before use. Do not adjust pH.
8. 6X SDS Sample Buffer (6X SB): In a 50 mL conical, add 15 mL of glycerol and 28 mL of stacking gel buffer (described previously) and warm to 37 °C. Then add 5 g of SDS and carefully mix until it completely dissolves. Add 600 µL of 0.5 M EDTA (pH 8.0) and 6 mg of Bromophenol Blue and carefully mix as before. Transfer 22.5 mL of the solution to a new 50 mL conical, and add 2.5 mL of 2-Mercaptoethanol (BME) (*see Note 7*). Aliquot 250 µL of the 6X SB into 0.5 mL tubes and immediately freeze at -20 °C. Once thawed and opened, do not reuse the aliquots.
9. Protein Ladder: PageRuler™ Prestained NIR Protein Ladder, Thermo Scientific™.
10. Mini-PROTEAN® Vertical Electrophoresis Cell, 4-gel, for 1.0 mm thick handcast gels: BIO-Rad (*see Note 8*).

2.7 LI-COR Odyssey® CLx Imaging System

1. Odyssey® CLx Imaging System: LI-COR Biosciences (*see Note 9*).

2.8 Coomassie Total Protein Stain

1. Coomassie stain: Add 100 mL of glacial acetic acid to 500 mL of sterile DI water. Add 400 mL of methanol. Add 1 g of Coomassie Brilliant Blue R250. Gently mix until the stain has dissolved. Gravity filter to remove particulates. Store at room temperature in a closed container.
2. Destain: Add 500 mL of sterile DI water to 500 mL of methanol. Then, add 100 mL of acetic acid. Store at room temperature in a closed container.

3 Methods

Carry out all procedures at room temperature unless otherwise specified.

3.1 Bacterial Harvest

1. Grow bacterial cells in batch culture with the appropriate media, carbon and energy source, and at the appropriate temperature to induce monooxygenase expression and activity (*see Note 10*).
2. Transfer culture to centrifuge tubes/bottles and centrifuge at 10,000 RCF for 5 min at 4 °C.
3. Carefully pour off the supernatant and resuspend the pellet in 20 mL of PB (*see Note 11*).
4. Centrifuge at 10,000 RCF for 5 min at 4 °C.
5. Repeat **steps 3** and **4** in sequence two more times.
6. Carefully remove supernatant. Resuspend cell pellet in PB to an approximate OD₆₀₀ of 4.0 (*see Note 12*).

3.2 Probing Active Monooxygenases

1. Transfer 1 mL of cell suspension from Subheading **3.1, step 6** to a 10 mL serum vial (or appropriate container) (*see Note 13*).
2. Add a volume of diyne probe to a final concentration of 1 mM (*see Notes 14* and *15*).
3. Cap the serum bottle with a butyl rubber stopper (or appropriate lid).
4. Place the bottle in a shaking incubator at 30 °C and 150 RPM for 1 h (*see Note 16*).
5. After incubation, unseal the serum bottle and transfer the cell suspension to a pre-labeled 2.0 mL screw cap tube (*see Note 17*). Add 1 mL of PB to the empty serum bottle and vortex for 15–30 s to recover any remaining cells (*see Note 18*). Transfer this wash to the 2.0 mL tube.
6. Centrifuge the tube at 10,000 RCF for 3 min at room temperature.

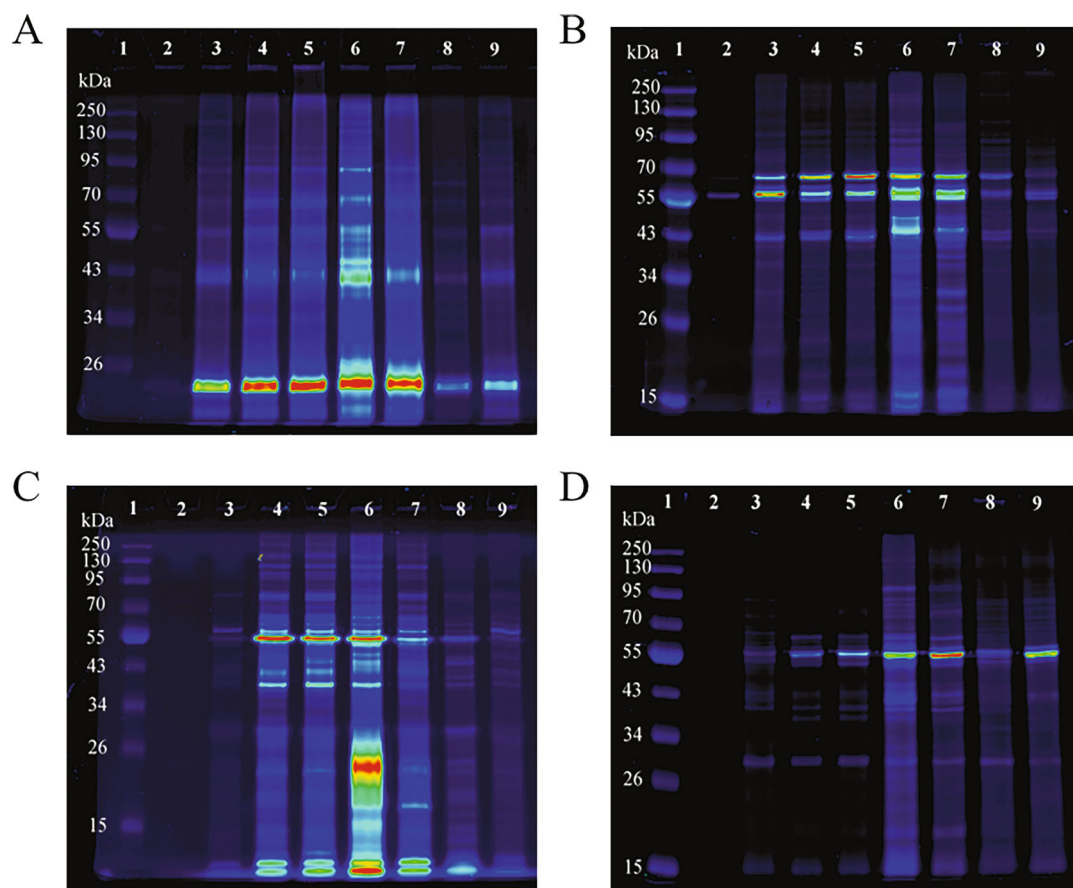


Fig. 2 Fluorescent labeling of polypeptides in diverse monooxygenase-expressing bacteria treated with diyne probes. The Figure shows SDS-PAGE/NIR analyses of the effect of a range of diyne probes on the fluorescent protein labeling patterns observed for several different bacteria grown on substrates known to support monooxygenase expression. The results demonstrate that the use of different probes can change both the molecular mass and labeling intensity of labeled polypeptides. Panel A: *N. europaea* grown on ammonia. Panel B: *Mycobacterium* sp. ELW1 grown on *n*-octane. Panel C: *Pseudomonas putida* GPo1 grown on *n*-octane. Panel D: *Mycobacterium* sp. NBB4 grown on propane. In each Panel the lanes are as follows: Lane 1, PageRuler™ Prestained NIR Protein Ladder. Lane 2, DMSO treated cells. Lane 3, 15HD treated cells. Lane 4, 16HD treated cells. Lane 5, 17OD treated cells. Lane 6, 18ND treated cells. Lane 7, 19DD treated cells. Lane 8, DEB treated cells. Lane 9, DIP treated cells. Bacteria were probed with 1 mM of the respective diyne listed and samples were labeled with Alexa Fluor™ 647 azide, as described in the Methods section. All lanes were loaded with 5 µg of protein

7. Carefully remove the supernatant and resuspend the cell pellet in 1 mL of PB.
8. Repeat **steps 6** and **7** two more times.
9. Carefully remove the supernatant.
10. Sample can be frozen at -20°C until ready for lysis or used immediately (Fig. 2).

3.3 Cell Lysis

1. Resuspend the pellet in 250 μ L of PB or, alternatively, 250 μ L of PB with 0.5%–1.0% SDS (*see Note 19* and *20*).
2. Add approximately 0.1500 g of lysis matrix B beads (*see Note 21*) to the 2.0 mL screw cap tube containing the resuspended cell pellet.
3. Place the tube on a bead mill for 120 s at 5 m/s (*see Note 22*), then chill for 5–10 min at 4 °C or on ice (*see Note 23*).
4. Repeat **step 2** for the desired number of times (*see Note 24*).
5. Briefly spin down the sample and store at –20 °C until ready for protein assay or use immediately (Fig. 3).

3.4 Protein Assay

1. Centrifuge the lysed cells at ~17,000 RCF for 10 mins at 4 °C (*see Note 25*).
2. Thaw a 2 mg/mL aliquot of BSA stock solution (*see Note 26*).
3. Create a protein assay standard curve using a serial dilution of the BSA stock mixed with DI water as follows: 2 mg/mL, 1.5 mg/mL, 1.0 mg/mL, 0.75 mg/mL, 0.5 mg/mL, 0.25 mg/mL, and 0.0 mg/mL (*see Note 27*).
4. Prepare 200 μ L of Working Solution (BCA Reagent 1 and BCA Reagent 2 at a 50:1 ratio) for all standards and samples (*see Note 28*).
5. Add 10 μ L of standard or sample to the prepared 200 μ L working solution (*see Note 29*). Vortex briefly to mix.
6. Heat standards and samples at 60 °C for 5 min, then allow to cool to room temperature (*see Note 30*).
7. Measure the samples using the Protein BCA application on a NanoDrop™ Eight Spectrophotometer (*see Note 31*), and determine the protein concentration.
8. Samples can be stored at –20 °C or used immediately (Fig. 4).

3.5 Fluorescent Labeling of Monooxygenases Using CuAAC (Click) Reactions

1. Perform the CuAAC labeling reaction (*see Note 32*) in 100 μ L total volume (*see Note 33*) with 25 μ g of total protein per reaction (*see Note 34*).
2. Add all reagents in the following order (*see Note 35*).
3. Add a volume of DI water that will bring the total volume to 100 μ L after the addition of protein and other reagents.
4. Add 25 μ g of protein (*see Note 36*).
5. Add 1.3 μ L of 0.588 mM Alexa Fluor™ 647 (*see Note 37*).
6. Combine 5 μ L of 10 mM THPTA (*see Note 38*) and 2 μ L of 5 mM copper solution (*see Note 39*) in a separate tube. This is a 5:1 ratio of THPTA to copper. Vortex briefly to mix. Add the 7 μ L of the THPTA/copper solution to the labeling reaction.
7. Add 4.4 μ L of 113.6 mM sodium ascorbate (*see Note 40*).

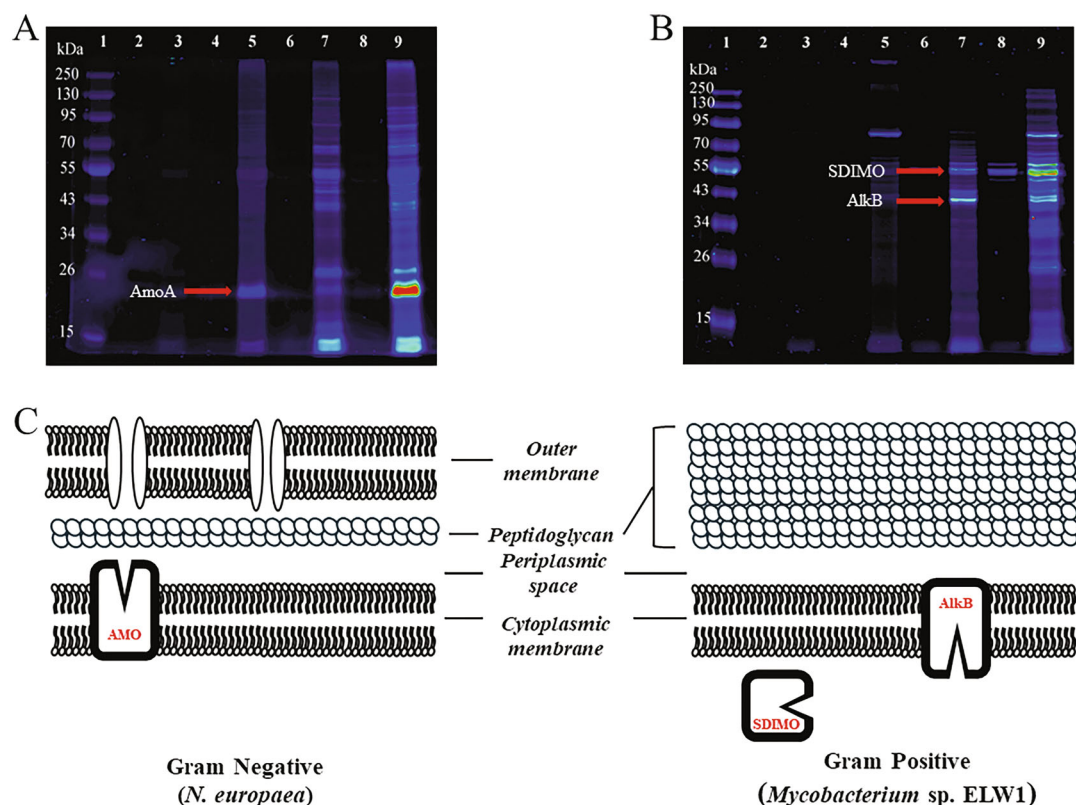


Fig. 3 Labeling monooxygenases based on cellular location and orientation. The figure shows the fluorescent protein labeling patterns for *N. europaea* and *Mycobacterium* sp. ELW1 observed after CuAAC labeling in intact or lysed cells. Panel A: SDS-PAGE/NIR analysis of fluorescent protein labeling of *N. europaea* grown on ammonia. Panel B: SDS-PAGE/NIR analysis of fluorescent protein labeling of *Mycobacterium* sp. ELW1 grown on *n*-octane. Panel C: Cartoon of cellular distribution and orientation of major monooxygenases in *N. europaea* and *Mycobacterium* sp. ELW1. Lanes are labeled 1–9. Lanes 2–5, CuAAC labeling reaction performed with intact cells. Lanes 6–9, CuAAC labeling reaction performed with lysed cells. Lane 1, PageRuler™ Prestained NIR Protein Ladder. Lane 2, treated with DMSO alone, no SDS during lysis. Lane 3, 170D probed, no SDS during lysis. Lane 4, DMSO probed, SDS during lysis. Lane 5, 170D probed, SDS during lysis. Lane 6, DMSO probed, no SDS during lysis. Lane 7, 170D probed, no SDS during lysis. Lane 8, DMSO probed, SDS during lysis. Lane 9, 170D probed, SDS during lysis. Samples were probed with 1 mM 170D and labeled with Alexa Fluor™ 647 azide, as described in the Methods section (with the following adjustments listed within this legend). All lanes were loaded with 24 μ L of the labeling reaction

8. Optional: Add 2 μ L of 250 mM aminoguanidine (*see* **Note 41**) and/or add 1 μ L of 100 mM diamide (*see* **Note 42**).
9. Cap the reaction mixture and vortex briefly to mix. Gently spin down the reaction for 2–3 s, and place in complete darkness at room temperature for 1 h (*see* **Note 43**).
10. The reaction can be stored in the dark at -20°C (*see* **Note 44**) or used immediately for SDS-PAGE (Figs. 5 and 6).

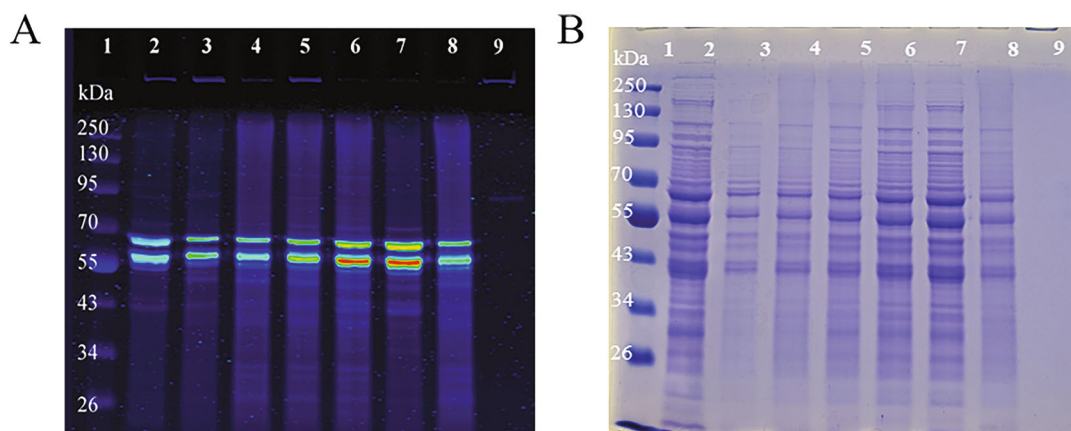


Fig. 4 Effect of different protein assays on fluorescent labeling of monooxygenase components in *Mycobacterium* sp. ELW1. The figure shows the results of an analysis of the effect of different protein assays on the fluorescent labeling of monooxygenases in *Mycobacterium* sp. ELW1. The bacterium was grown on isobutylene and probed with 1 mM 17OD. Dityne-treated samples were labeled with Alexa Fluor™ 647 azide using CuAAC reactions. All lanes were loaded with 15 µg of protein. Panel A: SDS-PAGE/NIR analysis of fluorescently labeled polypeptides. Panel B: The same gel after Coomassie staining for total protein. Lanes are labeled 1–9 and are described by Table 1

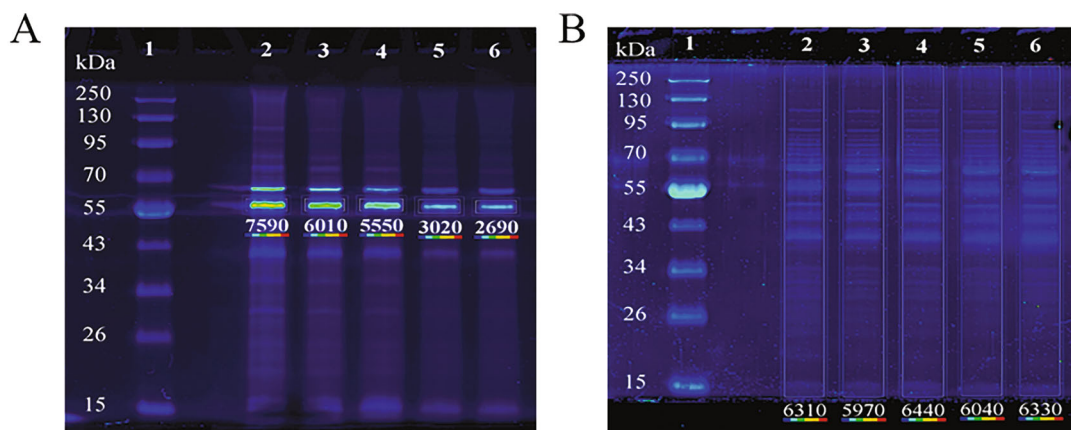


Fig. 5 The effect of protein concentration in CuAAC labeling reactions on fluorescent signal intensity. The figure shows the effect of changes in protein concentration of the fluorescent labeling of monooxygenases expressed by *Mycobacterium* sp. ELW1. The bacterium was grown on isobutylene and probed with 15HD. Dityne-treated cell extracts were labeled with Alexa Fluor™ 647 azide using CuAAC reactions, as described in the Methods section with the following protein concentration adjustments listed below. All lanes were loaded with 5 µg of protein. Panel A: SDS-PAGE/NIR scanning analysis of fluorescently labeled polypeptides. Panel B: The same gel after Coomassie staining for total protein. Lane 1, PageRuler™ Prestained NIR Protein Ladder; Lane 2, 25 µg protein; Lane 3, 37.5 µg protein; Lane 4, 50 µg protein; Lane 5, 62.5 µg protein; Lane 6, 75 µg protein. The fluorescence associated with the 55 kDa polypeptide in each sample was determined as described in the Methods section using Empiria Studio® software

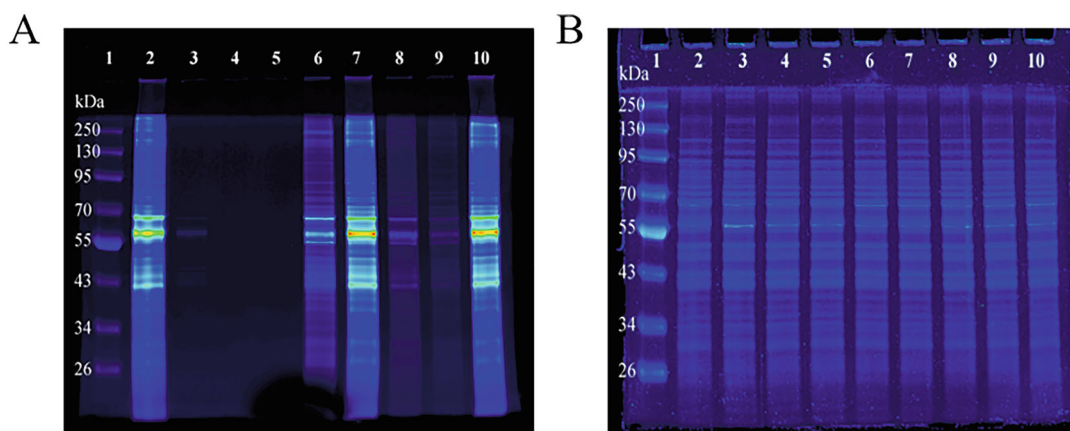


Fig. 6 Effects of CuAAC reagents and other additions on fluorescent labeling patterns. The figure shows the effects of individual CuAAC reagents on the fluorescent labeling of proteins in *Mycobacterium* sp. ELW1. The bacterium was grown on isobutylene and probed with 1 mM 17OD. The cells were lysed by bead beating and then labeled with Alexa Fluor™ 647 azide in CuAAC reactions as described in the Methods section with the changes noted below. All lanes were loaded with 10 μ g of total protein. Panel A: SDS-PAGE/NIR analysis of fluorescently labeled polypeptides. Panel B: The same gel after Coomassie staining. The samples in Lanes 3 to 8 did not contain diamide in the CuAAC reactions. Lane 1, PageRuler™ Prestained NIR Protein Ladder. Lane 2, all CuAAC reagents plus diamide. Lane 3, all CuAAC reagents except diamide (sample not probed with 17OD). Lane 4, all CuAAC reagents except Alexa Fluor™ 647 azide. Lane 5, all CuAAC reagents except copper. Lane 6, all CuAAC reagents except THPTA. Lane 7, all CuAAC reagents except aminoguanidine. Lane 8, all CuAAC reagents except sodium ascorbate. Lane 9, all CuAAC reagents plus diamide, and without sodium ascorbate. Lane 10, all CuAAC reagents plus diamide

3.6 10% Sodium Dodecyl Sulfate: Polyacrylamide Gel Electrophoresis

1. 10% resolving gel. Mix the following reagents in the order listed in a small beaker with a stir bar. Add 2.5 mL of 40% acrylamide (29:1), 2.5 mL of 1.5 M Tris-HCL, pH 8.8, 100 μ L 10% SDS, 4.85 mL of sterile DI water, and 100 μ L of 10% APS. Gently stir the solution to mix, being careful not to introduce bubbles. While mixing, add 7.5 μ L of TEMED.
2. Immediately transfer the solution to a 7.25 cm $\times$ 10 cm $\times$ 1.0 mm glass gel cassette (*see Note 45*), allowing space for the stacking gel (*see Note 46*). Gently overlay with sterile DI water (*see Note 47*).
3. 4% stacking gel. Mix the following reagents in the order listed in a small beaker with a stir bar. Add 0.5 mL of 40% acrylamide (29:1), 1.26 mL of 0.5 M Tris-HCL, pH 6.8, 50 μ L 10% SDS, and 3.18 mL of sterile DI water.
4. Once the resolving gel has solidified (*see Note 48*), carefully remove the overlay by pouring off the water. Wick away any residual water using clean filter paper, taking care to avoid touching the gel.

5. Finish the stacking gel by adding 50 μL of 10% APS. Gently stir the solution to mix, being careful not to introduce bubbles. While mixing, add 7.5 μL of TEMED.
6. Immediately transfer the stacking solution to the gel cassette and insert a 10-well gel comb without introducing air bubbles.
7. Once the gel has solidified, prepare the samples from Subheading 3.5, step 10 with 6X SB. Add 20 μL of 6X SB to each sample (*see* Note 49). Briefly vortex the sample to mix and briefly spin down with a mini centrifuge.
8. Set up the gel in the Mini-PROTEAN[®] Vertical Electrophoresis Cell, and add the appropriate volume of cold 1X SDS-PAGE running buffer.
9. Mix the protein ladder with an equal part of 1X SB, and load into a preferred well. The final volume loaded should be the same for all samples that are loaded on the gel.
10. Load the desired volume of samples (*see* Note 50). Load any empty well with an equal volume of 1X SB (this may prevent lane bloating).
11. Electrophorese at 25 mA until the protein ladder begins to separate (*see* Note 51), and then continue at 35 mA until the dye front (from the BPB dye in the samples) runs off the gel (*see* Note 52). At all times while running the gel, keep the gel box in complete darkness (*see* Note 53).
12. Optional (*see* Note 54). Place the gel box on ice or in a cold room while running.
13. Following electrophoresis, remove the gel from the plates and rinse with DI water (*see* Note 55). Keep the gel in sterile DI water in the dark and move to the imaging system.

3.7 Imaging with LICOR Odyssey[®] CLx System

1. Initiate the Empiria Studio[®] Software. (<https://www.licor.com/AnalysisQuickStart>).
2. Scan the gel with the following settings: Setup—Custom (10 $\times$ 9 panel scan area), No Analysis. Channels—Auto, (both 700 and 800 checked). Scan Controls—Scan resolution (169 μm), Scan Quality (low), Focus Offset (0.5 mm). Scan Area—Defaults. Odyssey CLx Imager: Scan a Gel ([licor.com](https://www.licor.com)) (*see* Note 56). Analysis of fluorescent band intensity can be performed with the same software tools.

3.8 Total Protein Staining with Coomassie Brilliant Blue

1. After imaging, place the gel in a clean microwave safe plastic container and fill with Coomassie stain (*see* Note 57) to about 1–1.5 cm deep.
2. Microwave the gel with the lid off for 1 min (*see* Note 58) and carefully monitor.

3. Place the gel on an orbital shaker at a low speed for 10 min.
4. After the incubation period, pour off the Coomassie stain (*see* **Note 59**) and rinse the gel with ~100 mL of DI water twice, then pour off any remaining water.
5. Fill the container with destain of about 1–1.5 cm deep.
6. Tie 4 Kimwipes into a knot and place in the corners of the container, being careful to not let the Kimwipes rest on the gel itself.
7. Microwave the gel with the lid off for 1 min.
8. Place the gel on an orbital shaker at a low speed for 10 mins.
9. Repeat **steps 4–8** once more.
10. After the incubation, rinse the gel with DI water and incubate on the orbital shaker again for 10 mins (*see* **Note 60**).
11. Image the gel as previously described in Subheading 3.7 (*see* **Note 61**).

4 Notes

1. The composition of the commercially sourced 1,5-hexadiyne (15HD) used in these experiments is a 50% [V/V] solution in *n*-pentane. (Thermo Scientific Chemicals). This volatile material is difficult to work with and our recommendation is to prepare a large stock volume (0.5 mL or greater) in DMSO. Our observation is that the *n*-pentane solvent preferentially volatilizes over time leading to increases in the 15HD concentration in DMSO. The presence of *n*-pentane in our probing reactions had no measurable effect on activity or covalent modification of 15HD to the proteins of interest (data not shown). When using 15HD as a probe, the effects of *n*-pentane on monooxygenase activity should be evaluated in the absence of 15HD.
2. Other probes may be available that are not listed here. Additionally, some of the reagents listed may have a damaging effect on cells leading to lysis. For example, exposure of cells of *N. europaea* to 1 mM 19DD results in a substantial reduction of measured total protein compared to cells similarly treated with other diynes.
3. This time may be extended but should be monitored closely. Stocks have also been stored at –20 °C for 6 months without issue.
4. Although the manufacturer states the Alexa Fluor™ 647 fluor is photostable, we have observed that repeated light exposure or imaging can reduce labeling intensity.

5. One aliquot is enough for a full 10 reactions using this protocol. Do not reuse leftover THPTA.
6. Ascorbic acid can also be used. Add 0.0199 g of ascorbic acid to 1 mL of sterile DI water and vortex to mix. Final concentration is 113.6 mM. Store at room temperature until ready to use. Must be prepared the day of use. Sodium ascorbate is the preferred reagent.
7. Other reducing agents such as dithiothreitol (DTT) or tris (2-carboxyethyl)phosphine (TCEP) can be used, but the same reductant should be used for all experiments if comparisons are to be made between different samples. Additionally, 6X SB is preferred so that greater loading volumes can be added to the gel without spillover between lanes.
8. Any SDS-PAGE gel electrophoresis system may be used. Hand cast or precast gels may also be used.
9. Other NIR imaging systems may be available or used.
10. The culture volumes described in this section are dependent on the bacterial species. For example, a 1 L culture of *N. europaea* can provide sufficient cells for multiple reactions or experimental conditions, whereas a tenfold smaller (100 mL) culture can provide sufficient material of cells that grow on more energetically favorable substrates such as hydrocarbons. Protein expression and enzymatic activity levels will vary by growth phase and species. We typically harvest cells in all cultures at the mid to late log phase.
11. Washing volume can be scaled up or down depending on the pellet size and the user's needs.
12. The resuspension volume listed here is arbitrary and is based on experimental conditions which can vary by bacterial species or desired downstream experiments. Optimization may be necessary for the user. An optical cell density (OD₆₀₀) of ~4.0 for *Mycobacterium* sp. ELW1 used in several of the experiments reported here results in a BCA protein assay measurement of ~1.0 mg cell protein per mL when performed using the methods described in Subheadings 3.2, 3.3 and 3.4. An OD₆₀₀ of ~3.0 for *N. europaea* results in a BCA protein assay measurement of ~1.5 mg cell protein per mL when performed using the methods described in Subheadings 3.2, 3.3 and 3.4 with the addition of 0.5% SDS. Standardization of reaction volumes/cell densities is highly recommended to maintain consistency throughout experiments and across gels due to protein measurement variations and the protein concentration-dependent nature of the labeling reaction as described in Subheadings 3.4 and 3.5.

13. This protein concentration is arbitrary and can be adjusted higher or lower as needed although if quantification of fluorescence is desired this concentration should be consistent across all samples in each experiment.
14. Higher or lower probe concentrations can be used, but at high concentrations (≥ 1 mM), some probes may cause cell lysis and a loss of cytoplasmic proteins into the reaction medium. Appropriate probe concentration should be determined by the user.
15. The most appropriate diyne probe for labeling specific monooxygenases depends on the user's needs. For example, 1,7-octadiyne (17OD) has been suggested to be a "universal" probe for ammonia-, methane-, and alkane monooxygenases [15]. However, there are well-defined differences between AMO in ammonia-oxidizing bacteria (AOB) and archaea (AOA) in terms of their reactivity toward *n*-alkynes as mechanism-based inactivators [16]. This suggests that 17OD is unlikely to be an effective probe for both forms of this enzyme. We have also demonstrated that aromatic probes, such as DEB, are particularly effective for selectively detecting phenol hydroxylase in phenol-grown cells of *Pseudomonas* sp. CF600 [17]. This selectivity likely reflects the structural similarity between the probe and this enzyme's natural substrates. In addition to enzyme specificity, our experience indicates longer carbon chain diyne probes, such as 18ND and 19DD, often produce high levels of nonspecific labeling compared to shorter diynes even though the level of enzyme inactivation is comparable in both cases (Fig. 2). The level of nonspecific fluorescent labeling likely reflects the difference in the half-life of the reactive intermediate generated during the catalytic activation of diynes by the target monooxygenase. A longer half-life for intermediates generated from 18ND or 19DD compared to diynes, such as 15HD, would enable these intermediates to migrate out of the monooxygenase's active site and covalently modify additional sites on the enzyme as well as other cellular components. This type of "off-target" labeling may be useful if the experimental aim is to detect and quantify monooxygenase-expressing intact cells based on fluorescence but will be detrimental if the experimental aim is to identify component polypeptides of specific monooxygenases.
16. These conditions may be adjusted and optimized for the user's needs. For example, in our experience, a 1 h incubation time is in excess of the amount of time required to obtain complete inactivation of most target monooxygenases. Variables that can impact the extent of monooxygenase inactivation and hence subsequent CuAAC-dependent labeling include the cell/enzyme concentration and the effect of sample agitation on

the availability of O₂ as a required monooxygenase substrate. Most of our experiments also do not include exogenous sources of reductant that are typically needed to support monooxygenase activity. Our experiments typically rely on internal endogenous sources of reductant to support monooxygenase activity and inactivation by diyne probes. Different culture conditions can impact the levels of these endogenous reductant sources, and it is important that the possible impacts of different culture conditions are considered in the design of ABPP studies of bacterial monooxygenases.

17. This tube type is only a recommendation and works well with the Bead Mill 4 Mini Homogenizer.
18. We have observed that 1,4-diethynylbenzene (DEB) can precipitate out of aqueous solution under some conditions. Care should be taken to avoid transferring this precipitate to the screw cap tube, and the user should try to remove it from the serum bottle prior to washing it.
19. This resuspension volume is arbitrary and based on expected protein concentrations. However, all experiments should be performed in the same manner to eliminate discrepancies during Subheadings 3.4 and 3.5. The use of SDS allows for a more efficient release of integral membrane proteins such as AMO from the cell membrane.
20. The cellular location and orientation of monooxygenases can potentially be important if intact diyne-treated cells are used in CuAAC labeling reactions and are subsequently used in either SDS-PAGE/NIR analyses or methods such as FACS, which rely on intact cells. In these cases, it is important to consider the membrane permeability of the reagents required for the two-step labeling approach described in this chapter. For example, the diyne probes we have used to characterize diverse bacterial monooxygenases are typically low-molecular-weight uncharged hydrophobic molecules that can confidently be expected to be highly membrane permeable. This suggests that with these types of probes, inactivation of susceptible monooxygenases will occur irrespective of the cellular location of these enzymes. In contrast, other reagents used for CuAAC reactions, such as charged Cu, ascorbate ions, or more bulky materials such as Alexa Fluor and THPTA, have little or no membrane permeability. The effects of cellular location of monooxygenases by CuAAC-dependent labeling of diyne-inactivated monooxygenases is illustrated in Fig. 3, where the effectiveness of CuAAC labeling reactions on whole and lysed cells of *N. europaea* (Panel A) and *n*-octane-grown *Mycobacterium* sp. ELW1 (Panel B) are compared. Ammonia-grown *N. europaea* cells express a single-monooxygenase, membrane-bound AMO. In contrast, *n*-octane grown

Mycobacterium sp. ELW1 expresses both a cytoplasmic SDIMO responsible for initiating isobutylene catabolism and a membrane-bound AlkB-containing alkane hydroxylase responsible for initiating *n*-octane catabolism (Panel C). Although both AMO and AlkB-containing alkane hydroxylases are both integral membrane-bound enzymes, the active site of AMO, and the presumed site of covalent modification by diyne inactivators, faces the periplasm. In contrast, the equivalent sites in alkane hydroxylases both face the cytoplasm [18]. The results shown in Fig. 3, Panel A, demonstrate that a low level of fluorescent labeling was observed for the AmoA component of AMO when 17OD-treated *N. europaea* were used in whole-cell CuAAC labeling reactions (Lane 5). A much higher level of AMO labeling was observed when the CuAAC reaction was conducted with the same 17OD-treated cells, which had either been lysed by bead beating alone (Lane 7) or beaten in the presence of SDS (Lane 9). This suggests that a suboptimal concentration of the required CuAAC reagents can migrate or be transported through the outer membrane into the periplasmic space of this Gram-negative strain so that they can react with the periplasmically facing active site of diyne-inactivated AMO. This limitation is overcome by lysing cells so that there is no barrier to the mixing of CuAAC reagents with the diyne-inactivated monooxygenase. In contrast, no labeling was observed of either the SDIMO or alkane hydroxylase when intact *n*-octane-grown cells of *Mycobacterium* sp. ELW1 were used in whole-cell CuAAC-labeling reactions (Fig. 3, Panel B, Lane 5). Again, the presumed hydroxylase components of both of these enzymes were effectively labeled when CuAAC reactions were conducted with diyne-treated cells, previously lysed by bead beating alone (Lane 7) or beaten in the presence of SDS (Lane 9). The lack of detection of fluorescently labeled polypeptides, when intact cells were used in CuAAC reactions, is therefore not due to a lack of prior monooxygenase inactivation by diyne probes but is likely a limitation caused by the inability of one or more of the CuAAC reagents to pass through the cell wall or cell membrane of *Mycobacterium* sp. ELW1.

21. The amount of beads indicated here allows for a sufficient volume of supernatant above the beads that can be used for protein assay and/or CuAAC labeling reactions. The amount of beads can be scaled up depending on the user's needs. Other methods of cell lysis can be used based on the user's preference or availability of equipment. Prior work performed by Bennett et al. (2016) used a freeze-thaw fractionation method to lyse cells in *N. europaea* [14].

22. The time and speed can be adjusted depending on the type of organism being lysed. Samples of Gram-negative bacteria are expected to require fewer rounds of bead beating compared to Gram-positive strains.
23. Cooling the sample is a precaution to limit any adverse effects of sample heating during the bead beating step.
24. The number of bead beating cycles required for full lysis is expected to be organism-specific. The extent of lysis can be determined by measuring protein concentration after each beating cycle and continuing with additional cycles until a constant level of protein has been released.
25. *Measuring protein concentration is the most important step in our two-step procedure if samples are to be compared to each other on a single gel or compared across multiple gels and experimental conditions.* The BCA protein assay is preferred due to low sample volume requirements. However, any protein assay can be used as long as it is consistently used across all samples and experiments. Different protein assays or variations within those assays may result in different estimates of protein concentration, and the effect of this is demonstrated in Fig. 4. A single sample of isobutylene-grown *Mycobacterium* sp. ELW1 was probed with 17OD, as described in the Methods section. The diyne-treated cells were then divided into eight separate equal aliquots, and the protein was measured as described in Table 1.

Table 1
Protein concentration determined by different protein assays and methods

Lane number	Protein assay method	Measured protein con. (mg/mL)
1	PageRuler™ prestained NIR protein ladder	NA
2	Bead beaten cells: pre-assay NaOH solubilization treatment; slow centrifugation (3000 RCF, 1 min); biuret assay	1.74
3	Bead beaten cells: Slow centrifugation (3000 RCF, 1 min); biuret assay	10.55
4	Bead beaten cells: slow centrifugation (3000 RCF, 1 min)—Quick spin of extract prior to measuring; biuret assay	3.20
5	Bead beaten cells: slow centrifugation (3000 RCF, 1 min); BCA assay	3.08
6	Bead beaten cells: pre-assay NaOH solubilization treatment; fast centrifugation (10,000 RCF, 5 min); biuret assay	1.55
7	Bead beaten cells: fast centrifugation (10,000 RCF, 5 min); biuret assay	1.78
8	Bead beaten cells: fast centrifugation (10,000 RCF, 5 min); BCA assay	3.35
9	Whole-cell, not bead beaten: Pre-assay NaOH solubilization treatment; fast centrifugation (10,000 RCF, 5 min); Biuret assay	3.20

In Fig. 4, Panel A, there are clear differences in fluorescent band intensities between the samples, despite that an equal amount of protein was added to the CuAAC labeling reaction and was subsequently analyzed by SDS-PAGE and NIR scanning. A total protein stain of the same gel using Coomassie Brilliant Blue is shown in Fig. 4, Panel B, and shows that the differences in protein measurement between the assays resulted in unequal loading of total protein. Another related consideration is the effect of dilution of lysed samples for subsequent protein assays. We have found that if the measured protein concentration falls outside the upper limit of a protein assay standard curve, and the sample needs to be diluted to be properly measured, then this will result in potentially significant differences between diluted and undiluted samples. In summary, if the experimental aim is to compare different samples using ABPP and fluorescent labeling, Subheading 3.2, 3.3, 3.4 and 3.5 should be the same for all samples.

26. It is ideal that the measured protein concentration from samples fall near the middle of the protein assay standard curve.
27. Dilutions listed are arbitrary and can be adjusted to the user's needs.
28. Performing this reaction in PCR strip tubes enables easy sample heating using a thermocycler.
29. If the protein concentration does not fall within the range of the protein assay standard curve, dilutions may be necessary. However, diluting samples with DI water may result in significant changes in measured protein concentration. *Measuring protein accurately and in a standardized fashion for all experiments will reduce variability in fluorescent band intensity across gels and samples.*
30. Heating is performed in a thermocycler as this decreases the time of the assay. However, samples can be incubated at room temperature for 1 h and achieve the desired effect.
31. Our laboratory uses the preprogrammed protein BCA application on the NanoDrop™ Eight Spectrophotometer following the user's guide. Other protein measurement methods and equipment can be used.
32. The CuAAC labeling reaction protocol described here is based on the work of Hong et al. (2009) [19].
33. It is convenient to use PCR tubes for the CuAAC labeling reactions; however, other tubes could be used.
34. *As indicated throughout this protocol, accurate and consistent measurement of protein concentration is necessary for meaningful comparisons of the fluorescent labeling of two or more sample conditions in the same experiment or between separate*

experiments. This requirement also extends to the amount of protein used in individual CuAAC reactions. An example of the effect of using differing protein concentrations in CuAAC reactions is demonstrated in Fig. 5, Panel A. In this instance, each CuAAC labeling reaction was conducted as described in Subheading 3.5. but with varying amounts of protein added from the same extract of 15HD-treated cells of isobutylene-grown *Mycobacterium* sp. ELW1. The amount of protein included in each CuAAC reaction is indicated in the Figure legend, while equal amounts of protein (5 µg) were subsequently analyzed by SDS-PAGE/NIR scanning. There was a nearly threefold decrease in the detected fluorescence of the ~55 kDa hydroxylase component of the SDIMO responsible for initiating isobutylene catabolism as the protein concentration in the CuAAC labeling reactions increased from 25 µg to 75 µg/reaction. This effect was not due to differences in the total amount of protein analyzed based on the level of protein staining observed for the same gel (Fig. 5, Panel B). The cause of this effect might reflect a limitation of one or more of the CuAAC reagents used (i.e., Alexa Fluor, THPTA, copper, ascorbate) and suggests that CuAAC reaction optimization will likely be required when different amounts of protein are used in CuAAC reactions and then analyzed by SDS-PAGE and NIR scanning.

35. It is recommended that ascorbate is the last reagent added to CuAAC reactions [20].
36. Ensure that the sample does not contain cellular debris or beads before extracting protein from the supernatant. Centrifuge the samples at ~17,000 RCF for 10 mins at 4 °C immediately prior to performing the labeling reaction (speed and time can be adjusted to 10,000 RCF and 5 mins if necessary).
37. Higher or lower concentrations may be used. Despite its reported photostability, we store the fluorophore in the dark for as long as possible before use. Additionally, other labels or tags (i.e., biotin azide or 6XHis azide) can be incorporated at this step, though optimization, protein precipitation, washing, and column purification may be necessary and are not discussed directly in this report.
38. THPTA is a water-soluble accelerating ligand that maintains the Cu(I) oxidation state necessary for CuAAC reactions and protects other biomolecules from damage during the labeling process [21, 22]. THPTA may interfere with whole-cell CuAAC labeling reactions of intact whole cells by chelating copper ions and preventing them from entering the outer cell membrane. This inhibitory effect of THPTA was previously reported for the labeling of AMO in 17OD-treated intact cells of *N. europaea* [14], but this effect is overcome by lysing cells prior to CuAAC reactions.

39. The concentration of copper described in this chapter is much lower than that originally described for labeling AmoA in *N. europaea* [14]. If possible, the use of a low concentration of copper is encouraged as high copper concentrations can cause protein aggregation and degradation, especially in the presence of sodium ascorbate [19, 23]. Sodium ascorbate in CuAAC reactions can enhance the rate of the formation of copper adducts of some amino acid side chains, thus increasing protein cross-linking and precipitation, [19]. Additionally, if attempting to label cell-surface molecules, the use of THPTA reduces the mortality of live cells by copper toxicity [24].
40. Sodium ascorbate is generally preferred as a source of reductant, though not in all CuAAC reactions [19].
41. The use of aminoguanidine may help alleviate protein adduct formation, cross-linking, and precipitation of proteins because of copper and sodium ascorbate-mediated reactions [19]. The effect of aminoguanidine on labeling can be seen in (Fig. 5, Panel A).
42. Diamide can be used to oxidize thiol-containing compounds to their disulfide form and thereby deplete cells of endogenous reductants that can potentially reduce Cu^{2+} in the absence of exogenous reductants such as ascorbate. Determining if this is a necessary reagent, including the appropriate concentration of diamide, must be tested by the user. The effect of diamide on labeling can be seen in (Fig. 5, Panel A).
43. Shorter incubation times are likely sufficient and may be adjusted but should be experimentally determined.
44. After CuAAC labeling reactions have been completed the mixtures can be frozen at $-20\text{ }^{\circ}\text{C}$ and then analyzed by SDS-PAGE at a later date. However, our experience indicates that frozen storage of labeled samples is not recommended if the experimental design involves comparisons of fluorescent intensities across samples or experiments.
45. Any size or brand of SDS-PAGE gel should work similarly.
46. The approximate distance from the well bottom and the resolving gel is $\sim 4\text{--}5\text{ mm}$ but can be adjusted to the user's needs.
47. We gently add 1 mL of DI water to the center of the mold. Isopropanol or ethanol can also be used as the overlay.
48. There is various guidance on how long it takes to completely polymerize an SDS-PAGE gel [25]. The resolving gels are typically usable after polymerization for 30 min at room temperature when using the methods described in this chapter.
49. Some recommendations suggest heating the samples before loading but sample heating is dependent on the protein of interest. For example, many SDS-PAGE protocols recommend

heating samples to high temperature in the presence of reducing agents to fully denature proteins and to ensure a consistent level of SDS binding. However, heating AMO-containing samples in the presence of 2-mercaptoethanol can result in selective aggregation of AmoA [25]. Sample heating was not performed on any of the bacterial samples mentioned in this chapter.

50. Loading ~5 µg of protein is enough to be visualized on a gel by fluorescent labeling or Coomassie staining.
51. The ladder listed in the Materials section of this chapter can be substituted for another of the user's preferences.
52. If using the ladder listed in the Materials section of this chapter, be aware that the smallest sizing polypeptide (11 kDa) often migrates with the SDS-PAGE dye front. We typically run the dye front off the gel because otherwise the lower portions of the gel can contain unreacted Alexa Fluor, which can interfere with quantification of protein-associated fluorescence.
53. Running the gel in the dark may prevent premature photodegradation of the fluorophore. Running the gel in the light is an option, and visualization of the labeled proteins will be seen. However, this is not recommended if comparisons between fluorescent intensities need to be made across samples or separate gels.
54. We run our gel electrophoresis units resting on ice in an ice bucket. This reduces gel heating during electrophoresis and diminishes protein band distortion, which can interfere with fluorescence quantification.
55. The gel can remain in the glass plates and be successfully scanned and captured. If you are using precast gels in plastic, the plastic should be removed as the plastic will distort the imaging process.
56. The focus offset will need to be adjusted depending on the thickness of the gel or if the gel is still in glass. It is advised to have the focal distance in the middle of the gel. Resolution, image quality, and other variables can be adjusted to the user's preference, as long as the same settings are used each time if labeling intensities are to be compared across samples or gels.
57. Other stains can be used to suit the user's needs but should be consistent throughout experiments.
58. Microwaving the gel in the stain will reduce the binding time of the stain to the proteins. However, the gel can be incubated at room temperature for 1 h for the same effect.
59. The stain can be reused 2–3 times without any apparent loss in staining capacity.

60. The gel can be left in DI water for an extended period, although it will continue to lose stain in the gel and eventually the protein bands themselves. The method described in this chapter is to standardize the technique as variations in staining time and destaining will cause total lane signal to vary across gels and experiments. If comparisons of fluorescent intensities of labeled proteins are to be made across samples and gels, then standardizing the staining and destaining process is paramount to ensure that total protein loading can be normalized across experiments. If this is not necessary, then the staining and destaining process described in this report are arbitrary and can be adjusted to the user's needs.
61. When quantifying protein fluorescence in SDS-PAGE/NIR analyses, it can be helpful to demonstrate that equal amounts of protein were analyzed for all samples. This can be achieved by using the same gel imaging system discussed in this chapter. When excited by NIR, Coomassie Brilliant Blue stain emits light in the 700–800 nm range [26]. Imaging software, such as Empiria Studio[®], can then be used to quantify the total fluorescence in a Coomassie Blue stained gel lane for sample normalization purposes. We have found the 700 channel to conflict with the 800 channel for normalization. Although not experimentally determined, residual fluorescence from labeled proteins by Alexa Fluor[™] 647 may contribute to a wider range of total lane fluorescence measurements when using the 700 nm channel compared to the 800 nm channel. It is recommended to experimentally determine the combined linear range (the range of protein to load on a gel that produces a linear response between total protein staining and fluorescent signal intensity of labeled proteins (see Western Blot Linear Range Is Important for Data Accuracy ([licor.com](https://www.licor.com)) for more guidance) if accurate comparisons are to be made across samples or gels.

Acknowledgments

Support for this work was provided by fellowships from the NC State University Genetics and Genomics Scholars Program as well as funding provided by the NC State University Molecular Biotechnology Training Program, NIH 1T32GM133366 to JJ.

References

1. Minerdi D, Loqui D, Sabbatini P (2023) Monooxygenases and antibiotic resistance: a focus on Carbapenems. *Biology* 12:1316
2. Torres Pazmiño DE, Winkler M, Glieder A et al (2010) Monooxygenases as biocatalysts: classification, mechanistic aspects and biotechnological applications. *J Biotechnol* 146:9–24

3. Smith TJ (2009) Monooxygenases, bacterial: oxidation of alkenes. In: Encyclopedia of industrial biotechnology. Wiley, pp 1–13
4. Heine T, van Berkel WJH, Gassner G et al (2018) Two-component FAD-dependent monooxygenases: current knowledge and biotechnological opportunities. *Biology* 7:42
5. Banerjee R, Jones JC, Lipscomb JD (2019) Soluble methane monooxygenase. *Annu Rev Biochem* 88:409–431
6. Groves JT, Feng L, Austin RN (2023) Structure and function of alkane monooxygenase (AlkB). *Acc Chem Res* 56:3665–3675
7. Hyman MR, Wood PM (1985) Suicidal inactivation and labelling of ammonia monooxygenase by acetylene. *Biochem J* 227:719–725
8. Porta EOJ, Steel PG (2023) Activity-based protein profiling: a graphical review. *Curr Res Pharmacol Drug Discov* 5:100164
9. Fuerst R, Breinbauer R (2021) Activity-Based Protein Profiling (ABPP) of oxidoreductases. *Chembiochem* 22:630–638
10. Hynes RK, Knowles R (1978) Inhibition by acetylene of ammonia oxidation in *Nitrosomonas europaea*. *FEMS Microbiol Lett* 4:319–321
11. Prior SD, Dalton H (1985) Acetylene as a suicide substrate and active site probe for methane monooxygenase from *Methylococcus capsulatus* (Bath). *FEMS Microbiol Lett* 29:105–109
12. McTavish H, Fuchs JA, Hooper AB (1993) Sequence of the gene coding for ammonia monooxygenase in *Nitrosomonas europaea*. *J Bacteriol* 175:2436–2444
13. van Kessel MAHJ, Speth DR, Albertsen M et al (2015) Complete nitrification by a single microorganism. *Nature* 528:555–559
14. Bennett K, Sadler NC, Wright AT et al (2016) Activity-based protein profiling of ammonia monooxygenase in *Nitrosomonas europaea*. *Appl Environ Microbiol* 82:2270–2279
15. Sakoula D, Smith GJ, Frank J et al (2022) Universal activity-based labeling method for ammonia- and alkane-oxidizing bacteria. *ISME J* 16:958–971
16. Wright CL, Schatteman A, Crombie AT et al (2020) Inhibition of ammonia monooxygenase from ammonia-oxidizing archaea by linear and aromatic alkynes. *Appl Environ Microbiol* 86: e02388–e02319
17. Oyarzun Mejia AP, Hyman MR (2023) Dinyne inactivators and activity-based fluorescent labeling of phenol hydroxylase in *Pseudomonas* sp. CF600. *FEMS Microbiol Lett* 370:fnad002
18. Guo X, Zhang J, Han L et al (2023) Structure and mechanism of the alkane-oxidizing enzyme AlkB. *Nat Commun* 14:2180
19. Hong V, Presolski SI, Ma C et al (2009) Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation. *Angew Chem Int Ed Engl* 48:9879–9883
20. Presolski SI, Hong VP, Finn MG (2011) Copper-catalyzed Azide-alkyne click chemistry for bioconjugation. *Curr Protoc Chem Biol* 3: 153–162
21. Meng G, Guo T, Ma T et al (2019) Modular click chemistry libraries for functional screens using a diazotizing reagent. *Nature* 574:86–89
22. Michaels HA, Zhu L (2011) Ligand-Assisted, Copper(II) Acetate-accelerated azide-alkyne cycloaddition. *Chem Asian J* 6:2825–2834
23. Zuily L, Lahrach N, Fassler R et al (2022) Copper induces protein aggregation, a toxic process compensated by molecular chaperones. *mBio* 13:e03251-21
24. Li L, Zhang Z (2016) Development and applications of the copper-catalyzed Azide-alkyne cycloaddition (CuAAC) as a bioorthogonal reaction. *Molecules* 21:1393
25. Kurien BT, Scofield RH (2012) Common artifacts and mistakes made in electrophoresis. *Methods Mol Biol (Clifton NJ)* 869:633–640
26. Butt RH, Coorssen JR (2013) Coomassie blue as a near-infrared fluorescent stain: a systematic comparison with Sypro ruby for in-gel protein detection. *Mol Cell Proteomics MCP* 12: 3834–3850



Activity-Based Profiling of Reactive Nucleophilic Cysteine Sites from Tissue Proteomes for Drug Discovery Applications

Hasini Kalpage, An Chi, and Smaranda Bodea

Abstract

Identifying engaged nucleophilic sites in a proteome-wide manner has become a powerful means to leverage activity-based protein profiling toward assessing the target landscape of covalent compounds ranging from fragments to approved drugs. In this chapter, we provide a detailed protocol for reactive cysteine profiling from tissue lysates using a low-input high throughput-compatible workflow based on multiplexing via TMT-16pro. We apply this method to a representative biological context, human heart tissue lysates treated with highly reactive and promiscuous electrophilic scout fragments. We detect over 11,000 cysteine sites and multiple engaged ligandable sites specific to each scout fragment, revealing the quantitative accuracy of this method. Due to the low input material requirements, this protocol can be applied to many different contexts relevant to drug discovery, including the profiling of tissues from animal models after in vivo treatment with covalent drugs.

Key words Covalent small molecule, Drug discovery, TMTpro, High-throughput, Mass spectrometry, Tissue proteome

1 Introduction

Cysteine is a redox-sensitive, unique amino acid with high nucleophilicity at physiological pH that plays important catalytic and regulatory roles in a range of enzymes such as kinases, phosphatases, oxidoreductases, proteases, deubiquitinating enzymes, and acyltransferases. Engaging cysteine residues with covalent (reversible or irreversible) electrophilic compounds has recently emerged as a powerful strategy for drug proteins that lack defined binding pockets and were previously thought to be “undruggable,” such as KRAS G12C [1–4]. This renewed interest in covalent ligands has accelerated the development and optimization of methods that can enable the discovery of chemical matter engaging nucleophilic sites or can characterize the selectivity of covalent compounds [5]. One

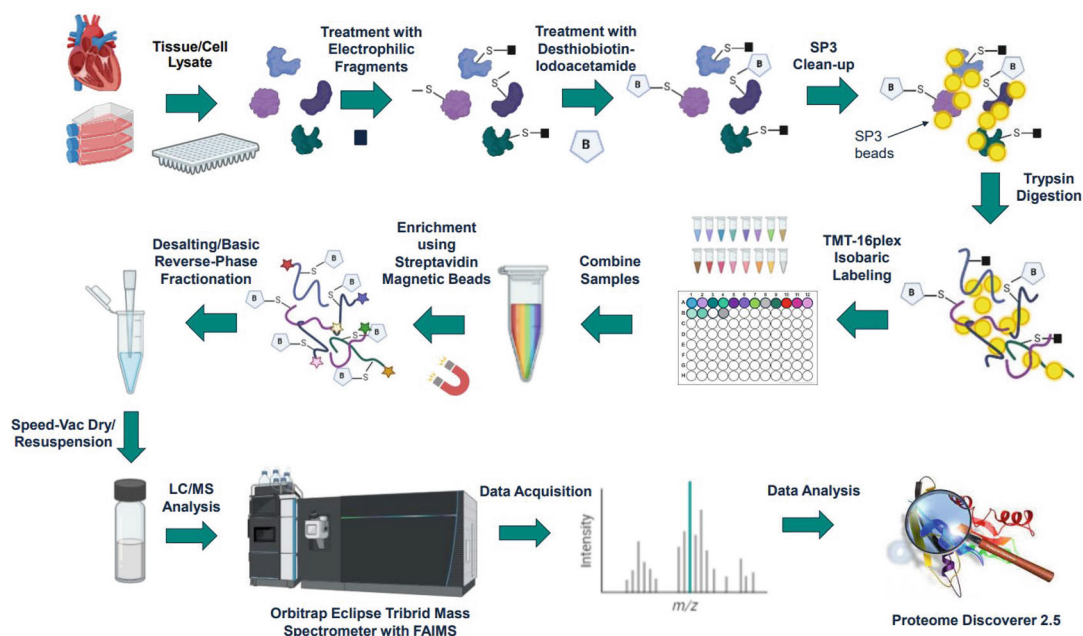


Fig. 1 General protocol for reactive cysteine profiling. Treatment with electrophiles can occur either in lysate or prior to lysate generation (in cell or in vivo). The protocol involves labeling of reactive cysteines with a desthiobiotin iodoacetamide probe, followed by single-pot, solid-phase-enhanced sample preparation (SP3) bead-based cleanup, digestion of proteins into peptides, multiplexing by TMTpro labeling, enrichment using magnetic streptavidin beads, optional off-line fractionation prior to LC-FAIMS-MS/MS analysis, and data analysis and interpretation with Proteome Discoverer Software

such method leveraged both for ligandability evaluation as well as target engagement and off-target identification is activity-based profiling of cysteine sites using discovery proteomics [1, 6].

In this chapter, we describe an ABPP chemoproteomic method to identify ligandable cysteines in the right ventricular human heart tissue proteome using low amounts of input material (from 100 to 30 μg total protein per condition afford similar results) (Fig. 1). We leveraged four previously reported promiscuous scout fragments (CL33, KB02, KB03, KB05) that covalently modify different cysteine residues across the proteome as proof-of-concept ligands that validated the quantitative accuracy of this approach [2, 7]. Ligand treatment of tissue lysates was followed by the addition of a reported desthiobiotin-iodoacetamide (DBIA) probe to label reactive cysteine sites and allow for streptavidin enrichment of peptides containing sites unbound by the covalent fragments [7, 8]. We also used single-pot, solid-phase enhanced sample preparation (SP3) beads for sample cleanup and improved recovery of modified-cysteine peptides after trypsin/Lys-C digestion [9]. Peptides were enriched using magnetic streptavidin beads that are also compatible with 96-well plates for enhanced

throughput and afford recovery comparable to previously published methods using agarose beads [9]. Labeled peptides were analyzed using an Orbitrap Eclipse Tribrid state-of-the-art mass spectrometer combined with high-field asymmetric waveform ion mobility spectrometry (FAIMS), where ions were further separated by their mobility in high and low electric fields, improving cysteine coverage as previously reported [9]. For quantification and greater throughput, peptides were 16-plex labeled by tandem mass tags (TMTpro) [7, 8]. As a result, using this TMT-based mass spectrometry workflow, we were able to obtain quantitative cysteine ligandability information across the complex biological system of heart tissue, profiling over 11,000 cysteine sites. Furthermore, this methodology optimized for tissue lysates can be used for target engagement and selectivity profiling of covalent compounds from various biological contexts ranging from cell lysates to tissues of preclinical models after in vivo drug treatment.

2 Materials

2.1 *General Equipment*

1. FastPrep-24 Classic bead beating grinder and lysis system.
2. Lysing matrix SS- for heart tissue.
3. Ultracentrifuge.
4. Eppendorf ThermoMixer C.
5. Plate sealer.
6. 96-well plate magnet.
7. Bath sonicator.
8. Sorvall centrifuge.
9. 1.5 mL Eppendorf tube magnet.
10. SpeedVac Vacuum Concentrator.
11. Eppendorf mini-spin centrifuge.
12. Orbitrap Eclipse Mass Spectrometer with FAIMS interfaced with a vanquish Neo UHPLC system.
13. EASY-Spray PepMap RSLC C18 analytical column.

2.2 *General Chemicals and Supplies*

1. 1 M HEPES solution.
2. 5 M Sodium Chloride Solution.
3. 1 M Magnesium Chloride Solution.
4. Surfact-Amps NP-40.
5. Glycerol.
6. Benzonase.

7. Complete ULTRA EDTA-free protease inhibitor cocktail tablets.
8. PhosSTOP phosphatase inhibitor cocktail tablets.
9. 100X Halt protease and phosphatase inhibitor cocktail, EDTA-free.
10. Pierce BCA protein assay.
11. 200 μ L Eppendorf PCR plates.
12. Protein LoBind Eppendorf 1.5 mL microcentrifuge tubes.
13. DMSO.
14. DTT.
15. Iodoacetamide.
16. 10% SDS.
17. 200 μ L, large orifice pipette tips.
18. 1000 μ L, large orifice pipette tips.
19. Sera-Mag speed beads (SP3) 15 mL.
20. Optima LC/MS-grade acetonitrile.
21. Ethanol-HPLC/spectrophotometric-grade.
22. Trypsin/Lys-C mix.
23. 1 M triethylammonium bicarbonate (TEAB) buffer for TMT experiments.
24. Calcium Chloride solution, BioUltra for molecular biology, 1 M in H₂O.
25. TMTpro 16-plex label Reagent Set.
26. Anhydrous acetonitrile.
27. 50% hydroxylamine.
28. Protein LoBind Eppendorf 2 mL microcentrifuge tubes.
29. Pierce streptavidin magnetic beads, 50% slurry.
30. 1X PBS, pH 7.2.
31. Urea.
32. Optima LC/MS-grade water.
33. Trifluoroacetic Acid, HPLC-grade.
34. C18 microspin columns.
35. Ammonium bicarbonate.
36. Methanol.
37. Formic acid, LC-MS grade.
38. Mass spectrometry microvials.

3 Methods

3.1 *Compound Treatment and Reactive Cysteine Labeling of Tissue Lysates*

1. Using a FastPrep24 classic bead beating and lysis system (MP Biomedicals), cryo-pulverize frozen tissue slices and lyse in 50 mM HEPES pH 7.5, 150 mM NaCl, 1.5 mM MgCl₂, 0.8% NP-40, 5% glycerol, 1 mM DTT, and 0.5 units/ μ L of benzonase supplemented with complete ULTRA EDTA-free protease inhibitor cocktail tablets (Roche) and PhosSTOP phosphatase inhibitor cocktail tablets (Roche).
2. Use lysing matrix SS (MP Biomedicals) for heart tissue with 2–3 lysing cycles (speed 6, 45 s/cycle).
3. Spin the heart lysates in the ultracentrifuge at 10000 $\times g$ for 1 h at 4 °C.
4. Quantify the heart lysate protein concentration using the Pierce BCA protein assay kit (Thermo Scientific) per the manufacturer's protocol.
5. Dilute the lysate in 50 mM HEPES pH 7.5, 150 mM NaCl, 1.5 mM MgCl₂, 5% glycerol, and 0.5 units/ μ L of benzonase supplemented with 1X EDTA-free protease and phosphatase inhibitor cocktail to result in 100 μ g tissue lysate per 60 μ L in each channel. Note: NP-40 content was adjusted to 0.4% in the diluted lysate.
6. Incubate the heart tissue lysates (100 μ g per condition) with 500 μ M of each scout fragment: CL33, KB02, KB03, KB05, and DMSO as a vehicle control in triplicate. Perform the treatment for 1 h at room temperature while shaking at 750 rpm in a ThermoMixer (Eppendorf) either in an Eppendorf PCR 96-well plate or protein LoBind 1.5 mL Eppendorf tubes. Note: The final % DMSO was kept at 0.5%.
7. Following compound treatment, add 500 μ M desthiobiotin iodoacetamide (DBIA) probe from a DMSO stock to each sample and incubate for 1 h in the dark at room temperature, while shaking at 750 rpm in a ThermoMixer to block the open cysteine sites unoccupied by the cysteine-targeting fragment compounds (CL33, KB02, KB03, KB05).

3.2 *Reduction/Alkylation*

1. To quench the excess DBIA and to reduce disulfide bonds, add 5 mM DTT (reconstituted fresh to obtain a 500 mM stock solution) to each reaction, and incubate for 30 min in the dark at room temperature with shaking at 750 rpm in a ThermoMixer.
2. Add 1% SDS (stock of 10% SDS) to each well using a multi-channel pipette. Heat the plate at 65 °C for 10 mins while shaking at 750 rpm in the ThermoMixer. Interim Stopping Point: The sample can be stored overnight at –80 °C (seal the plate with foil using a plate sealer).

3. If the samples are frozen overnight, heat the samples at 37 °C for 30 min while shaking at 750 rpm in a ThermoMixer to resolubilize any SDS that precipitated.
4. Subsequently, alkylate the reduced cysteine residues by adding 20 mM iodoacetamide for 30 min in the dark at 37 °C while shaking at 750 rpm in a ThermoMixer.

3.3 SP3-Bead-Based Cleanup and Digestion

1. Prepare a 1:1 mixture of Sera-mag SP3 (single-pot, solid-phase, enhanced sample preparation) speed bead mixture (#45152105050250; #65152105050250, GE Healthcare) reconstituted in LC-MS grade water. Note: Use a large orifice pipette tip when pipetting out beads.
2. Add 200 µg of SP3 beads for 100 µg of lysate (2:1 bead: protein ratio) per well. Higher ratios of bead to protein will provide an even better recovery upon digestion.
3. Add 100% LC/MS-grade acetonitrile at a volume equal to the current total volume (~80 µL) in each well. Using a multichannel pipette, mix by quickly pipetting up and down three times. Incubate the mixtures for 10 min. Beads will clump at this stage but it is normal.
4. Place 96-well PCR plate on a 96-well plate magnet, and incubate for 10 min (for subsequent steps, keep the plate on the magnet). Note: Carefully inspect if the beads are out of solution. If not, this might point to insufficient DNA shearing/lipid membrane degradation.
5. Carefully remove supernatant without touching the beads by using a multichannel pipette.
6. Add 200 µL of 70% ethanol to fill the well completely. Note: Adding 200 µL ensures that all remaining lysis buffer is rinsed from the walls of the well. Incubate for 15 s and remove 70% ethanol.
7. Repeat the 70% ethanol washing step twice for a total of three times and remove ethanol completely.
8. Remove the plate off the magnet and allow the beads to air-dry for about 1–2 min.
9. Resuspend Lys-C/trypsin mixture in 50 mM triethylammonium bicarbonate (TEAB), pH 8 supplemented with 1 mM CaCl₂ to afford 0.045 µg/µL solution.
10. Add 45 µL of Lys-C/trypsin per well, resulting in ~2 µg of enzyme per reaction.
Note: Do not pipette up and down at this step, as beads might stick to the tips.
11. Sonicate the plate using a Qsonica bath sonicator (amplitude: 20–40, 1 s on, 1 s off, 10 s duration) to break up aggregates

and resuspend the beads. Perform multiple rounds of sonication until beads are in solution. Note: If a sonicating water bath is not available, incubate the plate at 37 °C in a ThermoMixer for 10 min while shaking at 1000 rpm.

12. Following sonication or thermal shaking, seal the plate and centrifuge at 1000 g for 10 sec.
13. Incubate the PCR plate on a ThermoMixer at 37 °C at 750 rpm overnight. Interim Stopping Point: Samples can be stored at –80 °C until further use.

3.4 TMT Labeling

1. Immediately before use, equilibrate one pouch (0.8 mg) of the TMTpro 16-plex mass tag labeling reagents to room temperature for 30 min.
2. After overnight incubation, spin-down the plate briefly at 1000 g for 10 s before unsealing.
3. For each of the 800 µg TMT vials, add 41 µL of anhydrous acetonitrile, per TMT manufacturer's protocol.
4. Then, add ~200 µg (12 µL) of reconstituted reagent to each sample of digested peptides (1:2 protein: TMT ratio).
5. Incubate the reaction for 1 h at room temperature, with 700 rpm mixing.
6. Add 5% hydroxylamine (diluted from 50% solution) to each sample resulting in a final concentration of 0.5% and incubate for 15 min at room temperature in a ThermoMixer with 500 rpm shaking to quench the reaction.
7. Place the plate on a 96-well plate magnet to remove the SP3 beads out of the solution.
8. Combine all 16 labeled samples from plex into a new 2 mL low-protein binding tube (~700 µL) and vacuum dry in the SpeedVac concentrator to remove half of the acetonitrile.

3.5 DBIA-Labeled Cysteine Peptide Enrichment Using Streptavidin Beads

1. Use the 1.5-mL-tube magnet to remove any remaining SP3 beads, and transfer the supernatant to a new 2 mL low-protein binding Eppendorf tube.
2. Wash the beads 1 mL 1X PBS while placed on the magnet and combine with the previous supernatant in the 2 mL tube.
3. Use 4 mg of Pierce streptavidin magnetic bead 50% slurry (5 mL supplied at 10 mg/mL in water +0.05% sodium azide) per sample. Note: For any steps involving streptavidin beads, use tips with a large orifice.
4. Before use, wash the beads twice with 1 mL of 1X PBS, pH 7.2. Incubate for 5 mins on and off the magnet when washing.
5. Resuspend the magnetic bead slurry with 1X PBS, pH 7.2 by adding 100 µL per 4 mg of streptavidin beads.

6. Add 100 μ L of Pierce streptavidin magnetic bead 50% slurry to each tube containing DBIA-labeled cysteine peptides. Note: Use a large orifice pipette tip.
7. Top up each tube to 2 mL by adding 1X PBS, pH 7.2.
8. Incubate the peptide-streptavidin magnetic bead mixture by rotating end-over-end overnight at 4 °C to enrich for TMT-labeled, DBIA-conjugated cysteine peptides. Note: Alternatively, this step can be performed at room temperature for 2 h.
9. Following enrichment, place the beads on a 1.5-mL magnetic rack and allow to equilibrate for 5 min.
10. Wash beads to remove non-specific binding using the following steps:
 - Wash 3 times with 1 mL of 1X PBS, pH 7.2.
 - Wash 2 times with 0.8 mL of 1X PBS with 0.1% SDS, pH 7.2.
 - Wash 2 times with 1 mL of 1X PBS with 0.1 M urea, pH 7.2.
 - Finally, wash 3 times with 1 mL of LC/MS-grade water.

Note: Resuspend beads using a large orifice pipette between washes and place on the magnet between each wash.
11. To elute DBIA-conjugated cysteine peptides, add 500 μ L of 50% acetonitrile with 0.1% trifluoroacetic acid (TFA), and mix the beads at 1000 rpm for 10 min at room temperature in the ThermoMixer.
12. Place the tube on a 1.5-mL magnet and transfer the eluted peptides to a new protein low-binding Eppendorf tube, and additionally wash the beads with 200 μ L of 50% acetonitrile +0.1% TFA.
13. Again, mix the tube contents at 1000 rpm for 10 min at room temperature and combine with the original 500 μ L elution.
14. Vacuum dry the eluted sample to completion using a SpeedVac concentrator. Interim stopping point: Peptides can be stored at –80 °C until further processing.

3.6 Desalting and Basic pH Reversed Phase Fractionation

1. For this procedure, use SEM SS18V C18 microspin columns (Nest Group).
2. Resuspend the peptides in 200 μ L of 10 mM ammonium bicarbonate (pH 8.0) + 4% acetonitrile.
3. To condition the column, add 100 μ L of 100% methanol and centrifuge for 15 s at 110 $\times$ *g* (~800 rpm on an Eppendorf mini-centrifuge). Flush once with 100 μ L of 50% acetonitrile, 10 mM ammonium bicarbonate (pH 8.0), and once with

10 mM ammonium bicarbonate (pH 8.0), by spinning at 110 g for 15 sec each time.

4. To process the sample, place the column in a new 2 mL microcentrifuge tube. Add 100 μ L of the sample at a time and spin the column at 110 $\times g$ for 15 s.
5. Wash the column once with 100 μ L of 10 mM ammonium bicarbonate (pH 8.0) by spinning at 110 g for 15 s.
6. Elute bound peptides into four fractions (50 μ L each) with 25% acetonitrile +10 mM ammonium bicarbonate (pH 8.0), 35% acetonitrile +10 mM ammonium bicarbonate (pH 8.0), 50% acetonitrile +10 mM ammonium bicarbonate (pH 8.0), and 60% acetonitrile +10 mM ammonium bicarbonate (pH 8.0).
7. Dry the fractionated peptides to completion using a SpeedVac concentrator.
8. After drying ammonium bicarbonate, add 100 μ L of 70% acetonitrile +0.1% formic acid. Vacuum dry to completion using a SpeedVac concentrator.
9. Resuspend the peptides in 10 μ L of 4% acetonitrile +0.1% formic acid resuspension buffer. Optional: Briefly sonicate the tube for 30 s in a bath sonicator for improved resuspension.
10. Transfer the resuspended peptides to mass spectrometry microvials.

3.7 Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS)

1. Samples are analyzed using an Orbitrap Eclipse Mass Spectrometer with FAIMS interfaced with a Vanquish Neo LC system (Analytical column: EASY-Spray PepMap RSLC C18, 2 μ m particle size, 75 μ m $\times$ 50 cm). The FAIMS compensation voltages are: -35 CV, -55 CV, and -75 CV.
2. Inject 8 μ L of each fraction into the LC/MS system. Optional: Perform a 1 μ L test injection to determine the overall peptide intensity in the sample and, therefore, the suitable LC gradient length.
3. Peptides from each sample are separated at a flow rate of 300 nL/min using a 60-min or 120-min nonlinear gradient from 5% to 38% solvent B (0.1% formic acid in acetonitrile).
4. Spectra are acquired in data-dependent mode, automatically switching between MS1 and MS2 spectra. The spray voltage of the EASY-Spray source is set at 1900 V.
5. The MS1 scan is performed using the Orbitrap with a resolution of 120,000, scan range of 350–1600 m/z, RF lens of 30%, normalized AGC target of 150%, maximum injection time of 50 ms, and data type “profile” mode. Enable dynamic exclusion (repeat count 1, duration 60 s).

6. The MS2 scan is performed in quadrupole isolation mode (isolation window 0.7 m/z) of precursor ion followed by high-collision dissociation (HCD) in the orbitrap with normalized AGC target of 250%, normalized HCD collision energy of 38% and maximum injection time of 86 ms.

3.8 Mass Spectrometry Data Processing

1. The raw MS files are searched using Proteome Discoverer 2.5 with a FASTA formatted UniProt human database, including common contaminants and reversed sequences.
2. The following parameters are used for the search: precursor mass tolerance of 50 ppm and fragment mass tolerance of 0.02 Da. Static modifications are TMTpro16 (+304.207 Da) of lysine and peptide N-termini. Dynamic modifications are cysteine carbamidomethylation (+57.021 Da), methionine oxidation (+15.995 Da), and DBIA modification on cysteines (+296.185 Da). Protein N-terminus dynamic modifications are as follows: N-terminal acetylation (+42.011 Da), methionine-loss (−131.040 Da), and methionine-loss + acetylation (−89.030 Da).
3. Trypsin/Lys-C is selected as the proteolytic enzyme with up to two missed cleavages allowed. The minimum and maximum peptide length ranges from 6 to 144 amino acids.
4. The false discovery rate (FDR) target for peptides and PSMs is set at 1%.
5. The “peptide groups” table is filtered for peptides that contain “DBIA” modification, exported from Proteome Discoverer, and formatted. The results are imported into Perseus 1.6.15.0 as .txt files.
6. Log2 transformation, data filtering based on contaminants and missing values, normalization, imputation, annotation, and statistical analysis are performed in Perseus 1.6.15.0. Proteins listed as “contaminants” are filtered off the list. Column-based median normalization is performed. Peptides with more than 70% missing values in one plex were removed. Missing data were imputed for each condition by drawing random values based on the original distribution of each column (width 0.3, downshift 1.8). Statistical analysis was performed using a two-sided Student’s T-test with Benjamini-Hochberg correction (FDR = 0.05).
7. Volcano plots were generated to compare the cysteine site engagement by the compound fragments in comparison to the DMSO control. BH-corrected -log Student’s T-test p-value for DMSO vs. compound was plotted against the Student’s T-test difference for DMSO vs. compound.

4 Notes

1. TMTpro labeling efficiency can be tested as previously reported [10]. Raw files were searched using Proteome Discoverer 2.5 using Carbamidomethyl/+57.021 (C) as a static modification and TMTpro/+304.207 (K and peptide N-terminus) as dynamic modifications.
2. Increasing the amount of SP3 beads with respect to lysate (up to a ratio of 8:1) can lead to improved peptide recovery and increased depth of profiling.
3. The method described in this chapter has been deployed for irreversible covalent compounds. We have seen success with profiling the target landscape of certain reversible covalent compounds in cell lysates by decreasing the incubation time with the DBIA probe following compound treatment. We note that protocol optimization is needed in such cases and incubation times are dependent on the reactivity of the small molecule studied.

Acknowledgments

We thank Dylan Charles Mitchell and José Navarrete-Perea for helpful conversations in the streamlining of this workflow.

References

1. Maurais AJ, Weerapana E (2019) Reactive-cysteine profiling for drug discovery. *Curr Opin Chem Biol* 50:29–36
2. Backus KM et al (2016) Proteome-wide covalent ligand discovery in native biological systems. *Nature* 534(7608):570–574
3. Patricelli MP et al (2016) Selective inhibition of oncogenic KRAS output with small molecules targeting the inactive state. *Cancer Discov* 6(3):316–329
4. Boike L, Henning NJ, Nomura DK (2022) Advances in covalent drug discovery. *Nat Rev Drug Discov* 21(12):881–898
5. Lucas SCC et al (2024) Covalent hits and where to find them. *SLAS Discov* 29(3):100142
6. Weerapana E et al (2010) Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* 468(7325):790–795
7. Kuljanin M et al (2021) Reimagining high-throughput profiling of reactive cysteines for cell-based screening of large electrophile libraries. *Nat Biotechnol* 39(5):630–641
8. Yang K et al (2024) Accelerating multiplexed profiling of protein-ligand interactions: high-throughput plate-based reactive cysteine profiling with minimal input. *Cell Chem Biol* 31(3):565–576 e4
9. Yan T et al (2021) SP3-FAIMS Chemoproteomics for high-coverage profiling of the human Cysteinome*. *Chembiochem* 22(10):1841–1851
10. Li J et al (2020) TMTpro reagents: a set of isobaric labeling mass tags enables simultaneous proteome-wide measurements across 16 samples. *Nat Methods* 17(4):399–404



Activity-Based Protein Profiling for Functional Cysteines and Protein Target Identification

Tin-Yan Koo and Clive Yik-Sham Chung

Abstract

Activity-based protein profiling (ABPP) has been widely used for proteome-wide cysteine profiling to study functional cysteines in biological systems. ABPP can also be performed in a competitive manner for lead compound discovery and target identification of the lead compound. Recently, we reported a new class of acrylamide-based cysteine probe, NAIA, which has superior cysteine reaction kinetics to capture more functional cysteines. We further established the workflow of using NAIA for cysteine profiling by mass spectrometry (MS)-based ABPP. Here, we describe a step-by-step protocol using NAIA in the MS-based ABPP experiments for profiling functional cysteines in cancer cells/cell lysates and protein target identification of covalent ligands.

Key words Activity-based protein profiling, Cysteine profiling, Target identification, Activity-based probes, Chemoproteomics, Mass spectrometry, Cysteinome, Acrylamide

1 Introduction

Activity-based protein profiling (ABPP) is an emerging chemoproteomics approach to decipher biological questions unanswered by conventional proteomics approach, which focuses on the protein expression and modifications. By utilizing activity-based probes (ABPs) to covalently react with nucleophilic amino acid residues, ABPP enables a comprehensive investigation on protein reactivity and function. This provides important information for mining new targetable sites on disease-related proteins for therapy [1–3]. In addition, ABPP can couple with covalent ligand screening experiment to discover new drug candidates [4–6]. In covalent drug research and development, ABPP has been employed to determine drug target and drug selectivity profiles [7–9]. All these highlight the enormous values of ABPP for biological and biomedical applications.

ABPP has been developed to study function and reactivity of different amino acids, including cysteine [1–4, 10, 11], serine [12, 13], lysine [14], tyrosine [15], arginine [16], methionine [17, 18], histidine [19, 20], aspartate, and glutamate [21, 22]. Among these amino acids, cysteine and its proteome-wide profiling have gained the greatest attention because cysteine (1) is highly reactive owing to its high nucleophilicity, (2) plays versatile roles in biochemical and redox reactions, and (3) is a popular targeting site for covalent drugs. These reasons fuel the development of cysteine-reactive ABPs and further cysteinome profiling by coworkers [1–4, 10, 11].

ABPs for cysteine profiling utilize electrophilic warheads, including but not limited to maleimides, α,β -unsaturated ketones, and haloacetamides [23]. Among all these, iodoacetamide-based probes, such as iodoacetamide-alkyne (IA-alkyne) [1] and desthiobiotin iodoacetamide (DBIA) [4] with improved performance, are the most widely used ABPs for cysteine profiling. However, iodoacetamide-based probes are generally slow-reacting and demonstrate high cellular toxicity and low biostability [24]. They also have cross-reactivity with other thiols and nucleophilic amino acid (e.g., lysine and serine) [25]. Novel molecules with enhanced cysteine reactivity have been reported (e.g., maleimide-alkyne [26] and *N*-hydroxybenzimidoyl chloride [27]), but these compounds still suffer from the issues of selectivity and biostability [28]. Of note, in drug screening experiment, ABPs preferentially identify hit compounds with identical reactive group [29]. Among the cysteine-targeting covalent drugs approved by the FDA, acrylamide represents the predominant reactive warhead. Utilizing acrylamide-based probes for ABPP experiment should facilitate the development of acrylamide-containing lead compounds, which should have a higher potential for clinical applications. Yet, the very slow Cys reaction of typical acrylamide compounds hinders their application as ABPs in proteome-wide cysteine profiling. Taken all these together, we are motivated to develop acrylamide-based cysteine-reactive ABPs to address these concerns.

Recently, we have reported *N*-acryloylindole-alkyne (NAIA) as a novel acrylamide-based cysteine probe and established the workflow for cysteine profiling by mass spectrometry (MS)-based ABPP [28]. With delicate design of NAIA, its acrylamide warhead is activated by indole aromaticity, allowing NAIA to show superior reactivity toward cysteine, which cannot be found in typical acrylamide compounds. Therefore, unlike previous cysteine profiling study, the working concentration of NAIA can be reduced significantly (e.g., 10 μ M NAIA versus 100 μ M for IA-alkyne and 500 μ M for DBIA). Also, profiling of reactive cysteine in live cell is feasible with NAIA given its fast Cys reaction and low toxicity. Herein, we describe a step-by-step protocol for the application of NAIA in MS-based ABPP for cysteinome profiling and target identification

of covalent ligands. This includes sample preparation, biorthogonal conjugation by click chemistry, affinity enrichment of proteins with functional cysteines, MS data acquisition, and analysis.

2 Materials

All reagents are at analytical grade. Unless specified, Milli-Q water is used to prepare buffer solutions and reagent solutions.

2.1 Cell Culture

1. HepG2 cell (*see* **Note 1**).
2. Complete Dulbecco's Modified Eagle Medium (DMEM): DMEM supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin (PS) solution.
3. Trypsin-EDTA (0.25%).

2.2 Cysteine Probe Labelling

1. *N*-Hex-5-ynyl-2-iodoacetamide (IA-alkyne).
2. *N*-Acryloylindole-alkyne (NAIA) (*see* **Note 2**).
3. Anhydrous dimethyl sulfoxide (DMSO; *see* **Note 3**).

2.3 MS Preparation

1. 50 mM copper(II) sulfate pentahydrate ($\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$).
2. 1.7 mM tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine (TBTA) in DMSO/*tert*-butanol (1:4, v/v).
3. 5 mM desthiobiotin-PEG3-azide (DTB-PEG- N_3) in DMSO.
4. Tris(2-carboxyethyl)phosphine hydrochloride (TCEP). *Freshly prepared* before use. See Methods for concentration used in different steps.
5. Click master mix: 1 mM CuSO_4 , 100 μM TBTA, 100 μM DTB-PEG- N_3 and 1 mM TCEP (*see* Table 1).
6. Prechilled acetone.
7. Prechilled methanol.
8. 1.2% and 0.2% SDS (w/v%) in PBS solution.

Table 1
Recipe for Click master mix

	Stock concentration	Final concentration	Volume for $n = 1$
1. CuSO_4	50 mM	1 mM	40 μL
2. TBTA	1.7 mM (0.9 mg/mL)	100 μM	120 μL
3. DTB-PEG- N_3	5 mM	100 μM	40 μL
4. TCEP ^a	50 mM (14.4 mg/mL)	1 mM	40 μL

^aTCEP MUST be freshly made prior to use

9. Micro Bio-Spin Chromatography Columns.
10. Pierce™ Streptavidin Agarose beads.
11. 400 mM iodoacetamide (IA). *Freshly prepared* before use.
12. 6 M or 2 M urea in PBS. *Freshly prepared* before use.
13. 0.5 µg/µL sequencing grade modified trypsin resuspend in trypsin resuspension buffer.
14. Peptide elution buffer: 50% acetonitrile (MeCN) in H₂O with 0.1% formic acid (FA) at HPLC grade.
15. Pierce™ C18 Spin Tips.

2.4 Data Acquisition by LC-MS/MS

1. Solvent A: 0.1% FA in H₂O-MeCN solution mixture (98:2, v/v).
2. Solvent B: 0.1% FA in MeCN.

3 Methods

Unless specified, all procedures are carried out at room temperature.

3.1 Proteome-Wide Cysteine Profiling of Cell Lysates to Identify Functional Cysteines

A graphical summary of procedures is provided in Fig. 1.

3.1.1 Lysate Preparation

1. Culture 15-cm dishes of HepG2 cells in complete DMEM to approximately 80% confluency (*see Note 1*).
2. Handle the samples on ice for **steps 3–6**.
3. Aspirate media and gently wash the cells with 10 mL ice-cold PBS. Aspirate the washing PBS.
4. Add 3 mL of ice-cold PBS to each dish and collect cells by scraping gently with a cell lifter. Transfer cell suspension to a 15-mL centrifuge tube.
5. Lyse cells by probe sonication at 15% power for 15 s. Centrifuge the cell lysates to pellet cell debris at 5000×*g* for 10 min at 4 °C. Collect the soluble fraction which contains protein samples.
6. Normalize protein concentration to 2 mg/mL in PBS by BCA assay (see Manufacturer's information) or other quantification methods.

3.1.2 Probe Labeling and Click Chemistry

1. Aliquot 2 mL of lysates (2 mg/mL) for 1 sample in a 15-mL centrifuge tube.

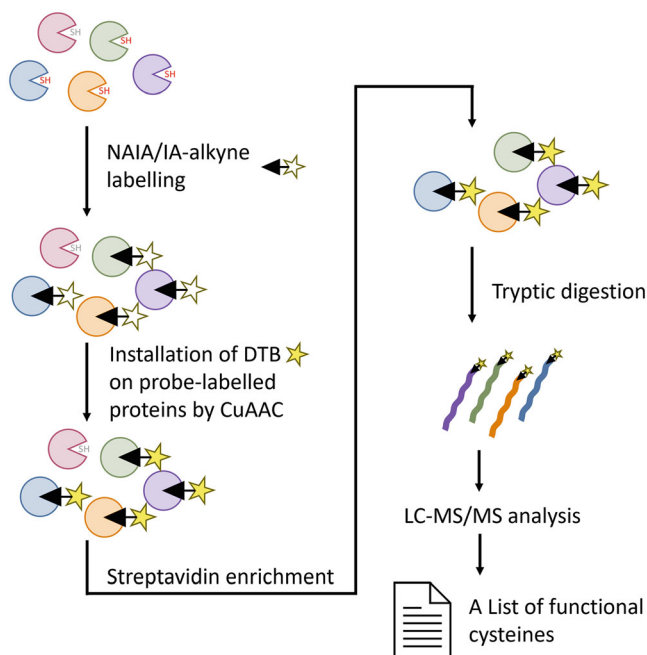


Fig. 1 Overview of ABPP for proteome-wide profiling of functional cysteines. Functional cysteines (red) are labeled by the cysteine probe (NAIA/IA-alkyne), but not for the unreactive one (gray). The probe-labeled proteins will then be installed with DTB by CuAAC click chemistry, enriched, and tryptic digested. The resulting peptides will then be analyzed by LC-MS/MS for the identification of functional cysteines and their proteins

2. *Freshly prepare* a 10 mM IA-alkyne (2.7 mg/mL) or 10 mM NAIA (3.2 mg/mL) in dry DMSO. Dilute the solutions ten-fold using anhydrous DMSO to 1 mM.
3. Add 20 μ L of 1 mM IA-alkyne or NAIA to each sample (final concentration of the cysteine-reactive probes: 10 μ M; *see Note 4*).
4. Incubate samples at room temperature in dark for 1 h, with vortex briefly every 15 min.
5. Prepare Click master mix for click reaction (aka copper-catalyzed azide-alkyne cycloaddition; CuAAC) for $n + 1$ samples. *See Table 1* for details and follow the addition sequence of the reagents. After the addition of TCEP, use immediately.
6. Add 240 μ L of Click master mix (Recipe *see Table 1*) to each sample for conjugation of DTB-PEG-N₃ onto probe-labeled proteins. Incubate samples at room temperature in dark for 1 h, with vortex briefly every 15 min.
7. Add 12 mL of cold acetone to each sample and allow protein to precipitate at -20°C overnight to remove unreacted probe and excess click reagent.

3.1.3 *Enrichment*

1. Pellet protein samples by centrifugation at $5000\times g$ for 10 min at 4 °C. Discard supernatant and re-suspend the protein pellet in 1 mL cold MeOH by probe sonication at 15% power for 15 s on ice.
2. (*Skip this step if you are not trying to compare the performance of IA-alkyne and NAIA for cysteine profiling*) Pairwise sample by combining IA-alkyne-labelled and NAIA-labelled protein solutions in a new 2-mL tube.
3. Centrifuge the samples at $6500\times g$ for 4 min at 4 °C. Save the pellet and repeat washing with MeOH once.
4. Add 1 mL of 1.2% SDS/PBS (w/v) to the protein pellet.
5. Re-suspend the pellet by probe sonication at 15% power for 30 s at room temperature. Complete solubilization of protein is needed. Longer sonication (for example another 30 s) may be required if large particles remain visible.
6. Boil the samples at 90 °C for 5 min.
7. Centrifuge samples at $6500\times g$ for 5 min at room temperature. There should only be a small blue pellet, while a large white pellet indicates inadequate solubilization. In the latter situation, repeat **steps 5–6** until the large white pellet is gone.
8. Re-suspend streptavidin agarose beads by swirling the bottle and aliquot bead slurry for $n + 1$ samples (100 μ L slurry per sample) into a Bio-Spin column (*see Note 5*). Wash beads by vacuum manifold with 3×1 mL PBS. Re-suspend beads in PBS for $n + 1$ samples (50 μ L PBS per sample).
9. Transfer 100 μ L washed streptavidin beads to 5 mL PBS in a 15-mL centrifuge tube.
10. Transfer the supernatant of the samples in 1.2% SDS/PBS (w/v) from **step 7** to the beads in the 15-mL centrifuge tubes.
11. Incubate samples with rotation at 4 °C overnight.

3.1.4 *Cysteine Capping and Protein Digestion*

1. Incubate the samples with gentle shaking at room temperature for 1 h to solubilize the precipitated SDS.
2. Centrifuge the tubes at $1400\times g$ for 3 min. Discard the supernatant.
3. Add 5 mL of 0.2% SDS/PBS (w/v) to the beads. Allow shaking at room temperature for 10 min. Centrifuge and discard the supernatant.
4. Transfer the beads to Bio-Spin column using two washes of 500 μ L PBS. Elute the PBS solution using a vacuum manifold.
5. Wash the beads on vacuum manifold with 3×1 mL PBS, followed by 3×1 mL H₂O.

6. Transfer beads to a new 1.5-mL tube using two washes of 250 μ L 6 M urea.
7. Add 25 μ L *freshly prepared* 110 mM TCEP (31.5 mg/mL) to the sample.
8. Incubate samples at 65 °C for 20 min. Re-suspend beads by gently tapping the tubes every 10 min or providing constant agitation.
9. Add 25 μ L *freshly prepared* 400 mM IA (74 mg/mL) to the sample.
10. Incubate samples at 37 °C for 30 min in dark. Re-suspend beads by gently tapping the tubes every 10 min or providing constant agitation.
11. Add 950 μ L PBS to dilute the reaction mixtures.
12. Centrifuge the tubes at 1400 $\times g$ for 3 min. Discard supernatant.
13. Re-suspend beads in 200 μ L of 2 M urea in PBS.
14. While on ice, add 40 μ L of Trypsin Resuspension Buffer to a vial of lyophilized, sequencing-grade trypsin (20 μ g). Allow 5 min rehydration of the trypsin.
15. Add 4 μ L of the trypsin solution to the beads in 2 M urea in PBS.
16. Digest the protein samples at 37 °C on a shaking incubator overnight.

3.1.5 Elution

1. Transfer the bead solution to a Bio-Spin column. Transfer any bead on the tube to the Bio-Spin column using one wash of 500 μ L PBS.
2. Aspirate the solution using vacuum manifold.
3. Wash beads on vacuum manifold with 3 $\times$ 1 mL PBS, followed by 3 $\times$ 1 mL H₂O.
4. Transfer the Bio-Spin column to a new 2-mL tube.
5. Add 200 μ L peptide elution buffer to the Bio-Spin column. Allow a 3-min incubation of the beads with the solution. Collect the eluent by centrifugation at 3000 $\times g$ for 3 min.
6. Repeat elution with 2 $\times$ 100 μ L elution buffer to collect a total 400 μ L of eluted sample.

3.1.6 Desalting

1. Dry the eluted peptide solution using a SpeedVac.
2. Re-suspend samples in 100 μ L H₂O with 1% FA by bath sonication (*see* **Note 6**).
3. Remove ionic salts in the samples using C18 desalting tips/columns (*see* Manufacturer's information).

Table 2
Recommended LC gradient to resolve peptides. The solvent B % varies linearly over time

Time/min	Duration/min	Solvent B %
0	–	2
100	100	30
105	5	44
105.5	0.5	95
114	8.5	95
114.5	0.5	2
120	5.5	2

3.1.7 Data Acquisition in LC-MS/MS

1. Dry the peptide solution using a SpeedVac.
2. Re-suspend samples in 10 μ L of solvent A.
3. Inject 10 μ L of the peptide solution to Aurora C18 UHPLC column (IonOpticks) coupled to timsTOF Pro mass spectrometer (Bruker). Peptides are separated by varying the composition of solvent A and B as described in Table 2 using a flow rate of 300 nL/min over 2 h.
4. Collect MS data over a m/z range of 100–1700 and an MS/MS range of 100–1700. During MS/MS data collection, each TIMS cycle is 1.1 s and includes 1 MS + an average of 10 PASEF MS/MS scans.

3.1.8 Data Processing and Analysis

1. Search cysteine modifications against UniProt human database using MaxQuant v2.0.3.0 (*see Note 7*). Specify the enzyme for digestion as trypsin and allow for a maximum of three missed cleavages. Peptide length is limited to 6–30 amino acids and max peptide mass is set to be 6000. Cysteine carboxyamido-methylation (+57.02146) is set as fixed modification, whereas methionine oxidation (+15.99491), N-terminal acetylation (+42.01056), cysteine modification by NAIA-5 (+681.38500), or cysteine modification by IA-alkyne (+494.32167) are set as variable modifications. Set the peptide false discovery rate to 1%.
2. Identify the functional cysteines and proteins probed by IA-alkyne or NAIA: After the completion of data searching, there are spreadsheet outputs named “IA-alkyne (C)Sites” and “NAIA-5 (C)Sites” (*see Note 8*). Open the files by MS Excel or other software. Filter out reverse sequence and contaminants. The position of the probed, functional cysteines on the proteins can be found from the column “Positions within

proteins.” For entries with more than 1 protein mapped, only the leading protein is considered. A sample dataset of cysteine profiling in HepG2 cells by NAIA vs IAA can be downloaded from PRIDE partner repository with the dataset identifier PXD041264 and analyzed data can be found in Supplementary Data 2 in our reported study [28].

3. (Optional) Compare the probe performance between IA-alkyne and NAIA based on spectral count in each run: Open the spreadsheet output named “modificationSpecific-Peptides.txt” by MS Excel or other software. Filter out reverse sequence and contaminants. Filter for the peptides with IA-alkyne or NAIA-5 modification (*see Note 9*). The spectral count for each identifier in every individual sample can be found from the column “Experiment X.” The dataset can also be analyzed by Perseus or other desired approaches (e.g., analysis by gene ontology, UniProt annotation, etc.).
4. *See Table 3* for troubleshooting in case of unexpected results.

Table 3
Troubleshooting of possible problems encountered and corresponding solutions

Problem	Possible solutions
Little probe-modified peptides being identified	1. Increase lysate input 2. Optimize probe concentration/streptavidin amount 3. Prepare a new stock solution of the probe from its solid powder, as water may be introduced to the DMSO stock solution after several freeze-thaw cycles and hence the probe can be hydrolyzed and inactivated 4. Ensure thorough washing of protein pellet in Subheading 3.1.3, steps 1–3 as residual unreacted DTB-PEG-N₃ could diminish the efficiency of enrichment of the probe-modified peptides 5. Ensure thorough washing of beads to minimize nonspecific binding of peptides in both Subheading 3.1.4, step 5 and Subheading 3.1.5, step 3, which could help improve signal-to-noise ratios
Huge variations between replicate groups	1. Carefully adhere to the protocol 2. Pay attention to the protein normalization to make sure the difference in starting material is minimum
Low MS signal intensity generally	1. Increase lysate input 2. Ensure a complete solvation of dried peptides in Subheading 3.1.6, step 2 by a longer bath sonication 3. Wash thoroughly and transfer all the beads to the Bio-Spin column/new Eppi tube in every transfer step. 4. Adjust the LC gradient (e.g., by increasing the maximum % of B) to ensure complete elution of the peptides in the LC-MS/MS analysis

3.2 Proteome-Wide Cysteine Profiling of Live Cells (i.e. In Situ) to Identify Functional Cysteines

3.2.1 In Situ Probe Labeling and Lysate Preparation

1. Culture 15-cm dishes of HepG2 cells in complete DMEM to approximately 80% confluency (*see* **Note 1**).
2. Prepare 100 μ M IA-alkyne or NAIA in complete DMEM while limiting DMSO content to be 1% to avoid toxicity.
3. Incubate live HepG2 cell with IA-alkyne or NAIA working solution for 30 min at 37 °C with 5% CO₂.
4. Prepare protein samples in PBS as described in Subheading 3.1.1, steps 2–6.

3.2.2 Click Chemistry, Enrichment, MS Preparation, Data Acquisition, and Analysis

1. Aliquot 2 mL of lysates (2 mg/mL) for 1 sample in a 15-mL centrifuge tube.
2. Perform CuAAC to conjugate DTB-PEG-N₃ onto probe-labeled proteins as described in Subheading 3.1.2, steps 5–7.
3. Subsequent procedures are identical to Subheading 3.1.3, 3.1.4, 3.1.5, 3.1.6, 3.1.7, 3.1.8.

3.3 Competitive ABPP for Target Identification of Cysteine-Reactive Covalent Ligands

A graphical summary of procedures is provided in Fig. 2a.

3.3.1 Compound Treatment

1. *For in situ treatment*, treat HepG2 cells with solvent vehicle (solvent control group) and cysteine-reactive covalent ligands (compound treatment group) in complete DMEM, respectively. Harvest cells and prepare cell lysate as described in Subheading 3.1.1, steps 2–6.
2. *For treatment on cell lysates*, incubate the cell lysates in PBS with solvent vehicle (solvent control group) and cysteine-reactive covalent ligands (compound treatment group), respectively.

3.3.2 Probe Labeling

1. Label the cell lysates with NAIA; perform CuAAC and acetone precipitation as described in Subheading 3.1.2, steps 5–7.

3.3.3 Enrichment, MS Preparation, and Data Acquisition

1. Pellet protein and wash with MeOH twice as described in Subheading 3.1.3 **EXCEPT** without pairing samples from solvent control group and compound treatment group (*see* **Note 9**), i.e., for two groups with $n = 3$ per group, there are 6 different tubes after the Enrichment steps.
2. Perform sample preparation following steps in Subheading 3.1.3, 3.1.4, 3.1.5, 3.1.6, 3.1.7, to acquire data.

3.3.4 Data Analysis

1. Search cysteine modifications against UniProt human database using MaxQuant v2.0.3.0 as described in Subheading 3.1.8, **step 1** **EXCEPT** using methionine oxidation (+15.99491),

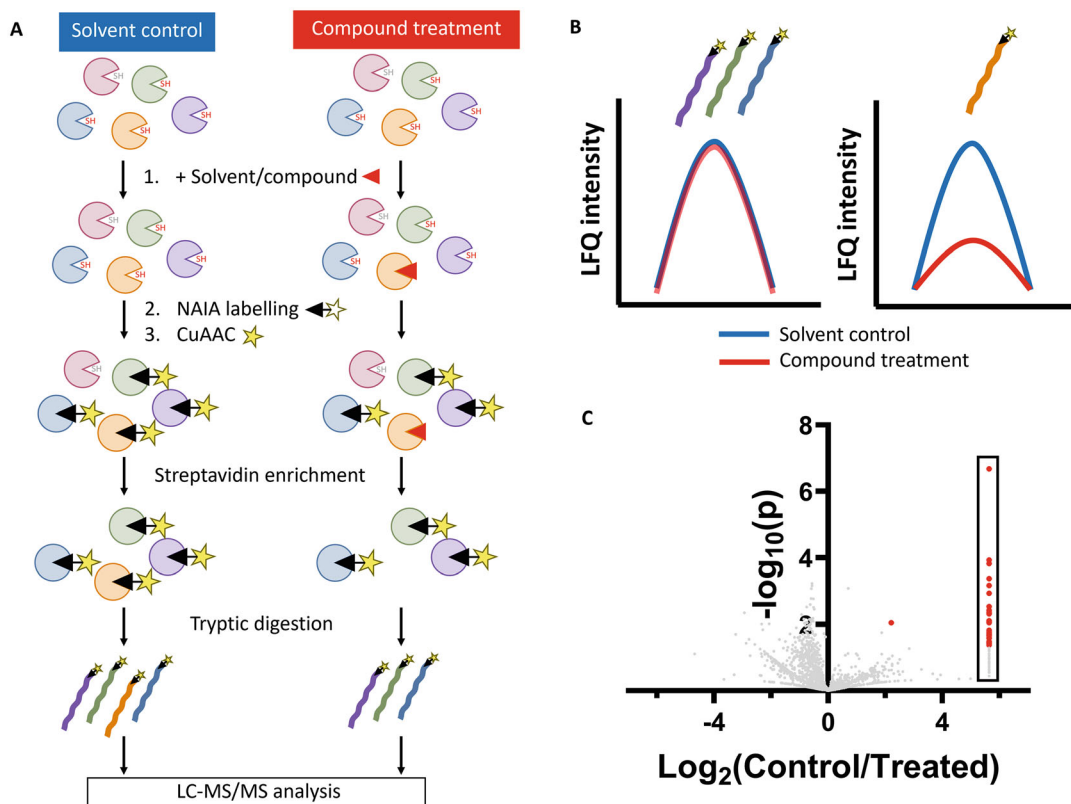


Fig. 2 Overview of competitive ABPP for target identification. (a) Schematic of the sample preparation and MS acquisition workflow. (b) Expected result of LFQ intensity of the peptide from protein target (orange) and nontargets (other colors) from the MS experiments. (c) Volcano plot showing the proteins bound by the compound (red dots), as found in the MS analysis for target identification of CL1 in our previous study [28]. Peptides with zero values in the treated samples but nonzero values in the control, i.e., the peptides bound strongly by the compound, are shown as dots inside the box

N-terminal acetylation (+42.01056), and cysteine modification by NAIA-5 (+681.38500) as variable modifications.

2. With Perseus software, load the “NAIA_CysSites.txt” output as matrix with intensity as reference (*see Note 8*). Filter out reverse sequence and contaminants. Group samples by categorizing different samples into the same group with identical naming (e.g., rename two solvent-treated samples “Treated 1” and “Treated 2” to “Treated”). Log-2 transform the LFQ intensity, and filter row based on minimum valid number of two in at least 1 group. Transform the $\log_2(\text{LFQ intensity})$ by antilog-2 and export the filter list. Note that MS Excel or other data processing software can also be used to analyze the data.
3. LFQ ratio of peptides can be calculated by (average LFQ intensity of solvent-treated sample)/(average LFQ intensity of compound-treated sample). p-value is calculated by two-sample t-test based on the LFQ intensity of each sample.

4. For peptides with nonzero LFQ intensity for all replicates in the solvent-treated sample and zero LFQ intensity for all replicates in the compound-treated sample, assign their LFQ ratios to be 50 (*see* **Note 10**). For other peptides with invalid LFQ ratios, filter out these peptides.
5. Summarize the fold change of identified peptide intensity by a volcano plot with $\log_2(\text{LFQ ratio})$ against $-\log_{10}(\text{p-value})$. Prism 9.0, or an alternative program (e.g., R studio), can be used. Cysteines which are binding sites of the compound should have a positive $\log_2(\text{LFQ ratio}) \geq 2$ (i.e., fourfold competition by the compound) with statistical significance (i.e., $-\log_{10}(p\text{-value}) > 1.3$). An example of the expected outcome is shown in Fig. 2, which is based on the analysis of our published data on target identification of CL1 (PRIDE ID: PXD041264) [28].
6. *See* Table 3 for troubleshooting in case of unexpected results.

4 Notes

1. ABPP experiment also works for other culturable cells.
2. Detailed synthesis of NAIA can be found in our previously reported study [28]. The NAIA probe can be provided upon request to Dr. Clive Yik-Sham Chung (cyschung@hku.hk).
3. To prevent undesirable degradation of the cysteine-reactive probes by water hydrolysis, dry DMSO should be used to make stock solution of the probes. Molecular sieves can be added to an opened bottle of DMSO to keep it dry.
4. The working concentration of NAIA (10 μM) for cysteine profiling of cell lysates is lower than that of IA-alkyne or other IA-based cysteine probes reported in the literature (100–500 μM). For profiling more functional cysteines, higher probe concentration can be used (100–500 μM).
5. All procedures with streptavidin beads should be handled with cut tip to avoid bead shearing.
6. SpeedVac of peptide solution often yields a nearly transparent patch, which is hard to confirm the status of re-suspension by visual inspection. Hence, sonication is recommended to ensure complete re-suspension of peptides.
7. The latest version of human proteome reference and software are recommended.
8. The example in this protocol is based on our published dataset (PRIDE ID: PXD041264), in which the modification be IA-alkyne and NAIA on cysteine for the profiling experiment

are named as IA-alkyne (C) and NAIA-5 (C), whereas the modification by NAIA on cysteine for the competition experiment is named as NAIA_Cys in MaxQuant search. The file naming corresponds to the input in MaxQuant v2.0.3.0.

9. Label-free quantification (LFQ) is used in this protocol. Procedure can be adapted to other MS approach such as SILAC to which include pairing of samples.
10. This is the maximum cap for the compound with strong binding onto the cysteine that results in zero LFQ intensity in the compound treatment group. The value of this maximum cap can be different for analysis of different dataset.

References

1. Weerapana E, Wang C, Simon GM, Richter F, Khare S, Dillon MB, Bachovchin DA, Mowen K, Baker D, Cravatt BF (2010) Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* 468(7325): 790–795. <https://doi.org/10.1038/nature09472>
2. Ward CC, Kleinman JI, Nomura DK (2017) NHS-esters as versatile reactivity-based probes for mapping proteome-wide Ligandable hotspots. *ACS Chem Biol* 12(6):1478–1483. <https://doi.org/10.1021/acscchembio.7b00125>
3. Zanon PRA, Lewald L, Hacker SM (2020) Isotopically labeled Desthiobiotin Azide (isoDTB) tags enable global profiling of the bacterial Cysteine. *Angew Chem Int Ed Engl* 59(7):2829–2836. <https://doi.org/10.1002/anie.201912075>
4. Kuljanin M, Mitchell DC, Schweppe DK, Gikandi AS, Nusinow DP, Bulloch NJ, Vinogradova EV, Wilson DL, Kool ET, Mancias JD, Cravatt BF, Gygi SP (2021) Reimagining high-throughput profiling of reactive cysteines for cell-based screening of large electrophile libraries. *Nat Biotechnol* 39(5):630–641. <https://doi.org/10.1038/s41587-020-00778-3>
5. Roberts AM, Ward CC, Nomura DK (2017) Activity-based protein profiling for mapping and pharmacologically interrogating proteome-wide ligandable hotspots. *Curr Opin Biotechnol* 43:25–33. <https://doi.org/10.1016/j.copbio.2016.08.003>
6. Spradlin JN, Zhang E, Nomura DK (2021) Reimagining druggability using chemoproteomic platforms. *Acc Chem Res* 54(7): 1801–1813. <https://doi.org/10.1021/acs.accounts.1c00065>
7. Lanning BR, Whitby LR, Dix MM, Douhan J, Gilbert AM, Hett EC, Johnson TO, Joslyn C, Kath JC, Niessen S, Roberts LR, Schnute ME, Wang C, Hulce JJ, Wei B, Whiteley LO, Hayward MM, Cravatt BF (2014) A road map to evaluate the proteome-wide selectivity of covalent kinase inhibitors. *Nat Chem Biol* 10(9): 760–767. <https://doi.org/10.1038/nchembio.1582>
8. Lyu P, Jiang K, Zhou Y, Hu J, Chang Y, Zhang Z, Huang M, Zhang ZM, Ding K, Hao P, Lin L, Li Z (2022) Proteome-wide identification of off-targets of a potent EGFR (L858R/T790M) mutant inhibitor. *ACS Med Chem Lett* 13(2):292–297. <https://doi.org/10.1021/acsmchemlett.1c00651>
9. Ye Z, Wang Y, Wu H, Song T, Li X, Liu Q, Wang C (2021) Chemoproteomic profiling of an Ibrutinib analogue reveals its unexpected role in DNA damage repair. *ChemBiochem* 22(1):129–133. <https://doi.org/10.1002/cbic.202000527>
10. Meng J, Fu L, Liu K, Tian C, Wu Z, Jung Y, Ferreira RB, Carroll KS, Blackwell TK, Yang J (2021) Global profiling of distinct cysteine redox forms reveals wide-ranging redox regulation in *C. elegans*. *Nat Commun* 12(1):1415. <https://doi.org/10.1038/s41467-021-21686-3>
11. Backus KM, Correia BE, Lum KM, Forli S, Horning BD, Gonzalez-Paez GE, Chatterjee S, Lanning BR, Teijaro JR, Olson AJ, Wolan DW, Cravatt BF (2016) Proteome-wide covalent ligand discovery in native biological systems. *Nature* 534(7608): 570–574. <https://doi.org/10.1038/nature18002>
12. Bachovchin DA, Ji T, Li W, Simon GM, Blankman JL, Adibekian A, Hoover H, Niessen S, Cravatt BF (2010) Superfamily-wide portrait

- of serine hydrolase inhibition achieved by library-versus-library screening. *Proc Natl Acad Sci USA* 107(49):20941–20946. <https://doi.org/10.1073/pnas.1011663107>
13. Adibekian A, Martin BR, Wang C, Hsu KL, Bachovchin DA, Niessen S, Hoover H, Cravatt BF (2011) Click-generated triazole ureas as ultrapotent in vivo-active serine hydrolase inhibitors. *Nat Chem Biol* 7(7):469–478. <https://doi.org/10.1038/nchembio.579>
 14. Hacker SM, Backus KM, Lazear MR, Forli S, Correia BE, Cravatt BF (2017) Global profiling of lysine reactivity and ligandability in the human proteome. *Nat Chem* 9(12):1181–1190. <https://doi.org/10.1038/nchem.2826>
 15. Hahm HS, Toroitich EK, Borne AL, Brulet JW, Libby AH, Yuan K, Ware TB, McCloud RL, Ciancone AM, Hsu KL (2020) Global targeting of functional tyrosines using sulfur-triazole exchange chemistry. *Nat Chem Biol* 16(2):150–159. <https://doi.org/10.1038/s41589-019-0404-5>
 16. Wanigasekara MSK, Huang X, Chakrabarty JK, Bugarin A, Chowdhury SM (2018) Arginine-selective chemical labeling approach for identification and enrichment of reactive arginine residues in proteins. *ACS Omega* 3(10):14229–14235. <https://doi.org/10.1021/acsomega.8b01729>
 17. Lin S, Yang X, Jia S, Weeks AM, Hornsby M, Lee PS, Nichiporuk RV, Iavarone AT, Wells JA, Toste FD, Chang CJ (2017) Redox-based reagents for chemoselective methionine bioconjugation. *Science* 355(6325):597–602. <https://doi.org/10.1126/science.aal3316>
 18. Gonzalez-Valero A, Reeves AG, Page ACS, Moon PJ, Miller E, Coulonval K, Crossley SWM, Xie X, He D, Musacchio PZ, Christian AH, McKenna JM, Lewis RA, Fang E, Dovala D, Lu Y, McGregor LM, Schirle M, Tallarico JA, Roger PP, Toste FD, Chang CJ (2022) An activity-based Oxaziridine platform for identifying and developing covalent ligands for functional allosteric methionine sites: redox-dependent inhibition of cyclin-dependent kinase 4. *J Am Chem Soc* 144(50):22890–22901. <https://doi.org/10.1021/jacs.2c04039>
 19. Cruite JT, Dann GP, Che J, Donovan KA, Ferrao S, Ficarro SB, Fischer ES, Gray NS, Huerta F, Kong NR, Liu H, Marto JA, Metivier RJ, Nowak RP, Zerfas BL, Jones LH (2022) Cereblon covalent modulation through structure-based design of histidine targeting chemical probes. *RSC Chem Biol* 3(9):1105–1110. <https://doi.org/10.1039/d2cb00078d>
 20. Wan C, Wang Y, Lian C, Chang Q, An Y, Chen J, Sun J, Hou Z, Yang D, Guo X, Yin F, Wang R, Li Z (2022) Histidine-specific bioconjugation via visible-light-promoted thioacetal activation. *Chem Sci* 13(28):8289–8296. <https://doi.org/10.1039/d2sc02353a>
 21. Bach K, Beerkens BLH, Zanoni PRA, Hacker SM (2020) Light-activatable, 2,5-disubstituted tetrazoles for the proteome-wide profiling of aspartates and glutamates in living bacteria. *ACS Cent Sci* 6(4):546–554. <https://doi.org/10.1021/acscentsci.9b01268>
 22. Ma N, Hu J, Zhang ZM, Liu W, Huang M, Fan Y, Yin X, Wang J, Ding K, Ye W, Li Z (2020) 2H-Azirine-based reagents for chemoselective bioconjugation at carboxyl residues inside live cells. *J Am Chem Soc* 142(13):6051–6059. <https://doi.org/10.1021/jacs.9b12116>
 23. Hoch DG, Abegg D, Adibekian A (2018) Cysteine-reactive probes and their use in chemical proteomics. *Chem Commun (Camb)* 54(36):4501–4512. <https://doi.org/10.1039/c8cc01485j>
 24. Schmidt MM, Dringen R (2009) Differential effects of iodoacetamide and iodoacetate on glycolysis and glutathione metabolism of cultured astrocytes. *Front Neuroenerg* 1:1. <https://doi.org/10.3389/neuro.14.001.2009>
 25. Nielsen ML, Vermeulen M, Bonaldi T, Cox J, Moroder L, Mann M (2008) Iodoacetamide-induced artifact mimics ubiquitination in mass spectrometry. *Nat Methods* 5(6):459–460. <https://doi.org/10.1038/nmeth0608-459>
 26. Zhang D, Devarie-Baez NO, Li Q, Lancaster JR Jr, Xian M (2012) Methylsulfonyl benzothiazole (MSBT): a selective protein thiol blocking reagent. *Org Lett* 14(13):3396–3399. <https://doi.org/10.1021/ol301370s>
 27. Chen FJ, Zheng M, Nobile V, Gao J (2022) Fast and cysteine-specific modification of peptides, proteins and bacteriophage using Chloro-oxyimes. *Chemistry* 28(20):e202200058. <https://doi.org/10.1002/chem.202200058>
 28. Koo TY, Lai H, Nomura DK, Chung CY (2023) N-Acryloylindole-alkyne (NAIA) enables imaging and profiling new ligandable cysteines and oxidized thiols by chemoproteomics. *Nat Commun* 14(1):3564. <https://doi.org/10.1038/s41467-023-39268-w>
 29. Yang F, Chen N, Wang F, Jia G, Wang C (2022) Comparative reactivity profiling of cysteine-specific probes by chemoproteomics. *Curr Res Chem Biol* 2:100024. <https://doi.org/10.1016/j.crchbi.2022.100024>



System-Wide Profiling of Ubiquitin E3 Ligase Activities with Quantitative Proteomics and Bioinformatics Analysis

Yue Chen and Yao Gong

Abstract

Ubiquitination is an essential protein posttranslational modification in eukaryotes governing protein homeostasis and ubiquitin-mediated cellular signaling. In human cells, the dynamics and substrate-specificity of ubiquitination are mediated by over 600 genetically encoded E3 ligases. Dysregulation of E3 ligase activities could disrupt protein turnover, gene expression, and signal transduction, which is widely implicated in diseases such as cancer and aging-related diseases. In this report, we presented a computational workflow, Ube3-APA, to profile E3 ligase activities based on quantitative proteomics data and bioinformatics analysis. We validated the strategy with published datasets from the analysis of cells and tissues with genetic knockout or knockdown of specific E3 ligases and demonstrated that the strategy can be applied to both data-dependent or data-independent quantitative proteomics datasets.

Key words Ubiquitination, E3 ligase, Quantitative proteomics, Ube3-APA, Activity-based profiling

1 Introduction

Ubiquitination is a widespread protein posttranslational modification (PTM) in eukaryotic cells highly conserved from yeast to human [1]. It is well characterized to regulate protein degradation, cell signaling and epigenetic gene expression. With relatively low stoichiometry, ubiquitination is a highly dynamic posttranslational modification mediated by the elegant E1-E2-E3 enzymatic cascade for ubiquitin conjugation and deubiquitinases for ubiquitin removal [2–4]. The substrate specificity of ubiquitination is mainly dependent on a large family of E3 ubiquitin ligases. In human cells, over 600 E3 ligases have been discovered and these E3 ligases are known to form specific protein-protein interaction networks for substrate recognition and ubiquitin conjugation [5]. Loss or gain of functions on E3 ligases are known to be involved in diverse pathological conditions including ageing, metabolic disorders and diseases such as cancer and neurological diseases [6–9]. Therefore,

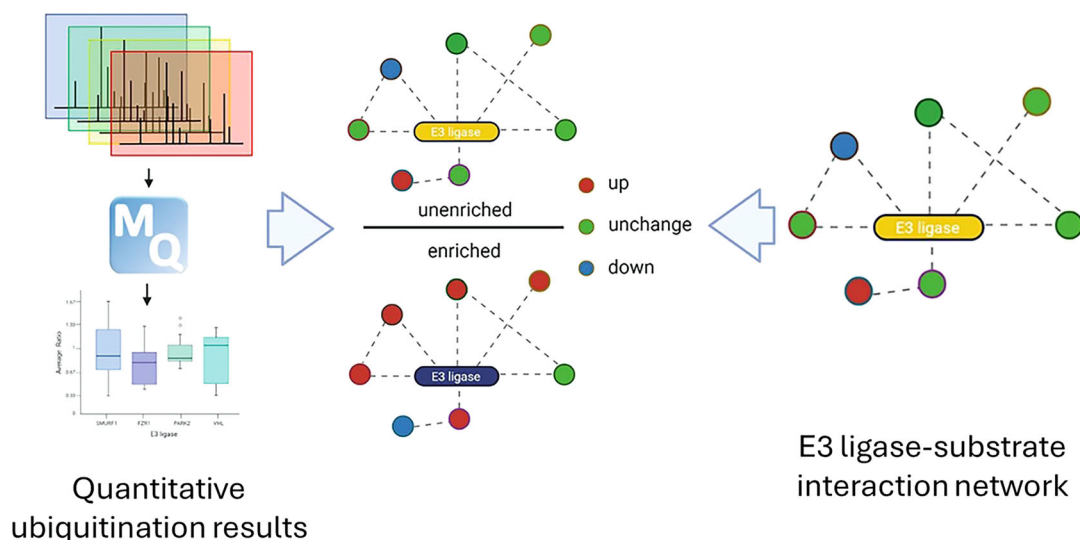


Fig. 1 A schematic representation of the dynamics of the ubiquitination proteome mediated by E1-E2-E3 enzymatic cascade and measured by quantitative ubiquitylome profiling analysis

measuring E3 ligase activities in a system-wide manner will provide critical insights into the disease pathogenesis and progression (Fig. 1).

Activity-based proteomic analysis is one of the functional proteomic strategies that aims to systematically profile the endogenous activities of proteins [10]. It can be classified into probe-based and probe-less analysis strategy. In probe-based analysis, a chemical probe is uniquely designed to recognize a specific protein group or enzyme family. Reactivities toward the probe can be measured in combination with quantitative analysis to reflect the changes in activities of the endogenous proteins. A great example of probe-based analysis is activity-based protein profiling (ABPP) pioneered by Benjamin Cravatt and others [11–14]. In probe-less analysis, a quantitative experiment is typically performed and the known targets for the protein of interest are monitored. The dynamic changes in abundance of the known targets can be measured to reflect the activities of the protein of interest. Classic strategies to monitor the activities of specific kinases using radioactive assay, Western blotting, or mass spectrometry are great examples [15]. Combining with advances in large-scale quantitative PTM proteomics analysis, several bioinformatic strategies have been developed to profile kinase activities based on phosphoproteomics datasets, including IKAP, KinasePA, KSEA, KEA3, PTMsigDB, and KAPA [16–21].

Measurement of endogenous activities of ubiquitin E3 ligases can be challenging due to the lack of characteristic flanking sequence motifs for individual E3 ligases and the dependence on the protein-based substrate recognition [22]. This is in distinct

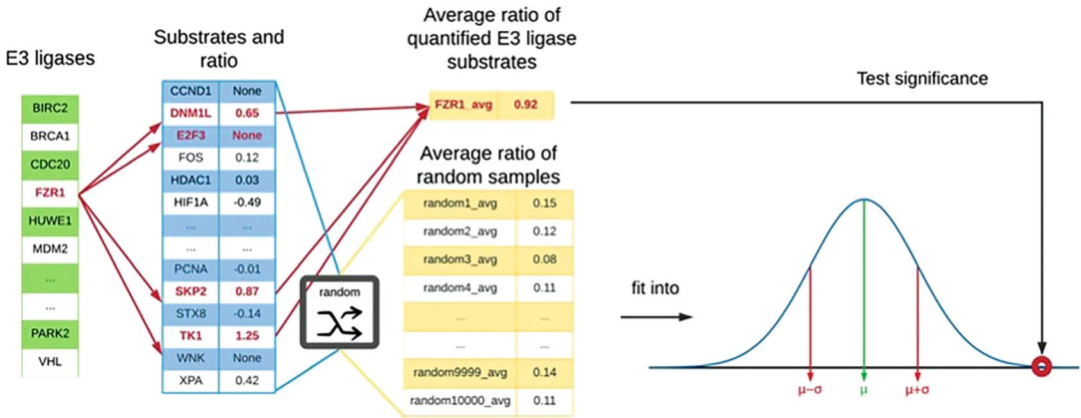


Fig. 2 A schematic representation of UbE3-APA analysis principles

contrast to phosphorylation analysis, where kinases target phosphorylation at the site-specific level [23]. Recent advances in protein-interaction network analysis and biochemical characterization of ubiquitin E3 ligase substrates led to the successful development of E3 ligase substrate databases [24–27]. Based on these community resources and the advances in mass spectrometry-based proteomics technology, we developed ubiquitin E3 ligase activity profiling analysis (UbE3-APA), a computational framework and python-based software, to globally discover significantly up- or downregulated ubiquitin E3 ligase activities [28] (Fig. 2).

UbE3-APA follows similar statistical principles as our previously developed kinase activity profiling analysis (KAPA) with some modifications for ubiquitination analysis. To develop UbE3-APA, we first integrated three recently published E3 ligase substrate databases into a single no-redundant E3 ligase substrate database. Then, we apply the software to analyze specific sets of quantitative proteomics experiments that measure the site-specific ubiquitination dynamics among different experimental or physiological conditions. From the quantitative proteomics experiment, the software extracts the quantitative ratios of ubiquitination sites and their corresponding proteins among different conditions based on the E3-ligase substrate database. Next, to determine the significant changes of E3 ligase activities, the software calculates the average ratios of known targets of the E3 ligase and performs bootstrap-based random sampling against all quantified ubiquitination sites or proteins. The number of random targets in each sampling matches the number of known substrate site or proteins for the E3 ligase quantified in the dataset. Based on the central limit theorem, the average ratios of the randomly sampled targets form a normal distribution. Comparing the average ratio of the known targets of the E3 ligase against the normal distribution gives statistical

confidence on whether the activity of a specific E3 ligase is significantly up- or downregulated in this quantitative experiment.

The confidence of the Ub E3 ligase activity profiling analysis depends on two aspects. First, the reliability and coverage of the E3 ligase-substrate interaction database is critical for accurate analysis of E3 ligase activities. Incorrect E3 ligase-substrate matches in the database increases background and reduces statistical significance of the results. Continuing efforts to expand the E3 ligase-substrate network with recent advances such as CRISPR screening [22] will enable in-depth discoveries of changes in E3 ligase activities under various patho- or physiological conditions. Second, the depth and quantification accuracy of ubiquitination proteomics analysis is critical to determine the E3 ligase activities. Our statistical model indicates that higher coverage of substrate site and protein quantification for each E3 ligase will lead to larger sampling size and smaller sampling variations, which will result in higher statistical significance in the analysis results. Recent advances in highly sensitive mass spectrometry analysis and new strategies for Ub stoichiometry analysis will further improve the Ub E3 ligase activity profiling analysis with more in-depth and confident discoveries [29, 30].

2 Materials

The UbE3-APA program requires a modern desktop computer or workstation running Windows or Unix-like operating system including Mac OS, Linux, or Unix. The computer should have Python version 3.9 or above installed and have access to the internet for the package installation. The program runs under python command line environment and requires standard python packages, including pathlib, pandas, random, statistics, SciPy, NumPy, and os, which will be installed automatically together with the installation of UbE3-APA package. To assist in data analysis and charting with volcano plot, it is recommended that the computer is installed with charting software such as Microsoft Excel or GraphPad Prism.

3 Methods

3.1 Installation of UbE3-APA

3.1.1 Installation of UbE3-APA Python Package

1. Open the command line window of the operating system. For example, on Windows computer, open the app “Command Prompt” and on Mac OS computer, open the app “Terminal.”
2. Execute the following code to install the ube3_apa package:

```
python -m pip install ube3_apa
```

3. Confirm the installation by showing the installation directory of ube3_apa package:

```
pip show ube3_apa
```

4. To obtain the software package with the testing data, download the UbE3-APA software package in a Zip file from the following GitHub repository:

<https://github.com/Chenlab-UMN/Ub-E3-ligase-Activity-Profiling-Analysis>

5. Extract the Zip file into the new folder and the testing data is located under the directory of “testdata.” The testing data was compiled based on quantitative ubiquitylome profiling data from Theurillat, J.P. et al, *Science* 2014 [31].

3.1.2 E3 Ligase and Substrate Interaction Database

1. E3 ligase and substrate interaction database is stored in a comma-delimited text file “E3dict.csv.” (**Note 1**). The file should be located under the installation directory of ube3_apa python package.
2. To update or compile a new database file, follow the example file format below, and save the new file as “E3dict.csv” to the installation directory of ube3_apa:

```
## column headers
```

```
e3ligase:
```

```
gene name of the E3 ligase
```

```
substrate:
```

```
gene name of a protein substrate
```

```
## An example—E3 ligase CBLC has two substrates—EGFR  
and SRC
```

```
CBLC,EGFR;SRC
```

```
.....
```

3.1.3 Input Data Preparation

1. The main input file of the UbE3-APA software is a comma-delimited text file containing the ratios of all quantifiable ubiquitination sites between the two conditions of interest. Follow the format below to prepare the site-ratio input file. The file name and location are determined by the user:

```
## column headers
```

```
id:
```

```
protein identifier
```

```
position:
```

```
position of the Ub site
```

siteratio:

quantification ratio of the site between two conditions of interests

An example—Ub site K128 on the protein with a Uniprot accession number of P18206 had a log2 transformed quantitative ratio of -0.37 in this experiment

P18206,128,-0.37

.....

2. An optional input file of the Ube3_APA software is a comma-delimited text file containing the ratios of all quantifiable proteins between the two conditions of interest. Follow the format below to prepare the protein-ratio input file (texts after *##* indicates notes and should not be included). The file name and location are determined by the user:

column headers

id:

protein identifier

ratio:

quantification ratio of the protein between two conditions of interests

An example—Protein with a Uniprot accession number of A5YKK6 has a log2 transformed quantitative ratio of -0.013 in this experiment

A5YKK6,-0.013

.....

3. Check to confirm that the E3 ligase-substrate interaction database “E3dict.csv” is in the ube3_apu installation directory.
4. Check to confirm that an optional support file “unigenedict.csv” is in the ube3_apu installation directory. The file contains one-on-one matches from Uniprot accession IDs to NCBI gene symbols with a format below (texts after *##* indicates notes and should not be included) (**Note 2**):

No column headers.

An example—Uniprot accession number A0A023J8Y7 is matched to gene ABO

A0A023J8Y7,ABO

.....

3.2 Data Analysis with Ube3-APA

1. From the command line window, start the python console:
>python
2. Import the ube3_apu package:
>>>import ube3_apu

3. Set the variable for the directory of the input data containing ubiquitination site-specific quantification ratios and optional protein quantification ratios using the directory of the testing data as an example (**Note 3**):

```
>>> input_directory = "C:/path/Ub-E3-ligase-Activity-
    Profiling-Analysis-master/testdata/"
```

4. Understand the input parameters of `e3enrich` function in the `ube3_ap` package:

```
e3enrich(siteratio_dir, input_type, output_dir, exp_label="",
    proratio_dir="None", grouped=False, ratio_output=False,
    log2trans=True)
```

siteratio_dir: string

The location of the CSV file that contains the quantification ratios of ubiquitination sites.

input_type: {"UniprotAC" or "protein", "gene symbol" or "gene"}

"UniprotAC" or "protein" example: Q00987, P40337, Q9HAU4, O43791

"gene symbol" or "gene" example: MDM2, VHL, SMUF2, SPOP

List of valid input and examples will be shown if it is an invalid value.

output_dir: string

An existing directory where the analysis result files will be written in.

exp_label: string, default ""

A string attached to the end of the output file name that can be customized to distinguish results from different batches of analysis.

proratio_dir: string, default ""

The location of the CSV file that contains the quantification ratios of proteins. A valid directory input will enable the normalization of site ratios by the corresponding protein ratios before E3 activity profiling analysis.

grouped: bool, default False

The parameter is set to False by default. If True, the program will group the E3 ligases in the activity profiling results in the output file based on the substrates quantified in this experiment for each E3 ligase. For example, if in this experiment, E3 ligase "AAA" has three quantified substrates ("aaa," "bbb," "ccc"). E3 ligase "BBB" has two quantified substrates ("bbb," "ccc") which is a just subset of substrates of "AAA," the program will put "AAA" and

“BBB” in one E3 ligase group and assign “AAA” as the leading E3 ligase of this group. The output file will only list the activity enrichment analysis of the leading E3 ligases in each group. This mode is useful to reduce redundancy and increase confidence of the analysis results.

ratio_output: bool, default False

The parameter is set to False by default. If True, files that contain the ratio of each ubiquitylation site and the average ratio of each ubiquitinated protein will be generated as reference and useful checkpoints for the users.

log2trans: bool, default True

The parameter is set to True by default. If True, the program will perform log₂ transformation ($y = \log_2(x)$) to the ratios in the input site and/or protein quantification data before the E3 ligase enrichment analysis. If False, the ratio data from the input files has already been transformed and can be used for enrichment analysis directly.

5. Perform UbE3-APA analysis with the testing data in the standard mode where each E3 ligase is analyzed individually (**Note 4**):

```
>>>ube3_ap.e3enrich(siteratio_dir=input_directory+"siteratio_testdata.csv", input_type="UniprotAC", output_dir="C:/output/", exp_label="1", grouped=False, ratio_output=True, log2trans=False)
```

6. Or perform UbE3-APA analysis with the testing data in the grouped mode where activity enrichment analysis results of each E3 ligase were grouped based on the substrates of each E3 ligase quantified in this experiment as explained in **step 4**:

```
>>>ube3_ap.e3enrich(siteratio_dir=input_directory+"siteratio_testdata.csv", input_type="UniprotAC", output_dir="C:/output/", exp_label="2", grouped=True, ratio_output=True, log2trans=False)
```

7. Or perform UbE3-APA analysis with the testing data in the standard mode with protein-level normalization (**Note 5**):

```
>>>ube3_ap.e3enrich(siteratio_dir=input_directory+"siteratio_testdata.csv", input_type="UniprotAC", output_dir="C:/output/", exp_label="3", prorate_dir=input_directory+"proteinratio_testdata.csv", grouped=False, ratio_output=True, log2trans=False)
```

3.3 Interpretation of the UbE3-APA Analysis Results

1. If ratio_output parameter of the e3enrich function is set as True, the program generates two CSV files—“site_level_ratio” and “protein_level_ratio,” which contains the log₂

transformed quantification ratios for each ubiquitination site and the averaged ratios for each protein. These data are used in the statistical significance analysis by the program to generate main enrichment analysis outputs.

2. The site_level_ratio CSV file follows the format below:

column headers

genename:

NCBI gene name

proteinid:

Uniprot protein accession ID

position:

Ub site position

siteratio:

log2 transformed quantification ratio of this Ub site

An example—Ub site K138 in RPL4 with Uniprot accession ID P36578 has a log2 transformed quantification ratio of 0.22

RPL4,P36578,138,0.22

.....

3. The protein_level_ratio CSV file follows the format below:

column headers

genename:

NCBI gene name

proteinid:

Uniprot protein accession ID

positions:

a string of positions of quantified Ub sites

average_siteratio:

averaged log2 transformed ratios of the protein based on all quantified sites,

siteratios:

a string of log2 transformed ratios of all quantified Ub sites

number_of_sites:

the number of quantified Ub sites

An example - gene CTHR1 with Uniprot accession ID E9PJD7 and quantified Ub sites K58, K167 and K366, has an averaged log2 ratio of 0.693 based on the site log2 ratios of 0.45, 0.88 and 0.75 from 3 sites.

CYHR1,E9PJD7,58;167;366,0.693,0.45;0.88;0.75,3.0

.....

4. The program also outputs the E3 ligase and substrate interaction network involved in this study “e3_sub_network3.csv” as a reference for the user and the file follows the format below:

column headers

e3ligase:

gene name of the E3 ligase

substrate_protein:

gene name of one of its substrate proteins

An example - E3 ligase MDM2 targets TP53

MDM2,TP53

.....

5. One of the main outputs of the UbE3-APA analysis is the “E3enrichment_site_level” CSV file (**Note 6**). The file name also contains the analysis settings. For example, if the experiment label in the input parameter is set as “3” and protein-level normalization is enabled, the site-level analysis file name would be:

“E3enrichment_site_level_normalized_by_protein3.csv”

The site-level analysis file follows the format below:

column headers

e3ligase:

Gene name of the E3 ligase

site_count:

Total number of quantified Ub sites for the substrates of the E3 ligase in this experiment

site_avg:

The averaged log2 ratio of quantified Ub sites of the E3 ligase substrates

sample_avg:

The average for the ratio distribution of the randomly sampled Ub sites

sample_std:

The standard deviation for the ratio distribution of the randomly sampled Ub sites

p_value:

Cumulative p value of the averaged log2 ratio of the E3 ligase target Ub sites in the ratio distribution of randomly sampled Ub sites

substrates:

Gene names of the quantified E3 ligase substrates in this experiment delimited by semicolon

substrate_ratio:

Log2 ratios of quantified Ub sites of the E3 ligase substrates in this experiment delimited by semicolon

An example—E3 ligase RBBP6 with two quantified Ub sites has an averaged Ub site ratios of 0.365. The random sampling of all quantified Ub sites with two sites in each sampling gives a normal distribution with an average ratio of 0.1148 and standard deviation of 0.268. The cumulative p-value of the averaged Ub target site ratio of 0.365 for RBBP6 in this normal distribution is 0.35. The only substrate of RBBP6 quantified in this experiment is YBX1 with two sites K65 and K178 and log2 ratios of 0.59 and 0.14, respectively.

RBBP6,2,0.365,0.1148,0.268,0.35,YBX1,YBX1_65:0.59;
YBX1_178:0.14

- Another main output of the UbE3-APA analysis is the “E3enrichment_protein_level” CSV file (Fig. 3a). The file name also contains the analysis settings similar to the site_level CSV file, depending on the settings of the input variables. Different from the site_level analysis, for protein_level analysis, the software first collects all nonredundant Ub proteins based on the Ub site quantification data and calculates the average

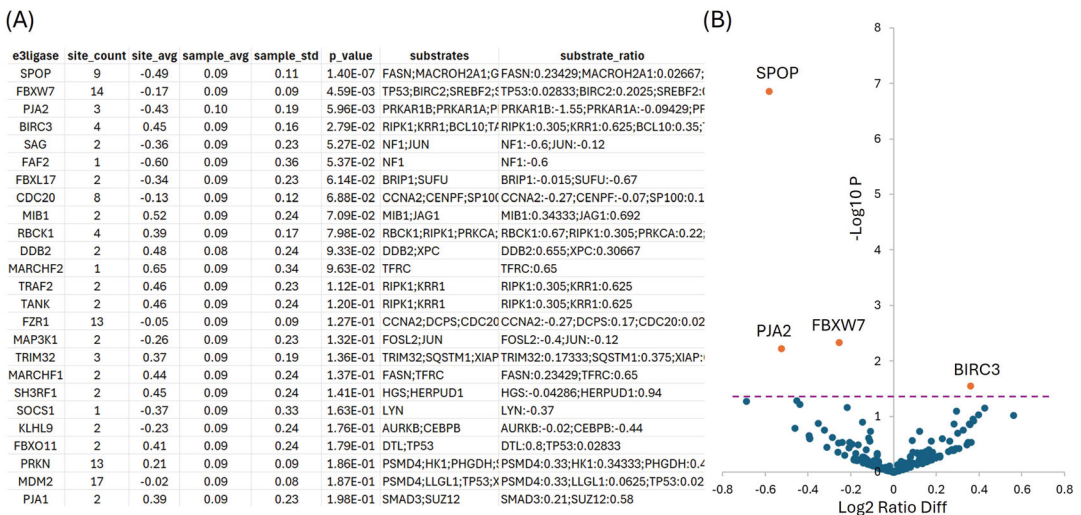


Fig. 3 Examples of UbE3-APA analysis outputs. (a) a screenshot of the output CSV file from the protein-level UbE3-APA enrichment analysis of the testing data opened in Microsoft Excel; (b) a volcano scatter plot drawn in Microsoft Excel based on the protein-level UbE3-APA enrichment analysis of the testing data. Dashed line represents $p = 0.05$ and red dots represent E3 ligases whose activity profiles are significantly changed

log2 ratio of each Ub protein. Then, the software proceeds with the statistical analysis using the ratios of substrate proteins instead of Ub sites for each E3 ligase following the same statistical method.

The protein_level result file follows the format below:

column headers

e3ligase:

Gene name of the E3 ligase

site_count:

Total number of quantified Ub substrate proteins of the E3 ligase in this experiment

site_avg:

The averaged log2 ratio of quantified Ub proteins of the E3 ligase substrates

sample_avg:

The average for the ratio distribution of the randomly sampled Ub proteins

sample_std:

The standard deviation for the ratio distribution of the randomly sampled Ub proteins

p_value:

Cumulative p value of the averaged log2 ratio of the E3 ligase substrate proteins in the ratio distribution of randomly sampled Ub proteins

substrates:

Gene names of the quantified E3 ligase substrate proteins in this experiment delimited by semicolon

substrate_ratio:

Log2 ratios of quantified Ub substrate proteins of the E3 ligase in this experiment delimited by semicolon

An example—E3 ligase RAD18 with three quantified Ub substrate proteins has an averaged Ub protein log2 ratios of 0.24. The random sampling of all quantified Ub proteins with three proteins in each sampling gives a normal distribution with an average ratio of 0.089 and standard deviation of 0.19. The cumulative p-value of the averaged Ub substrate protein ratio of 0.24 for RAD18 in this normal distribution is 0.4256. The three substrates of RAD18 quantified in this experiment are PCNA, RAD18 and FANCD2 with log2 ratios of 0.244, 0.29955 and 0.18, respectively.

RAD18,3,0.24,0.089,0.19,0.4256,PCNA;RAD18;FANCD2,
PCNA:0.244;RAD18:0.29955;FANCD2:0.18

7. Draw a volcano plot for the E3 ligase activity profiling analysis using scatter plot function in the charting software like Excel or Prism (Fig. 3b). The x axis of the scatter plot is the difference between *site_avg* column and *sample_avg* column. The y axis of the scatter plot is the $-\log_{10}$ transformation of the *p_value* column.

4 Notes

1. Current E3 ligase-substrate interaction database is limited to human genes, but the database can be easily updated following the format specified in Subheading 3.1.2. It is important to incorporate new research data and resources to ensure accurate and comprehensive coverage of E3 ligase-substrate information relevant to the study.
2. Uniprot-to-gene name matches stored in the “unigenedict.csv” file should be updated periodically from the Uniprot database using the link below to avoid obsolete records:
https://ftp.uniprot.org/pub/databases/uniprot/knowledgebase/idmapping/by_organism/
3. In the python console and in the input parameter, be sure to use forward slash “/” instead of backslash “\” for all directory names. Be sure to have “/” in the end of the input_directory, so the software can correctly find the input data with the data filenames.
4. In the python console and in the input parameter, make sure to avoid using curly smart quotes such as “” from Microsoft word as quotation marks and please always use ASCII code " as quotation mark.
5. It is important to consider the types of quantification ratios for the Ub site and optional protein input data. In a typical quantitative proteomics analysis workflow based on Stable Isotope Labeling of Amino acid in Cell culture (SILAC), analysis software like MaxQuant provides two different quantification ratios—normalized SILAC ratios and original SILAC ratios. Original SILAC ratios are the SILAC ratios the software directly measures based on the original peak intensities in the mass spectrometry analysis. Normalized SILAC ratios are based on the original SILAC ratios with the normalization of median ratio being shifted to 1 by the MaxQuant software. If the user plans to analyze Ub E3 ligase activities based on Ub ratios normalized by its corresponding protein ratios, the user should input the original ratios for both site and protein input data.

The caveat of this approach is that a portion of the Ub proteome may not have corresponding protein quantification for normalization. Alternatively, the user may choose to analyze Ub E3 ligase activities based on median-normalized Ub site ratios only, and in this case, the user should input normalized ratios of Ub sites as input data.

6. UBE3-APA analysis offers site-level and protein-level analysis as two main outputs. The site-level analysis tends to offer higher statistical significance compared to the protein-level analysis. This is because a higher number of total Ub sites are typically quantified for each E3 ligase than the number of Ub proteins. However, it is important to take into precautions that site-level analysis results can be distorted by heavily modified Ub proteins. Therefore, it would be a useful practice to compare site-level and protein-level analysis results to identify more confident and significant changes in Ub E3 activities.

Acknowledgments

The development of UBE3-APA analysis workflow was supported by the National Institute of Health (R35GM124896 to Y.C.) and the National Science Foundation (CHE-1753154 to Y.C.).

References

1. Pickart CM (2001) Mechanisms underlying ubiquitination. *Annu Rev Biochem* 70:503–533. <https://doi.org/10.1146/annurev.biochem.70.1.503>
2. Pickart CM, Eddins MJ (2004) Ubiquitin: structures, functions, mechanisms. *Biochim Biophys Acta* 1695(1–3):55–72. <https://doi.org/10.1016/j.bbamcr.2004.09.019>
3. Komander D, Rape M (2012) The ubiquitin code. *Annu Rev Biochem* 81:203–229. <https://doi.org/10.1146/annurev-biochem-060310-170328>
4. Dikic I, Schulman BA (2023) An expanded lexicon for the ubiquitin code. *Nat Rev Mol Cell Biol* 24(4):273–287. <https://doi.org/10.1038/s41580-022-00543-1>
5. Yang Q, Zhao J, Chen D, Wang Y (2021) E3 ubiquitin ligases: styles, structures and functions. *Mol Biomed* 2(1):23. <https://doi.org/10.1186/s43556-021-00043-2>
6. Liu N, Lin MM, Wang Y (2023) The emerging roles of E3 ligases and DUBs in neurodegenerative diseases. *Mol Neurobiol* 60(1):247–263. <https://doi.org/10.1007/s12035-022-03063-3>
7. Hughes DC et al (2022) Ubiquitin ligases in longevity and aging skeletal muscle. *Int J Mol Sci* 23(14). <https://doi.org/10.3390/ijms23147602>
8. Chen X, Ma J, Wang ZW, Wang Z (2024) The E3 ubiquitin ligases regulate inflammation in cardiovascular diseases. *Semin Cell Dev Biol* 154(Pt C):167–174. <https://doi.org/10.1016/j.semcdb.2023.02.008>
9. Senft D, Qi J, Ronai ZA (2018) Ubiquitin ligases in oncogenic transformation and cancer therapy. *Nat Rev Cancer* 18(2):69–88. <https://doi.org/10.1038/nrc.2017.105>
10. Erber LN, Chen Y (2021) Functional proteomics. American Chemical Society
11. Liu Y, Patricelli MP, Cravatt BF (1999) Activity-based protein profiling: the serine hydrolases. *Proc Natl Acad Sci USA* 96(26):14694–14699. <https://doi.org/10.1073/pnas.96.26.14694>
12. Barglow KT, Cravatt BF (2007) Activity-based protein profiling for the functional annotation of enzymes. *Nat Methods* 4(10):822–827. <https://doi.org/10.1038/nmeth1092>

13. Cravatt BF, Wright AT, Kozarich JW (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* 77:383–414. <https://doi.org/10.1146/annurev.biochem.75.101304.124125>
14. Nomura DK, Dix MM, Cravatt BF (2010) Activity-based protein profiling for biochemical pathway discovery in cancer. *Nat Rev Cancer* 10(9):630–638. <https://doi.org/10.1038/nrc2901>
15. Haubrich BA, Swinney DC (2016) Enzyme activity assays for protein kinases: strategies to identify active substrates. *Curr Drug Discov Technol* 13(1):2–15. <https://doi.org/10.2174/1570163813666160115125930>
16. Mischnik M et al (2016) IKAP: a heuristic framework for inference of kinase activities from Phosphoproteomics data. *Bioinformatics* 32(3):424–431. <https://doi.org/10.1093/bioinformatics/btv699>
17. Yang P et al (2016) KinasePA: phosphoproteomics data annotation using hypothesis driven kinase perturbation analysis. *Proteomics* 16(13):1868–1871. <https://doi.org/10.1002/pmic.201600068>
18. Wiredja DD, Koyuturk M, Chance MR (2017) The KSEA app: a web-based tool for kinase activity inference from quantitative phosphoproteomics. *Bioinformatics* 33(21):3489–3491. <https://doi.org/10.1093/bioinformatics/btx415>
19. Krug K et al (2019) A curated resource for phosphosite-specific signature analysis. *Mol Cell Proteomics* 18(3):576–593. <https://doi.org/10.1074/mcp.TIR118.000943>
20. Erber LN et al (2021) Iron deficiency reprograms phosphorylation signaling and reduces O-GlcNAc pathways in neuronal cells. *Nutrients* 13(1). <https://doi.org/10.3390/nut13010179>
21. Kuleshov MV et al (2021) KEA3: improved kinase enrichment analysis via data integration. *Nucleic Acids Res* 49(W1):W304–W316. <https://doi.org/10.1093/nar/gkab359>
22. Timms RT et al (2023) Defining E3 ligase-substrate relationships through multiplex CRISPR screening. *Nat Cell Biol* 25(10):1535–1545. <https://doi.org/10.1038/s41556-023-01229-2>
23. Bradley D, Beltrao P (2019) Evolution of protein kinase substrate recognition at the active site. *PLoS Biol* 17(6):e3000341. <https://doi.org/10.1371/journal.pbio.3000341>
24. Nguyen VN et al (2016) UbiNet: an online resource for exploring the functional associations and regulatory networks of protein ubiquitylation. *Database (Oxford)* 2016. <https://doi.org/10.1093/database/baw054>
25. Li Y et al (2017) An integrated bioinformatics platform for investigating the human E3 ubiquitin ligase-substrate interaction network. *Nat Commun* 8(1):347. <https://doi.org/10.1038/s41467-017-00299-9>
26. Chen D et al (2019) A multidimensional characterization of E3 ubiquitin ligase and substrate interaction network. *iScience* 16:177–191. <https://doi.org/10.1016/j.isci.2019.05.033>
27. Li Z et al (2021) UbiNet 2.0: a verified, classified, annotated and updated database of E3 ubiquitin ligase-substrate interactions. *Database (Oxford)* 2021. <https://doi.org/10.1093/database/baab010>
28. Gong Y, Chen Y (2022) UbE3-APA: a bioinformatic strategy to elucidate ubiquitin E3 ligase activities in quantitative proteomics study. *Bioinformatics* 38(8):2211–2218. <https://doi.org/10.1093/bioinformatics/btac069>
29. Hansen FM et al (2021) Data-independent acquisition method for ubiquitinome analysis reveals regulation of circadian biology. *Nat Commun* 12(1):254. <https://doi.org/10.1038/s41467-020-20509-1>
30. Li Y et al (2019) A quantitative chemical proteomics approach for site-specific stoichiometry analysis of ubiquitination. *Angew Chem Int Ed Engl* 58(2):537–541. <https://doi.org/10.1002/anie.201810569>
31. Theurillat JP et al (2014) Prostate cancer. Ubiquitylome analysis identifies dysregulation of effector substrates in SPOP-mutant prostate cancer. *Science* 346(6205):85–89. <https://doi.org/10.1126/science.1250255>



Recent Overview of Protein Palmitoylation and Profiling Methodologies

Changyan Jin, Qiaoling Yuan, and Zhipeng Tao

Abstract

Protein palmitoylation is a reversible posttranslational modification in which a palmitoyl group (a 16-carbon saturated fatty acid) is covalently attached to cysteine residues on proteins, typically through a thioester bond. This modification affects the protein's hydrophobicity, influencing its membrane association, localization, stability, trafficking, and overall function. Dysregulation of palmitoylation has been implicated in diseases such as cancer, neurodegenerative diseases, and cardiovascular disorders. In this review, we summarize the recent findings related to protein palmitoylation and its biological functions. More importantly, we examine proteomic studies that utilize active-based protein profiling (ABPP) to design novel probes or inhibitors aimed at enhancing the accuracy and efficiency of large-scale analyses of protein palmitoylation. These advancements will facilitate the findings of novel therapeutic targets and the designing of targeted therapies, providing increasingly critical insights into the role of this modification in health and diseases.

Key words Palmitoylation, Active-based protein profiling, Chemical probes

1 Recent Findings of Protein Palmitoylation and Functions

Protein palmitoylation is a reversible posttranslational modification where a fatty acid, typically palmitic acid, is covalently attached to cysteine residues via a thioester bond [1]. This lipid modification plays a crucial role in the regulation of protein localization, stability, and function, impacting a variety of cellular processes such as signaling, trafficking, and membrane association [2–6]. For example, Gasdermin D (GSDMD) palmitoylation is orchestrated by reactive oxygen species (ROS) and inflammasome activation [3], which interferes with pyroptosis and the release of cytokine [7–9]. Similarly, protein palmitoylation can also affect the activity of inflammasome. NOD-, LRR-, and pyrin domain-containing protein 3 (NLRP3) is a pivotal component of inflammasome, and it is essential for the control of inflammatory responses [10, 11]. A previous study found that S-palmitoylation at conserved cysteine

844 mediated by the zinc finger DHHC-type palmitoyltransferase 12 (ZDHHC12) promotes NLRP3 degradation through chaperone-mediated autophagy, resulting in the regulation of NLRP3 and effective switch off of inflammasome activity [12]. However, other studies have indicated that NLRP3 palmitoylation, particularly at cysteine 898, is regulated by ZDHHC5 and fatty acid synthesis, which are essential for inflammasome activation [13, 14]. These findings highlight the complex regulation of protein palmitoylation under physiological conditions. Wu et al. found that palmitoylation of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike (S) protein, facilitated by ZDHHC5 and Golgin A7 (GOLGA7), is crucial for membrane fusion and viral entry [15]. Inhibiting this process reduces infectivity by disrupting S-trimer formation [15]. Another study also emphasized the posttranslational modification of viral proteins and palmitoylation of the envelope (E) protein of SARS-CoV-2 at 40/43/44 cysteine residues via ZDHHC3, 6, 12, 15, and 20 are essential for its stability and interaction with structure proteins of the virus [16]. Mutations in these sites impair virus assembly, indicating that palmitoylation is critical for viral entry, infectivity, and particle formation in SARS-CoV-2 [16]. In addition, the roles of palmitoylation in major signaling pathways have been studied. Firstly, palmitoylation can modulate protein kinase B (AKT) signaling both directly and indirectly. AKT is activated by phosphatidylinositol-3-kinase (PI3K) [17–19]. It is essential for cell growth, proliferation, metabolism, and survival, with upregulated activity linked to cancer [20]. Palmitoylation of AKT at Cys344 is critical for its activation and protection from degradation. Mutation at this site reduces AKT phosphorylation and its recruitment to lysosomes [20, 21]. In breast cancer cells, ZDHHC22-mediated palmitoylation of the mammalian target of rapamycin (mTOR) suppresses AKT signaling by destabilizing mTOR [20, 22]. Conversely, in liver cancer, proprotein convertase subtilisin/kexin type 9 (PCSK9) is an enzyme for modulating cholesterol balance [23, 24]. Palmitoylation promotes phosphatase and tensin homolog (PTEN) degradation, removing its inhibition of AKT, and enhancing AKT signaling [25]. Therefore, AKT regulation by palmitoylation is complex and varies depending on cellular and physiological context [20, 25]. Secondly, Wnt signaling is fine-tuned by protein palmitoylation. Wnt signaling is crucial for organ development and maintenance [26, 27]. Both low-density lipoprotein receptor-related proteins 5 and 6 (LRP5/6) co-receptors activate the β -catenin pathway, while dickkopf1 (DKK1) inhibits Wnt by degrading LRP6 [28]. Combined to form a complex of DKK1 and LRP6 with cytoskeleton-associated protein 4 (CKAP4), a palmitoylated endoplasmic reticulum (ER) protein, which influences mitochondrial structure and function by binding to voltage-dependent anion-selective channel protein 2 (VDAC2) on the

outer mitochondrial membrane. Knockout of CKAP4 in HeLaS3 cells increases ER-mitochondria contacts, raises mitochondrial Ca^{2+} , and reduces membrane potential, leading to impaired cell growth. Palmitoylation at Cys100 is essential for CKAP4's role in maintaining mitochondrial function as well as inhibiting Wnt signaling by DKK1 [28]. Similarly, focal adhesion kinase (FAK) is a critical factor in tumor metastasis, requiring its S-palmitoylation for membrane localization and activation, which was catalyzed by ZDHHC5. Consequently, inhibition of FAK palmitoylation affects tumor cell invasion and proliferation [29]. Additionally, insulin-like growth factor-1 (IGF-1) signaling triggers palmitoylation turnover of flotillin-1, dynamically regulated by acyl protein thioesterases-1 (APT-1) and ZDHHC-19, which prevents IGF-1 receptor desensitization in cervical cancer cells [30]. Therefore, blocking palmitoylation turnover decreased epithelial-mesenchymal transition (EMT) and cancer cell invasion [30]. Hence, this evidence revealed the critical role of protein palmitoylation in mediating protein function in the signaling pathway. Furthermore, current studies have also demonstrated that modulation of protein palmitoylation may potentially provide therapeutic potential and targets for the treatment of specific diseases and cancer. Zheng et al. demonstrated that ZDHHC11 mediates S-acylation of adipose triacylglyceride lipase (ATGL), affecting lipid droplets (LDs) size and abundance via regulating ATGL, which is a rate-limiting enzyme for hydrolysis of triacylglycerol in lipid droplets [31]. Palmitoylation of ATGL at Cys15 is associated with enhancing LD homeostasis via decreasing LD size and abundance as well as regulating ATGL function [31]. In breast cancer (BrCa), ZDHHC22-regulated mTOR palmitoylation blocked the AKT signaling pathway and diminished malignant proliferation and endocrine resistance in BrCa cells [22]. Regarding the effect of palmitoylation on neurological disorders, synaptotagmin-11 (Syt11), a vesicle-trafficking protein associated with Parkinson's disease (PD), may interact with α -synuclein and cause abnormal aggregation that defines PD pathology [32]. Ho et al. demonstrated that Syt11 is palmitoylated in both mouse and human brain tissue and cultured neurons. The modification at Cys39 and Cys40 protects Syt11 from breakdown, increasing its levels and promoting α -synuclein's binding to intracellular membranes. As a result, the physiological tetrameric form of α -synuclein decreases, while its aggregation-prone monomeric form increases, suggesting that Syt11 palmitoylation, could trigger the monomeric form of α -synuclein aggregation in PD [32]. Moreover, in another neurological disease—Huntington disease (HD), ML348 stimulates the palmitoylation of brain-derived neurotrophic factor (BDNF) which, in turn, improves the activity of transport of BDNF in addition to restoring axonal transport, synaptic status, and survival signals toward wild-type level via preventing the depalmitoylation activity of acyl- protein thioesterase

1 (APT1) [33]. Additionally, nucleotide oligomerization domain-containing protein 2 (NOD2) is an intracellular immunoreceptor that detects bacterial peptidoglycans [34, 35]. Though mostly in the cytosol, part of NOD2 is associated with the plasma membrane and endosomes for microbial surveillance. Palmitoylation by ZDHHC5 facilitates the recruitment of NOD2 to membranes, triggering proinflammatory and antimicrobial responses. A depalmitoylation processing performed by proteins containing α/β -hydrolase structural domains (ABHD17A, 17B, 17C) releases NOD2 back to the cell membrane, allowing for a dynamic cycle. However, the inhibition of ABHD17 depalmitoylation activity increases membrane localization of wild-type NOD2 and Crohn's disease-associated variants, boosting NF- κ B activation and cytokine production. Hence, ABHD17 inhibition may restore function in some NOD2 variants relevant to Crohn's disease [36].

In recent years, the posttranslational modification, palmitoylation, has emerged as an important area in the ongoing search for diagnostic and therapeutic targets for cancer, immune disorders, and neurodegenerative diseases. Therefore, identifying additional palmitoylated proteins using high-throughput methodologies is crucial.

2 Active-Based Protein Profiling for Protein Palmitoylation

The challenge posed by the lack of antibodies specific to palmitoylated proteins limits our understanding of how dynamic lipid modification affects cell signaling and regulation. The development of proteomics techniques and chemical probes to study protein palmitoylation has significantly advanced our knowledge in the field [20, 37]. First, Studies have improved the efficiency of enrichment and detection methods compared to traditional methodologies over the years. For instance, Chen et al. developed a simplified method called the “soluble gel-based single-tube treatment (DGOTT)” that combines soluble tube gels with the classical ABE method to reduce sample loss due to multiple methanol-chloroform precipitations. This method allows alkylation and substitution reactions to be carried out in a single tube and simplifies the removal of interferences by gel washing, thereby reducing complexity and shortening processing time. Using DGOTT, 1625 potential S-acylated proteins have been identified from 800 μ g of mouse cerebral cortex proteins [38]. Mariscal et al. used a new metabolically labelled, cysteine-focused method, low background acyl biotin exchange (LB-ABE), to analyze the palmitoyl proteome of both large-extracellular vesicles (L-EVs) and small-extracellular vesicles (S-EVs) in prostate cancer cells. L-EVs and S-EVs were shown to contain proteins associated with different biological functions and origins. Six transmembrane epithelial

antigens of the prostates (STEAP1, STEAP2) and ATP binding cassette subfamily C member 4 (ABCC4) were defined as prostate cancer-specific palmitoyl proteins enriched in both EV types via LB-ABE coupled LC-MS/MS analysis. Notably, inhibition of palmitoylation in matrices reduces the presence of these proteins in EVs. Therefore, the modification may regulate the classification of EV-bound secretomes and contribute toward the assay of pathological biomarkers [39]. Certainly, these improvements are not only present in the ABE assay but also in the metabolic labeling enrichment assay. Indirect identification of modified proteins with unattached fatty acids has limitations, while accurate identification of S-acylation sites for attached fatty acids is complicated. Hence, to analyze the heterogeneity and turnover of S-acylation, Wu et al. developed a bioorthogonal tag that is cleavable in metabolic labeling approaches (ssMLCC) to analyze intact S-acylated peptides [40]. The tag enables selective enrichment and efficient MS analysis of intact S-acylated peptides, allowing direct mapping of S-acylated sites and their bound fatty acids while reducing false positives. In HeLa cells, 606 S-palmitoylation (C16:0) sites have been identified in 441 proteins. In addition, 85 fast cycling sites have been identified in 79 proteins by monitoring S-palmitoylation turnover rates using a pulse-chase approach [40]. For the identification of protein palmitoylation, a challenge is the absence of a reliable high-throughput screening (HTS) method. Zheng et al. synthesized and characterized 2-bromohexadec-15-ynoic acid and 2-bromooctadec-17-ynoic acid, which are analogs of 2-bromopalmitate (2-BP), as activity-based probes for PATs as well as other palmitoylating and 2-BP-binding enzymes; these probes will serve as new chemical tools for activity-based protein profiling to explore palmitoyl acyltransferases (PATs) [41]. TEA domain transcription factor (TEAD) and regulatory factor X3 (RFX3) are the identified autopalmitoylated proteins by using these probes [5, 42]. Since then, the autopalmitoylation of TEAD proteins have been served as a platform for small molecule designing and screening to be used in the treatment of cancers [43–47]. An inhibitor similar to 2-bromopalmitate (2-BP) has been reported, cyanoacrylamide (CMA) improves the non-specificity of 2-BP. CMA showed little inhibition of APT1 and APT2 even at higher concentrations (up to 50 μ M), whereas 2-BP significantly inhibited both enzymes, especially at concentrations of 25 μ M and less. This suggests that CMA is an important tool for studying the dynamics and kinetics of the acylation process without interfering with APT-mediated diacylation [48, 49]. It has an acrylamide warhead instead of the α -halogenated fatty acid of 2-BP, inhibits the DHHC family of proteins in cells, reduces toxicity, and impedes suppression of depalmitoylases [50]. CMA has been shown to bind to DHHC family proteins in cells, inhibiting protein S-acylation and disrupting DHHC-regulated cellular processes. It is a versatile chemical tool

for unwinding the complexity of the cellular signaling mediated via protein S-acylation by DHHC [50]. After that, the same laboratory team reported palmitoyl transferase probes (PTPs), advanced pro-fluorescent molecules that mimic the palmitoyl-CoA substrate of DHHC proteins [51]. These PTPs were used to create and validate a high-throughput screening (HTS) assay with a high Z-factor. Using this assay, they conducted a pilot screen of 1687 acrylamide-based compounds against zDHHC20. Although no potent non-lipid compounds have been identified so far, future larger-scale HTS using this probe has the potential to identify more drug-like DHHC inhibitors and perhaps isozyme-selective inhibitors [51]. Similarly, regarding the modeling drug therapy in response to protein palmitoylation, traditional transmembrane Förster resonance energy transfer (FRET) is inefficient due to donors and acceptors being on opposite membrane sides. Researchers illustrated a down-membrane multiple FRET (multi-FRET) method to enhance the visualization of palmitoylated proteins using metabolic labeling and target recognition. Azidopalmitic acid (azido-PA) is metabolized into palmitoylated proteins and then reacts with dibenzocyclooctyne-modified Cy5 (DBCO-Cy5) via copper-free click chemistry. The multi-FRET probe binds specifically to the target protein, initiating a hybridization chain reaction (HCR) for signal amplification. This cis-membrane approach shortens the donor-acceptor distance, boosting FRET efficiency. HCR further promotes FRET, enabling improved imaging of palmitoylated programmed death-ligand 1 (PD-L1) and detection of drug-induced changes in its palmitoylation [52]. Meanwhile, Salaun et al. also used a FRET-based high-throughput screening assay to identify compounds that inhibit the autoacylation of ZDHHC2, a neuronal pathway enzyme. By screening more than 350,000 compounds, they found two tetrazolium-containing inhibitors (TTZ-1 and TTZ-2) that were effective in both cell-free systems and human embryonic kidney 293 T cells, showing isoenzyme selectivity in inhibiting substrate S-acylation (synaptosomal-associated protein, 25 kDa (SNAP25), postsynaptic density protein 95 (PSD95)) [53]. Interestingly, a yeast-based high-throughput method has been reported that could avoid toxicity and permeability of cells. This method uses a reporter gene, HIS3, which can monitor the acylation state of a palmitoylation substrate linked to a transcription factor as well as been shown to work for heterologous PATs like human ZDHHC20, mouse ZDHHC21, and a PAT from *Giardia lamblia* [54]. In addition, 3200 compounds were screened for inhibitors of palmitoylation of yeast casein kinase 2 (Yck2) by the yeast PAT aldo-keto reductase family 1 (Akr1), and the successful identification of a candidate molecule (4-(3-bromophenyl)-2-(methylthio)-6-oxo-1,4,5,6-tetrahydropyridine-3-carbonitrile (P15G3)) demonstrates the validity of the methodology. However, it also needs further study to validate due to Yck2 being

palmitoylated by various PATs, making it difficult to identify inhibitors specific to a single PAT [54]. Finally, one recent study developed an advanced chemical genetic approach to directly label and recognize substrates for specific ZDHHC enzymes in cells using a stereo-complementation method (formation of “bump and hole”) [55]. Based on the stereo-complementation method, five mutant/probe pairs have been identified, which are ZDHHCs 3, 7, 11, 15, and 20, respectively. And selective activity of ZDHHCs has been demonstrated, with higher specificity than the wild-type enzyme. A combination of stereo-complementation and chemical proteomics allowed them to identify more than 300 potential ZDHHC-specific substrates and S-acylation sites in various cell lines, resulting in the expansion of the substrate mapping network for ZDHHC 7, 15, and 20 [55]. In the future, researchers can also use cysteine set enrichment analysis (CSEA), which is a computational analysis method based on gene-centric enrichment score algorithms. If the ligandable cysteines identified in the study overlap with the palmitoylation site, then these cysteines could act as dual functional sites. Drugs designed to bind these cysteines have the potential to disrupt cancer-related signaling pathways regulated by protein palmitoylation [56]. Proteomic approaches and active probes deliver a robust method for studying PTMs across the proteome, providing the capability to identify and quantify the hundreds of proteins that occur with specific modifications in complex biological systems.

In summary, emerging evidence has revealed the critical role of protein palmitoylation in posttranslational modification and physiological and pathological conditions, especially metabolic diseases, neurodegenerative diseases, and cancer. The development of innovative technologies for identifying and detecting protein palmitoylation would address key emerging challenges in the field and toward revolutionizing the study of protein palmitoylation. Therefore, identifying S-acyl proteomics through the use of novel probes and inhibitors would play important roles in understanding disease mechanisms and discovering novel biomarkers and therapeutic targets.

Acknowledgments

This work is partially supported by Texas Woman's University Research Enhancement Program Award #4191 to Z.T.

References

1. Chen B, Sun Y, Niu J, Jarugumilli GK, Wu X (2018) Protein lipidation in cell signaling and diseases: function, regulation, and therapeutic opportunities. *Cell Chem Biol* 25(7):817–831
2. Linder ME, Deschenes RJ (2007) Palmitoylation: policing protein stability and traffic. *Nat Rev Mol Cell Biol* 8(1):74–84

3. Du G, Healy LB, David L, Walker C, El-Baba TJ, Lutowski CA, Goh B, Gu B, Pi X, Devant P (2024) ROS-dependent S-palmitoylation activates cleaved and intact gasdermin D. *Nature* 630:1–3
4. Liu Z, Li S, Wang C, Vidmar KJ, Bracey S, Li L, Willard B, Miyagi M, Lan T, Dickinson BC (2024) Palmitoylation at a conserved cysteine residue facilitates gasdermin D-mediated pyroptosis and cytokine release. *Proc Natl Acad Sci* 121(29):e2400883121
5. Chan P, Han X, Zheng B, DeRan M, Yu J, Jarugumilli GK, Deng H, Pan D, Luo X, Wu X (2016) Autopalmitoylation of TEAD proteins regulates transcriptional output of the Hippo pathway. *Nat Chem Biol* 12(4):282–289
6. Dong G, Adak S, Spyropoulos G, Zhang Q, Feng C, Yin L, Speck SL, Shyr Z, Morikawa S, Kitamura RA (2023) Palmitoylation couples insulin hypersecretion with β cell failure in diabetes. *Cell Metab* 35(2):332–344. e337
7. Liu Z, Li S, Wang C, Vidmar KJ, Bracey S, Li L, Willard B, Miyagi M, Lan T, Dickinson BC, Osme A, Pizarro TT, Xiao TS (2024) Palmitoylation at a conserved cysteine residue facilitates gasdermin D-mediated pyroptosis and cytokine release. *Proc Natl Acad Sci* 121(29):e2400883121. <https://doi.org/10.1073/pnas.2400883121>
8. Balasubramanian A, Hsu AY, Ghimire L, Tahir M, Devant P, Fontana P, Du G, Liu X, Fabin D, Kambara H (2024) The palmitoylation of gasdermin D directs its membrane translocation and pore formation during pyroptosis. *Sci Immunol* 9(94):eadn1452
9. Zhang N, Zhang J, Yang Y, Shan H, Hou S, Fang H, Ma M, Chen Z, Tan L, Xu D (2024) A palmitoylation–depalmitoylation relay spatiotemporally controls GSDMD activation in pyroptosis. *Nat Cell Biol* 26(5):757–769
10. Xu J, Núñez G (2023) The NLRP3 inflammasome: activation and regulation. *Trends Biochem Sci* 48(4):331–344
11. Jin C, Flavell RA (2010) Molecular mechanism of NLRP3 inflammasome activation. *J Clin Immunol* 30:628–631
12. Wang L, Cai J, Zhao X, Ma L, Zeng P, Zhou L, Liu Y, Yang S, Cai Z, Zhang S, Zhou L, Yang J, Liu T, Jin S, Cui J (2023) Palmitoylation prevents sustained inflammation by limiting NLRP3 inflammasome activation through chaperone-mediated autophagy. *Mol Cell* 83(2):281–297. e210. <https://doi.org/10.1016/j.molcel.2022.12.002>
13. Leishman S, Aljadeed NM, Qian L, Cockcroft S, Behmoaras J, Anand PK (2024) Fatty acid synthesis promotes inflammasome activation through NLRP3 palmitoylation. *Cell Rep* 43(8):114516
14. Zheng S, Que X, Wang S, Zhou Q, Xing X, Chen L, Hou C, Ma J, An P, Peng Y (2023) ZDHHC5-mediated NLRP3 palmitoylation promotes NLRP3-NEK7 interaction and inflammasome activation. *Mol Cell* 83(24):4570–4585. e4577
15. Wu Z, Zhang Z, Wang X, Zhang J, Ren C, Li Y, Gao L, Liang X, Wang P, Ma C (2021) Palmitoylation of SARS-CoV-2 S protein is essential for viral infectivity. *Signal Transduct Target Ther* 6(1):231. <https://doi.org/10.1038/s41392-021-00651-y>
16. Wang Z, Qiu M, Ji Y, Chai K, Liu C, Xu F, Guo F, Tan J, Liu R, Qiao W (2024) Palmitoylation of SARS-CoV-2 Envelope protein is central to virus particle formation. *J Virol* 98:e01072–e01024
17. Lee M-S, Jeong M-H, Lee H-W, Han H-J, Ko A, Hewitt SM, Kim J-H, Chun K-H, Chung J-Y, Lee C (2015) PI3K/AKT activation induces PTEN ubiquitination and destabilization accelerating tumorigenesis. *Nat Commun* 6(1):7769
18. Luo J, Manning BD, Cantley LC (2003) Targeting the PI3K-Akt pathway in human cancer: rationale and promise. *Cancer Cell* 4(4):257–262
19. Tao Z, Cheng Z (2023) Hormonal regulation of metabolism—recent lessons learned from insulin and estrogen. *Clin Sci* 137(6):415–434
20. Li M, Zhang L, Chen C-W (2023) Diverse roles of protein palmitoylation in cancer progression, immunity, stemness, and beyond. *Cells* 12(18):2209
21. Blaustein M, Piegari E, Martínez Calejman C, Vila A, Amante A, Manese MV, Zeida A, Abrami L, Veggetti M, Guertin DA (2021) Akt is S-palmitoylated: a new layer of regulation for Akt. *Front Cell Dev Biol* 9:626404
22. Huang J, Li J, Tang J, Wu Y, Dai F, Yi Z, Wang Y, Li Y, Wu Y, Ren G (2022) ZDHHC22-mediated mTOR palmitoylation restrains breast cancer growth and endocrine therapy resistance. *Int J Biol Sci* 18(7):2833
23. Ioannou GN, Lee SP, Linsley PS, Gersuk V, Yeh MM, Chen YY, Peng YJ, Dutta M, Mascarinas G, Molla B (2022) Pcsk9 deletion promotes murine nonalcoholic steatohepatitis and hepatic carcinogenesis: role of cholesterol. *Hepatol Commun* 6(4):780–794
24. Fasolato S, Pigozzo S, Pontisso P, Angeli P, Ruscica M, Savarino E, De Martin S, Lupo MG, Ferri N (2020) PCSK9 levels are raised

- in chronic HCV patients with hepatocellular carcinoma. *J Clin Med* 9(10):3134
25. Sun Y, Zhang H, Meng J, Guo F, Ren D, Wu H, Jin X (2022) S-palmitoylation of PCSK9 induces sorafenib resistance in liver cancer by activating the PI3K/AKT pathway. *Cell Rep* 40(7):111194
 26. Steinhart Z, Angers S (2018) Wnt signaling in development and tissue homeostasis. *Development* 145(11):dev146589
 27. Clevers H, Loh KM, Nusse R (2014) An integral program for tissue renewal and regeneration: Wnt signaling and stem cell control. *Science* 346(6205):1248012
 28. Harada T, Sada R, Osugi Y, Matsumoto S, Matsuda T, Hayashi-Nishino M, Nagai T, Harada A, Kikuchi A (2020) Palmitoylated CKAP4 regulates mitochondrial functions through an interaction with VDAC2 at ER-mitochondria contact sites. *J Cell Sci* 133(21). <https://doi.org/10.1242/jcs.249045>
 29. Wang Y, Shen N, Yang Y, Xia Y, Zhang W, Lu Y, Wang Z, Yang Z, Wang Z (2024) ZDHHC5-mediated S-palmitoylation of FAK promotes its membrane localization and epithelial-mesenchymal transition in glioma. *Cell Commun Signal* 22(1):46. <https://doi.org/10.1186/s12964-023-01366-z>
 30. Kwon H, Choi M, Ahn Y, Jang D, Pak Y (2023) Flotillin-1 palmitoylation turnover by APT-1 and ZDHHC-19 promotes cervical cancer progression by suppressing IGF-1 receptor desensitization and proteostasis. *Cancer Gene Ther* 30(2):302–312. <https://doi.org/10.1038/s41417-022-00546-2>
 31. Zheng Y, Chen J, Macwan V, Dixon CL, Li X, Liu S, Yu Y, Xu P, Sun Q, Hu Q, Liu W, Raught B, Fairn GD, Neculai D (2024) S-acylation of ATGL is required for lipid droplet homeostasis in hepatocytes. *Nat Metab* 6(8):1549–1565. <https://doi.org/10.1038/s42255-024-01085-w>
 32. Ho GPH, Wilkie EC, White AJ, Selkoe DJ (2023) Palmitoylation of the Parkinson's disease-associated protein synaptotagmin-11 links its turnover to α -synuclein homeostasis. *Sci Signal* 16(772):eadd7220. <https://doi.org/10.1126/scisignal.add7220>
 33. Virlogeux A, Scaramuzzino C, Lenoir S, Carpentier R, Louessard M, Genoux A, Lino P, Hinckelmann M-V, Perrier AL, Humbert S, Saudou F (2021) Increasing brain palmitoylation rescues behavior and neuropathology in Huntington disease mice. *Science*. *Advances* 7(14):eabb0799. <https://doi.org/10.1126/sciadv.abb0799>
 34. Keestra-Gounder AM, Tsolis RM (2017) NOD1 and NOD2: beyond peptidoglycan sensing. *Trends Immunol* 38(10):758–767
 35. Girardin SE, Travassos LH, Hervé M, Blanot D, Boneca IG, Philpott DJ, Sansonetti PJ, Mengin-Lecreulx D (2003) Peptidoglycan molecular requirements allowing detection by Nod1 and Nod2. *J Biol Chem* 278(43):41702–41708
 36. Dixon CL, Martin NR, Niphakis MJ, Cravatt BF, Fairn GD (2023) Attenuating ABHD17 enhances S-palmitoylation, membrane localization and signal transduction of NOD2 and Crohn's disease-associated variants. *bioRxiv*:2023.2012.2020.572362. <https://doi.org/10.1101/2023.12.20.572362>
 37. Wang Y, Yang W (2020) Proteome-scale analysis of protein S-acylation comes of age. *J Proteome Res* 20(1):14–26
 38. Chen W, He Y, Fang C, Lu H (2024) A rapid and convenient sample treatment method based on the dissolvable polyacrylamide gel for S-acylation proteomics. *Anal Methods* 16(27):4590–4598. <https://doi.org/10.1039/D4AY00937A>
 39. Mariscal J, Vagner T, Kim M, Zhou B, Chin A, Zandian M, Freeman MR, You S, Zijlstra A, Yang W, Di Vizio D (2020) Comprehensive palmitoyl-proteomic analysis identifies distinct protein signatures for large and small cancer-derived extracellular vesicles. *J Extracell Vesicles* 9(1):1764192. <https://doi.org/10.1080/20013078.2020.1764192>
 40. Wu R, Ji G, Chen W, Zhang L, Fang C, Lu H (2024) Simultaneous and site-specific profiling of heterogeneity and turnover in protein S-acylation by intact S-acylated peptide analysis with a cleavable bioorthogonal tag. *Analyst* 149(4):1111–1120. <https://doi.org/10.1039/D3AN02059B>
 41. Zheng B, DeRan M, Li X, Liao X, Fukata M, Wu X (2013) 2-Bromopalmitate analogues as activity-based probes to explore palmitoyl acyltransferases. *J Am Chem Soc* 135(19):7082–7085
 42. Chen B, Niu J, Kreuzer J, Zheng B, Jarugumilli GK, Haas W, Wu X (2018) Auto-fatty acylation of transcription factor RFX3 regulates ciliogenesis. *Proc Natl Acad Sci* 115(36):E8403–E8412
 43. Tang TT, Konradi AW, Feng Y, Peng X, Ma M, Li J, Yu F-X, Guan K-L, Post L (2021) Small molecule inhibitors of TEAD auto-palmitoylation selectively inhibit proliferation and tumor growth of NF2-deficient mesothelioma. *Mol Cancer Ther* 20(6):986–998

44. Holden JK, Crawford JJ, Noland CL, Schmidt S, Zbieg JR, Lacap JA, Zang R, Miller GM, Zhang Y, Beroza P (2020) Small molecule dysregulation of TEAD lipidation induces a dominant-negative inhibition of hippo pathway signaling. *Cell Rep* 31(12):107809
45. Tang TT, Seely JC, Bleavins MR, Post L (2024) Comparing TEAD palmitoylation inhibitors with differential TEAD selectivity in combination efficacy with targeted therapies and in renal safety. *Cancer Res* 84(6_Supplement):7282–7282
46. Sun Y, Hu L, Tao Z, Jarugumilli GK, Erb H, Singh A, Li Q, Cotton JL, Greninger P, Egan RK (2022) Pharmacological blockade of TEAD–YAP reveals its therapeutic limitation in cancer cells. *Nat Commun* 13(1):6744
47. Li Y, Li Y, Ning C, Yue J, Zhang C, He X, Wang Y, Liu Z (2022) Discovering inhibitors of TEAD palmitate binding pocket through virtual screening and molecular dynamics simulation. *Comput Biol Chem* 98:107648
48. Pedro MP, Vilcaes AA, Tomatis VM, Oliveira RG, Gomez GA, Daniotti JL (2013) 2-Bromopalmitate reduces protein deacylation by inhibition of acyl-protein thioesterase enzymatic activities. *PLoS One* 8(10):e75232
49. Qiu T, Kathayat RS, Cao Y, Beck MW, Dickinson BC (2018) A fluorescent probe with improved water solubility permits the analysis of protein S-depalmitoylation activity in live cells. *Biochemistry* 57(2):221–225
50. Azizi S-A, Lan T, Delalande C, Kathayat RS, Banales Mejia F, Qin A, Brookes N, Sandoval PJ, Dickinson BC (2021) Development of an acrylamide-based inhibitor of protein S-acylation. *ACS Chem Biol* 16(8):1546–1556
51. Qiu T, Azizi S-A, Brookes N, Lan T, Dickinson BC (2022) A high-throughput fluorescent turn-on assay for inhibitors of DHHC family proteins. *ACS Chem Biol* 17(8):2018–2023
52. Fu Y, Qian H, Yang Y, Li J, Xie G (2024) Enhanced imaging of protein-specific palmitoylation with HCR-based cis-membrane multi-FRET. *Talanta* 266:124972. <https://doi.org/10.1016/j.talanta.2023.124972>
53. Salaun C, Takizawa H, Galindo A, Munro KR, McLellan J, Sugimoto I, Okino T, Tomkinson NCO, Chamberlain LH (2022) Development of a novel high-throughput screen for the identification of new inhibitors of protein S-acylation. *J Biol Chem* 298(10):102469. <https://doi.org/10.1016/j.jbc.2022.102469>
54. Coronel Arrechea C, Giolito ML, García IA, Soria G, Valdez Taubas J (2021) A novel yeast-based high-throughput method for the identification of protein palmitoylation inhibitors. *Open Biol* 11(8):200415
55. Ocasio CA, Baggelaar MP, Siphthorp J, Losada de la Lastra A, Tavares M, Volarić J, Soudy C, Storck EM, Houghton JW, Palma-Duran SA, MacRae JI, Tomić G, Carr L, Downward J, Eggert US, Tate EW (2024) A palmitoyl transferase chemical-genetic system to map ZDHHC-specific S-acylation. *Nat Biotechnol* 42:1548. <https://doi.org/10.1038/s41587-023-02030-0>
56. Takahashi M, Chong HB, Zhang S, Yang T-Y, Lazarov MJ, Harry S, Maynard M, Hilbert B, White RD, Murrey HE (2024) DrugMap: a quantitative pan-cancer analysis of cysteine ligandability. *Cell* 187(10):2536–2556. e2530

INDEX

A

- Acrylamide 5, 10, 34, 45, 53, 68, 78, 88, 183, 184, 189, 191, 298, 304, 332, 365
- Activity-based probes (ABPs) 1, 2, 4, 58–60, 63, 69, 119–135, 139–153, 181, 182, 184, 188–190, 196, 217–229, 331, 332, 365
- Activity-based protein profiling (ABPP) 1–27, 58–61, 63–66, 68, 69, 120–123, 128, 129, 131, 132, 139–142, 145–150, 152, 155–178, 181, 182, 199, 202, 203, 208, 209, 212, 213, 217, 294–296, 309, 312, 320, 331–343, 346, 365
- Activity proteomics 2, 4
- Anticancer drugs 265, 268

B

- Bacteria 58–70, 141, 150, 294, 295, 300, 303, 304, 311
- Binding site resolution 155–178
- Bioprocess development 201–215

C

- C12FDG 257–260, 263
- Cancer 1, 93, 217, 232, 265, 345, 362
- Chemical biology 1
- Chemical probes 120, 203, 204, 213, 214, 218, 219, 232, 346, 364
- Chemical proteomics 120, 156, 231–252, 266, 367
- Chemoproteomics 58, 73–90, 131, 320, 331
- Chromatin association 74–76, 79, 84
- Click-chemistry 58, 59, 123, 125, 129, 134, 140, 182, 189, 190, 218, 221, 226, 232, 234, 235, 239, 240, 250, 333–335, 340, 366
- Covalent kinase probe 250
- Covalent small molecule 319
- Cysteine 1, 2, 4, 25, 52, 74, 87, 89, 265–272, 319–329, 331–343, 361, 362, 367
- Cysteine profiling 320, 332, 334–341
- Cysteinome 321, 332

D

- Data integration 95, 107

- DeISGylating enzyme 1–27
- DeISGylation 2
- DIA-PASEF MS-analysis 276
- Diphenyl phosphonate (DPP) 218–223, 227
- Diyne probes 294–296, 299, 300, 308–310
- Drug discovery 122, 123, 266, 276, 319–329
- Drug selectivity 331

E

- EC₅₀ 246
- E3 ligase 345–358
- Excipient stability 202
- Exosomes 255–263

F

- Fluorescent imaging 226
- Fluorophosphonate 60, 122, 127, 134, 141, 143, 146, 150, 181, 204

G

- Genetic diseases 31–54

H

- Hepatic lipase (LIPC) 183
- High-throughput 2–4, 7, 8, 16–20, 23, 105, 121, 123, 364–366

I

- In-gel fluorescence 132, 182, 183, 186, 188, 189, 194, 221, 226, 227
- In situ target engagement 231–252
- Inhibitors 1–27, 34, 41, 50, 53, 58, 60, 63, 68, 69, 119–123, 128, 129, 131, 132, 134, 135, 139–153, 155, 156, 168, 170, 174, 181, 183, 184, 195, 196, 218, 226, 232–235, 238, 239, 244, 246–250, 257, 267, 279, 280, 322, 323, 365–367
- Interferon-stimulated gene 15 (ISG15) 1, 2, 4, 9, 17

K

- Kinome profiling 231

L

LC-MS 34, 36, 37, 52, 65,
66, 80, 85, 133, 161, 176, 177, 186, 202, 204,
211, 243, 275, 322, 324, 327

M

Mass spectrometry (MS) 2, 32, 58, 74,
94, 131, 141, 176, 183, 204, 219, 244, 276, 295,
321, 332, 346
Metabolomics 32, 33, 36–40
Microplates 123, 124, 128, 129,
131, 134, 240, 251
MicroRNA (miRNA) 182–186, 188, 195
Monooxygenases 293–296, 299–303,
306, 308–310
Multi-omics 31–54, 94, 95, 98, 105–107
Mycobacteria 139–153

N

Near-infrared (NIR) 295, 298, 300,
302–304, 307, 309, 311–313, 316
Nuclear cysteine ligandability 74
Nuclear localization 73–90

P

Palmitoylation 361–367
Pharmacology 1
Phosphopeptide enrichment 275, 276,
278, 281, 288
Phosphoproteomics 275–290, 346
Phosphorylation events 276, 278, 284, 287
Plate-based assays 121, 132
Post-translational modifications (PTMs) 32,
181, 266, 345, 346, 367
Preadipocytes 255–263
Product stability 202
Protein kinases 231, 276, 362
Proteogenomics 93–107
Proteomics 23, 32, 74, 94,
133, 139, 157, 202, 219, 251, 269, 275, 294,
320, 331, 346, 364
Proximity labeling 74–76
Pulldown 197, 198

Q

Quantitative proteomics 277, 345–358

R

Reactive oxygen species (ROS) 265–272, 361

S

SARS-CoV-2 276, 278, 280, 288, 362
Screening 1–27, 74, 77, 82,
120, 121, 123, 131, 132, 157, 159, 165–170,
223, 266, 331, 332, 348, 365, 366
Senescence 255, 256, 263
Serine hydrolases 58–70, 120–122,
140–142, 150, 181, 183, 196, 201–215, 247
Serine proteases 202, 203, 217–229
Sodium dodecyl sulfate-polyacrylamide gel
electrophoresis (SDS-PAGE) 9, 10, 12,
58, 60–65, 69, 78, 82, 88, 140, 143–147, 149,
150, 157, 183, 184, 221, 226, 251, 295, 298,
300, 302–305, 307, 309, 312–316
Stable isotope labeling of amino acid in cell culture
(SILAC) 141, 145, 149,
151, 153, 343, 353
Subcellular localization 87, 141,
158, 159, 171, 172

T

Target identification 140, 141, 156,
157, 332, 340–342
Target protein purification 172–175, 201
Tissue proteomes 319–329
TMTpro 321, 322, 325, 329
Tuberculosis 141, 142, 145
TurboID 74

U

Ubiquitination 345–349, 351, 353
Ubiquitin E3 ligase activity profiling analysis (Ube3-
APA) 347–352, 354, 355, 358

V

Vascular normalization 155–178
Viral infections 182, 183, 186, 275–290
Virus 182, 185, 186,
195, 203, 276–279, 289, 362
Virus–host interactions 181, 199

W

Western blotting 3, 4, 10, 24,
27, 65, 69, 78, 182–184, 197, 261, 346

X

XO44 232–234, 238, 239, 246–248, 250