

**UCLA**

**UCLA Electronic Theses and Dissertations**

**Title**

The Development and Application of Proteomic Approaches to Assess Consequences of Amino Acid Modifications Within the Proteome

**Permalink**

<https://escholarship.org/uc/item/9v99m7t4>

**ISBN**

9798277475522

**Author**

Shikwana, Flowreen

**Publication Date**

2026-03-20

**Supplemental Material**

<https://escholarship.org/uc/item/9v99m7t4#supplemental>

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

The Development and Application of Proteomic Approaches to Assess  
Consequences of Amino Acid Modifications Within the Proteome

A dissertation submitted in partial satisfaction  
of the requirements for the degree Doctor of  
Philosophy in Biochemistry, Molecular and Structural Biology

by

Flowreen Shikwana

2026



© Copyright by

Flowreen Shikwana

2026

## ABSTRACT OF THE DISSERTATION

The Development and Application of Proteomic Approaches to Assess  
Consequences of Amino Acid Modifications Within the Proteome

by

Flowreen Shikwana

Doctor of Philosophy in Biochemistry, Molecular and Structural Biology

University of California, Los Angeles, 2026

Professor Keriann M. Backus, Chair

A central challenge of biochemistry is determining the consequences of modifications to proteinaceous amino acids as they pertain to protein function and broader cellular function. Mass spectrometry (MS)-based proteomics enables proteome-wide identification and quantification of modified peptides either through chemical modifications to amino acids such as covalent probe addition or even genetic modifications such as single amino acid variants (SAAVs). Chemical modifications are often assayed through chemoproteomic workflows in order to identify specific ligandable sites by covalent probes designed to be reactive towards certain amino acids, such as cysteine residues in the case of cysteine chemoproteomics. While chemoproteomic workflows are great tools to discover targetable sites by many different types of covalent probes, the scope for which the direct and indirect effects of liganding specific cysteines contribute to reported modes-of-action remains under-characterized. Knowledge of such effects are especially important for prioritizing compounds for follow up studies when performing high throughput screens with multiple scout fragments. Similarly, when using MS-based proteomics for the study

of genetic modifications to amino acids, proteogenomic workflows often only address the identification of modifications such as SAAVs and their relative abundance within cells but fail to stratify or establish functional significance of each of these variants. In order to determine functional significance, follow up biochemical studies are often needed which is a time-intensive and low-throughput process. In this work, we first address improving the throughput of cysteine chemoproteomics through the adaptation of the single-pot solid-phase-enhanced sample preparation (SP3) workflow paired with our isotopic capture reagents for cysteine chemoproteomic profiling. We scaled the workflow to accommodate a 96-well plate sample preparation and in-plate peptide enrichment, along with identifying low-cost magnetic beads capable of comparable cysteine coverage, which allowed us to profile 12,265 and identify 2633 ligandable cysteines. We also demonstrated its use for screening multiple compounds, including atropisomers, at several concentrations to better determine unique cysteine and protein targets of such molecules. Next, we investigated a panel of cysteine reactive fragments in order to delineate observed degrader effects on our protein of interest, the SARS-CoV-2 non-structural protein 14 (nsp14), from proteome-wide proteostasis dysregulation. We found that direct covalent modification, ubiquitylation of nsp14, as well as the inactivation of protein disulfide isomerases (PDIs) 3 and 6 are necessary for nsp14 depletion.

In addition, we implicated cysteine-reactive electrophiles, used at screening concentrations, in generalized ubiquitylation, proteasome activation, widespread aggregation, and the formation of stress granules upon cysteine-reactive electrophile screening. These unexpected effects on cellular proteostasis led us to develop a framework for identifying wide ranging cellular effects that should be monitored alongside screens which use protein depletion as a readout for identifying cysteine-reactive degrader compounds. Turning to SAAVs, we sought to use thermal proteome profiling (TPP) as a high-throughput proteomic methodology to stratify variants of functional significance as read out by shifts in thermal stability. We first established the use of TPP to characterize alterations in thermal stability between wild-type (WT) and variant

proteoforms using B-Raf and its more well-known SAAVs as a model system for this platform. We validated the use of TPP both at the protein and peptide level for assaying variant thermal stability first within an overexpression system, then within a pooled B-Raf variant cell line. Finally, we extended our variant TPP methodology to quantify 77 endogenous variants in Molt-4 leukemia cells through proteogenomic workflows and identify stabilizing and destabilizing SAAVs.

The dissertation of Flowreen Shikwana is approved.

Joseph A. Loo

Margot E. Quinlan

James A. Wohlschlegel

Keriann M. Backus, Committee Chair

University of California, Los Angeles

2026

## DEDICATION

*To*

*My parents, Farid and Linda,*

*My sister, Rebecca,*

*My husband Jake,*

*And all of the friends and family*

*who have supported me along my life adventures*

## TABLE OF CONTENTS

|                                                                                                                                     |     |
|-------------------------------------------------------------------------------------------------------------------------------------|-----|
| LIST OF FIGURES.....                                                                                                                | ix  |
| LIST OF SCHEMES.....                                                                                                                | xi  |
| ACKNOWLEDGEMENTS.....                                                                                                               | xii |
| CIRRICULUM VITAE.....                                                                                                               | xiv |
| Chapter 1: Introduction.....                                                                                                        | 1   |
| References.....                                                                                                                     | 7   |
| Chapter 2: CySP3-96 Enables Scalable, Streamlined, and Low-Cost Sample Preparation for<br>Cysteine Chemoproteomic Applications..... | 14  |
| Abstract.....                                                                                                                       | 15  |
| Introduction.....                                                                                                                   | 16  |
| Experimental Section.....                                                                                                           | 19  |
| Results.....                                                                                                                        | 29  |
| Discussion.....                                                                                                                     | 47  |
| References.....                                                                                                                     | 50  |
| Supporting Information.....                                                                                                         | 61  |
| Chapter 3: Delineating Cysteine-Reactive Compound Modulation of Cellular Proteostasis<br>Processes.....                             | 119 |
| Abstract.....                                                                                                                       | 120 |
| Introduction.....                                                                                                                   | 120 |
| Results.....                                                                                                                        | 122 |
| Discussion.....                                                                                                                     | 145 |
| Experimental Section.....                                                                                                           | 157 |
| References.....                                                                                                                     | 178 |
| Supporting Information.....                                                                                                         | 186 |

|                                                                                     |     |
|-------------------------------------------------------------------------------------|-----|
| Chapter 4: Assessing opportunities and challenges in proteogenomic thermal proteome |     |
| profiling.....                                                                      | 253 |
| Abstract.....                                                                       | 253 |
| Introduction.....                                                                   | 254 |
| Results.....                                                                        | 259 |
| Discussion.....                                                                     | 281 |
| References.....                                                                     | 283 |
| Supporting Information.....                                                         | 289 |
| Experimental Section.....                                                           | 305 |
| Chapter 5: Conclusion.....                                                          | 314 |



## LIST OF FIGURES

### Chapter 2:

|                                                                                                                                                                        |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1. SP3 peptide cleanup volume scaledown results in overall decreased coverage.....                                                                              | 30 |
| Figure 2. Skipping peptide cleanup results in comparable sample coverage and a more streamlined sample enrichment.....                                                 | 33 |
| Figure 3. Streamlined SP3 preparation seamlessly transitions into 96 deep well plate format for screening of cysteine reactive electrophiles.....                      | 40 |
| Figure 4. Concentration screen of two atropisomers BEH-1-( <i>S<sub>a</sub></i> ) and BEH-1-( <i>R<sub>a</sub></i> ) illustrates stereoselectivity across targets..... | 43 |
| Figure 5. Identifying a low-cost magnetic option for SP3.....                                                                                                          | 46 |

### Chapter 3:

|                                                                                                                                                              |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 1. Chemoproteomic profiling of SARS-CoV-2 nonstructural proteins with electrophilic scout fragments identifies JC19 as a putative nsp14 degrader..... | 124 |
| Figure 2. Both the chloroacetamide and sulfonyl fluoride electrophiles are required for the rapid depletion of nsp14.....                                    | 126 |
| Figure 3. JC19-induced nsp14 degradation is mediated by ubiquitylation and the proteasome.....                                                               | 129 |
| Figure 4. Direct labeling of nsp14 and the protein disulfide isomerases (PDIs) potentiates nsp14 degradation.....                                            | 134 |
| Figure 5. Multimodal proteomics implicates global rewiring of the proteome by electrophilic compounds.....                                                   | 139 |
| Figure 6. Cysteine-reactive electrophilic fragments induce global proteome remodeling through protein aggregation and stress granule formation.....          | 144 |
| Extended Data Figure 1. Compound structures.....                                                                                                             | 148 |
| Extended Data Figure 2. JC19 analogs have varying degrees of nsp14 depletion efficiency...                                                                   | 149 |

|                                                                                                                                             |     |
|---------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Extended Data Figure 3. JC19 induces linkage of predominantly K48-linked polyubiquitin chains on nsp14.....                                 | 150 |
| Extended Data Figure 4. Common cell stress-sensing pathways are not responsible for nsp14 depletion.....                                    | 151 |
| Extended Data Figure 5. PDIA6-His pulldown.....                                                                                             | 152 |
| Extended Data Figure 6. High-ratio cysteines as identified by isoTOP–ABPP belong to proteins spanning various subcellular compartments..... | 153 |
| Extended Data Figure 7. Delineation of aggregation and degradation properties of JC19, EN450 and NB001.....                                 | 154 |
| Extended Data Figure 8. Various cysteine-reactive electrophiles induce stress granule formation.....                                        | 155 |
| Extended Data Figure 9. Genetic knockdown of PDIA3 and PDIA6 lead to increases in aggregated proteins (aggresomes).....                     | 156 |
| Extended Data Figure 10. Recommendations for best practices.....                                                                            | 157 |
| <b>Chapter 4:</b>                                                                                                                           |     |
| Figure 1. Summary of TPP workflow for variant TPP.....                                                                                      | 261 |
| Figure 2. Blot-based CETSA analysis for individually overexpressed B-Raf SAAVs.....                                                         | 264 |
| Figure 3. Establishing variant TPP analysis with model B-Raf variants.....                                                                  | 269 |
| Figure 4. Peptide $T_m$ mapped to active and inactive B-Raf states in 3D highlights variant-induced stability changes.....                  | 272 |
| Figure 5. Variant TPP analyses of a pooled B-Raf COSMIC variant cell line.....                                                              | 277 |
| Figure 6. Variant TPP analyses within the Molt-4 cell line.....                                                                             | 280 |

## LIST OF SCHEMES

### **Chapter 2:**

|                                                     |    |
|-----------------------------------------------------|----|
| Scheme 1. Synthesis of atropisomeric compounds..... | 36 |
|-----------------------------------------------------|----|

## ACKNOWLEDGEMENTS

First and foremost, I would like to thank my parents, Farid and Linda, for leaving behind a life they knew in exchange for an unknown one in pursuit of a better and safer future for my sister and me. Without everything they have done and sacrificed, I would not have had the wonderful opportunities which led me here. My parents and my sister Rebecca have supported me throughout my entire academic journey and encouraged me to work hard while also reminding me to enjoy and live life along the way. For that I am forever grateful.

I would also like to extend my love and gratitude to my wonderful husband, Jake, who really supported me and kept me going all throughout my scientific training (and would sometimes label some of my sample vials just to be helpful). From my time as a postbac, on the opposite sides of the country, to starting and finishing grad school, he believed in me and was so incredibly patient in this long journey. Thank you for always keeping me leveled throughout the ups and downs of grad school and always reminded me of what my priorities in life were.

I'd also like to thank my undergraduate mentors, Dr. Jessica Bell and Dr. J. Ellis Bell, for getting me excited about research and especially Dr. Jessica Bell for helping me obtain amazing research opportunities as a postbac – I'm not sure I would have gotten here without her.

To my PI and PhD advisor, Dr. Keri Backus, thank you for giving me the opportunity to become the scientist I am today. For allowing me to explore my interests, learn proteomics, and chase interesting problems. Thank you for pushing me to become a better scientist, to write more, and to be more rigorous.

To all of the Backus lab members, previous & current, that I have had the pleasure to meet and work with throughout my time within the lab: Maria Palafox, Jian Cao, Jose Castellon, Sunny Yan, Heta Desai, Lisa Boatner, Ashley Julio, Nik Burton, Miranda Villanueva, Andrew Becker, Alex Turmon, Eli Bilech, Nithesh Perumal, Halime Yilmaz, Adi Sivakumar, Hannah Goldberg, Katrina Andrews, Chau Ngo, Sho Takechi, Sam Ofori, Lara Holoidovsky, and Alex Sun. Thank you all for being the best colleagues, always having you all to bounce ideas off of or talk through science or

even tolerate my chaos. Thank you all for your friendship — which was absolutely essential throughout the long days of experiments or LC/mass spec maintenance! I don't think I could've asked for a better lab environment throughout my PhD — the vibes were always amazing because of you all! Special thank you to Cindy Truong, my undergraduate mentee, who helped me and taught me a lot when it came to computational workflows.

To my committee members and all who have served on my committee: Joseph Loo, James Wohlshlegel, Margot Quinlan, and Todd Yates, for all of your support, advice, and contributions throughout my PhD, from my oral qualifying exam, throughout my yearly committee meetings, and through the end of my PhD. I'd also like to thank the UCLA grad office and the biological chemistry department for all their support throughout the PhD program. I'd also like to thank the UCLA department of graduate education for their financial support through the Dissertation Year Award (2024-2025).

And finally, to myself, for choosing to follow all the things that bring me joy and that I love — which have led me to chase science I am passionate about while surrounding myself with the people I love. Without all that I wouldn't have finished my PhD without losing myself or my sanity.

Chapter 2 is a version of Shikwana, F., Heydari, B., Ofori, S., Truong, C., Turmon, A., Darrouj, J., Holoidovsky, L., Gustafson, J. & Backus, K.M. CySP3-96 enables scalable, streamlined, and low-cost sample preparation for cysteine chemoproteomic applications. *Mol. Cell. Proteomics* 100898 (2024). doi:10.1016/j.mcpro.2024.100898. Reprinted with permission. Copyright © 2024 Elsevier Inc on behalf of American Society for Biochemistry and Molecular Biology

Chapter 3 is a version Shikwana, F.\*, Julio, A. R.\*, Truong, C., Burton, N. R., Dominguez, E. R., Turmon, A. C., Cao, J. & Backus, K. M. Delineating cysteine-reactive compound modulation of cellular proteostasis processes. *Nat. Chem. Biol.* (2024). doi:10.1038/s41589-024-01760-9. of Reprinted with permission. Copyright © 2024 Springer Nature America, Inc.

Chapter 4 is a version of Shikwana, F.; Backus, K.M. Assessing opportunities and challenges in proteogenomic thermal proteome profiling. **2026** Manuscript in preparation.

## CIRRICULUM VITAE

### EDUCATION

Master | UNIVERSITY OF CALIFORNIA, LOS ANGELES | 2020.09-2022.03  
Master of Science in the Biochemistry, Molecular, and Structural Biology

Undergraduate | UNIVERSITY OF SAN DIEGO | 2014.09-2018.05  
Bachelor of Arts in Biochemistry, *magna cum laude*

### PUBLICATION

CySP3-96 enables scalable, streamlined, and low-cost sample preparation for cysteine chemoproteomic applications.

**Shikwana, F.**, Heydari, B., Ofori, S., Truong, C., Turmon, A., Darrouj, J., Holoidovsky, L., Gustafson, J. & Backus, K.M. *Mol. Cell. Proteomics* 100898 (2024).  
doi:10.1016/j.mcpro.2024.100898

Delineating cysteine-reactive compound modulation of cellular proteostasis processes.

**Shikwana, F.\***, Julio, A. R.\*, Truong, C., Burton, N. R., Dominguez, E. R., Turmon, A. C., Cao, J. & Backus, K. M. (\*contributed equally) *Nat. Chem. Biol.* (2024). doi:10.1038/s41589-024-01760-9

CIAA: Integrated Proteomics and Structural Modeling for Understanding Cysteine Reactivity with Iodoacetamide Alkyne.

Boatner, L. M., Eberhardt, J., **Shikwana, F.**, Holcomb, M., Lee, P., Houk, K. N., Forli, S. & Backus, K. M. *ACS Chem. Biol.* (2025). doi:10.1021/acschembio.5c00225

Chemoproteogenomic stratification of the missense variant cysteinome.

Desai, H., Andrews, K. H., Bergersen, K. V., Ofori, S., Yu, F., **Shikwana, F.**, Arbing, M. A., Boatner, L. M., Villanueva, M., Ung, N., Reed, E. F., Nesvizhskii, A. I. & Backus, K. M. *Nat. Commun.* **15**, 9284 (2024).

Generating Cysteine-Trypsin Cleavage Sites with 2-Chloroacetamide Capping.

Ofori, S., Desai, H. S., **Shikwana, F.**, Boatner, L. M., Dominguez, E., Castellón, J. O. & Backus, K. M. *Chem. Commun.* (2024). doi:10.1039/D4CC01583E

Solid-Phase Compatible Silane-Based Cleavable Linker Enables Custom Isobaric Quantitative Chemoproteomics.

Burton, N. R., Polasky, D. A., **Shikwana, F.**, Ofori, S., Yan, T., Geiszler, D. J., Veiga Leprevost, F. da, Nesvizhskii, A. I. & Backus, K. M. *J. Am. Chem. Soc.* **145**, 21303–21318 (2023).

Bivalent recognition of fatty acyl-CoA by a human integral membrane palmitoyltransferase.

Lee, C. J., Stix, R., Rana, M. S., **Shikwana, F.**, Murphy, R. E., Ghirlando, R., Faraldo-Gómez, J. D., & Banerjee, A. (2022). *Proceedings of the National Academy of Sciences of the United States of America*, 119(7).

Predicting and validating a model of suppressor of IKKepsilon through biophysical characterization.

Machek, M. L., Sonnenschein, H. A., Graham, S. I., **Shikwana, F.**, Kim, S. L., Garcia DuBar, S., Minzer, I. D., Wey, R. and Bell, J. K. (2019). *Protein Science*, 28: 1423-1436.  
doi:10.1002/pro.3640

## CONFERENCES

2025 73<sup>rd</sup> ASMS annual conference poster presenter  
American Society for Mass Spectrometry, Baltimore, MD

2024 72<sup>nd</sup> ASMS annual conference poster presenter  
American Society for Mass Spectrometry, Anaheim, CA

2023 71<sup>st</sup> ASMS annual conference poster presenter  
American Society for Mass Spectrometry, Houston, TX

2018 ASBMB Undergraduate poster competition poster presenter  
American Society for Biochemistry and Molecular Biology, San Diego, CA

## AWARDS

2025 ASMS Graduate Student Travel Award

2024 Dissertation Year Award

2018 University of San Diego Research Excellence Award

## Chapter 1: Introduction

Assigning functionality to proteins and amino acid residues under various protein states is a central goal of biochemistry. Traditionally, studying protein modifications under distinct states has been a relatively low-throughput process, focusing on individual proteins rather than conducting proteome-wide measurements. Over the last few decades, advances in protein mass spectrometry have allowed for the field of proteomics to study the proteome in a sensitive, quantitative, and high throughput manner. There are many applications of bottom-up mass spectrometry-based proteomics including bulk label-free proteomics to monitor relative protein abundance within the cell, proteomics to detect and quantify post-translational modifications, chemoproteomics to identify and quantify small chemical modifications, proteogenomics to detect protein variants throughout the proteome, and many more. This work centers around the use of mass spectrometry-based proteomics to identify functional consequences of amino acid modifications, either through chemical modification or through genetic modification. We first address the chemoproteomic subset of proteomics with the advancement of cysteine chemoproteomic screening in Chapter 2 and assessment of non-specific cellular effects in response to cysteine reactive compound treatment in Chapter 3. This work then broadens out to apply proteomic methodology to identify and functionally stratify genetic amino acid modifications that occur within the proteome in Chapter 4.

Chemoproteomics is used to study a wide range of nucleophilic amino acids such as cysteine<sup>1-3</sup>, lysine<sup>4,5</sup>, serine<sup>6-8</sup>, tyrosine<sup>9-11</sup>, methionine<sup>12,13</sup>, glutamate, and aspartate<sup>14-16</sup> that can be targeted by small electrophilic molecules. In the quest towards drugging the yet to be drugged portion of the proteome, cysteine is a favorable residue to target mainly due to its multi-faceted roles which contribute to protein structure, through disulfide bonds and coordination of metal binding sites<sup>17</sup>, as well as protein function, which involves cysteine in catalysis, redox sensing, and post-translational modifications (PTMs)<sup>18</sup>. The development of cysteine-reactive drugs has shown to



produce successful blockbuster drugs such as the Gly12Cys GTPase KRas inhibitors<sup>19</sup> and the SARS-CoV-2 protease inhibitor nirmatrelvir<sup>20</sup>. Cysteine chemoproteomics is typically the standard for determining the ligandability of cysteines by certain drug molecules. Workflows such as isotopic Tandem Orthogonal Proteolysis – Activity-Based Protein Profiling (isoTOP-ABPP)<sup>3</sup> allow for monitoring cysteine ligandability through competition<sup>1</sup>. A competition ratio (CR) can be measured for each cysteine by competing the cysteine-reactive molecule of interest with a pan-reactive probe, such as an iodoacetamide alkyne (IAA), or a matched alkynylated probe, which can then accommodate an isotopically differentiated enrichable handle<sup>3</sup> such as an azido-biotin through copper-catalyzed azide-alkyne cycloaddition (CuACC)<sup>21</sup>. Control samples not labeled with the cysteine-reactive molecule of interest are also labeled with IAA, followed by CuACC to a different isotopically differentiated azido-biotin, then combined with the labeled sample and enriched on avidin resin. The differences in abundance for each peptide with the isotopically differentiated tag can be measured via mass spectrometry across thousands of peptides which is a powerful way to determine specificity or off-target liganding events proteome-wide.

Discovering new drugs and cysteine targets requires the screening of multiple scout compounds across multiple chemotypes against the entire proteome. Proteomic workflows and instrumentation allow for the assessment of the whole proteome in one experiment; however, screening multiple chemotypes and compounds remains a bottleneck due to the number of samples that need to be taken through sample preparation as well as the time needed for sample acquisition. Additionally, many of these compounds are screened at singular or high/low mid-micromolar concentrations which may have unintended effects<sup>22,23</sup>. Even at later stages, screening compounds with more elaborated chemical structures may require an expanded concentration range for better understanding compound specificity before moving into more time-consuming mechanistic studies of these compounds. Various proteomics plate-based workflows have been developed to address issues of throughput for both labeled and unlabeled samples. These methods incorporate various proteomic clean up techniques such as in-stage tip<sup>24</sup>, S-

Trap<sup>25</sup>, and single-pot, solid-phase-enhanced, sample preparation (SP3)<sup>26,27</sup>. Each of these plate-based approaches have their strengths for different types of sample input. SP3 has been shown to be useful for chemoproteomic workflows, as illustrated by earlier work from our lab<sup>28</sup> as well as TMT- and plate-based chemoproteomic platform developed by the Gygi lab<sup>29</sup>. Previous plate-based cleanup methods focused on lower sample inputs<sup>30</sup> than is typically required for cysteine chemoproteomics and/or utilized additional or different reagents than would be utilized in a non-plate-based workflow, often making it more difficult or expensive to transition to a higher throughput platform.

In Chapter 2 we develop a higher throughput workflow for cysteine chemoproteomics, termed CySP3-96, in which up to 96 compounds can be screened within one deep-well plate. The CySP3-96 platform is compatible with multiple cysteine capture reagents, either custom or commercially available, allowing for a more seamless transition for chemoproteomic workflows. Additionally, all the proteomic sample preparation steps can be fully done within a plate, including cysteine enrichment at the peptide level. We utilized the CySP3-96 workflow to screen 16 different compounds, including concentration ranges for two electrophilic atropisomer compounds, which allowed us to identify 2633 ligandable cysteines in aggregate. For the atropisomer compounds we were able to observe stereoselective and concentration dependent labeling across 66 cysteines. We also establish that the CySP3-96 workflow is compatible with lower cost generic magnetic DNA cleanup beads that can be used in place of Cytiva SP3 beads, allowing for reduced costs with the sample preparation scale up, and without reducing the number or types of proteins or peptides captured.

While covalent compound screening provides information about residue-level labeling, modifying cysteines throughout the proteome can have a wide range of effects depending on the functionality of that cysteine. Typically, the intended effect for probe/drug development is that the modification's effects are specific to the protein and or cysteine being targeted. Often, efficient screening strategies for scout fragments rely on prioritizing compounds by phenotypic readouts

such as protein depletion or attenuation of protein activity, usually for certain prioritized protein targets. In the context of potential covalent degrader molecules, depletion of a protein of interest (POI) via direct labeling and involvement of the proteasome may not be the full story. Investigating non-specific cellular responses to cysteine labeling events<sup>22,23</sup> is useful when assessing specificity of a compound at multiple concentrations or even for identifying other targetable proteins that may be of interest in the future. Within this work, we explore these generalized effects of cysteine reactive covalent probes through a focused case study with the SARS-CoV-2 nonstructural protein 14 (nsp14) and cysteine reactive compound JC19. We initially observed the depletion of nsp14 upon JC19 treatment which was dependent on direct labeling of C11 and C272. However, upon further investigation, it appeared that more than one mechanism could be responsible for nsp14 depletion. While direct labeling of nsp14 itself was required for depletion by JC19, inactivation of protein disulfide isomerases (PDIs) 3 and 6—presumably by JC19 labeling—were also responsible for the depletion of nsp14. We also observed more generalized effects such as formation of stress granules and aggresome structures upon treatment with JC19 as well as more broadly with mid-micromolar concentrations of various cysteine reactive compounds. Together, Chapter 3 provided a framework of best practices for identifying ways to monitor these cell stress responses when screening cysteine reactive compounds and elucidating their effects on cell biology.

Turning from chemical modifications to genetic modifications of amino acids, mass spectrometry is a versatile tool that is also used to identify single amino acid variants (SAAVs) within protein samples through proteogenomics workflows<sup>31</sup>. SAAVs are usually first identified at genomic or transcriptomic level of analyses; however, since there is often low correlation between the global transcriptome and the proteome, proteomic analyses are needed to determine if the SAAVs become functional proteins within the cell. Typically, proteomic analyses utilize search databases that consist of canonical protein sequences which overlook SAAVs, splice variants, and various other non-canonical protein isoforms. In order to identify these non-canonical proteins, variant

databases are made based on variants detected within certain cell lines. This data can come from accessible public databases such as DepMap<sup>32</sup> or the Catalogue of Somatic Mutations in Cancer (COSMIC)<sup>33</sup>, which contain previously identified variants categorized by cell line or cell type. Alternatively, custom variant protein databases can be generated by using data from whole genome sequences (WGS) and transcriptome sequencing of the cell line of interest. To control for the standard 1% false discovery rate (FDR) within proteomic searches<sup>34</sup>, these datasets are first searched against the canonical protein database, as the canonical sequences are more abundant and more likely to match the majority of peptide spectra. The spectra that do not match against the canonical database then get searched in a second pass against the variant database to identify and quantify variant peptides. While proteogenomics allows for the identification of variant protein sequences through mass spectrometry, the functional relevance or effect of those variant proteins cannot be understood or even stratified within this context. The traditional, and most comprehensive, methods to determine functional relevance of protein variants are in depth biochemical studies. While these types of in-depth studies are eventually required for variants of interest, it remains difficult to prioritize which variants should have dedicated follow up studies. Computational tools such as SIFT<sup>35</sup>, PolyPhen<sup>36</sup>, and CADD<sup>37</sup> have been developed to computationally predict pathogenicity of amino acid substitutions within proteins. Despite these tools, there still exist variants that have ambiguous classifications—these are often termed variants of unknown significance (VUS)<sup>38</sup>. Mass spectrometry and proteogenomics offers a way to profile these variants in a higher throughput manner and help stratify their relevance to disease. Proteomic methods such as thermal proteome profiling (TPP)<sup>39</sup> allow for proteome-wide measurement of gain or loss of interactions that is measured through protein thermal stability changes. TPP has mainly been used to assess drug binding across the entire proteome, however, it has also been deployed to assess protein-protein interactions, protein-nucleic acid binding, protein-metabolite binding, and even effects of PTMs such as phosphorylation<sup>40</sup>.

In Chapter 4 we leverage the TPP workflow to interrogate protein SAAVs by assessing changes in thermal stability between WT and SAAV peptides. We first establish that changes in protein thermal stability can be detected between WT and variant proteins using a model system of the serine/threonine protein-kinase B-Raf and a few of its more well characterized variants: class I V600E, class II L597Q, and class III N581S. This B-Raf model system was analyzed with blot-based cellular thermal shift assays (CETSA<sup>s</sup>)<sup>41</sup> as well as mass spectrometry-based TPP analysis. We found that subtle differences can be detected for certain variants and that peptide level TPP analysis is representative of the protein level TPP analysis. We then moved into a more complex system of pooled B-Raf COSMIC variants to deploy our variant TPP workflow on a system more representative of endogenous variants within cells. Of the 47 B-Raf variants we detected and 19 that we quantified, 2 variants appeared to be stabilized and 2 variants exhibited destabilization. We finally apply variant TPP to Molt-4 leukemia cells with a custom variant database based on previously generated WGS and transcriptome data to identify 77 variants. We were able to quantify 56 of those variants through peptide-level comparisons and peptide-protein level comparisons to find destabilizing variants and 4 stabilizing variants. Together these data gave insight into the information that can be obtained from variant TPP analysis as well as highlighted some of the challenges towards improving variant quantification through mass spectrometry proteogenomic analyses.

## Chapter 1 References:

1. Backus, K. M., Correia, B. E., Lum, K. M., Forli, S., Horning, B. D., González-Páez, G. E., Chatterjee, S., Lanning, B. R., Teijaro, J. R., Olson, A. J., Wolan, D. W. & Cravatt, B. F. Proteome-wide covalent ligand discovery in native biological systems. *Nature* **534**, 570–574 (2016).
2. Takahashi, M., Chong, H. B., Zhang, S., Yang, T.-Y., Lazarov, M. J., Harry, S., Maynard, M., Hilbert, B., White, R. D., Murrey, H. E., Tsou, C.-C., Vordermark, K., Assaad, J., Gohar, M., Dürr, B. R., Richter, M., Patel, H., Kryukov, G., Brooijmans, N., Alghali, A. S. O., Rubio, K., Villanueva, A., Zhang, J., Ge, M., Makram, F., Griesshaber, H., Harrison, D., Koglin, A.-S., Ojeda, S., Karakyriakou, B., Healy, A., Popoola, G., Rachmin, I., Khandelwal, N., Neil, J. R., Tien, P.-C., Chen, N., Hosp, T., van den Ouweland, S., Hara, T., Bussema, L., Dong, R., Shi, L., Rasmussen, M. Q., Domingues, A. C., Lawless, A., Fang, J., Yoda, S., Nguyen, L. P., Reeves, S. M., Wakefield, F. N., Acker, A., Clark, S. E., Dubash, T., Kastanos, J., Oh, E., Fisher, D. E., Maheswaran, S., Haber, D. A., Boland, G. M., Sade-Feldman, M., Jenkins, R. W., Hata, A. N., Bardeesy, N. M., Suvà, M. L., Martin, B. R., Liao, B. B., Ott, C. J., Rivera, M. N., Lawrence, M. S. & Bar-Peled, L. DrugMap: A quantitative pan-cancer analysis of cysteine ligandability. *Cell* **187**, 2536-2556.e30 (2024).
3. Weerapana, E., Wang, C., Simon, G. M., Richter, F., Khare, S., Dillon, M. B. D., Bachovchin, D. A., Mowen, K., Baker, D. & Cravatt, B. F. Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **468**, 790–795 (2010).
4. Abbasov, M. E., Kavanagh, M. E., Ichu, T.-A., Lazear, M. R., Tao, Y., Crowley, V. M., Am Ende, C. W., Hacker, S. M., Ho, J., Dix, M. M., Suci, R., Hayward, M. M., Kiessling, M.

- L. L. & Cravatt, B. F. A proteome-wide atlas of lysine-reactive chemistry. *Nat. Chem.* **13**, 1081–1092 (2021).
5. Hacker, S. M., Backus, K. M., Lazear, M. R., Forli, S., Correia, B. E. & Cravatt, B. F. Global profiling of lysine reactivity and ligandability in the human proteome. *Nat. Chem.* **9**, 1181–1190 (2017).
6. Simon, G. M. & Cravatt, B. F. Activity-based proteomics of enzyme superfamilies: serine hydrolases as a case study. *J. Biol. Chem.* **285**, 11051–11055 (2010).
7. Shenoy, V. M., Thompson, B. R., Shi, J., Zhu, H.-J., Smith, D. E. & Amidon, G. L. Chemoproteomic Identification of Serine Hydrolase RBBP9 as a Valacyclovir-Activating Enzyme. *Mol. Pharm.* **17**, 1706–1714 (2020).
8. Li, W., Blankman, J. L. & Cravatt, B. F. A functional proteomic strategy to discover inhibitors for uncharacterized hydrolases. *J. Am. Chem. Soc.* **129**, 9594–9595 (2007).
9. Chen, Y., Craven, G. B., Kamber, R. A., Cuesta, A., Zherish, S., Moroz, Y. S., Bassik, M. C. & Taunton, J. Direct mapping of ligandable tyrosines and lysines in cells with chiral sulfonyl fluoride probes. *Nat. Chem.* **15**, 1616–1625 (2023).
10. Hahm, H. S., Toroitich, E. K., Borne, A. L., Brulet, J. W., Libby, A. H., Yuan, K., Ware, T. B., McCloud, R. L., Ciancone, A. M. & Hsu, K.-L. Global targeting of functional tyrosines using sulfur-triazole exchange chemistry. *Nat. Chem. Biol.* **16**, 150–159 (2020).
11. Mortenson, D. E., Brighty, G. J., Plate, L., Bare, G., Chen, W., Li, S., Wang, H., Cravatt, B. F., Forli, S., Powers, E. T., Sharpless, K. B., Wilson, I. A. & Kelly, J. W. “inverse drug discovery” strategy to identify proteins that are targeted by latent electrophiles as exemplified by aryl fluorosulfates. *J. Am. Chem. Soc.* **140**, 200–210 (2018).
12. Lin, S., Yang, X., Jia, S., Weeks, A. M., Hornsby, M., Lee, P. S., Nichiporuk, R. V., Iavarone, A. T., Wells, J. A., Toste, F. D. & Chang, C. J. Redox-based reagents for chemoselective methionine bioconjugation. *Science* **355**, 597–602 (2017).

13. Gonzalez-Valero, A., Reeves, A. G., Page, A. C. S., Moon, P. J., Miller, E., Coulonval, K., Crossley, S. W. M., Xie, X., He, D., Musacchio, P. Z., Christian, A. H., McKenna, J. M., Lewis, R. A., Fang, E., Dovala, D., Lu, Y., McGregor, L. M., Schirle, M., Tallarico, J. A., Roger, P. P., Toste, F. D. & Chang, C. J. 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**, 22890–22901 (2022).
14. Gonzalez-Valero, A., Reeves, A. G., Page, A. C. S., Moon, P. J., Miller, E., Coulonval, K., Crossley, S. W. M., Xie, X., He, D., Musacchio, P. Z., Christian, A. H., McKenna, J. M., Lewis, R. A., Fang, E., Dovala, D., Lu, Y., McGregor, L. M., Schirle, M., Tallarico, J. A., Roger, P. P., Toste, F. D. & Chang, C. J. 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**, 22890–22901 (2022).
15. Ma, N., Hu, J., Zhang, Z.-M., Liu, W., Huang, M., Fan, Y., Yin, X., Wang, J., Ding, K., Ye, W. & Li, Z. 2H-Azirine-Based Reagents for Chemoselective Bioconjugation at Carboxyl Residues Inside Live Cells. *J. Am. Chem. Soc.* **142**, 6051–6059 (2020).
16. Cheng, K., Lee, J.-S., Hao, P., Yao, S. Q., Ding, K. & Li, Z. Tetrazole-Based Probes for Integrated Phenotypic Screening, Affinity-Based Proteome Profiling, and Sensitive Detection of a Cancer Biomarker. *Angew. Chem. Int. Ed* **56**, 15044–15048 (2017).
17. Giles, N. M., Watts, A. B., Giles, G. I., Fry, F. H., Littlechild, J. A., & Jacob, C. (2003). Review Metal and Redox Modulation of Cysteine Protein Function metal ions such as Fe <sup>2/3</sup> , Zn <sup>2</sup> , Cd <sup>2</sup> , and Cu [6], oxidized forms of cysteine, such as sulfenic and sulfinic acids, have recently been implicated in the formation of. *Chemistry & Biology*, *10*, 677–693. <https://doi.org/10.1016/S>



18. Fra, A., Yoboue, E. D., & Sitia, R. (2017). Cysteines as redox molecular switches and targets of disease. In *Frontiers in Molecular Neuroscience* (Vol. 10). Frontiers Media S.A. <https://doi.org/10.3389/fnmol.2017.00167>
19. Ostrem, J. M., Peters, U., Sos, M. L., Wells, J. A. & Shokat, K. M. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* **503**, 548–551 (2013).
20. Owen, D. R., Allerton, C. M. N., Anderson, A. S., Aschenbrenner, L., Avery, M., Berritt, S., Boras, B., Cardin, R. D., Carlo, A., Coffman, K. J., Dantonio, A., Di, L., Eng, H., Ferre, R., Gajiwala, K. S., Gibson, S. A., Greasley, S. E., Hurst, B. L., Kadar, E. P., ... Zhu, Y. (2021). An oral SARS-CoV-2 Mpro inhibitor clinical candidate for the treatment of COVID-19. *Science*, 374(6575), 1586–1593. <https://doi.org/10.1126/science.abl4784>
21. Speers, A. E., Adam, G. C., & Cravatt, B. F. (2003). Activity-Based Protein Profiling in Vivo Using a Copper(I)-Catalyzed Azide-Alkyne [3 + 2] Cycloaddition. *Journal of the American Chemical Society*, 125(16), 4686–4687. <https://doi.org/10.1021/ja034490h>
22. Dickson, P., Abegg, D., Vinogradova, E., Takaya, J., An, H., Simanski, S., Cravatt, B. F., Adibekian, A. & Kodadek, T. Physical and functional analysis of the putative rpn13 inhibitor RA190. *Cell Chem. Biol.* **27**, 1371-1382.e6 (2020).
23. Gamayun, I., O’Keefe, S., Pick, T., Klein, M.-C., Nguyen, D., McKibbin, C., Piacenti, M., Williams, H. M., Flitsch, S. L., Whitehead, R. C., Swanton, E., Helms, V., High, S., Zimmermann, R. & Cavalié, A. Eeyarestatin Compounds Selectively Enhance Sec61-Mediated Ca<sup>2+</sup> Leakage from the Endoplasmic Reticulum. *Cell Chem. Biol.* **26**, 571-583.e6 (2019).
24. Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N. & Mann, M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* **11**, 319–324 (2014).

25. Zougman, A., Selby, P. J. & Banks, R. E. Suspension trapping (STrap) sample preparation method for bottom-up proteomics analysis. *Proteomics* **14**, 1006–1000 (2014).
26. Moggridge, S., Sorensen, P. H., Morin, G. B. & Hughes, C. S. Extending the compatibility of the SP3 paramagnetic bead processing approach for proteomics. *J. Proteome Res.* **17**, 1730–1740 (2018).
27. Hughes, C. S., Moggridge, S., Müller, T., Sorensen, P. H., Morin, G. B. & Krijgsveld, J. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat. Protoc.* **14**, 68–85 (2019).
28. Yan, T., Desai, H. S., Boatner, L. M., Yen, S. L., Cao, J., Palafox, M. F., Jami-Alahmadi, Y. & Backus, K. M. SP3-FAIMS Chemoproteomics for High-Coverage Profiling of the Human Cysteinome\*. *Chembiochem* **22**, 1841–1851 (2021).
29. Kuljanin, M., Mitchell, D. C., Schweppe, D. K., Gikandi, A. S., Nusinow, D. P., Bulloch, N. J., Vinogradova, E. V., Wilson, D. L., Kool, E. T., Mancias, J. D., Cravatt, B. F. & Gygi, S. P. Reimagining high-throughput profiling of reactive cysteines for cell-based screening of large electrophile libraries. *Nat. Biotechnol.* **39**, 630–641 (2021).
30. Müller, T., Kalxdorf, M., Longuespée, R., Kazdal, D. N., Stenzinger, A. & Krijgsveld, J. Automated sample preparation with SP3 for low-input clinical proteomics. *Mol. Syst. Biol.* **16**, e9111 (2020).
31. Sheynkman, G. M., Shortreed, M. R., Frey, B. L., Scalf, M., & Smith, L. M. (2014). Large-Scale Mass Spectrometric Detection of Variant Peptides Resulting from Nonsynonymous Nucleotide Differences. *Journal of Proteome Research*, 13(1), 228–240. <https://doi.org/10.1021/pr4009207>
32. Arafeh, R., Shibue, T., Dempster, J. M., Hahn, W. C., & Vazquez, F. (2025). The present and future of the Cancer Dependency Map. *Nature Reviews Cancer*, 25(1), 59–73. <https://doi.org/10.1038/s41568-024-00763-x>

33. Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., Boutselakis, H., Cole, C. G., Creatore, C., Dawson, E., Fish, P., Harsha, B., Hathaway, C., Jupe, S. C., Kok, C. Y., Noble, K., Ponting, L., Ramshaw, C. C., Rye, C. E., ... Forbes, S. A. (2019). COSMIC: The Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Research*, 47(D1), D941–D947. <https://doi.org/10.1093/nar/gky1015>
34. Nesvizhskii, A. I. (2014). Proteogenomics: Concepts, applications and computational strategies. In *Nature Methods* (Vol. 11, Issue 11, pp. 1114–1125). Nature Publishing Group. <https://doi.org/10.1038/NMETH.3144>
35. Ng, P. C., & Henikoff, S. (2003). SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Research*, 31(13), 3812–3814. <https://doi.org/10.1093/nar/gkg509>
36. Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, 7(4), 248–249. <https://doi.org/10.1038/nmeth0410-248>
37. Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., & Kircher, M. (2019). CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Research*, 47(D1), D886–D894. <https://doi.org/10.1093/nar/gky1016>
38. Rodriguez, H., Zenklusen, J. C., Staudt, L. M., Doroshow, J. H., & Lowy, D. R. (2021). The next horizon in precision oncology: Proteogenomics to inform cancer diagnosis and treatment. *Cell*, 184(7), 1661–1670. <https://doi.org/https://doi.org/10.1016/j.cell.2021.02.055>
39. Savitski, M. M., Reinhard, F. B. M., Franken, H., Werner, T., Savitski, M. F., Eberhard, D., Martinez Molina, D., Jafari, R., Dovega, R. B., Klaeger, S., Kuster, B., Nordlund, P., Bantscheff, M. & Drewes, G. Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* **346**, 1255784 (2014).

40. Mateus, A., Kurzawa, N., Becher, I., Sridharan, S., Helm, D., Stein, F., Typas, A. & Savitski, M. M. Thermal proteome profiling for interrogating protein interactions. *Mol. Syst. Biol.* **16**, e9232 (2020).
41. Martinez Molina, D., Jafari, R., Ignatushchenko, M., Seki, T., Larsson, E. A., Dan, C., Sreekumar, L., Cao, Y. & Nordlund, P. Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science* **341**,84–87 (2013).

## Chapter 2: CySP3-96 Enables Scalable, Streamlined, and Low-Cost Sample Preparation for Cysteine Chemoproteomic Applications

Reprinted with permission from

**Shikwana, F.**, Heydari, B., Ofori, S., Truong, C., Turmon, A., Darrouj, J., Holoidovsky, L., Gustafson, J. & Backus, K.M. CySP3-96 enables scalable, streamlined, and low-cost sample preparation for cysteine chemoproteomic applications. *Mol. Cell. Proteomics* 100898 (2024). doi:10.1016/j.mcpro.2024.100898

Copyright © 2024 Elsevier Inc on behalf of American Society for Biochemistry and Molecular Biology

Flowreen Shikwana<sup>1,2</sup>, Beeta S. Heydari<sup>3</sup>, Samuel Ofori<sup>1</sup>, Cindy Truong<sup>2</sup>, Alexandra C. Turmon<sup>1,2</sup>, Joelle Darrouj<sup>2</sup>, Lara Holoidovsky<sup>1</sup>, Jeffrey L. Gustafson<sup>3,4</sup>, Keriann M. Backus<sup>1,2,5,6,7,8\*</sup>

1. Biological Chemistry Department, David Geffen School of Medicine, UCLA, Los Angeles, CA, 90095, USA; 2. Department of Chemistry and Biochemistry, UCLA, Los Angeles, CA, 90095, USA; 3. Department of Chemistry and Biochemistry, San Diego State University, San Diego, CA, 92182, USA; 4. Department of Chemistry, Stony Brook University, Stony Brook, New York, USA; 5. Molecular Biology Institute, UCLA, Los Angeles, CA, 90095, USA; 6. DOE Institute for Genomics and Proteomics, UCLA, Los Angeles, CA, 90095, USA; 7. Jonsson Comprehensive Cancer Center, UCLA, Los Angeles, CA, 90095, USA; 8. Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, UCLA, Los Angeles, CA, 90095, USA; \*Keriann M. Backus, Biological Chemistry Department, David Geffen School of Medicine, UCLA, Los Angeles, CA, 90095, USA, E-mail: kbackus@mednet.ucla.edu.

## **Abstract**

Cysteine chemoproteomic screening platforms are widely utilized for chemical probe and drug discovery campaigns. Chemoproteomic compound screens, which use a mass spectrometry-based proteomic readout, can interrogate the structure activity relationship for thousands of proteins in parallel across the proteome. The versatility of chemoproteomic screens has been demonstrated across electrophilic, nucleophilic, and reversible classes of molecules. However, a key bottleneck that remains for these approaches is the low throughput nature of the most established sample preparation workflows, which rely on many time-intensive and often error prone steps. Addressing these challenges, here we establish a novel workflow, termed CySP3-96, that pairs single-pot, solid-phase-enhanced, sample preparation (SP3) with a customized 96-well sample cleanup workflow to achieve streamlined multiplexed sample preparation. Our CySP3-96 method addresses prior volume limitations of SP3, which allows for seamless 96-well chemoproteomic sample preparation, including for large input amounts that are incompatible with prior methods. By deploying CySP3-96 to screen a focused set of 16 cysteine-reactive compounds, we identify 2633 total ligandable cysteines, including 21 not captured in CysDB. Chemoproteomic analysis of a pair of atropisomeric electrophilic kinase inhibitors reveals striking stereoselective cysteine ligandability for 67 targets across the proteome. When paired with our innovative budget friendly magnetic resin, CySP3-96 represents a versatile, low cost, and highly reproducible screening platform with widespread applications spanning all types of chemoproteomic studies.

## **Introduction**

Mass spectrometry–based chemoproteomics has emerged as a powerful technology capable of pinpointing potential druggable sites proteome-wide. Illustrating the broad impact of chemoproteomics, screening platforms are now available to assess nearly all nucleophilic amino acid side chains (1), inclusive of cysteine (2–5), lysine (6, 7), serine (8–10), tyrosine (11–13), methionine (14, 15), glutamate, and aspartate (16–18). Cysteine remains a favored residue to target, due to its high nucleophilicity, important structural and functional roles within proteins, and proven therapeutic relevance, as exemplified by FDA-approved cysteine-reactive drugs such as the Gly12Cys GTPase KRas inhibitors (19, 20). Thus, a key objective of cysteine chemoproteomic studies is to enable the discovery of scout or pathfinder molecules (3, 21–37) for all potentially druggable, or “ligandable,” cysteine residues.

Towards achieving this goal, chemoproteomic screening platforms apply the same general strategy. First, cells or lysates are treated with electrophilic compounds or vehicle, followed by capping of all unreacted cysteines with a pan-cysteine reactive probe, such as iodoacetamide alkyne (IAA) or iodoacetamide desthiobiotin. After control and treatment groups are isotopically differentiated, either via heavy/light biotin-reagents (5, 38) or with isobaric tags (21, 39, 40), samples are then subject to sequence-specific proteolysis and enrichment on avidin resin along with multiple sample cleanup steps throughout this process. Compound-modified cysteines are then identified via liquid-chromatography tandem mass-spectrometry (LC-MS/MS) based on compound-induced decreases in either precursor or reporter ion abundance. Demonstrating their broad utility, these workflows have now been widely implemented both by our group (3, 38, 40–43) and many others (4, 21, 39, 44), enabling discovery of covalent tool compounds that engage a wide range of protein classes, spanning transcription factors (45), RNA binding proteins (35), kinases (36), proteasome regulators (34), proteases (3, 46), and even classes of posttranslationally modified cysteines (47).

One important consideration for chemoproteomic screens is library composition. As chemoproteomic screens remain comparatively low throughput, selection of a focused and

diverse set of electrophilic fragments and more elaborated compounds has become a go-to strategy to maximize the coverage of potential targets (2, 44). Alongside structural diversity, the addition of enantioenriched libraries to chemoproteomic screening decks provides added value to facilitate delineation of high confidence compound–cysteine interactions and on target phenotypes (28, 34, 45, 48–51). While a number of studies have now reported chemoproteomic datasets using pairs of enantioprobes, the target profile of compounds featuring other types of chirality remains largely unexplored. Atropisomerism is a conformational chirality that occurs when there is constrained rotation about a bond that results in the ‘rotamers’ being isolable enantiomers. Atropisomerism is becoming increasingly ubiquitous throughout modern drug discovery (52). However, to the best of our knowledge, there have been no chemoproteomic studies for electrophilic atropisomer pairs.

With the increasing availability of comparatively large 3200+ electrophilic compound libraries available for purchase and new straightforward chemistries for rapid electrophilic compound library synthesis (53), the number of easily accessible electrophilic compounds available to researchers far outstrips current throughput of chemoproteomic screening platforms, both in terms of sample preparation and data acquisition. Methodological innovations that result in streamlined sample preparation offer the potential to improve sample throughput. Many groups have developed approaches to address these issues, such as SP2 for rapid and reproducible proteomics (54), implementation of microflow LC-MS/MS analysis for improved reproducibility (55), development of intelligent acquisition software (56), and utilizing a label-free approach to speed up acquisition and throughput (57). Transitioning from 1.5 ml microfuge tubes to 96-well plates is one approach to improve sample reproducibility and throughput, which allows for reduced manual sample manipulation, decreased points for human error, and compatibility with automation. Sample cleanup methods, such as in-stage tip (58), S-Trap (59), and single-pot, solid-phase-enhanced sample preparation (SP3) (60, 61) are integral parts of most plate-based sample workflows, as illustrated by their applications to phosphoproteomics (62), clinical



proteomics (63, 64), identification of stereo-specific targets (45), and discovery of off target proteins (39). Our recent work revealed that SP3 cleanup is well-suited to chemoproteomics, yielding improved coverage and benefitting from decreased input material (41, 42). SP3 cleanup is also featured in the recently reported TMT- and 96-well-plate–based chemoproteomic platform by Gygi et al. (21) and has been shown to have high reproducibility between experimental replicates (65), which further demonstrates the utility of SP3 in transitioning cysteine chemoproteomics to streamlined automated sample preparation.

Despite these considerable advances, several unaddressed challenges remain. First, as demonstrated by us and others (41, 66), SP3 cleanup does not scale well, with decreased yields at larger input materials—this limitation is particularly relevant for less reactive chemical probes, for which more sample is required to achieve higher coverage and for compatibility with microflow acquisition methods that are distinguished by lower failure rates than nanoflow (55, 67). Second, and looking beyond SP3-specific workflows, many cleanup methods are quite costly, hindering implementation for large-scale screening applications. And lastly, for chemoproteomics, the avidin-enrichment step remains a bottleneck for fully transitioning an entire workflow into a 96-well format, with most prior reports conducting peptide capture manually in plastic containers.

Here we report the CySP3-96 method, which implements an unprecedented scalable SP3 cleanup and a low cost SP3-alternative resin to enable robust and high coverage cysteine ligandability screens in a fully 96-well plate format. To establish CySP3-96, we first determined the source of the decreased coverage for SP3 samples generated with higher protein input, namely increased peptide concentration in the tryptic digest. We addressed this limitation by bypassing the SP3 peptide cleanup step entirely, which allowed us to maintain volumes compatible with deep-well plates. Benchmarking the CySP3-96 method to our established cysteine-SP3 platform revealed comparable coverage together with at least a 90-min per sample decrease in processing time with increasing time saved proportional to increase in the number of samples. These improvements allowed for high coverage screening in a 96-well plate format of

16 total known and novel electrophiles, including unprecedented atropisomeric compounds, which identified 2633 total ligandable cysteines, of which 21 have not been reported previously in CysDB (68). Distinguished by scalability, low cost, and ease of implementation, we expect CySP3-96 to prove widely useful for a broad range of cysteine chemoproteomic applications.

## **Experimental Procedures**

**Cell Lines, Culture Conditions.** Cells were cultured in Roswell Park Memorial Institute (Gibco RPMI 1640 Medium, 11875119) 1640 media supplemented with 1% Penicillin-Streptomycin (Gibco, Penicillin-Streptomycin 10,000 U/ml, 15140122) and 10% fetal bovine serum (Avantor, Seradigm, Premium Grade Fetal Bovine Serum, Cat. No. 97068-085). The fetal bovine serum used with our cell culture media was obtained from Avantor Seradigm (lot #214B17). The Jurkat cell line was obtained from ATCC. Cells were grown at 37 °C and 5% CO<sub>2</sub> in a humidified incubator. Cell line was tested monthly with the Mycoplasma detection kit (InvitroGen).

**Cell Harvesting and Lysis.** Cells were centrifuged at 300g for 5 min and media was aspirated. Cells were washed twice with cold 1X PBS twice. Cells were resuspended in cold 1X PBS prior to lysis which was done by probe sonication on ice (1 s of sonication at 2 amplitude with 10 s of rest on ice for a total of 10 cycles). Lysates were clarified by centrifugation at 13,000g for 5 min and the supernatant transferred to a new cold microfuge tube. The protein concentration of the lysate was determined by BioRad DC protein assay kit (Cat. No. 5000112 BioRad Life Science). Lysates were then diluted to stated concentrations for experiments.

**IA-Rhodamine Gel.** Jurkat whole cell lysates were normalized to 1.5 mg/ml. Each lysate was treated with either DMSO or compound (500  $\mu$ M or at stated concentration) for 1 h at ambient temperature. Treatment was followed by labeling with 5  $\mu$ M of IA Rhodamine for 20 min at ambient temperature. Then, 4 $\times$  SDS loading dye was added to each sample at a final concentration of 1 $\times$ . Samples were then loaded into a BioRad Bis/Tris gel and run until the dye front ran off at 150V. Gel was imaged with a BioRad ChemiDoc using the Rhodamine channel. Coomassie (Abcam

InstantBlue Coomassie Protein Stain, Cat. No. ab119211) was obtained by rocking gel with Coomassie overnight at 4 °C and washing gel with MilliQ water. Gel was imaged with Coomassie gel setting on the BioRad ChemiDoc.

**Cysteine Chemoproteomic Sample Preparation Using Biotin-Azide Capture Reagents, TEV-Cleavable Biotin-Azide Tags, or sCIP Biotin Capture Reagents.** Jurkat whole cell lysates were normalized to either 1.5 mg/ml or 2 mg/ml (as stated). Each lysate was treated with either DMSO or compound (500  $\mu$ M or at stated concentration) for 1 h at ambient temperature. Treatment was followed by labeling with 200  $\mu$ M IAA for 1 h at ambient temperature. Then, samples were subjected to bioorthogonal copper(I)-catalyzed azide-alkyne cycloaddition ‘click’ conjugation to heavy (DMSO treated) or light (compound treated) biotin azide tags at ambient temperature for 1 h. The click reaction was performed for 200  $\mu$ l sample volumes by adding 24  $\mu$ l of a premixed solution of click reagents which included TBTA (12  $\mu$ l of 1.7 mM stock in DMSO/t-butanol 1:4 per sample, 100  $\mu$ M final concentration), CuSO<sub>4</sub> (4  $\mu$ l of 50 mM stock in water per sample, 1 mM final concentration), TCEP (4  $\mu$ l of fresh 50 mM stock in water per sample, 100  $\mu$ M final concentration), and biotin azide (for biotin-azide and silane-based cleavable isotopically labeled proteomic (sCIP) capture reagents: 4  $\mu$ l of 20 mM stock in DMSO per sample, 400  $\mu$ M final concentration; for tobacco etch virus (TEV) tags: 4  $\mu$ l of 5 mM stock per sample, final concentration = 100  $\mu$ M). After click, protein was solubilized by the addition of SDS at a final concentration of SDS. Samples were then treated with benzonase (Novagen Benzonase Nuclease, Purity >90% MilliporeSigma 707464) at 37 °C for 30 min. Heavy and light samples were combined at equal volumes and subjected to cleanup.

**Protein Cleanup Using SP3.** Sera-Mag SpeedBeads Carboxyl Magnetic Beads (GE Healthcare, 65152105050250, 50  $\mu$ g/ $\mu$ l, total 1 mg and GE Healthcare, 45152105050250, 50  $\mu$ g/ $\mu$ l, total 1 mg) were combined in equal volumes for a total of 60 to 80  $\mu$ l per sample. The bead slurry was washed 3 times with ultrapure water and resuspended in 60 to 80  $\mu$ l of water per sample. The washed beads were then added to each sample (600–800  $\mu$ g protein input) and incubated at

ambient temperature with shaking (1000 rpm) for 5 min. Six hundred microliters of absolute ethanol (for  $\geq 55\%$  ethanol v/v) was added to each sample and samples were then incubated with shaking (1000 rpm) for an additional 5 to 10 min at ambient temperature. Beads were then separated from the solution using a magnetic rack (Sergi Lab Supplies, Cat. No. 1005a) and washed 3 times using 80% ethanol. Beads were then resuspended in 200  $\mu$ l of 2M urea in PBS with 0.5% SDS. DTT was then added at a final concentration of 10 mM and samples were incubated at 65 °C for 15 min. Following this, IA was added at a final concentration of 20 mM and samples were incubated for 30 min at 37 °C. Then, 500  $\mu$ l of absolute ethanol was added to each sample and incubated at ambient temperature with shaking (1000 rpm) for 5 to 10 min. A magnetic rack was again used to separate the beads and washed 3 times with 80% ethanol. The beads were then resuspended in 150  $\mu$ l of 2M urea in PBS containing 4 mM  $\text{CaCl}_2$  and 2  $\mu$ l of 1 mg/ml trypsin (Worthington Biochem LS003740). Samples were digested at 37 °C with shaking (300 rpm) for 12 to 16 h.

**Peptide Cleanup Using SP3.** Acetonitrile was added to the tryptic peptides and SP3 beads at  $\geq 95\%$  of the final volume (3.8 ml acetonitrile for the 150  $\mu$ l digest solution; refer to Supplemental Table S1 for exact volumes corresponding to Fig. 1) and incubated at ambient temperature for 10 min with shaking (1000 rpm). The beads were then separated from the solution using a magnetic rack and washed 3 times with 1 ml of 100% acetonitrile. The beads were then resuspended in 100  $\mu$ l of 2% DMSO in ultrapure water and incubated at 37 °C with shaking (1000 rpm) for 30 min. This elution was collected in a separate low-bind tube (Fisherbrand, Low-Retention Microcentrifuge Tubes Cat. No. 02-681-320) and the beads were resuspended in an additional 100  $\mu$ l of 2% DMSO in ultrapure water at the same conditions as the first elution, except for 45 min. The second elution was collected and combined with the first elution.

**Peptide Cleanup Through Neutravidin Enrichment.** Note: Neutravidin enrichment was used for the isotopically labeled biotin azide enrichment reagents which have no cleavable moiety. These reagents were used primarily throughout this work unless otherwise stated. For each

sample, 50 µl of agarose neutravidin beads (Thermo Fisher Scientific 29200) were washed 3 times with IAP buffer (50 mM MOPS-NaOH, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 50 mM NaCl, pH 7.2), then resuspended in a slurry of 400 µl of IAP buffer per sample. The 400 µl neutravidin bead slurry was added to each sample of tryptic peptide (as obtained through the method described above “Protein cleanup using single-pot solid-phase-enhanced sample preparation (SP3)”) and let incubate for 2 h at ambient temperature (for microfuge tubes, samples were rotated; for plate, samples were left shaking at 600 rpm). The 2 M urea digest samples were diluted 4-fold during this step. Then, beads were centrifuged at 1800g for 3 min and the supernatant was aspirated. The beads were then washed 3 times with PBS followed by another three washes with ultrapure water. The beads were then resuspended in 60 µl of 80% acetonitrile in water with 0.1% formic acid to elute peptides from beads. This first elution was incubated at ambient temperature for 10 min. Beads were separated by centrifugation (1800g for 3 min) and the elution was collected in a new low-bind microfuge tube. The beads were again resuspended in another 60 µl of the elution solution and let incubate at 72 °C for 10 min. The second elution was collected and combined with the first. The beads were once again resuspended with 60 µl of the same elution solution and this last elution was also collected and combined with the prior elutions. Samples were then dried (SpeedVac) and reconstituted with mass spectrometry buffer (5% acetonitrile and 1% formic acid in ultrapure water) and analyzed by LC-MS/MS.

**Protein Cleanup in a 96-Deep Well Plate Using SP3 Beads.** Samples were prepared in the same way as described in the methods above “Isotop-ABPP sample preparation” and “Protein cleanup using SP3” with the following exceptions: All lysates were aliquoted into each well of a 96-deep well plate (MilliporeSigma Supelco SPE 96-Deep Square Well Collection Plate, Cat. No. 11-100-3690) and all treatments and click were carried out within the wells of the plate. Plate was sealed with a sealing mat (MilliporeSigma Supelco 96 Square Well Pierceable Cap Mats, Cat. No. 11-100-3884) during each incubation. The magnetic beads were separated with a magnetic stand-96 apparatus (Thermo Fisher Scientific, Cat no. AM10027). Beads were resuspended

within the wells of the plate by brief (30 seconds-1 minute) shaking at 1000 rpm on an Eppendorf ThermoMixerC (Cat. No. 5382000023) with plate attachment (Eppendorf Smartblock Thermoblock Cat. No. 5363000039). Multichannel pipettes were used for the addition of lysates and reagents to the wells. All volumes were kept under 1.1 ml within the wells.

**Peptide Cleanup in a 96-Deep Well Plate Using SP3 Beads.** Digest solutions post-digest were transferred into a new 96-deep well plate and diluted 2 fold with PBS. For each sample, 50  $\mu$ l of magnetic neutravidin beads (Cytiva Cat No. 78152104010150) were washed 3 times with IAP buffer (50 mM MOPS-NaOH, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 50 mM NaCl, pH 7.2) and then resuspended in a slurry of 300  $\mu$ l of IAP buffer per sample. Plate was then sealed with a sealing mat. Samples were incubated with the neutravidin beads for 2 h at ambient temperature with shaking at 1000 rpm on the thermomixer with plate attachment. Beads were then separated using the magnetic stand-96 apparatus and washed 3 times with PBS followed by another three washes with ultrapure water. Beads were then resuspended in 60  $\mu$ l of 80% acetonitrile in water with 0.1% formic acid for the elution of peptides at ambient temperature for 10 min. For resuspension of beads, plate was subjected to brief (30 seconds-1 min) of shaking at 1000 rpm on plate shaker. Three total elutions were performed and collected as described in the section “Peptide cleanup through neutravidin enrichment” with the modifications of the incubations done with mixing at 1000 rpm and the 72 °C incubation done on the thermomixer with an Eppendorf ThermoTop (Cat. No. 5308000003). Samples were then vacuum dried (SpeedVac) and reconstituted with mass spectrometry buffer (5% acetonitrile and 1% formic acid in ultrapure water) and analyzed by LC-MS/MS.

**Peptide Cleanup Through Streptavidin Enrichment.** Note: Streptavidin enrichment used for the following biotin capture reagents: TEV cleavable isotopic Tandem Orthogonal Proteolysis-ABPP (isoTOP-ABPP)-TEV biotin azide reagents and sCIP biotin azide reagents containing acid cleavable dialkoxydiphenylsilane moiety. These reagents were used to establish compatibility of method with streptavidin enrichment as well as neutravidin enrichment. For each sample, 50  $\mu$ l of

agarose streptavidin beads (Thermo Fisher Scientific 20353) were washed 3 times with PBS and then resuspended in a slurry of 400  $\mu$ l of beads in PBS per sample. The 400  $\mu$ l neutravidin bead slurry was added to each sample of tryptic peptide (as obtained through the method described above "Protein cleanup using single-pot SP3") and let incubate for 2 h at ambient temperature (for microfuge tubes, samples were rotated; for plate, samples were left shaking at 600 rpm). The 2 M urea digest samples were diluted 4-fold during this step. Then, beads were centrifuged at 1800g for 3 min and the supernatant was aspirated. The beads were then washed three times with PBS followed by another three washes with ultrapure water. For isoTOP-ABPP-TEV (5) reagents, beads were resuspended in TEV buffer (50 mM Tris, pH 8, 0.5 mM EDTA, 1 mM DTT) with 1  $\mu$ l TEV protease (2 mg/ml; MacroLab) and left rotating in microfuge tubes overnight. Elution was collected by centrifugation of beads (1800g for 3 min) followed by collection of supernatant, then one 20  $\mu$ l wash of the beads followed by another round of centrifugation and collection. Elutions were then zip-tipped using Pierce C18 tips (Thermo Fisher Scientific 87784) according to the manufacturer's protocol. Samples were then dried (SpeedVac) and reconstituted with mass spectrometry buffer (5% Optima LC/MS grade acetonitrile and 1% Optima LC/MS formic acid in ultrapure water) and analyzed by LC-MS/MS.

For MS1 sCIP reagents, the beads were resuspended in 200  $\mu$ l of 2% formic acid in water and the elution was incubated at ambient temperature for 30 min. Beads were separated by centrifugation (1800g for 3 min) and the elution was collected in a new low-bind microfuge tube. A second elution was carried out with 80% acetonitrile in water for 5 min at ambient temperature. The second elution was collected and combined with the first. Samples were then dried (SpeedVac) and reconstituted with mass spectrometry buffer (5% Optima LC/MS grade acetonitrile and 1% Optima LC/MS formic acid in ultrapure water) and analyzed by LC-MS/MS.

**Protein Cleanup Using Standard DNA Cleanup Magnetic Beads.** DNA cleanup magnetic beads obtained from SergiLabSupplies (Cat. No. 1040, Lot S22122801) were used at the stated volumes of bead slurry. The exact procedures for protein cleanup with SP3 beads were followed.

**Peptide Cleanup Using Standard DNA Cleanup Magnetic Beads.** The exact procedures for peptide cleanup with SP3 beads were followed.

**Measurement of Peptide Concentrations.** Peptide concentrations were obtained through the use of the Pierce Quantitative Colorimetric Peptide Assay (Thermo Fisher Scientific, Cat no. 23275). Manufacturer's protocol was followed.

**Assessing Bead Concentration and Size.** A Thermo Fisher Countess II was used to measure bead concentration and size for each bead type. Beads were diluted as stated and the final reported count was adjusted based on the initial dilution.

**LC-MS/MS Acquisition.** The samples were analyzed by LC-MS/MS using a Thermo Scientific Orbitrap Eclipse Tribrid mass spectrometer. Peptides were fractionated online using an 18 cm long, 100  $\mu$ M inner diameter (ID) fused silica capillary packed in-house with bulk C18 reversed phase resin (particle size, 1.9  $\mu$ m; pore size, 100 Å; Dr Maisch GmbH). The 70-min or 120-min water/acetonitrile gradient was delivered using a Thermo Scientific EASY-nLC 1200 system at different flow rates (buffer A: water with 3% DMSO and 0.1% formic acid and buffer B: 80% acetonitrile with 3% DMSO and 0.1% formic acid). Detailed gradient information for each experiment is outlined in Supplemental Table S3. Data was collected with charge exclusion (1, 8, >8). Data was acquired using a data-dependent acquisition method consisting of a full MS1 scan (Resolution = 120,000) followed by sequential MS2 scans (Resolution = 15,000) to utilize the remainder of the 1 s cycle time. Precursor isolation window was set as 1.6 and normalized collision energy was set as 30%. Files of MS data with corresponding experiments can be found in Supplemental Table S8.

**Methods for Covalent Docking.** The covalent docking function of Molecular Operating Environment (MOE) 2020 was used to dock BEH-1 into a cocrystal structure of serine/threonine-protein kinase Chk2 (CHEK2) (PDB ID: 2XBJ) using the Amber10:EHT forcefield. The protein was first prepared using the “QuickPrep” function in MOE with its default settings, and BEH-1 was put in the correct protonation state at pH 7. Cys231 was identified in the crystal structure and was



used to define the active site of CHEK2. Solvent atoms were not taken into account during the docking. The reaction that was used to generate the covalent bond was “Michael acceptor, 1,4-addition, beta-mercapto carbonyl,” which was picked among the list of reactions MOE provided (Functional group, Class, and Product areas respectively). The induced fit postplacement refinement method was used, in which the ligand was allowed to move ~6 Å away from the defined binding site, and the side chains were free to move. The final scoring methodology used was the Generalized Born Volume Integral/Weighted Surface Area dG. The top five scoring (“S score” or “MOE predicted affinity value” on MOE) nonredundant poses were retained. Each pose was further minimized using the “Minimize” function on MOE to obtain the final docked structures.

**Data Compilation and Statistics.** Raw data collected by LC-MS/MS were searched with MSFragger and FragPipe (version 19.0) (69–71). MSFragger isoDTB-ABPP workflow, with the following parameters and all other settings as default. Protease was set to strict trypsin with cut sites at K and R with no cuts if P prior. Precursor and fragment mass tolerance was set as 20 ppm. Missed cleavages were allowed up to 2. Peptide length was set 7 to 50 and peptide mass range was set 500 to 5000. MS1 intensity ratio of heavy and light labeled cysteine peptides and compiled for individual proteins were reported. Chemoproteomics datasets were searched with variable modifications for both carbamidomethylation (+57.02146), and for biotin azide reagents, MS1 labeling quant was enabled with light set as C+463.2366 and heavy set as C+469.2742. For isoTOP-ABPP experiments with TEV reagents, MS1 labeling quant was enabled with light set as C+521.3074 and heavy set as C+527.3213. For multicysteine-containing peptides, MSFragger modification localization sites were reported (72). For chemoproteomic experiments with sCIP reagents, MS1 labeling quant was enabled with light set as C+493.3007 and heavy set as C+499.3145. MS1 intensity ratio of heavy- and light-labeled cysteine peptides was reported. Peptide and protein level FDR were set to 1%. Database searches were conducted against a reverse concatenated UniProtKB Human database corresponding to the 18432 CCDS release with a release date of 01-01-2020 with 18,555 entries and 18,555 decoys. Calibrated and

deisotoped spectrum files produced by FragPipe were retained and reused for this analysis. RAW files were searched for initial coverage studies with MSConvert (73, 74) used to generate .mzML files for ratio-based analysis with IonQuant (71). For experiments utilizing MS1 quantitation (e.g., isoTOP-ABPP), custom python scripts were implemented to compile labeled peptide and protein datasets. Unique proteins, unique cysteines, and unique peptides were quantified for each dataset. Unique proteins were established based on UniProtKB protein IDs. Unique peptides were found based on sequences containing a modified cysteine residue. Unique cysteines were classified by an identifier consisting of a UniProtKB protein ID and the amino acid number of the modified cysteine (ProteinID\_C#); residue numbers were found by aligning the peptide sequence to the corresponding UniProt protein sequence. When there are multiple cysteines in one peptide, all the cysteine residue numbers will be reported as ProteinID\_C#\_C# though it is undistinguishable on which cysteine the modification is. Cysteine residue identifiers in this format were compared to CysDB v1.5 (68) where stated. Ligandability was assessed based on Log<sub>2</sub>(H/L) ratios of  $\geq 2$  unless otherwise stated. The total ligandable cysteines stated excluded cysteine residues liganded by SO56 and SO59 due to their high proportion of “ligandable” cysteines (i.e., apparent high reactivity). For Figure 2H analysis, percentages report the fraction of CysDB ligandable cysteines out of the total cysteines liganded by each condition (21 total conditions). Higher percentages indicate more promiscuously liganded compounds in CysDB. Statistical values including the exact n, statistical test, and significance are reported in the Figure Legends. Statistical significance was defined as p-value < 0.05 and unless indicated otherwise determined by an unpaired 2-tailed Student’s t test.

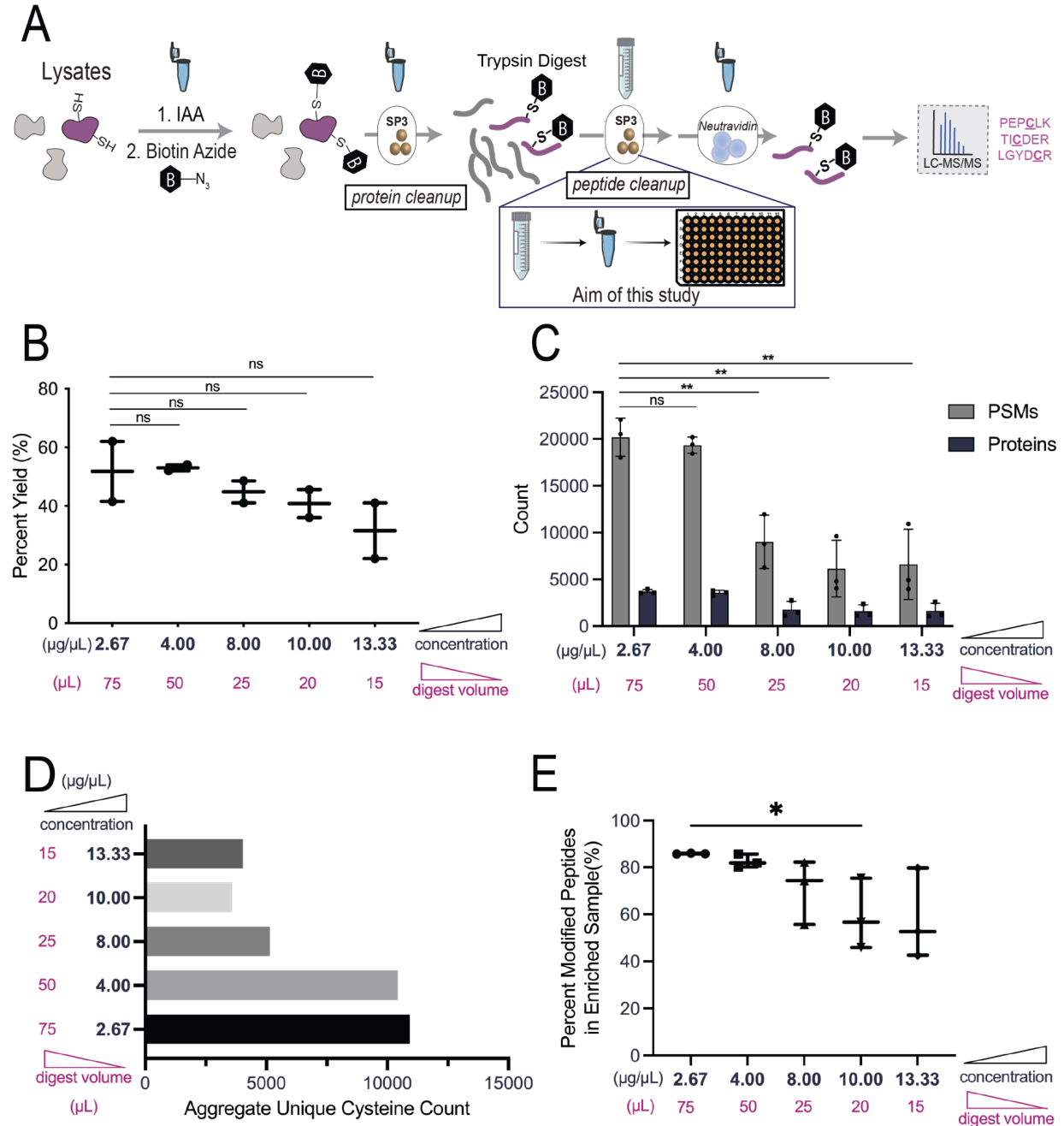
**Statistics.** Statistical analysis was performed using GraphPad Prism (v10.1.0) for Mac (GraphPad Software). Statistical values including the exact n, statistical test, and significance are reported in the Figure Legends. Statistical significance was defined as p-value < 0.05 and unless indicated otherwise determined by an unpaired 2-tailed Student’s t test.

**Experimental Design and Statistical Rational.** All experiments were performed with at least  $n = 2$  replicates. Figure 1B consisted of two replicates and Figure 1, C–E as well as Figure 3, B–D consisted of three replicates which was deemed sufficient for the assessment of our modified workflow and its effect on peptide/protein/cysteine coverage. For Figures 1, C–E and 2, B–D, a student's unpaired  $t$  test was used to determine statistical significance, where applicable. Controls employed within these experiments are methodological controls with already published and established workflows performed side-by-side with the newly modified workflow as a benchmark. The experiment corresponding to Figure 3, E and F consisted of  $n = 4$  replicates which was deemed necessary for assessing the reliability of the ratios obtained from both workflows in the context of electrophilic fragment screening. For the assessment of ratio correlation, only cysteines identified in two or more replicates in both experiments were considered for assessing correlation as to increase the confidence in the data. For experiments corresponding to the plate screen in Figures 2 and 4, each compound was screened in duplicate with a DMSO control in each H/L pair. These should be considered biological replicates as the lysate treatments were all done separately. The filtering criteria for these datasets to assess ligandability was based on  $\text{Log}_2(\text{H/L})$  ratios of  $\geq 2$  unless otherwise stated. Cysteine identifiers per compound needed to have H/L ratios in both replicates. Data was further filtered so that ligandability criteria required both replicates to have a  $\text{Log}_2(\text{H/L})$  ratio of  $\geq 2$  if the SD exceeded 2. The total ligandable cysteines stated excluded cysteine residues liganded by SO56 and SO59 due to their high proportion of "ligandable" cysteines (i.e., apparent high reactivity). For Figure 5 experiments, duplicates were deemed sufficient for benchmarking the low-cost magnetic beads to the standard SP3 magnetic beads (control) in terms of coverage.

## Results

### Increased Tryptic Digest Concentration Fails to Enable Lower Volume Sample Preparation

The first step to transition our cysteine chemoproteomic workflow into a more high throughput platform was to scale down the volumes of key steps of our established method. In our previous SP3-enabled cysteine chemoproteomic sample preparation workflow (Fig. 1A), the maximum volume reached during SP3 cleanup is 4 ml, which far exceeds the maximum volume available for a 96-well plate (2 ml for a 96-well deep well plate)—this large volume is required during the acetonitrile (>95% v/v) peptide-binding step that occurs after the on-SP3 tryptic digest (Supplemental Table S1). Furthermore, the shaking steps required for peptide binding to beads also limits the volume allowed in a deep-well plate, with excess volume leading to inadvertent mixing between wells during shaking. Thus, to transition our method from 1.5 ml microcentrifuge tubes to a plate format, we first focused on the scaledown of this step. We first postulated that scaledown could easily be achieved by implementing an ultra-high concentration and low volume tryptic digest (15  $\mu$ l). We tested the impact of this scaledown for cysteine chemoproteomic samples prepared using our v1.0 workflow (Fig. 1A) with varying trypsin digest volumes, for Jurkat whole cell lysate (200  $\mu$ g total material) labeled with IAA followed by click conjugation to biotin azide (Supplemental Fig. S1). Disappointingly, but consistent with prior reports of high protein concentration negatively impacting proteomic coverage (75), we observed a decrease in peptide recovery with increased digest concentration (Fig. 1B). Coverage of enriched biotinylated cysteine peptide spectral matches and unique cysteines identified also decreased dramatically (~50% loss) when the digest protein concentration exceeded 4 mg/ml (Fig. 1, C and D). Alongside this decrease in cysteine peptide coverage, we also observed an increased capture of background peptides lacking cysteine-biotin modifications for higher digest concentrations (Fig. 1E).



**Figure 1. SP3 peptide cleanup volume scaledown results in overall decreased coverage.**

**A)** Shows v1.0 cysteine chemoproteomics workflow in which SP3 resin is used for both protein- and peptide-level sample cleanup, with scaledown of trypsin digestion sample volume to achieve compatibility with 96-well plate. **B-E)** 200  $\mu\text{g}$  of Jurkat whole lysate capped with IAA and biotin azide and subjected to v1.0 cysteine chemoproteomic sample preparation varying the trypsin digest volume and concentration. **B)** Calculated peptide recovery using peptide assay varying digest concentration and digest volume.  $n=2$ . **C)** Peptide spectral matches (PSMs) and protein coverage for digest volume scaledown.  $n=3$ . **D)** Counts of unique cysteines across all replicates ( $n=3$ ) in aggregate for each digest volume used for sample preparation. **E)** Percent modified peptides in comparison to all detected peptides for each digest concentration.  $n=3$ . For B, C, E, data represent mean values and standard deviation. Statistical significance was calculated with

unpaired Student's t-tests, ns, not significant, \*  $p < 0.05$ , \*\*  $p < 0.005$ , NS  $p > 0.05$ . All MS data is available in **Table S4**.

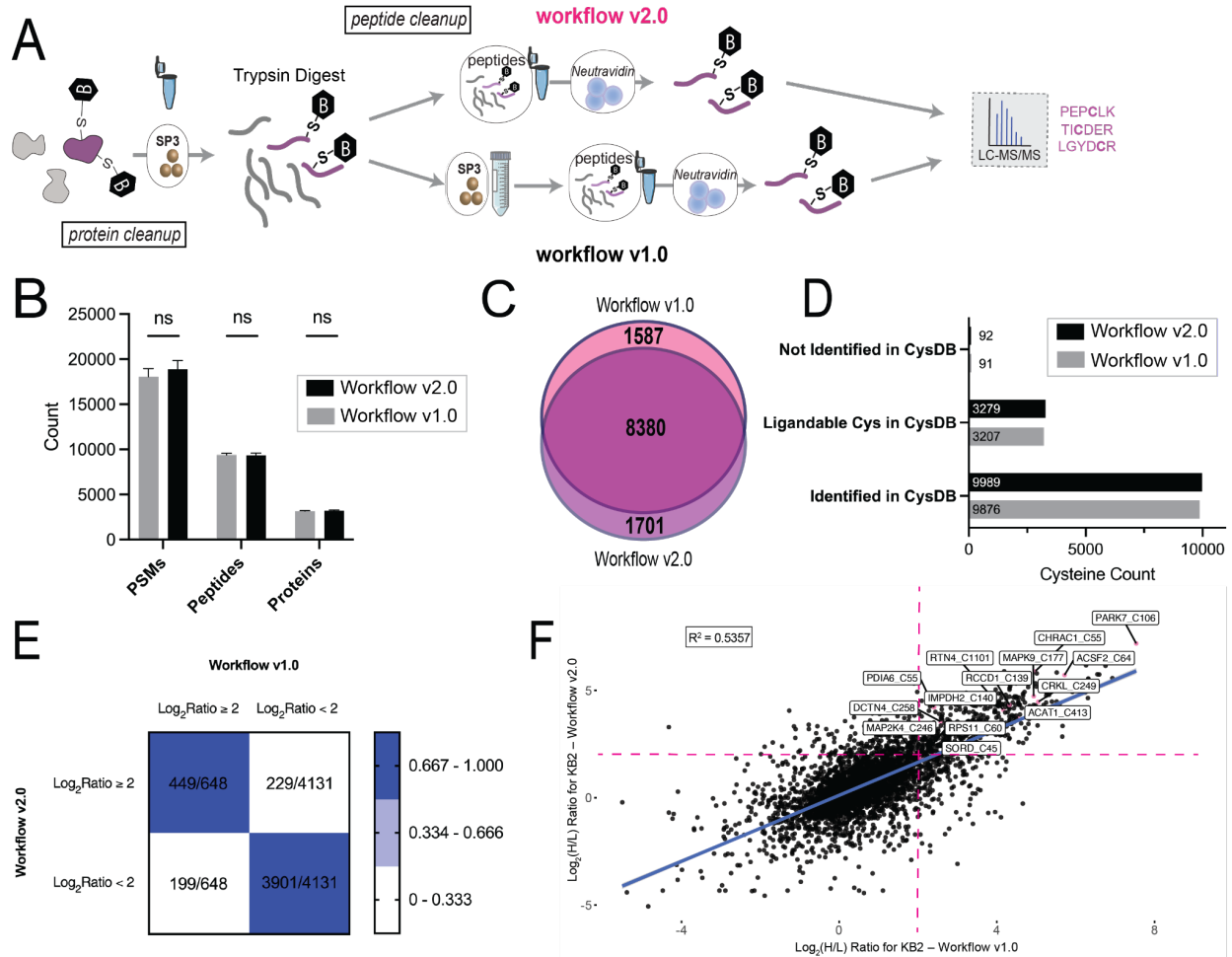
### **Bypassing Peptide Cleanup Step Enables High Coverage and Low Volume Chemoproteomics**

While our prior studies had indicated that SP3-cleanup is highly beneficial for increasing coverage by removing trace biotin contaminant (41, 42), we were inspired by prior studies that have shown that the SP3 peptide-level cleanup step can be skipped for various downstream applications, including phosphoproteomics (62) and TMT labeling (4, 63, 76–78). Therefore, we next opted to test whether the peptide cleanup step could be omitted while maintaining high coverage (Fig. 2A). Gratifyingly, no significant change in coverage was observed for samples prepared with the omission of the peptide cleanup step (Fig. 2B). We term this new cysteine chemoproteomic workflow, which omits the peptide cleanup step on SP3 prior to avidin enrichment, workflow v2.0, and the original workflow outlined in Yan et al. 2021, as workflow v1.0. Further illustrating comparable performance, a substantial overlap for cysteines identified was observed when both workflows were done side by side (Fig. 2C), and similar coverage when compared to CysDB annotations of identified cysteines (68) (Fig. 2D). Notably, peptide neutravidin eluants were directly analyzed with no further cleanup—to ensure that samples contained no substantial residual contaminants from the tryptic digest, we incorporated additional resin washes during the neutravidin capture step (see Experimental Procedures). We observed no significant difference in the proportion of enriched modified peptides between the two workflows (Supplemental Fig. S3). The elimination of a peptide-cleanup step on SP3 beads cuts the sample preparation time by at least 90 min and minimizes sample loss by reducing sample transfer steps. To ensure successful implementation of the v2.0 workflow, we recommend maintaining the digestion protein concentration within a range of 2.67 to 4 mg/ml, which corresponds to a volume ~100 to 150  $\mu$ l, and avoids excessively concentrated protein digests that can negatively impact digest efficiency, as reported previously (75).

### **Streamlined SP3 Cleanup is Compatible with Small Molecule Screening**

Pleased with the overall coverage of our newly modified and more streamlined SP3-chemoproteomic sample preparation method, we next turned to testing its utility for covalent ligand discovery. Following the workflow shown in Supplemental Fig. S4, we subjected Jurkat whole cell lysates to our benchmarking scout fragment KB2 (Supplemental Fig. S2), for which we (3, 40, 79) and others (4, 21, 22, 68) have generated many high coverage prior datasets. For quantitative chemoproteomic target engagement analysis, we utilized isotopically labeled (heavy and light) biotin azide enrichment tags that can be easily clicked to IAA (Supplemental Fig. S1), enriched on neutravidin resin, and eluted from the resin under mild acid conditions. These tags have been shown to provide reliable ratios when used for quantification (38). We tested both our biotin azide-based activity protein profiling (ABPP) under the workflow v1.0 cysteine chemoproteomic conditions (Fig. 1A) and our newly optimized workflow v2.0 (Fig. 2A) in parallel to further evaluate the scope of our newly modified cleanup strategy in the context of cysteine chemoproteomics (Supplemental Fig. S4). We observe similar coverage to our prior studies for both methods, with high overlap in coverage (74% of identified cysteines shared) (Supplemental Fig. S5) as well as similar ligandability profiles (Supplemental Fig. S6, A and B) and types of cysteines captured (as annotated by CysDB) (Supplemental Fig. S6, C and D). We observe generally strong concordance between liganded cysteines identified between both workflows (Fig. 2E) and those previously reported as engaged by KB2, both for our group's prior datasets and for those generated more broadly, as reported by CysDB (68) (Fig. 2F and Supplemental Fig. S6). To assess whether our newly modified SP3 workflow is also compatible with other reagents besides our isotopically tagged biotin azide capture reagents, we compared them against isoTOP-ABPP TEV cleavable biotin reagents (5) and MS1 compatible sCIP reagents (40) using the same electrophile KB2. Both reagents require streptavidin enrichment followed by either TEV or acid-assisted cleavage from the resin. We observed similar cysteine coverage across all reagents

(Supplemental Fig. S7), illustrating the generalizability of this workflow to various custom capture reagents and enrichment types.



**Figure 2. Skipping peptide cleanup results in comparable sample coverage and a more streamlined sample enrichment.** **A)** Shows comparison of v1.0 to v2.0 workflows, with the latter omitting SP3 peptide cleanup post-digestion. Instead, in v2.0 the tryptic digest is immediately enriched onto avidin beads without extra cleanup steps. For the head-to-head comparison of workflows v1.0 and v2.0, the digest volumes were maintained at 150  $\mu$ L. **B-D)** Jurkat whole cell lysate (400  $\mu$ g) in microcentrifuge tubes labeled with IAA (200  $\mu$ M) and clicked to biotin-azide processed using either workflow. n=3 per condition. **B)** Comparison of labeled PSM, peptide, and protein coverage between workflow v1.0 and workflow v2.0. **C)** Aggregate unique and shared cysteines across three replicates for each workflow. **D)** Comparison of coverage to CysDB, including both identified and ligandable unique cysteines. **E,F)** Jurkat whole cell lysate (400  $\mu$ g) in microcentrifuge tubes treated with either KB2 (500  $\mu$ M) or vehicle followed by IAA (200  $\mu$ M), clicked to biotin-azide, and processed using either workflow, n=4 per condition. **E)** Concordance plot comparing distribution of Log<sub>2</sub>(H/L) ratios for workflows v1.0 and 2.0 for KB2-treated Jurkat whole cell lysates. Cysteine identifiers found in more than one replicate in both workflows considered. **F)** Concordance of ligandability ratios for each workflow for KB2-treated Jurkat whole cell lysates with highlighted known KB2 liganded cysteines. Pink dashed lines indicate threshold



for cysteine ligandability ( $\text{Log}_2$  HL ratio  $\geq 2$ ). Cysteine identifiers found in more than 1 replicate in both workflows considered. For B, data represent mean values and standard deviation. Statistical significance was calculated with unpaired Student's t-tests, ns, not significant  $p > 0.05$ . All MS data is available in **Table S5**.

### **Transitioning Our Enhanced SP3 Method into a 96-Well Plate Format**

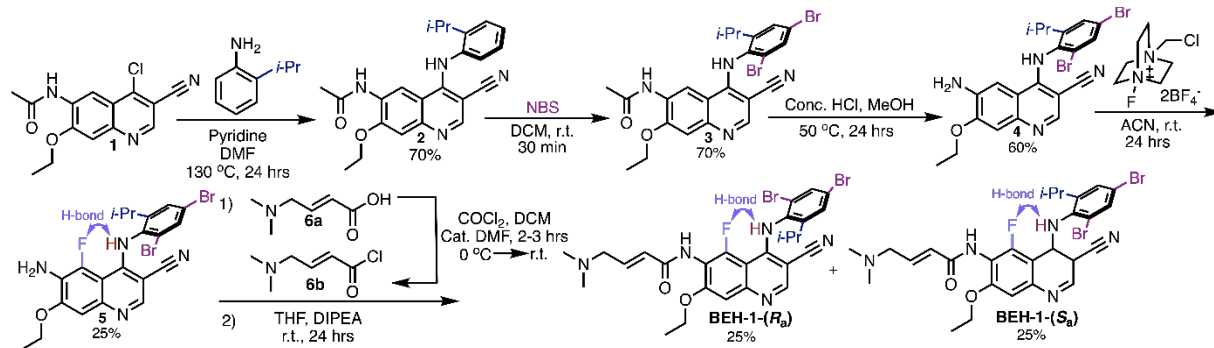
The key overarching goal of our study was to achieve streamlined well-plate sample preparation, with a particular focus on compatibility with large sample inputs that are not amenable to the state-of-the-art SP3-TMT methods (21, 39). Therefore, our next step was to transition our method from 1.5 ml microcentrifuge tubes to a 96-well plate. To enable compatibility with  $>200\ \mu\text{g}$  sample input, we selected a 2 ml deep well plate as the optimal container to match anticipated workflow volumes (Supplemental Table S2). We then purchased two models of plate (round U-bottom and conical V-bottom) and compared capture of SP3 resin by two widely used magnets (80–82), Magnetic Stand-96 versus DynaMag 96. We find that the conical bottom plates paired with the Magnetic Stand-96 afford improved ease of handling and similar coverage as compared to other options assessed (Supplemental Fig. S8). Visual inspection rationalized the increased performance of the Magnetic Stand-96, which features round magnets, compared to the DynaMag 96 side skirted plate magnet, with the former affording improved SP3 resin capture (Fig. 3, A–C and Supplemental Fig. S8). Lastly, our choice of plate was also guided by compatibility with reusable plate seals that are compatible with the vigorous shaking steps in the SP3 workflow.

Benchmarking against our microcentrifuge-based approach revealed comparable cysteine chemoproteomic coverage for our newly established well-plate-based platform (Supplemental Fig. S9, A–C). Further illustrating the robustness of our method, we find that similar coverage was observed for the enrichment on magnetic neutravidin, with magnet-based resin cleanup, compared to conventional neutravidin, with centrifugation-based cleanup (Supplemental Fig. S9, A–C). Notably, this advance allows for full plate-based sample preparation, which substantially minimizes both manual sample manipulation and preparation time.

Showcasing the versatility of our approach, we observe high coverage of cysteine peptides across a range of protein inputs, spanning 100 to 800  $\mu\text{g}$ , with a slight decrease in coverage at lower 50  $\mu\text{g}$  input amounts (Supplemental Fig. S10A). While this plate-based platform is compatible with larger volumes, for cysteine chemoproteomics, 200  $\mu\text{g}$  proved sufficient to achieve high coverage (Supplemental Fig. S10, A and B). We note that the compatibility of our method with larger volumes could be useful for studies using less reactive probes (66). These findings align with our prior report for down-scaling of microcentrifuge-based sample preparation (41), and we expect that, when paired with isobaric multiplexing, further scale down should be feasible, consistent with the recent report by Gygi et al. (39). Additionally, when compared to more dilute chemoproteomic samples, we observed that higher concentration and lower volume samples afforded a modest increase in coverage (Supplemental Fig. S10B), which further highlights the value of our reduced volume approach.

### **CySP3-96 Enables Rapid Screening of Focused Sets of Cysteine-Reactive Electrophilic Compounds**

With our plate-based method established, we next opted to deploy our approach to screen a focused library of cysteine reactive electrophiles in Jurkat whole cell lysates. We assembled a panel of two previously screened electrophilic compounds, the aforementioned KB2 together with KB7 (3, 40, 79), and 12 newly synthesized fragment-like molecules that span a range of electrophilic chemotypes (Supplemental Fig. S1). We additionally included two more elaborated atropisomeric compounds into our screening library (BEH-1-(Sa) and BEH-1-(Ra)), which were designed based on the covalent receptor tyrosine-protein kinase erbB-2 (HER2) and epidermal growth factor receptor kinase inhibitor neratinib that is used to treat HER2+ breast cancer (91) and synthesized in 25% yield (Scheme 1).



**Scheme 1.** Synthesis of atropisomeric compounds.

Then, following the scheme shown in Figure 3D (detailed workflow shown in Supplemental Fig. S11), we subjected Jurkat whole cell lysates to each compound or DMSO vehicle. Following compound incubation and click conjugation to isotopically enriched biotin-azide capture reagents, the samples were combined pairwise across the plate to generate 48 unique samples for competitive chemoproteomic analysis (Supplemental Fig. S4 workflow). To further showcase the versatility of our workflow and to better pinpoint high affinity targets, we also generated dose-response datasets, screening BEH-1-(Sa) and BEH-1-(Ra) at a range of compound concentrations spanning 5  $\mu$ M to 100  $\mu$ M. In total, across all datasets, 12,265 unique cysteines from 4375 proteins were quantified. This coverage is comparable to that reported for other recent high coverage chemoproteomic screening platforms (21, 83). Coverage was highly consistent across all wells and compounds analyzed (Fig. 3, E and F and Supplemental Fig. S12), illustrating that the CySP3-96 workflow does not introduce plate position-based biases or other significant sources of variability.

Prior studies have revealed that apparent cysteine-reactivity can vary considerably across different electrophiles and compounds of varying electronics and lipophilicities (1, 3). Therefore, we next assessed the relative compound reactivity, as inferred from the fraction of cysteines liganded ( $\text{Log}_2(\text{H/L}) \geq 2$ ). We find that the topmost reactive compounds are SO56 and SO59, which have the same scaffold and adduct 23% and 28% of the identified cysteines (Fig. 3F). By contrast, most other compounds (other than KB2 and KB7) showed attenuated reactivity <12%,

including the two atropisomeric compounds. To further corroborate these findings, we performed gel-based analysis and assessed the relative competition of IARho labeling by each compound, which revealed labeling patterns generally consistent with our calculated proteomic reactivity (Supplemental Fig. S13). Thus, CySP3-96 can capture the relative reactivity of different cysteine reactive electrophiles across the proteome.

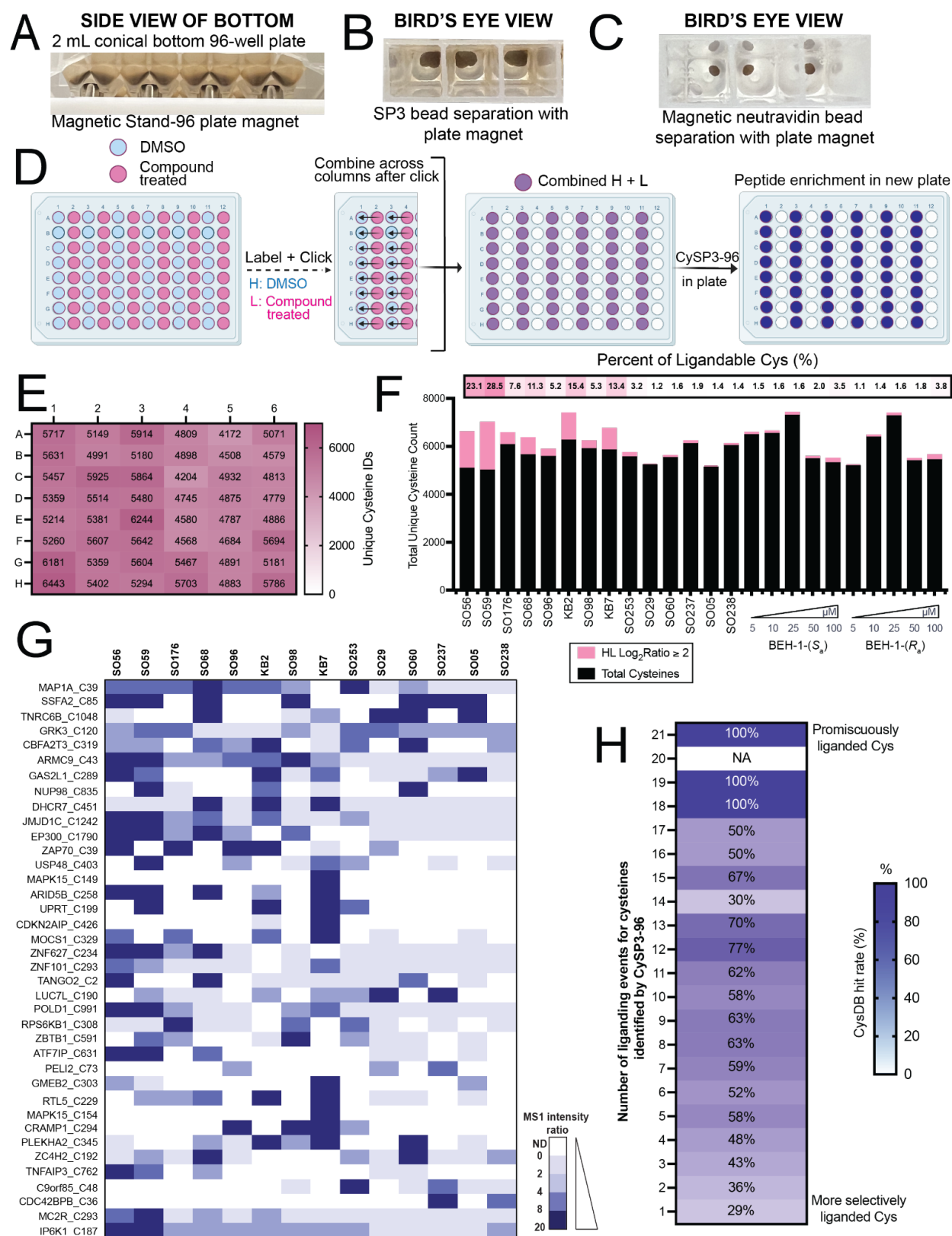
To further vet our CySP3-96 screening data, we compared the liganded cysteines ( $\text{Log}_2(\text{H/L}) \geq 2$ ) identified for widely used compounds, KB2 and KB7, to those identified previously. We observe a number of shared targets, including peroxiredoxin-5 mitochondrial (PRDX5) Cys 204 and cyclin-G-associated kinase (GAK) Cys 190 for KB2 as well as mitogen-activated protein kinase 9 (MAPK9) Cys177 for KB7 and KB2 (3, 4) (Supplemental Fig. S14). Looking more broadly across our data, we identified 2280 total ligandable cysteines across all compounds screened. A number of highly ligandable cysteines stand out, including Cys1101 in reticulon-4 (RTN4) and Cys114 in fructose-2,6-bisphosphate TIGAR (TIGAR) (Supplemental Fig. S14), which are also identified as very ligandable by prior studies (3, 68). These examples illustrate the capacity of CySP3-96 to identify bona fide cysteine labeling sites. Additionally, we found generally good concordance for the ratios measured for KB2 samples prepared in microfuge tubes using workflow v1.0 to the KB2 samples generated in the 96-well plate (Supplemental Fig. S15).

Looking beyond these well-characterized cysteines, we were also pleased to observe that our dataset captured 607 cysteines not previously identified in CysDB, including Cys 120 in G protein-coupled receptor kinase 3 (GRK3), Cys 835 in nuclear pore complex protein Nup98-Nup96 (NUP98), and Cys 403 in ubiquitin carboxyl-terminal hydrolase 48 (USP48) (Fig. 2G). Consistent with bona fide liganding events, we note that many of these newly observed cysteines observed SAR across the compound library. For example, while Cys 2 in transport and golgi organization protein 2 homolog (TANGO2) was generally insensitive to most compounds tested, this cysteine was sensitive to both of the highly structurally related compounds SO60 and SO68 (Fig. 3G). Similarly, for SO60 and SO68, both were observed to engage Cys 319 in protein CBFA2T3

(CBFA2T3) and Cys 192 in zinc finger C4H2 domain-containing protein (ZC4H2). Furthermore, the structurally similar compounds SO237 and SO238, which showed very attenuated proteome-wide reactivity, both liganded Cys 36 in serine/threonine-protein kinase MRCK beta (CDC42BPB) (Fig. 3G). Analysis of the cysteines not previously reported in CysDB revealed that 24% of the new liganded cysteines (engaged by 2 or more compounds) belong to the transcription factor and/or regulator tough-to-target class of proteins (Supplemental Fig. S16). The next most liganded groups were the proteins classified under the enzyme group, the scaffolding, modulator, adaptor group, and the uncategorized group, each containing 17% of the cysteines in the group to be liganded by two or more of the compounds used for screening (Supplemental Fig. S16) We ascribe newfound ligandability both to the novel compounds in our screening library together with the robust coverage afforded by our CySP3-96 platform. To further understand how our dataset compares to prior studies, we compared the relative ligandability (fraction of compounds screened that label a cysteine) for CysDB data compared to our CySP3-96 screen. Gratifyingly, the relative ligandability trends corroborate similar performance of our platform to prior studies, with greater ligandability in CysDB matching that observed in our CySP3-96 screen (Fig. 3H).

In addition to these newly ligandable cysteines, we also note the presence of cysteines that show compound-induced increased labeling by IAA, as indicated by negative  $\log_2(H/L)$  ratios. While we cannot exclude the possibility that some of these IDs could stem from some degree of stochasticity, we rationalize these negative values as likely stemming from so-called “anti-ligandable” cysteines, which are cysteines that show an increase in IAA labeling concurrent with a parallel liganding event at another cysteine in the same protein, as reported recently by Takahashi et al. (4). Consistent with this likelihood, we note a number of anti-ligandable cysteines are shared between our KB2 dataset and the prior report, including Cys 249 in small ribosomal subunit protein RACK1 (RACK1), Cys 40 in UMP-CMP kinase 2 mitochondrial (CMPK2), Cys 111 in ubiquitin-conjugating enzyme E2 D2 (UBE2D2), and Cys 98 in 14-3-3 protein epsilon (YWHAE).

Taken together, these examples illustrate the capacity of CySP3-96 to identify established and novel ligandable cysteines and to enable high coverage screening of focused electrophilic libraries.



**Figure 3. Streamlined SP3 preparation seamlessly transitions into 96 deep well plate format for screening of cysteine reactive electrophiles.** A) Deep-well plate with Magnetic

Stand-96 magnet used for bead separation. **B)** Visualization of the magnetic bead separation in each well. **C)** Visualization of Cytiva magnetic neutravidin bead separation in each well for enrichment. **D)** Workflow for the plate-based screening starting with in-plate lysate treatment with compounds or vehicle, labeling with IAA, click conjugate to either heavy or light biotin azide enrichment reagents, and proteomic sample preparation. **E)** Number of unique cysteines identified per well across the plate for cysteine electrophilic compound screen. **F)** Cysteine ligandability per compound as compared to overall coverage. **G)** Shows the SAR of liganded cysteines not previously identified in CysDB. MS1 intensity ratios are calculated as H/L, with N.D. indicating that the cysteine was not detected. **H)** Comparison of the hit rate (percentage of total compounds screened that engage a cysteine in CysDB) to the number of liganding events reported by CySP3-96, with a max of 21 liganding events for the 16 compounds screened, including dose response assessments. For panels E-H, Jurkat whole cell lysate (300  $\mu$ g) in 96-well plate treated with either compound (500  $\mu$ M unless otherwise stated) or vehicle followed by IAA (200  $\mu$ M), clicked to either light or heavy biotin-azide, and processed via chemoproteomics workflow v2.0, n=2 per condition. All MS data is available in **Table S6**.

### **Atropisomeric Electrophiles Show Distinct Proteome-wide SAR**

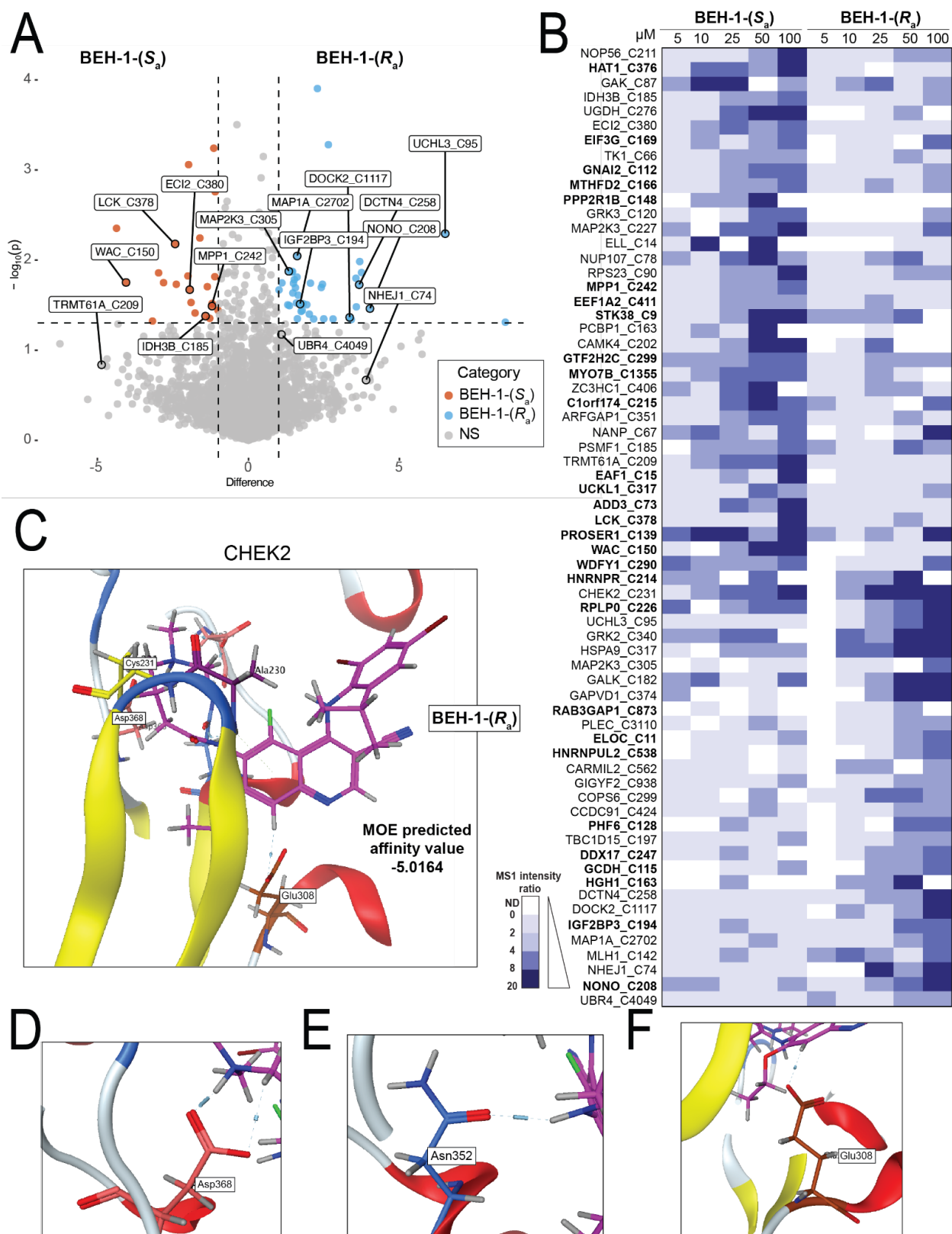
As atropisomers feature conformationally stable enantiomers (84), we hypothesized that BEH-1-(Sa) and BEH-1-(Ra) compounds would each engage unique cysteines. To test this hypothesis, we first stratified the cysteine ligandability ratios across the concentration range screened for each compound and compared the ratios obtained for the BEH-1-(Sa) and BEH-1-(Ra) compounds (Fig. 4A). Consistent with our hypothesis, we find that 66 total cysteines are preferentially labeled by one atropisomer. Of these, 30 cysteines had not been found to be ligandable in CysDB, which are highlighted with bolded identifiers in Figure 4B. As the general reactivity of each compound was observed to be comparable (Fig. 3F), we expected that these differences were reflective of bona fide SAR.

These newly liganded cysteine proteins include DNA/RNA-binding proteins, such as Cys208 in non-POU domain-containing octamer-binding protein and Cys169 in eukaryotic translation initiation factor 3 subunit. We were also pleased to observe that both atropisomers engaged many kinase cysteines, including residues such as Cys317 in uridine-cytidine kinase-like 1 (Sa), Cys9 in serine/threonine-protein kinase 38 (Sa), and Cys340 in beta-adrenergic receptor kinase 1 (Ra). This observation aligns with the nature of these compounds as analogs of known kinase inhibitors. Further highlighting enantioselective SAR reported by our platform, we find that BEH-1-(Sa)



shows increased labeling of transferase/kinase annotated proteins as compared to BEH-1-(Ra) (Supplemental Fig. S17). Illustrating this activity, we see that the active-site adjacent calcium/calmodulin-dependent protein kinase type IV Cys202 was preferentially engaged by BEH-1-(Sa) (Fig. 4B). Highlighting possibilities for distinct binding modes, we observed that Cys231 in CHEK2 was preferentially engaged by BEH-1-(Ra), particularly at lower compound concentrations (Fig. 4B). These compelling examples illustrate the utility of screening atropisomeric compounds for target hunting. We do also note that we did not identify cysteines from the neratinib targets epidermal growth factor receptor and HER2, likely due to very low levels of protein expression in Jurkat cells and the lack of inhibition of these kinases observed for BEH-1-(Sa) and BEH-1-(Ra), in an established kinase activity assay (Supplemental Fig. S18).

Intrigued by the preferential labeling of CHEK2 by the R atropisomer, we next subjected CHEK2 (PDB ID: 2XBJ) to covalent molecular docking using MOE software, with the goal of assessing whether the chemoproteomic data could be corroborated *in silico*. Consistent with the chemoproteomic-preference for BEH-1-(Ra), we find that most (4/5) of the binding poses observed for BEH-1 adopt the R conformation (Fig. 4C and Supplemental Fig. S19). The best R enantiomer-binding pose was also observed to exhibit the lowest overall binding energy, with an MOE predicted affinity score of -5.0164 (Fig. 4C). In addition to the covalent linkage, multiple contacts were observed between CHEK2 and the BEH-1-(Ra) atropisomer within this pocket (Fig. 4, D, E and F). Taken together, these docking studies corroborate the chemoproteomic reported enantioselective labeling at CHEK2.



**Figure 4. Concentration screen of two atropisomers BEH-1-(S<sub>a</sub>) and BEH-1-(R<sub>a</sub>) illustrates stereoselectivity across targets. A,B) Jurkat whole cell lysate (300 μg) in 96-well plate treated**

with either BEH-1-( $S_a$ ) and BEH-1-( $R_a$ ) at the indicated concentration or vehicle followed by IAA (200  $\mu$ M), clicked to either light or heavy biotin-azide, and processed via chemoproteomics workflow v2.0, n=2 per condition. **A)** Volcano plot of the fold change difference of the Log<sub>2</sub>(H/L) ratios for Jurkat whole cell lysates treated with BEH-1-( $S_a$ ) and BEH-1-( $R_a$ ) (100  $\mu$ M) followed by cysteine chemoproteomic analysis. **B)** Heatmap of atropisomer targets across a range of concentrations between 5 and 100  $\mu$ M. Bolded sites indicate those not previously identified as ligandable in CysDB. **C)** CHEK2 (PDB ID: 2XBJ) Cys231 docked to BEH-1-( $R_a$ ) with an MOE predicted affinity value of -5.0164. Four out of five poses observed with BEH-1-( $R_a$ ) atropisomer. **D)** A zoom in of C: Asp368 in the DFG motif of CHEK2 making hydrogen bonding interactions with the dimethylamino moiety of BEH-1-( $R_a$ ) as well as its adjacent CH<sub>2</sub> group. **E)** A zoom in of C: Asn352 in the hinge region of CHEK2 making a hydrogen bond with the hydrogen of the amine group in the covalent linker moiety of BEH-1-( $R_a$ ). **F)** A zoom in of C: Glu308 in the hinge region of CHEK2 making a hydrogen bond with the C8 hydrogen of the quinoline moiety of BEH-1-( $R_a$ ). All docking figures generated with CHEK2 PDB structure 2XBJ in MOE 2020. All MS data is available in **Table S6**.

### Establishing a Low-Cost Option for Magnetic Bead-Based Sample Cleanup

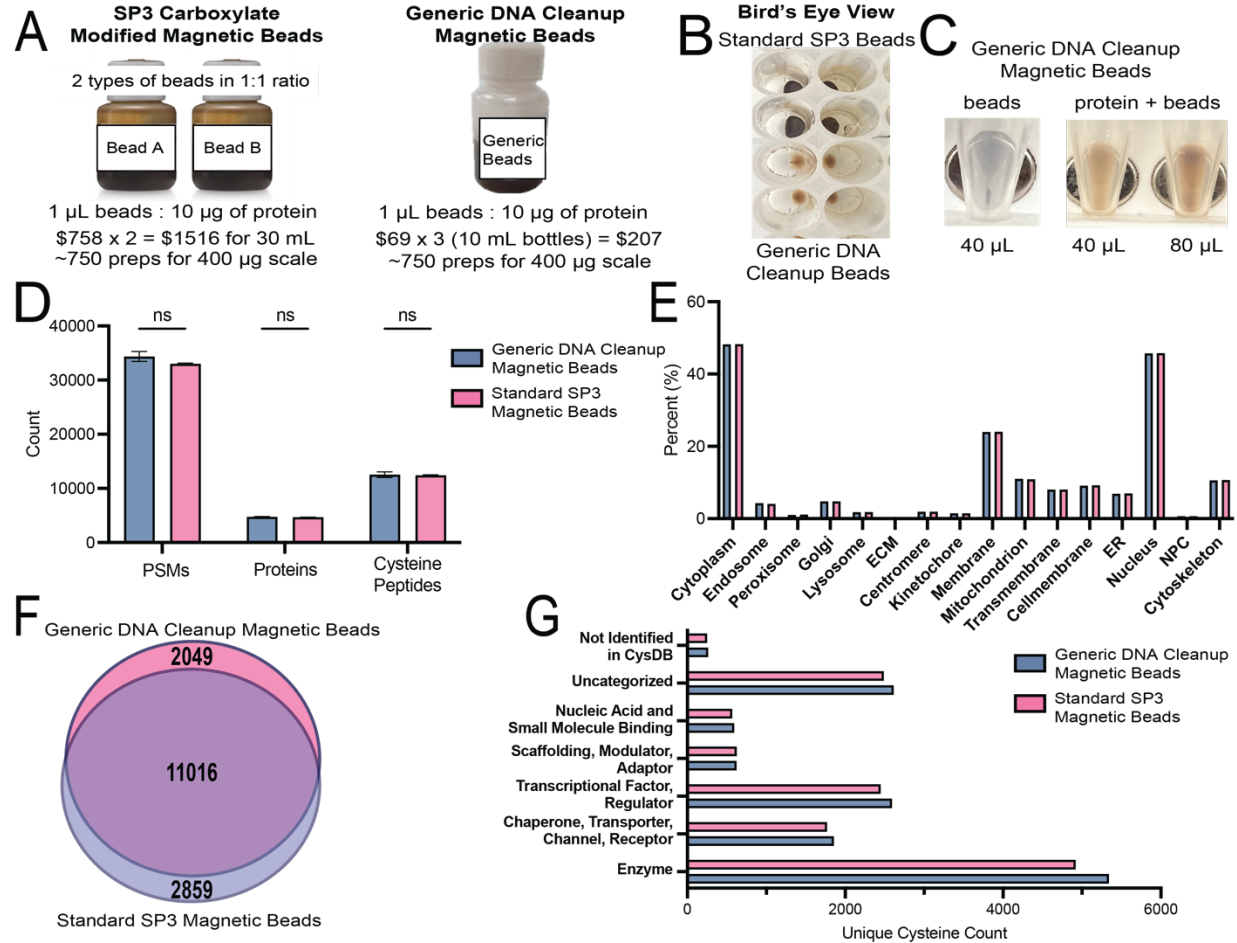
Achieving cost savings was a key goal of our study. Therefore, we next sought to identify budget friendly SP3 resin. We were inspired by recent studies that had shown, for protein cleanup, low cost glass beads function comparably to higher cost magnetic resin (85), albeit using centrifugation to isolate the resin. Other studies have also shown that the surface chemistry of beads is largely irrelevant to sample preparation, which is facilitated by protein precipitation rather than true binding to beads (86). Thus, we hypothesized that other low-cost magnetic resins would function comparable to the widely used SP3 resin (standard SP3 beads). To test this hypothesis, we selected a low-cost, commercially available resin (generic DNA cleanup beads) as a reagent for oligonucleotide cleanup (Fig. 5A). Visual inspection of the appearance of each resin slurry revealed that the low cost resin appeared more dilute, which was confirmed by the lower compacted bead volume (approximately 1/5th the size for the same volume taken, upon visual inspection) (Fig. 5, B and C). Analysis of each bead type revealed that the standard SP3 beads had a concentration of  $\sim 5 \times 10^8$  beads per mL of resin slurry while the generic DNA cleanup beads had a concentration of  $\sim 1 \times 10^8$  beads per mL of resin slurry, corresponding to approximately 4.4 $\times$  difference in concentration. Thus, we opted to adjust the volume of slurry used to match the compacted bead volume between the two resins for our initial proteomic sample

preparation. Gratifyingly, head-to-head chemoproteomic coverage comparison revealed no significant difference in coverage between the two resins (Fig. 5D).

As the scaled-up slurry volume would diminish the potential cost-savings, we also tested the performance of lower slurry volumes (Supplemental Fig. S20). We find the coverage is unchanged when the magnetic beads were used in the same ratio as the SP3 beads: 1  $\mu$ g of protein to 10  $\mu$ l of slurry (Fig. 5D and Supplemental Fig. S20). Additional inspection of the size of the beads revealed that the standard SP3 beads ( $\sim$ 4.23  $\mu$ m) are slightly larger than the generic DNA cleanup beads ( $\sim$ 3.11  $\mu$ m) (Supplemental Fig. S21). We hypothesize that the slightly smaller size of the generic DNA cleanup beads allows for more surface area for proteins to bind and/or aggregate onto, enabling the use of smaller bead volume. Visual inspection of these samples revealed a striking advantage of the low-cost resin: the smaller bead volume of the magnetic beads allows for more space within the well for aspirating supernatant throughout this workflow (Fig. 5B). The low bead volume also afforded a visual cue of efficient protein capture, namely an increase in the whitish hue of resin that occurs after protein precipitation (Fig. 5C). As a note of caution, we do acknowledge that ultra-low resin volumes ( $<40$   $\mu$ l bead slurry), particularly for small volume samples, may result in protein pellets that are challenging to visualize. Therefore, we do not recommend using bead slurry volumes of less than 40  $\mu$ l.

To further vet this resin, we also compared coverage of cysteines identified. We find that  $>80\%$  of cysteines are captured by both resins with no significant difference in the classes of proteins observed for each resin (Fig. 5, E and F). We ascribe the handful of cysteines preferentially identified by each resin to the stochastic nature of data-dependent acquisition, although we cannot rule out differences in resin affinities. Lastly, we also assessed whether this new resin could also prove useful for peptide-level sample cleanup post-tryptic digest—while not required for the CySP3-96 method, peptide cleanup is highly useful for other methods, such as our redox proteomic platform SP3-Rox (43). We find comparable peptide recovery is achieved for both resins analyzed (Supplemental Fig. S22). In sum, these data provide compelling evidence of the

general utility of the generic DNA cleanup resin, which affords savings of >85% across 750 sample preparations equates to over \$1300 (Fig. 5A).



**Figure 5. Identifying a low-cost magnetic option for SP3.** **A)** Cost analysis of using SP3 resin (standard SP3 beads) versus commercially available magnetic beads (generic DNA cleanup beads) usually used for oligonucleotide purification. **B)** Shows deep-well plate setup with generic DNA cleanup magnetic beads (bottom 2 rows) or (standard SP3 beads (top two rows) to show differences in bead volume. Smaller bead volumes allow for more free space for tips when aspirating washes. **C)** Shows generic DNA cleanup beads without protein and when protein is immobilized onto beads. **D)** Comparison of PSM, protein, and cysteine coverage between the same volume of generic magnetic beads as compared to standard SP3 beads, n=2. **E)** Classes of proteins detected, based upon UniProtKB annotations, from sample preparations using the magnetic beads or SP3 beads. **F)** Overlap of unique cysteines identified by each bead type. **G)** Categorization of each cysteine identifier from CysDB for each bead type. Statistical significance was calculated by performing an unpaired Student's t-test.  $p < 0.05$  for significance cutoff. All MS data is available in **Table S7**.

## Discussion

In this study, we developed a 96-well SP3 workflow, CySP3-96, for the high throughput processing of chemoproteomic samples. Taken together, CySP3-96 chemoproteomics performed comparably to other cutting-edge platforms, identifying in aggregate 12,265 unique cysteines, of which 607 had not been previously captured by CysDB.

To establish CySP3-96, we first address the scalability limitation of SP3, to enable sample preparations with large amounts of input material. We expect that this methodological innovation should prove particularly useful for applications where coverage is tightly linked to sample amount, such as chemoproteomics using less reactive probes. Likely the scalability of our approach will also prove beneficial in addition to workflows that pool samples at the protein level, thereby increasing sample amount per experiment. Such applications include for cysteine redox proteomics (SP3-Rox) (43) and for isobaric labeling studies using our sCIP reagents (40).

After extensive benchmarking to ensure the robustness of our plate-based workflow, we deployed our platform to screen a focused library of cysteine-reactive electrophilic compounds. Our screening library contains both previously reported and novel electrophilic compounds, including an elaborated pair of atropisomeric analogs of neratinib. In aggregate, we identify 2280 ligandable cysteines, including 67 residues with atropisomeric selectivity. As demonstrated by recent work using stereoprobe libraries (45), we expect that atropisomeric compounds should prove similarly useful in guiding the discovery of high confidence and actionable cysteine-compound pairs.

The increased sample throughput offered by CySP3-96 (between 1.5 h to 12 h decreased active sample manipulation time, Supplemental Fig. S23) facilitates testing dose-response across a range of compound concentrations. As shown for the pair of atropisomeric compounds tested here, testing a range of compound concentrations is valuable for the delineation of high versus low affinity labeling events. We expect that dose ranges should also prove useful for delineating on-target labeling from compound-induced changes to protein abundance caused by generalized cell stress—of note, our recent work revealed that many electrophilic compounds activate cellular

stress responses, including stress granules and aggresomes when compound concentrations exceed low micromolar concentrations (87).

The many available, reliable, high performance, and high-throughput SP3 platforms, including for cysteine chemoproteomics (21, 39), highlight the burgeoning enthusiasm for SP3-based screening methods. Our CySP3-96 method is not intended to supplant these other impactful methods but rather to provide several innovations that are highly complementary to these existing platforms. Thus, we encourage broad consideration of the relative strengths of different approaches for those interested in establishing such SP3-based screening methods. We see our platform working well for those looking to reduce cost and for applications in which increased protein input is beneficial, for example, applications using less reactive probes in which scaled up protein input along with incorporation of offline fractionation will likely boost coverage. We envision a key area where the CySP3-96 method could prove impactful is for streamlining sample preparation via pooling of samples at the protein level, as has been recently used with custom, enrichable isobaric tags (40, 88) which are methods that necessitate larger-scale SP3 cleanup. Our method also incorporates plate-based avidin enrichment, which offers the opportunity for a fully automatable cleanup and enrichment in the future. We do also acknowledge the cumbersome pairwise sample pooling is a clear drawback of the current implementation of CySP3-96, which we hope to address via either the utilization of the aforementioned isobaric protein capture reagents or through implementation of label-free quantification strategies, most notably data independent acquisition.

Looking beyond chemoproteomics, we expect that our low-cost magnetic resin-based scalable high throughput cleanup workflow should prove useful for a wide range of applications that require increased sample input as well as increased number of samples. Our envisioned applications include for SP3-based phosphoproteomics workflows (62), SP3-based immunopeptidomics (89), and for upscaling for microflow-based sample separation (55, 67). Lastly, building on prior reports that revealed the surface chemistry flexibility of bead-based peptide cleanup, we anticipate that a

range of magnetic resins will likely function comparable for both protein- and peptide-based sample decontamination.

### **Data Availability**

The MS data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (90) with the dataset identifier PXD054043. File details can be found in Supplemental Table S8.

### **Code Availability**

The scripts used for this work are available at <https://github.com/BackusLab>.

### **Supplemental data**

This article contains supplemental data (5, 40, 41, 68, 91, 92).

### **Conflict of interest**

K. M. B. is a member of the advisory board at Matchpoint Therapeutics. All other authors declare no competing interests.

### **Acknowledgments**

We additionally thank all members of the Backus lab members for helpful suggestions.

### **Funding and additional information**

This study was supported by a Packard Fellowship 2020-71388 (K.M.B.), Ono Breakthrough Science Initiative Award (K. M. B.), and NIH NIGMS 5R35GM124637 (J. L. G.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

### **Author contributions**

F. S., J. L. G., and K. M. B. writing—review and editing; F. S. and K. M. B. writing—original draft; F. S. and B. S. H. visualization; F. S., A. C. T., J. D., and L. H. validation; F. S. and J. L. G. methodology; F. S., B. S. H., S. O., A. C. T., J. D., and L. H. investigation; F. S., B. S. H., and C. T. formal analysis; F. S. and K. M. B. data curation; F. S. and K. M. B. conceptualization; B. S. H.,



S. O., and J. L. G. resources; C. T. software; J. L. G. and K. M. B. supervision; J. L. G. and K. M. B. project administration; J. L. G. and K. M. B. funding acquisition.

## References

1. [preprint] Zanon, P. R. A., Yu, F., Musacchio, P., Lewald, L., Zollo, M., Krauskopf, K., et al. (2021) Profiling the proteome-wide selectivity of diverse electrophiles. ChemRxiv. <https://doi.org/10.26434/chemrxiv.14186561.v1>
2. Lin, Z., Wang, X., Bustin, K. A., Shishikura, K., McKnight, N. R., He, L., et al. (2021) Activity-based hydrazine probes for protein profiling of electro-philic functionality in therapeutic targets. ACS Cent. Sci. 7, 1524–1534
3. Backus, K. M., Correia, B. E., Lum, K. M., Forli, S., Horning, B. D., González-Paéz, G. E., et al. (2016) Proteome-wide covalent ligand discovery in native biological systems. Nature 534, 570–574
4. Takahashi, M., Chong, H. B., Zhang, S., Yang, T. Y., Lazarov, M. J., Harry, S., et al. (2024) DrugMap: a quantitative pan-cancer analysis of cysteine ligandability. Cell 187, 2536–2556.e30
5. Weerapana, E., Wang, C., Simon, G. M., Richter, F., Khare, S., Dillon, M. B. D., et al. (2010) Quantitative reactivity profiling predicts functional cysteines in proteomes. Nature 468, 790–795
6. Abbasov, M. E., Kavanagh, M. E., Ichu, T. A., Lazear, M. R., Tao, Y., Crowley, V. M., et al. (2021) A proteome-wide atlas of lysine-reactive chemistry. Nat. Chem. 13, 1081–1092
7. Hacker, S. M., Backus, K. M., Lazear, M. R., Forli, S., Correia, B. E., and Cravatt, B. F. (2017) Global profiling of lysine reactivity and ligandability in the human proteome. Nat. Chem. 9, 1181–1190
8. Simon, G. M., and Cravatt, B. F. (2010) Activity-based proteomics of enzyme superfamilies: serine hydrolases as a case study. J. Biol. Chem. 285, 11051–11055

9. Shenoy, V. M., Thompson, B. R., Shi, J., Zhu, H. J., Smith, D. E., and Amidon, G. L. (2020) Chemoproteomic identification of serine hydrolase RBBP9 as a valacyclovir-activating enzyme. *Mol. Pharm.* 17, 1706–1714
10. Li, W., Blankman, J. L., and Cravatt, B. F. (2007) A functional proteomic strategy to discover inhibitors for uncharacterized hydrolases. *J. Am. Chem. Soc.* 129, 9594–9595
11. Chen, Y., Craven, G. B., Kamber, R. A., Cuesta, A., Zherish, S., Moroz, Y. S., et al. (2023) Direct mapping of ligandable tyrosines and lysines in cells with chiral sulfonyl fluoride probes. *Nat. Chem.* 15, 1616–1625
12. Hahm, H. S., Toroitich, E. K., Borne, A. L., Brulet, J. W., Libby, A. H., Yuan, K., et al. (2020) Global targeting of functional tyrosines using sulfur- triazole exchange chemistry. *Nat. Chem. Biol.* 16, 150–159
13. Mortenson, D. E., Brighty, G. J., Plate, L., Bare, G., Chen, W., Li, S., et al. (2018) “inverse drug discovery” strategy to identify proteins that are tar- geted by latent electrophiles as exemplified by aryl fluorosulfates. *J. Am. Chem. Soc.* 140, 200–210
14. Lin, S., Yang, X., Jia, S., Weeks, A. M., Hornsby, M., Lee, P. S., et al. (2017) Redox-based reagents for chemoselective methionine bioconjugation. *Science* 355, 597–602
15. Gonzalez-Valero, A., Reeves, A. G., Page, A. C. S., Moon, P. J., Miller, E., Coulonval, K., et al. (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, 22890–22901
16. Bach, K., Beerken, B. L. H., Zanon, P. R. A., and Hacker, S. M. (2020) Light- activatable, 2,5-disubstituted tetrazoles for the proteome-wide profiling of aspartates and glutamates in living bacteria. *ACS Cent. Sci.* 6, 546–554
17. Ma, N., Hu, J., Zhang, Z. M., Liu, W., Huang, M., Fan, Y., et al. (2020) 2H- Azirine-Based reagents for chemoselective bioconjugation at Carboxyl residues inside live cells. *J. Am. Chem. Soc.* 142, 6051–6059

18. Cheng, K., Lee, J. S., Hao, P., Yao, S. Q., Ding, K., and Li, Z. (2017) Tetrazole-based probes for integrated phenotypic screening, affinity-based proteome profiling, and sensitive detection of a cancer biomarker. *Angew. Chem. Int. Ed.* 56, 15044–15048
19. Ostrem, J. M., Peters, U., Sos, M. L., Wells, J. A., and Shokat, K. M. (2013) K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* 503, 548–551
20. Fell, J. B., Fischer, J. P., Baer, B. R., Blake, J. F., Bouhana, K., Briere, D. M., et al. (2020) Identification of the clinical development candidate MRTX849, a covalent KRASG12C inhibitor for the treatment of cancer. *J. Med. Chem.* 63, 6679–6693
21. Kuljanin, M., Mitchell, D. C., Schweppe, D. K., Gikandi, A. S., Nusinow, D. P., Bulloch, N. J., et al. (2021) Reimagining high-throughput profiling of reactive cysteines for cell-based screening of large electrophile libraries. *Nat. Biotechnol.* 39, 630–641
22. Vinogradova, E. V., Zhang, X., Remillard, D., Lazar, D. C., Suciu, R. M., Wang, Y., et al. (2020) An activity-guided map of electrophile-cysteine interactions in primary human T cells. *Cell* 182, 1009–1026.e29
23. Bar-Peled, L., Kemper, E. K., Suciu, R. M., Vinogradova, E. V., Backus, K. M., Horning, B. D., et al. (2017) Chemical proteomics identifies druggable vulnerabilities in a genetically defined cancer. *Cell* 171, 696–709.e23
24. Abegg, D., Frei, R., Cerato, L., Prasad Hari, D., Wang, C., Waser, J., et al. (2015) Proteome-wide profiling of targets of cysteine reactive small molecules by using ethynyl benziodoxolone reagents. *Angew. Chem.* 127, 11002–11007
25. Grossman, E. A., Ward, C. C., Spradlin, J. N., Bateman, L. A., Huffman, T. R., Miyamoto, D. K., et al. (2017) Covalent ligand discovery against druggable hotspots targeted by anti-cancer natural products. *Cell Chem. Biol.* 24, 1368–1376.e4
26. Wang, C., Weerapana, E., Blewett, M. M., and Cravatt, B. F. (2014) A chemoproteomic platform to quantitatively map targets of lipid-derived electrophiles. *Nat. Methods* 11, 79–85

27. Zambaldo, C., Vinogradova, E. V., Qi, X., Iaconelli, J., Suciu, R. M., Koh, M., et al. (2020) 2-Sulfonylpyridines as tunable, cysteine-reactive electrophiles. *J. Am. Chem. Soc.* 142, 8972–8979
28. [preprint] Cravatt, B., Njomen, E., Hayward, R., DeMeester, K., Ogasawara, D., Dix, M., et al. (2023) Comprehensive mapping of electrophilic small molecule-protein interactions in human cells. *ChemRxiv*. <https://doi.org/10.26434/chemrxiv-2023-s446n>
29. Zhang, J., Simpson, C. M., Berner, J., Chong, H. B., Fang, J., Ordulu, Z., et al. (2023) Systematic identification of anticancer drug targets reveals a nucleus-to-mitochondria ROS-sensing pathway. *Cell* 186, 2361–2379. e25
30. Henning, N. J., Manford, A. G., Spradlin, J. N., Brittain, S. M., Zhang, E., McKenna, J. M., et al. (2022) Discovery of a covalent FEM1B recruiter for targeted protein degradation applications. *J. Am. Chem. Soc.* 144, 701–708
31. Henning, N. J., Boike, L., Spradlin, J. N., Ward, C. C., Liu, G., Zhang, E., et al. (2022) Deubiquitinase-targeting chimeras for targeted protein stabilization. *Nat. Chem. Biol.* 18, 412–421
32. Toriki, E. S., Papatzimas, J. W., Nishikawa, K., Dovala, D., Frank, A. O., Hesse, M. J., et al. (2023) Rational chemical design of molecular glue degraders. *ACS Cent. Sci.* 9, 915–926
33. Boike, L., Cioffi, A. G., Majewski, F. C., Co, J., Henning, N. J., Jones, M. D., et al. (2021) Discovery of a functional covalent ligand targeting an intrinsically disordered cysteine within MYC. *Cell Chem. Biol.* 28, 4–13. e17
34. Lazear, M. R., Remsberg, J. R., Jaeger, M. G., Rothamel, K., Her, H. L., DeMeester, K. E., et al. (2023) Proteomic discovery of chemical probes that perturb protein complexes in human cells. *Mol. Cell* 83, 1725–1742. e12
35. Kathman, S. G., Koo, S. J., Lindsey, G. L., Her, H. L., Blue, S. M., Li, H., et al. (2023) Remodeling oncogenic transcriptomes by small molecules targeting NONO. *Nat. Chem. Biol.* 19, 825–836

36. Darabedian, N., Ji, W., Fan, M., Lin, S., Seo, H. S., Vinogradova, E. V., et al. (2023) Depletion of creatine phosphagen energetics with a covalent creatine kinase inhibitor. *Nat. Chem. Biol.* 19, 815–824
37. Shuken, S. R., McAlister, G. C., Barshop, W. D., Canterbury, J. D., Bergen, D., Huang, J., et al. (2023) Deep proteomic compound profiling with the Orbitrap ascend Tribrid mass spectrometer using tandem mass tags and real-time search. *Anal. Chem.* 95, 15180–15188
38. [preprint] Desai, H., Ofori, S., Boatner, L., Yu, F., Villanueva, M., Ung, N., et al. (2023) Multi-omic stratification of the missense variant cysteinome. *bioRxiv*. <https://doi.org/10.1101/2023.08.12.553095>
39. Yang, K., Whitehouse, R. L., Dawson, S. L., Zhang, L., Martin, J. G., Johnson, D. S., et al. (2024) Accelerating multiplexed profiling of protein- ligand interactions: high-throughput plate-based reactive cysteine profiling with minimal input. *Cell Chem. Biol.* 31, 565–576.e4
40. Burton, N. R., Polasky, D. A., Shikwana, F., Ofori, S., Yan, T., Geiszler, D. J., et al. (2023) Solid-phase compatible silane-based cleavable linker enables custom isobaric quantitative chemoproteomics. *J. Am. Chem. Soc.* 145, 21303–21318
41. Yan, T., Desai, H. S., Boatner, L. M., Yen, S. L., Cao, J., Palafox, M. F., et al. (2021) SP3-FAIMS chemoproteomics for high-coverage profiling of the human cysteinome. *Chembiochem* 22, 1841–1851
42. Cao, J., Boatner, L. M., Desai, H. S., Burton, N. R., Armenta, E., Chan, N. J., et al. (2021) Multiplexed CuAAC suzuki-miyaura labeling for tandem activity-based chemoproteomic profiling. *Anal. Chem.* 93, 2610–2618
43. Desai, H. S., Yan, T., Yu, F., Sun, A. W., Villanueva, M., Nesvizhskii, A. I., et al. (2022) SP3-Enabled rapid and high coverage chemoproteomic identification of cell-state-dependent redox-sensitive cysteines. *Mol. Cell. Proteomics* 21, 100218
44. Fu, L., Jung, Y., Tian, C., Ferreira, R. B., Cheng, R., He, F., et al. (2023) Nucleophilic covalent ligand discovery for the cysteine redoxome. *Nat. Chem. Biol.* 19, 1309–1319

45. [preprint] Won, S. J., Zhang, Y., Reinhardt, C. J., MacRae, N. S., DeMeester, K. E., Njomen, E., et al. (2024) Redirecting the pioneering function of FOXA1 with covalent small molecules. *bioRxiv*. [https://doi.org/ 10.1101/2024.03.21.586158](https://doi.org/10.1101/2024.03.21.586158)
46. Castellón, J. O., Ofori, S., Burton, N. R., Julio, A. R., Turmon, A. C., Armenta, E., et al. (2024) Chemoproteomics identifies state-dependent and proteoform-selective caspase-2 inhibitors. *J. Am. Chem. Soc.* 146, 14972–14988
47. Akter, S., Fu, L., Jung, Y., Conte, M. L., Lawson, J. R., Lowther, W. T., et al. (2018) Chemical proteomics reveals new targets of cysteine sulfinic acid reductase. *Nat. Chem. Biol.* 14, 995–1004
48. Wang, Y., Dix, M. M., Bianco, G., Remsberg, J. R., Lee, H. Y., Kalocsay, M., et al. (2019) Expedited mapping of the ligandable proteome using fully functionalized enantiomeric probe pairs. *Nat. Chem.* 11, 1113–1123
49. Liu, Z., Remsberg, J. R., Li, H., Njomen, E., DeMeester, K. E., Tao, Y., et al. (2024) Proteomic ligandability maps of spirocycle acrylamide stereoprobes identify covalent ERCC3 degraders. *J. Am. Chem. Soc.* 146, 10393–10406
50. Feldman, H. C., Merlini, E., Guijas, C., DeMeester, K. E., Njomen, E., Kozina, E. M., et al. (2022) Selective inhibitors of SARM1 targeting an allosteric cysteine in the autoregulatory ARM domain. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2208457119
51. Njomen, E., Hayward, R. E., DeMeester, K. E., Ogasawara, D., Dix, M. M., Nguyen, T., et al. (2024) Multi-tiered chemical proteomic maps of tryptophan acrylamide-protein interactions in cancer cells. *Nat. Chem.* 16, 1592–1604. <https://doi.org/10.1038/s41557-024-01601-1>
52. Basilaia, M., Chen, M. H., Secka, J., and Gustafson, J. L. (2022) Atropisomerism in the pharmaceutically relevant realm. *Acc. Chem. Res.* 55, 2904–2919
53. Cheng, Y., Li, G., Smedley, C. J., Giel, M. C., Kitamura, S., Woehl, J. L., et al. (2022) Diversity oriented clicking delivers  $\beta$ -substituted alkenyl sulfonyl fluorides as covalent human neutrophil elastase inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2208540119

54. Waas, M., Pereckas, M., Jones Lipinski, R. A., Ashwood, C., and Gundry, R. L. (2019) SP2: rapid and automatable contaminant removal from peptide samples for proteomic analyses. *J. Proteome Res.* 18, 1644–1656
55. Bian, Y., Zheng, R., Bayer, F. P., Wong, C., Chang, Y. C., Meng, C., et al. (2020) Robust, reproducible and quantitative analysis of thousands of proteomes by micro-flow LC-MS/MS. *Nat. Commun.* 11, 157
56. Budayeva, H. G., Ma, T. P., Wang, S., Choi, M., and Rose, C. M. (2024) Increasing the throughput and reproducibility of activity-based proteome profiling studies with hyperplexing and intelligent data acquisition. *J. Proteome Res.* 23, 2934–2947
57. [preprint] Biggs, G. S., Cawood, E. E., Vuorinen, A., McCarthy, W. J., Wil- ders, H., Riziotis, I. G., et al. (2024) Robust proteome profiling of cysteine- reactive fragments using label-free chemoproteomics. *bioRxiv*. [https:// doi.org/10.1101/2024.07.25.605137](https://doi.org/10.1101/2024.07.25.605137)
58. Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014) Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* 11, 319–324
59. Zougman, A., Selby, P. J., and Banks, R. E. (2014) Suspension trapping (STrap) sample preparation method for bottom-up proteomics analysis. *Proteomics* 14, 1006–1000
60. Hughes, C. S., Moggridge, S., Müller, T., Sorensen, P. H., Morin, G. B., and Krijgsveld, J. (2019) Single-pot, solid-phase-enhanced sample prepara- tion for proteomics experiments. *Nat. Protoc.* 14, 68–85
61. Moggridge, S., Sorensen, P. H., Morin, G. B., and Hughes, C. S. (2018) Extending the compatibility of the SP3 paramagnetic bead processing approach for proteomics. *J. Proteome Res.* 17, 1730–1740
62. Chang, A., Leutert, M., Rodriguez-Mias, R. A., and Ville´n, J. (2023) Auto- mated enrichment of phosphotyrosine peptides for high-throughput proteomics. *J. Proteome Res.* 22, 1868–1880

63. Müller, T., Kalxdorf, M., Longuespée, R., Kazdal, D. N., Stenzinger, A., and Krijgsveld, J. (2020) Automated sample preparation with SP3 for low- input clinical proteomics. *Mol. Syst. Biol.* 16, e9111
64. Yu, Y., Bekele, S., and Pieper, R. (2017) Quick 96FASP for high throughput quantitative proteome analysis. *J. Proteomics* 166, 1–7
65. Sielaff, M., Kuharev, J., Bohn, T., Hahlbrock, J., Bopp, T., Tenzer, S., et al. (2017) Evaluation of FASP, SP3, and ist protocols for proteomic sample preparation in the low microgram range. *J. Proteome Res.* 16, 4060–4072
66. [preprint] Takechi, S., Ngo, C., Burton, N., Villanueva, M., Boatner, L., Yu, F., et al. (2024) Silyl ether enables high coverage chemoproteomic interaction site mapping. *ChemRxiv*. <https://doi.org/10.26434/chemrxiv-2024-21r7b>
67. Bian, Y., Bayer, F. P., Chang, Y. C., Meng, C., Hoefer, S., Deng, N., et al. (2021) Robust microflow LC-MS/MS for proteome analysis: 38 000 runs and counting. *Anal. Chem.* 93, 3686–3690
68. Boatner, L. M., Palafox, M. F., Schweppe, D. K., and Backus, K. M. (2023) CysDB: a human cysteine database based on experimental quantitative chemoproteomics. *Cell Chem. Biol.* 30, 683–698.e3
69. Teo, G. C., Polasky, D. A., Yu, F., and Nesvizhskii, A. I. (2021) Fast deisotoping algorithm and its implementation in the msfragger search engine. *J. Proteome Res.* 20, 498–505  
*Mol Cell Proteomics* (2025) 24(9) 100898 17
70. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D., and Nesvizhskii, A. I. (2017) MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat. Methods* 14, 513–520
71. Yu, F., Haynes, S. E., and Nesvizhskii, A. I. (2021) IonQuant enables accurate and sensitive label-free quantification with FDR-controlled match- between-runs. *Mol. Cell. Proteomics* 20, 100077



72. Geiszler, D. J., Kong, A. T., Avtonomov, D. M., Yu, F., Leprevost, F. V., and Nesvizhskii, A. I. (2021) PTM-shepherd: analysis and summarization of post-translational and chemical modifications from open search results. *Mol. Cell. Proteomics* 20, 100018
73. Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., et al. (2012) A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* 30, 918–920
74. Holman, J. D., Tabb, D. L., and Mallick, P. (2014) Employing proteowizard to convert raw mass spectrometry data. *Curr. Protoc. Bioinformatics* 46, 13. 24.1–13.24.9
75. Mansuri, M. S., Bathla, S., Lam, T. T., Nairn, A. C., and Williams, K. R. (2024) Optimal conditions for carrying out trypsin digestions on complex proteomes: from bulk samples to single cells. *J. Proteomics* 297, 105109
76. Wu, X. N., Xi, L., Pertl-Obermeyer, H., Li, Z., Chu, L. C., and Schulze, W. X. (2017) Highly efficient single-step enrichment of low abundance phosphopeptides from plant membrane preparations. *Front. Plant Sci.* 8, 1673
77. Muneer, G., Chen, C.-S., Lee, T.-T., Chen, B.-Y., and Chen, Y.-J. (2024) A rapid one-pot workflow for sensitive microscale phosphoproteomics. *J. Proteome Res.* 23, 3294–3309
78. [preprint] Yang, D., Harry, S. A., Chong, H. B., Zhang, E., Nordenfelt, N. S., Chen, N., et al. (2024) Systematic targeting of protein complexes with molecular COUPLrs. *bioRxiv*. <https://doi.org/10.1101/2024.07.16.603666>
79. [preprint] Julio, A. R., Shikwana, F., Truong, C., Burton, N. R., Dominguez, E., Turmon, A. C., et al. (2023) Pervasive aggregation and depletion of host and viral proteins in response to cysteine-reactive electrophilic compounds. *bioRxiv*. <https://doi.org/10.1101/2023.10.30.564067>
80. Bekker-Jensen, D. B., Bernhardt, O. M., Hogrebe, A., Martinez-Val, A., Verbeke, L., Gandhi, T., et al. (2020) Rapid and site-specific deep phosphoproteome profiling by data-independent acquisition without the need for spectral libraries. *Nat. Commun.* 11, 787

81. Kawakami, K., Fujita, Y., Kato, T., Horie, K., Koie, T., Umezawa, K., et al. (2021) Diagnostic potential of serum extracellular vesicles expressing prostate-specific membrane antigen in urologic malignancies. *Sci. Rep.* 11, 15000
82. Shubin, A. V., Kollar, B., Dillon, S. T., Pomahac, B., Libermann, T. A., and Riella, L. V. (2019) Blood proteome profiling using aptamer-based technology for rejection biomarker discovery in transplantation. *Sci. Data* 6, 314
83. Burton, N. R., and Backus, K. M. (2024) Functionalizing tandem mass tags for streamlining click-based quantitative chemoproteomics. *Commun. Chem.* 7, 80
84. Vaidya, S. D., Heydari, B. S., Toenjes, S. T., and Gustafson, J. L. (2022) Approaches toward atropisomerically stable and conformationally pure diarylamines. *J. Org. Chem.* 87, 6760–6768
85. Johnston, H. E., Yadav, K., Kirkpatrick, J. M., Biggs, G. S., Oxley, D., Kramer, H. B., et al. (2022) Solvent precipitation SP3 (SP4) enhances recovery for proteomics sample preparation without magnetic beads. *Anal. Chem.* 94, 10320–10328
86. Batth, T. S., Tollenaere, M. A. X., Rüther, P., Gonzalez-Franquesa, A., Prabhakar, B. S., Bekker-Jensen, S., et al. (2019) Protein aggregation capture on microparticles enables multipurpose proteomics sample preparation. *Mol. Cell. Proteomics* 18, 1027–1035
87. Julio, A. R., Shikwana, F., Truong, C., Burton, N. R., Dominguez, E. R., 3rd, Turmon, A. C., et al. (2024) Delineating cysteine-reactive compound modulation of cellular proteostasis processes. *Nat. Chem. Biol.* [https:// doi.org/10.1038/s41589-024-01760-9](https://doi.org/10.1038/s41589-024-01760-9)
88. Ma, T. P., Izrael-Tomasevic, A., Mroue, R., Budayeva, H., Malhotra, S., Raisner, R., et al. (2023) AzidoTMT enables direct enrichment and highly multiplexed quantitation of proteome-wide functional residues. *J. Proteome Res.* 22, 2218–2231
89. Stopfer, L. E., Mesfin, J. M., Joughin, B. A., Lauffenburger, D. A., and White, F. M. (2020) Multiplexed relative and absolute quantitative immunopeptidomics reveals MHC I repertoire alterations induced by CDK4/6 inhibition. *Nat. Commun.* 11, 2760

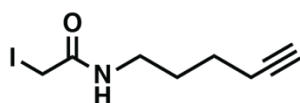
90. Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., et al. (2022) The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* 50, D543–D552
91. Collins, D. M., Conlon, N. T., Kannan, S., Verma, C. S., Eli, L. D., Lalani, A. S., et al. (2019) Preclinical characteristics of the irreversible pan-HER kinase inhibitor neratinib compared with lapatinib: implications for the treatment of HER2-positive and HER2-mutated breast cancer. *Cancers (Basel)* 11, 737
92. Yan, T., Palmer, A. B., Geiszler, D. J., Polasky, D. A., Boatner, L. M., Burton, N. R., et al. (2022) Enhancing cysteine chemoproteomic coverage through systematic assessment of click chemistry product fragmentation. *Anal. Chem.* 94, 3800–3810

## Chapter 2

### Supporting Information

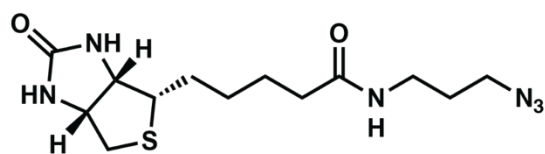
#### Supplementary Figures

A

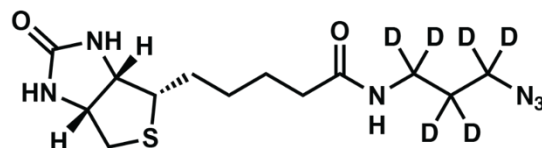


Iodoacetamide Alkyne  
(IAA)

B



(Light) Biotin Azide



Heavy Biotin Azide

**Figure S1. Structures of capture reagents used in this study.** A) Structure of pan-reactive iodoacetamide alkyne (IAA)<sup>1</sup>. B) Structures of biotin azide enrichment handles<sup>2,3</sup>. Biotin azide and light biotin azide refers to top structure and is used in all experiments. Heavy biotin azide is used for ABPP experiments.

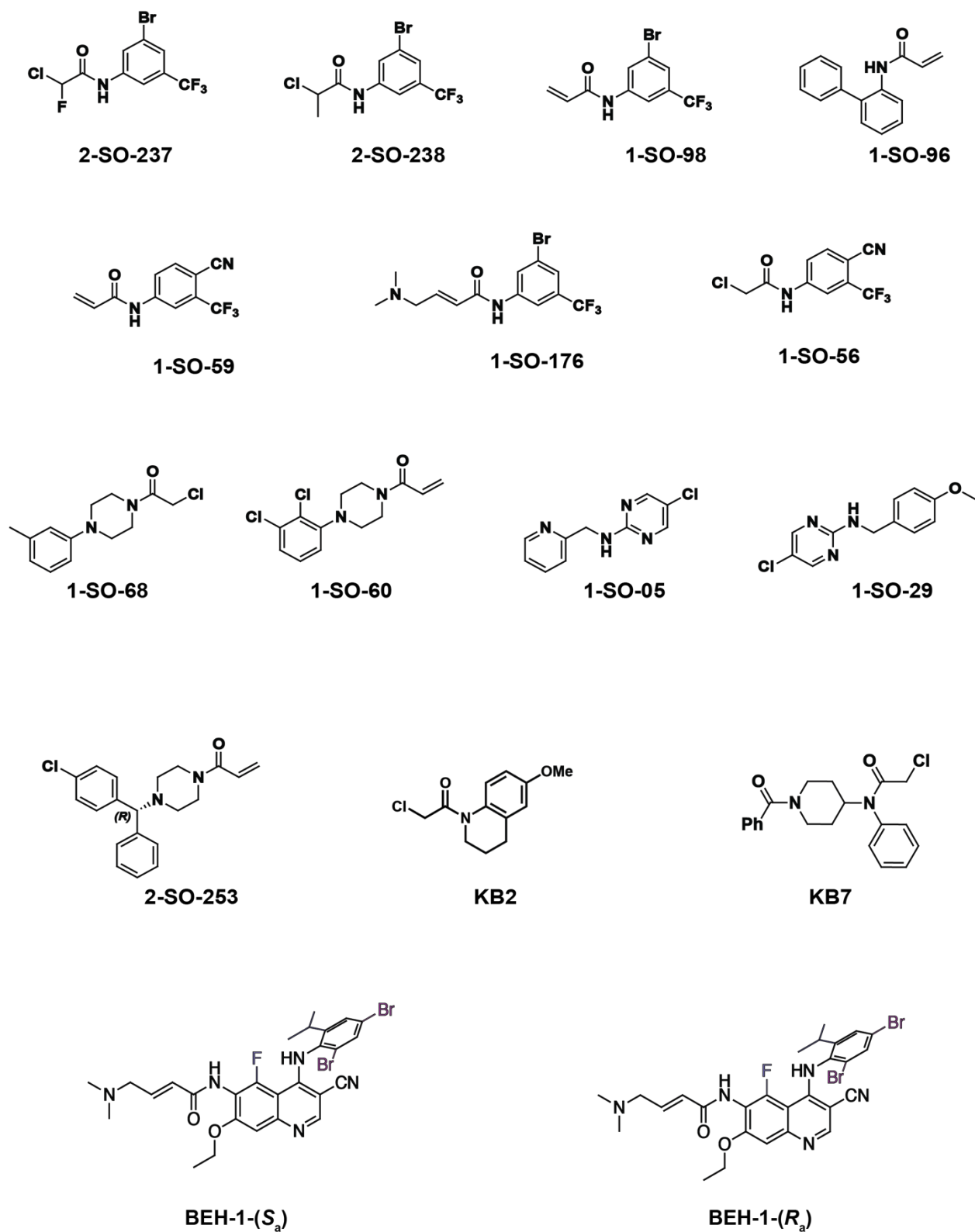
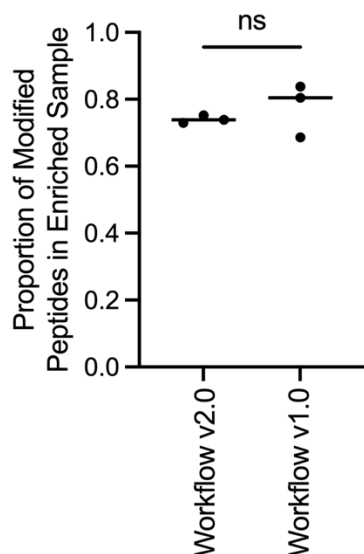
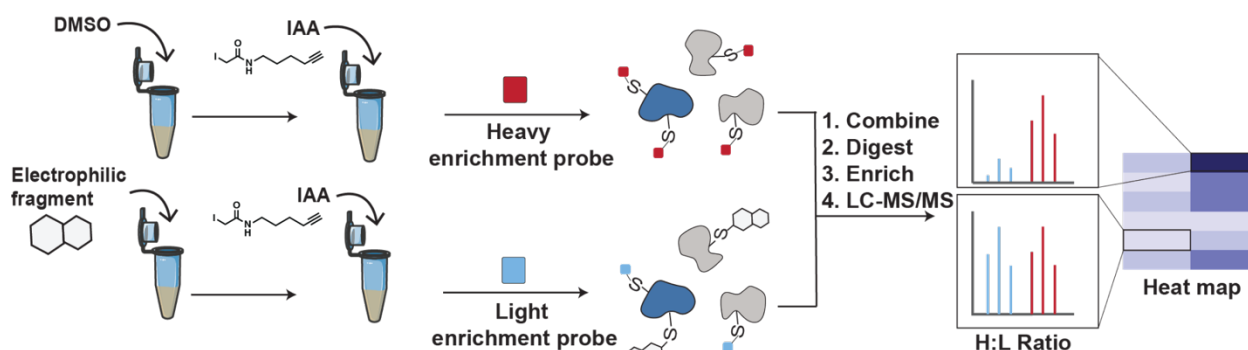


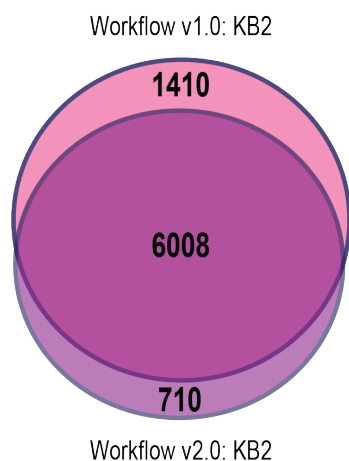
Figure S2. Structures of electrophilic compounds used in this study.



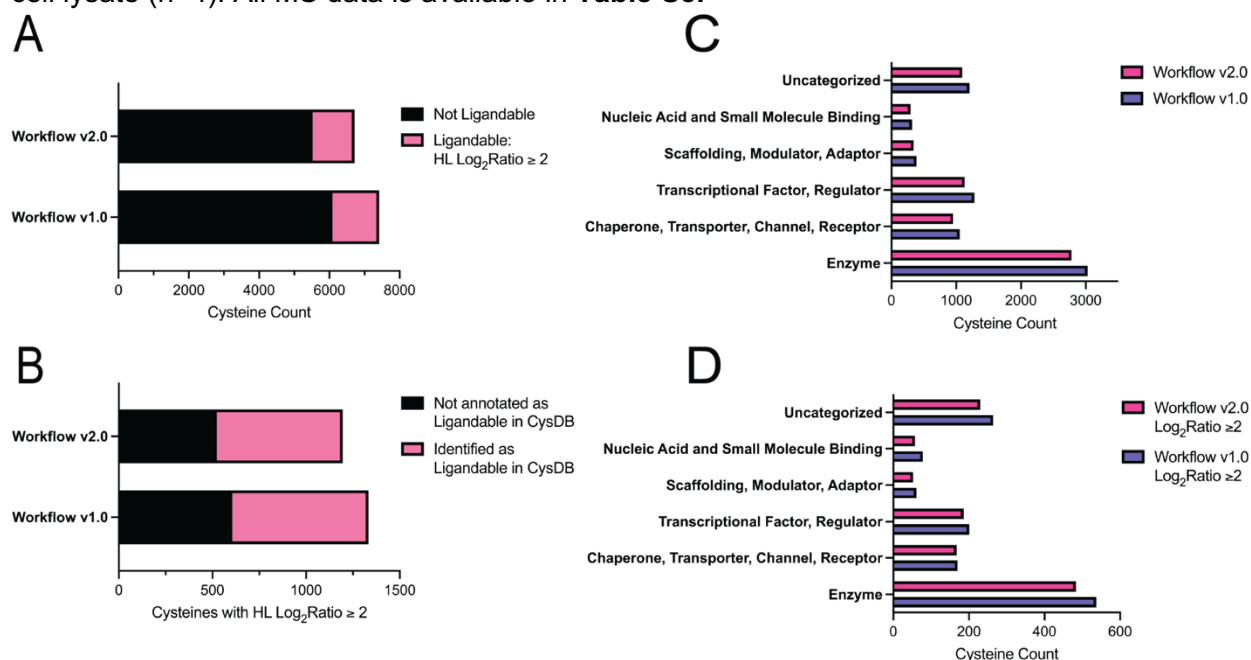
**Figure S3. Proportion of enriched modified peptides.** Analysis for proportion of modified peptides captured by enrichment for each workflow. Statistical significance was calculated by performing an unpaired Student's t-test.  $p < 0.05$  for significance cutoff,  $n=3$ . All MS data is available in **Table S4**.



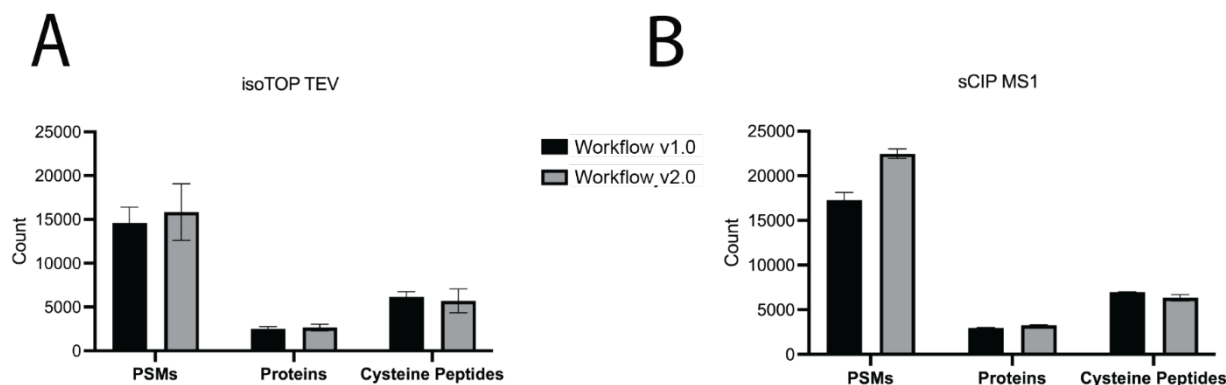
**Figure S4. Cysteine chemoproteomic electrophilic compound screening workflow.** Jurkat cell lysates are treated with either vehicle (DMSO) or electrophile-containing compound, for a short incubation period (1h) followed by labeling (1h) with the pan-cysteine reactive probe iodoacetamide alkyne (IAA). After alkyne functionalization, the samples are subjected to click conjugation with isotopically enriched biotin azide capture agents, pooled, digested with trypsin, enriched on neutravidin resin, and captured peptides analyzed by LC-MS/MS.



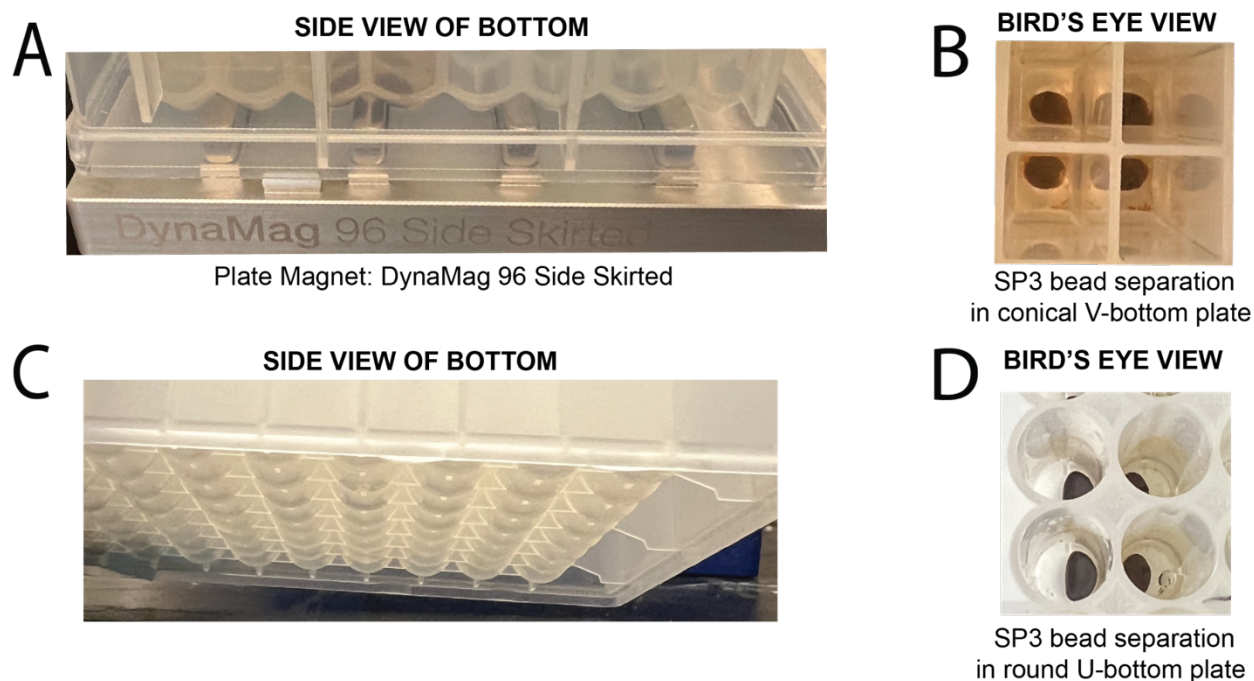
**Figure S5. Comparison of coverage obtained using v1.0 versus v2.0 workflows for cysteine chemoproteomics ligandability studies.** A) Analysis of unique cysteine identifier coverage between both workflows for KB2 (500  $\mu$ M) competitive chemoproteomic analysis in whole Jurkat cell lysate (n=4). All MS data is available in **Table S5**.



**Figure S6. Comparison of workflows for cysteine chemoproteomics ligandability studies with model compound KB2.** A) Analysis for proportion of unique cysteines showing a  $\text{Log}_2(\text{H/L})$  Ratio of 2 or greater for each workflow in relation to all unique cysteines identified in each experiment. B) Analysis for proportion of cysteines annotated in CysDB<sup>4</sup> as ligandable for each workflow in relation to all ligandable cysteines ( $\text{Log}_2(\text{H/L})$  Ratio  $\geq 2$ ) identified in each experiment. C) Analysis for proportion of cysteine types as categorized by CysDB for both workflow conditions. D) Analysis for proportion of ligandable cysteine types as categorized by CysDB for both workflow conditions. Dataset used for these analyses are of KB2 (500  $\mu$ M) treated competitive chemoproteomic analysis in whole Jurkat cell lysate, n=4 for each group. All MS data is available in **Table S5**.

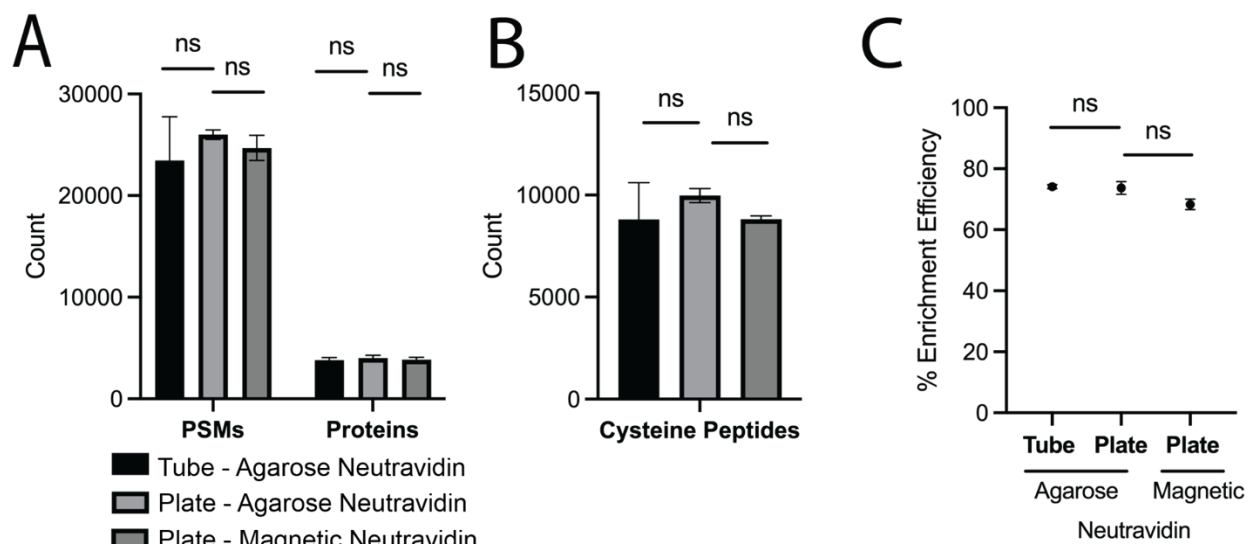


**Figure S7. Cysteine chemoproteomic workflow is compatible with multiple enrichment-handle reagents. A,B)** Comparison of coverage for samples prepared with either v1.0 versus v2.0 workflows for Jurkat cell lysates treated with KB2 (500  $\mu$ M) or vehicle (DMSO),  $n=2$ , with A) samples prepared using previously reported SP3-cleanup<sup>5</sup> enabled workflow for isoTOP-ABPP<sup>1</sup> using TEV-protease cleavable biotin-azide (TEV tags), with enrichment on streptavidin resin and cleaved off by TEV protease. **B)** Samples prepared using silane-based cleavable isotopically labeled proteomics (sCIP) reagents<sup>6</sup> with enrichment on streptavidin resin followed by low pH elution. All MS data is available in **Table S5**.

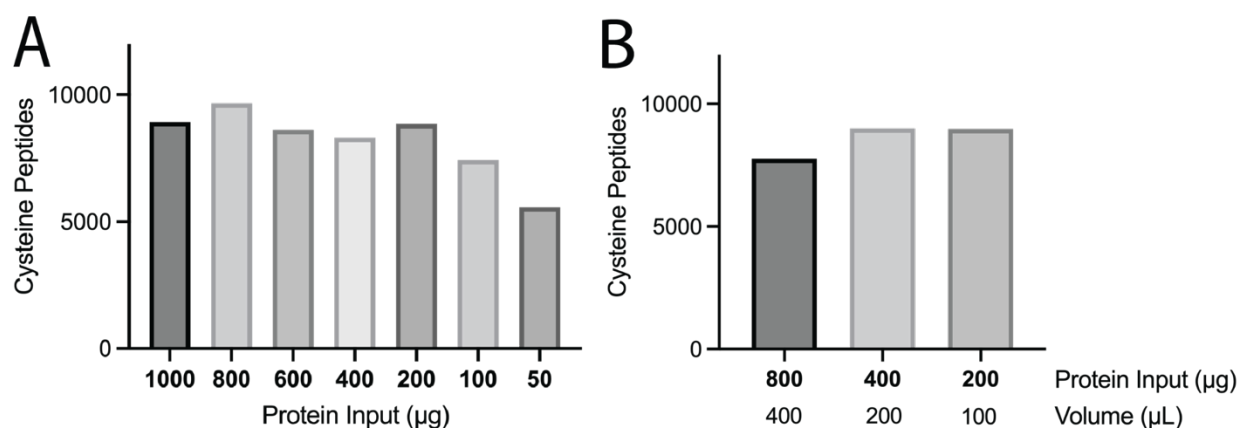


**Figure S8. Comparison of plates and magnets for CySP-96. A)** Conical V-bottom 96-well plate with a DynaMag 96 side skirted plate magnet (Invitrogen Cat. No. 12027). **B)** Birds eye view of the bead separation in a conical V-bottom plate using a Dynamag 96 side skirted plate magnet. **C)** Round U-bottom 96-well plate (MilliporeSigma™ Supelco™ SPE 96-Deep Square Well Collection Plate, Cat. No. 11-100-3690). **D)** Birds eye view of the bead separation in a round U-bottom plate using a Dynamag 96 side skirted plate magnet.



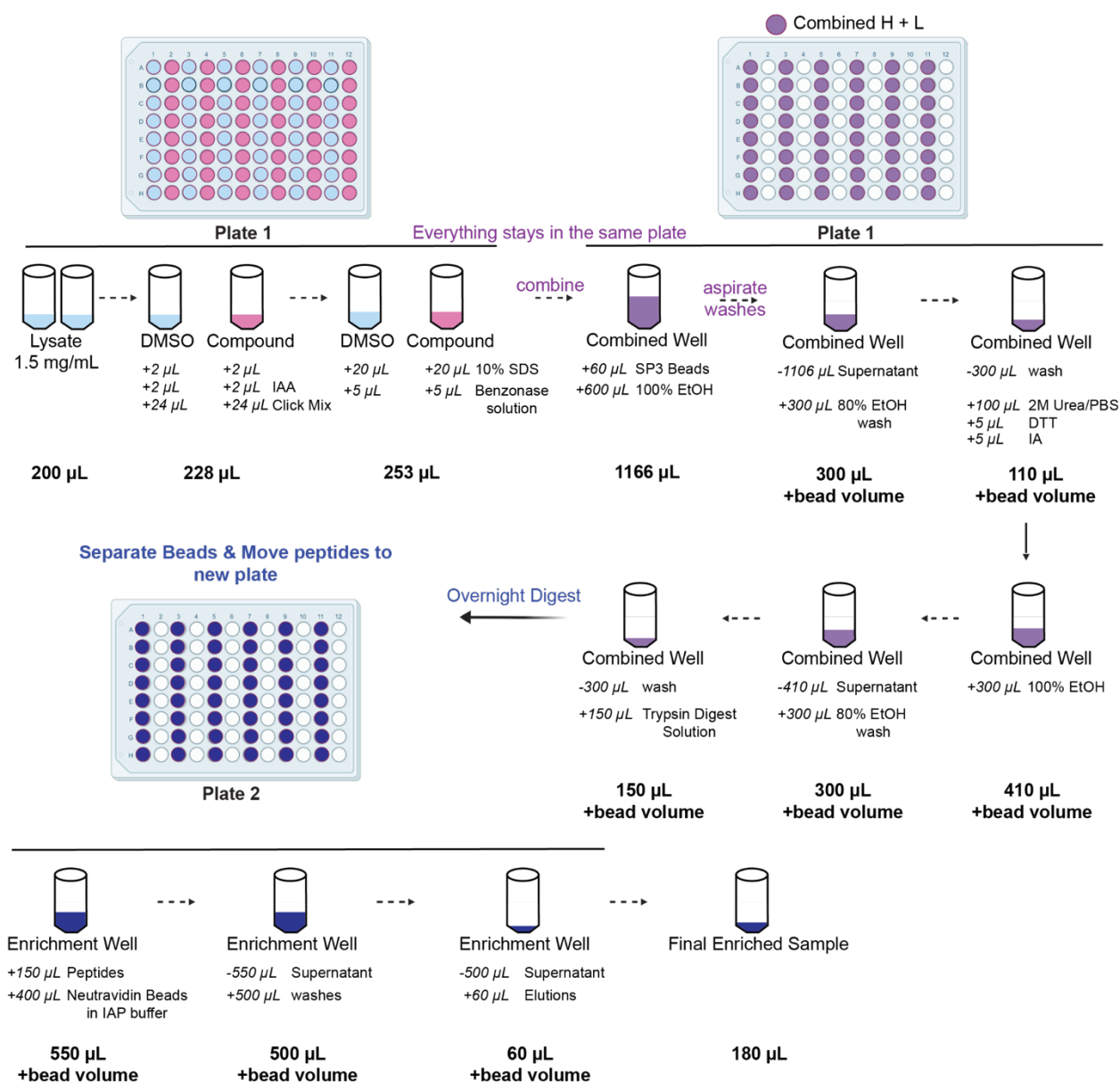


**Figure S9. Comparing coverage for samples prepared in microcentrifuge tubes versus 96-well plate format, including for neutravidin enrichment. A-C)** Comparison of coverage obtained using the indicated workflows, with “tube” corresponding to samples prepared with workflow v2.0 in 1.5 mL snap cap microcentrifuge tube and “plate” corresponding to samples prepared using v2.0 workflow in a 96-well plate. To test the impact of also performing neutravidin enrichment in the 96-well plate, samples were also prepared comparing enrichment plate-based sample preparation followed by neutravidin enrichment in snap cap microcentrifuge tube versus complete sample prep in the 96-well plate, including enrichment using magnetic neutravidin. **A)** Comparison of PSMs and proteins identified, **B)** Comparison of biotinylated cysteine peptides. **C)** Comparison of enrichment efficiency. Statistical significance was calculated by performing an unpaired Student’s t-test.  $p < 0.05$  for significance cutoff,  $n=2$ . All MS data is available in **Table S6**.

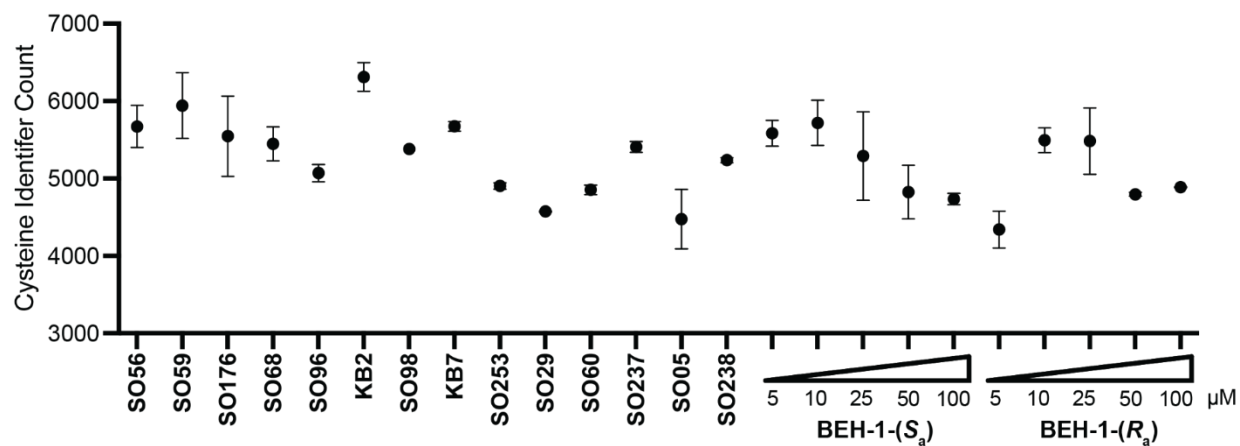


**Figure S10. Assessing the impact of protein input on cysteine coverage. A)** Coverage comparison of cysteine peptides enriched, for cysteine chemoproteomic samples prepared using the indicated protein input (Jurkat whole cell lysates, 50-1000 µg) at constant volume (200 µL) were labeled with IAA (200 µM) followed by click conjugation to biotin azide conjugation and chemoproteomic sample preparation analyzed for cysteine coverage,  $n=1$ . **B)** Coverage comparison of cysteine peptides for samples prepared using the indicated protein input (Jurkat whole cell lysates, 200-800 µg) while varying the input volume to maintain constant protein concentration (2 mg/mL). Input labeled with IAA (200 µM) followed by click conjugation to biotin

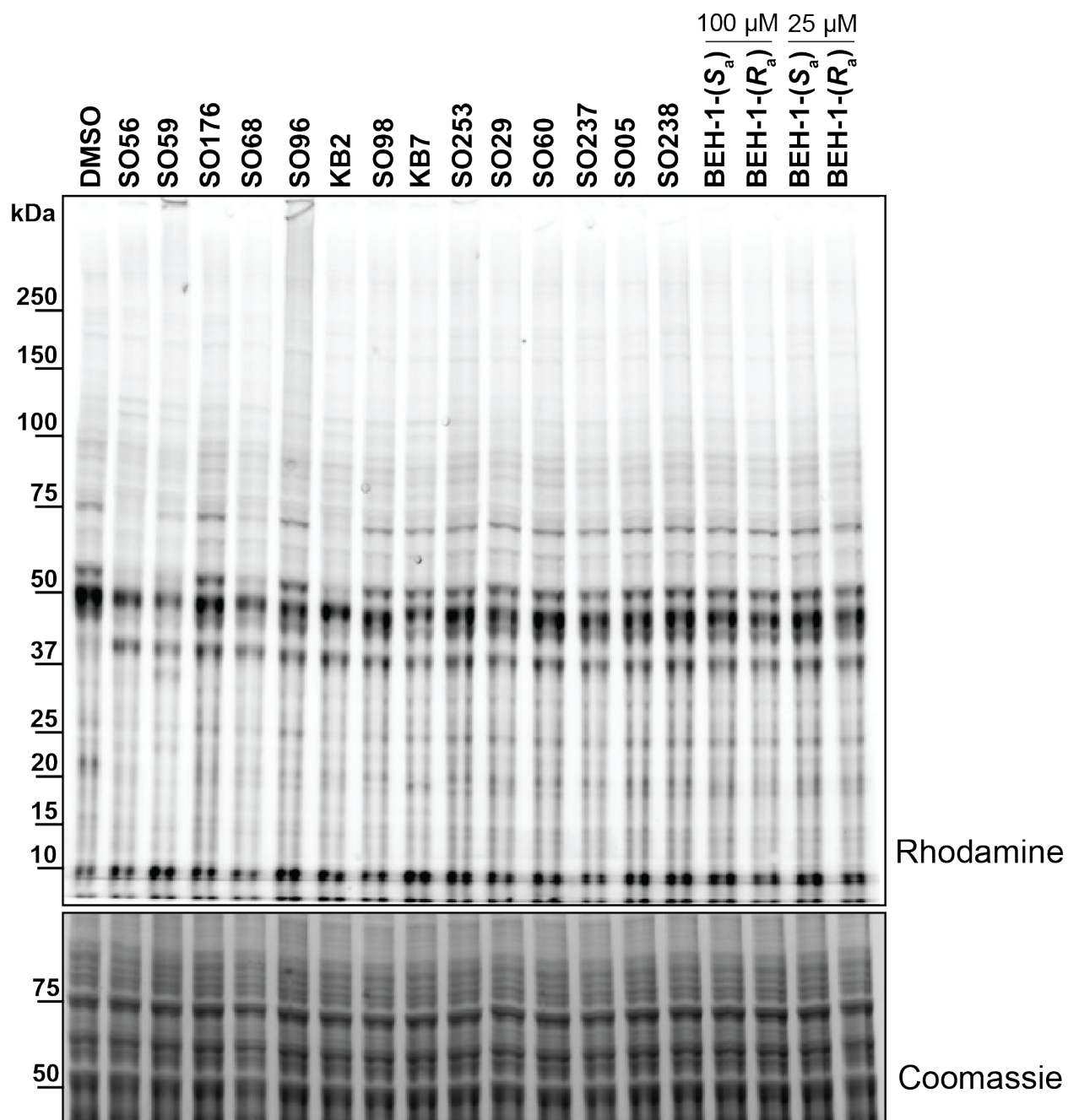
azide conjugation and chemoproteomic sample preparation analyzed for cysteine coverage, n=1. All MS data is available in **Table S6**.



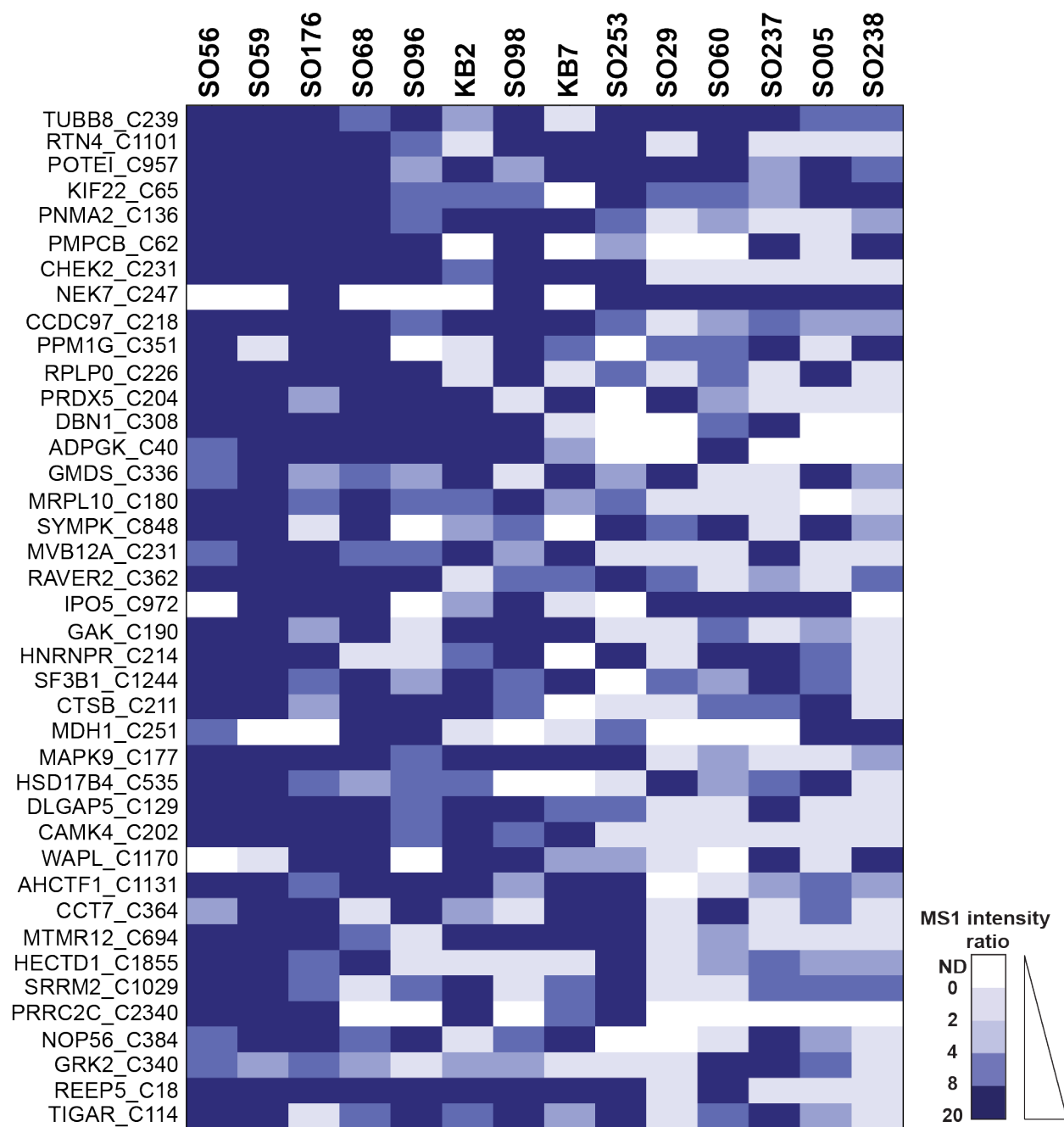
**Figure S11. Detailed workflow for CySP3-96 in plate.** Total volume in the well is shown for each step in bold and plate changes are noted where they occur. Note that for multiple washes & for multiple elutions only one is shown to represent the volume used per wash/elution.



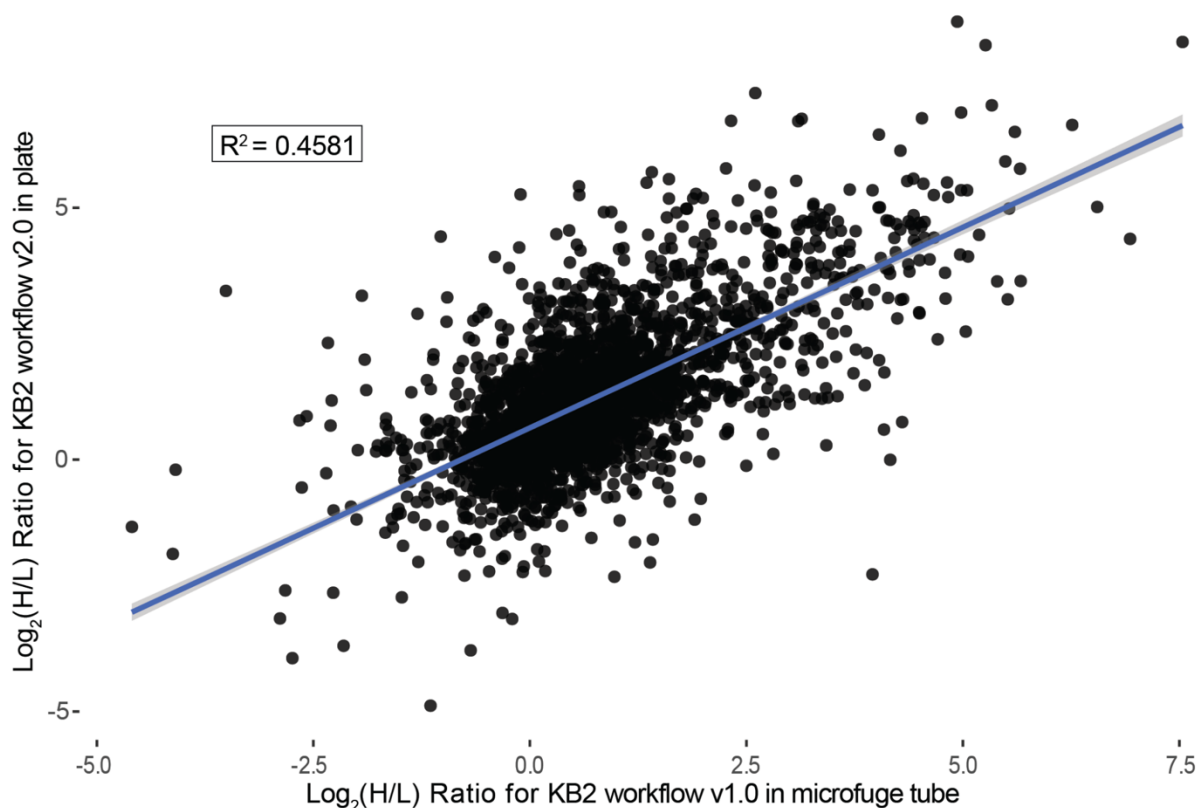
**Figure S12. Coverage and variability of unique cysteine identifiers for all compounds screened in plate.** All compounds screened at 500  $\mu$ M (unless otherwise stated) in Jurkat whole cell lysates in replicates of 2. All MS data is available in **Table S6**.



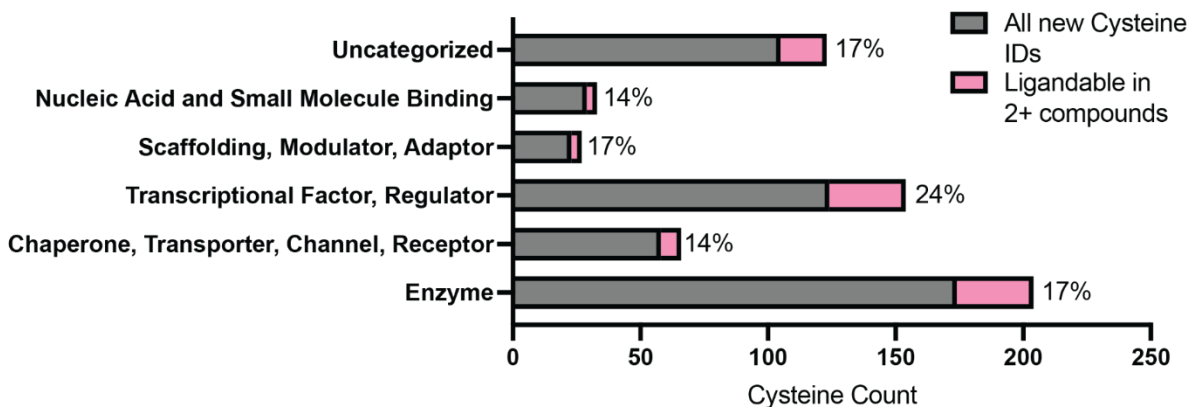
**Figure S13. General reactivity of each screened compound as assessed by in-gel fluorescence.** Jurkat whole cell lysates were subjected to the indicated compounds (500 μM, 1h, unless otherwise stated), followed by IA-Rhodamine (5 μM, 20 min), separation by SDS-PAGE gel and visualization by in-gel fluorescence and coomassie InstantBlue stain.



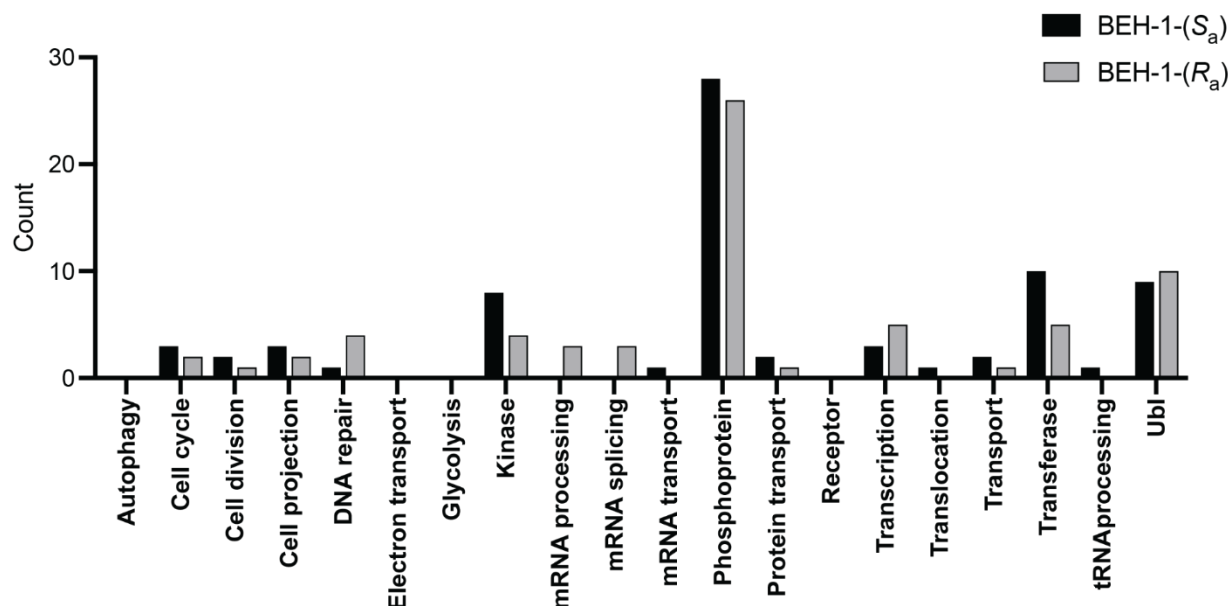
**Figure S14. Top reactive cysteines across compounds screened in plate.** All compounds treated at 500  $\mu$ M (unless otherwise stated) for 1 hour in Jurkat whole cell lysates. Jurkat whole cell lysates were subjected to the indicated compounds (500  $\mu$ M, 1h) followed by chemoproteomic sample preparation. All MS data is available in **Table S6**.



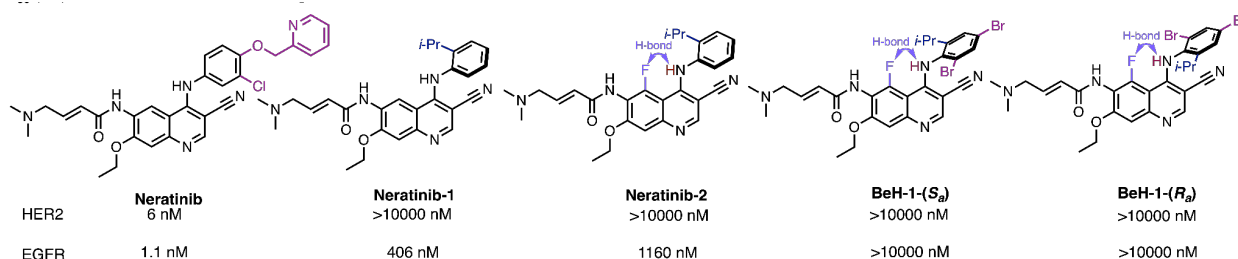
**Figure S15. Assessing concordance of ratios for samples prepared with plate workflow v2.0 versus microcentrifuge tube with workflow v1.0.** Jurkat cell whole cell lysate treated with KB2 (500  $\mu\text{M}$ ), comparing plate-generated data to workflow v1.0 in microfuge tube dataset previously shown in Figure 2. Data was filtered for cysteine identifiers present in both replicates and for each cysteine identifier ratio with a standard deviation of less than two. For workflow v1.0 prepared in a microcentrifuge tube,  $n=4$ . For workflow v2.0 in plate  $n=2$ . All MS data is available in **Table S5** (microfuge tube) and **Table S6** (plate).



**Figure S16. Stratification of protein categories for cysteines not previously identified in CysDB<sup>4</sup>.** Protein categorization was conducted as reported previously<sup>4</sup> for all newly identified cysteines and the subset that are liganded ( $\text{Log}_2(\text{H/L})$  ratio  $\geq 2$ ) by 2 or more compounds. All MS data is available in **Table S6**.

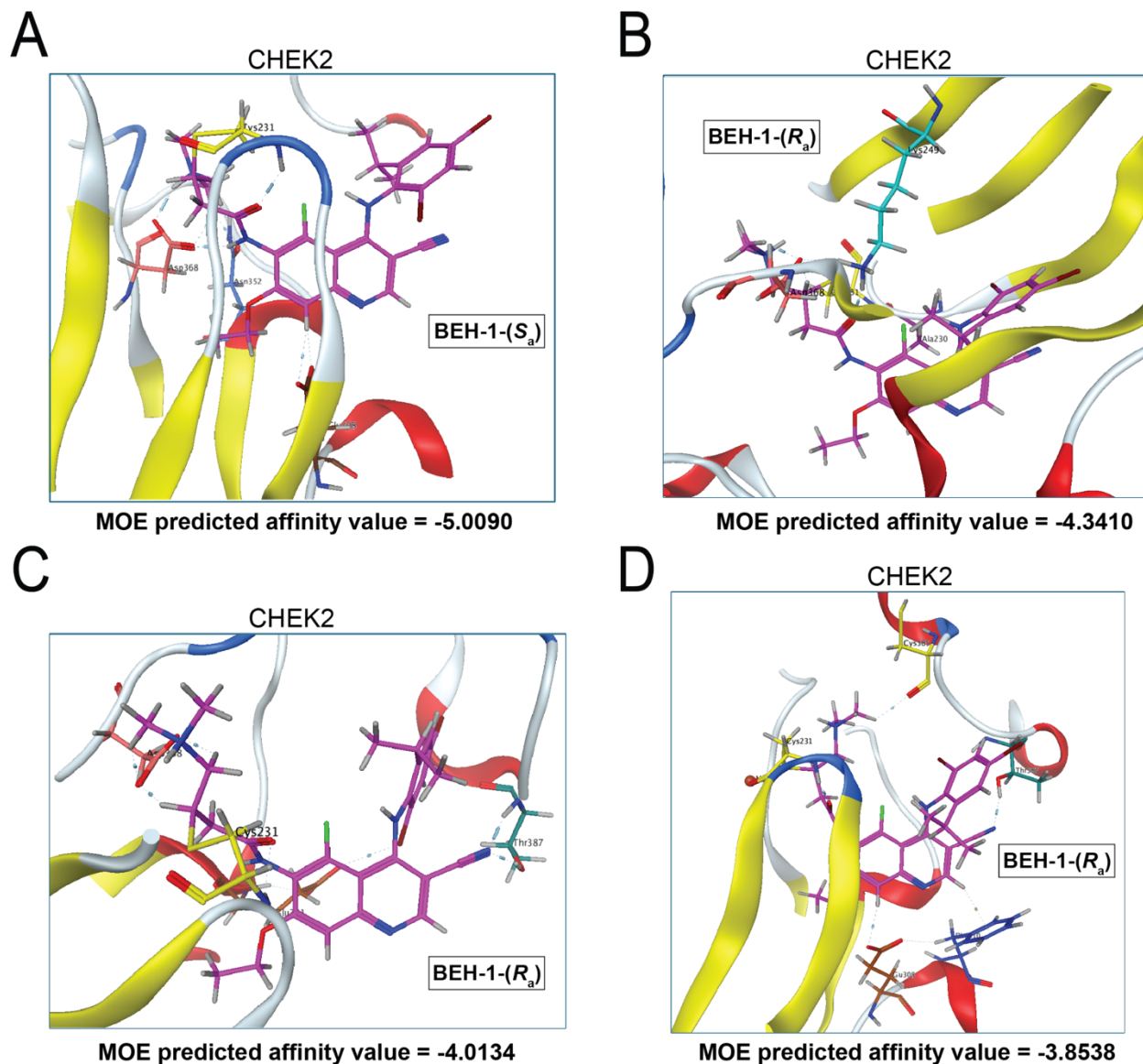


**Figure S17 . UniProtKB annotations for types of proteins showing stereoselectivity with atropisomers BEH-1-(S<sub>a</sub>) and BEH-1-(R<sub>a</sub>).** Protein categorization based on UniprotKB annotations for cysteines liganded by either BEH-1-(S<sub>a</sub>) (black bars) or BEH-1-(R<sub>a</sub>) (gray bars). All MS data is available in **Table S6**.



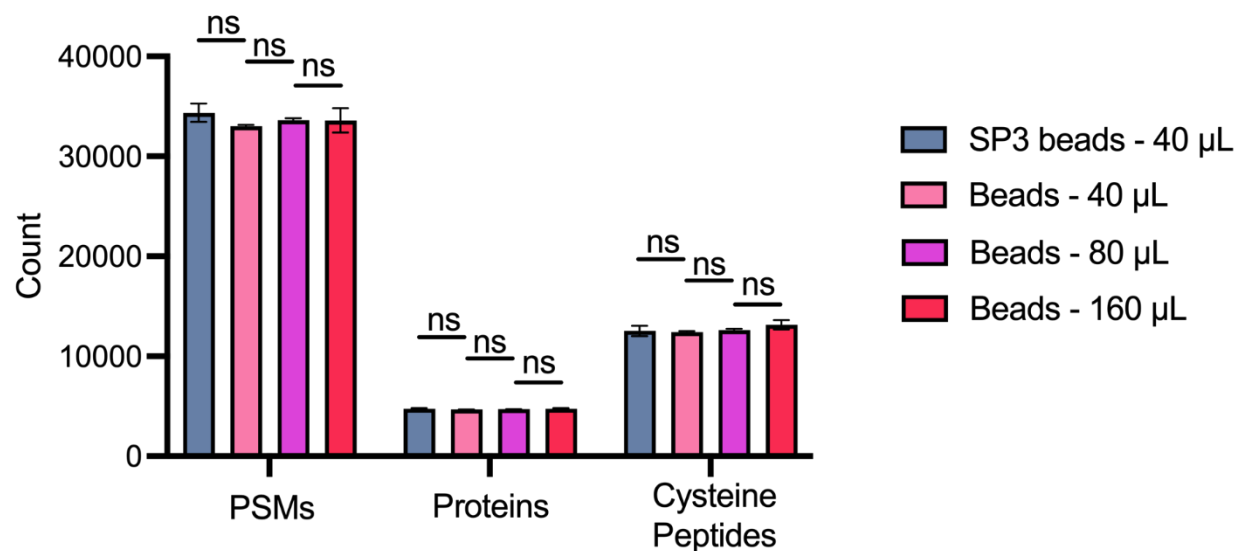
**Figure S18. IC<sub>50</sub> values for Neratinib and analogs for known kinase targets.** The IC<sub>50</sub> values for Neratinib's primary targets were determined from the literature<sup>7</sup>. The rest of the IC<sub>50</sub> values were obtained from Thermo Fisher Scientific using their Z'-LYTE assay.



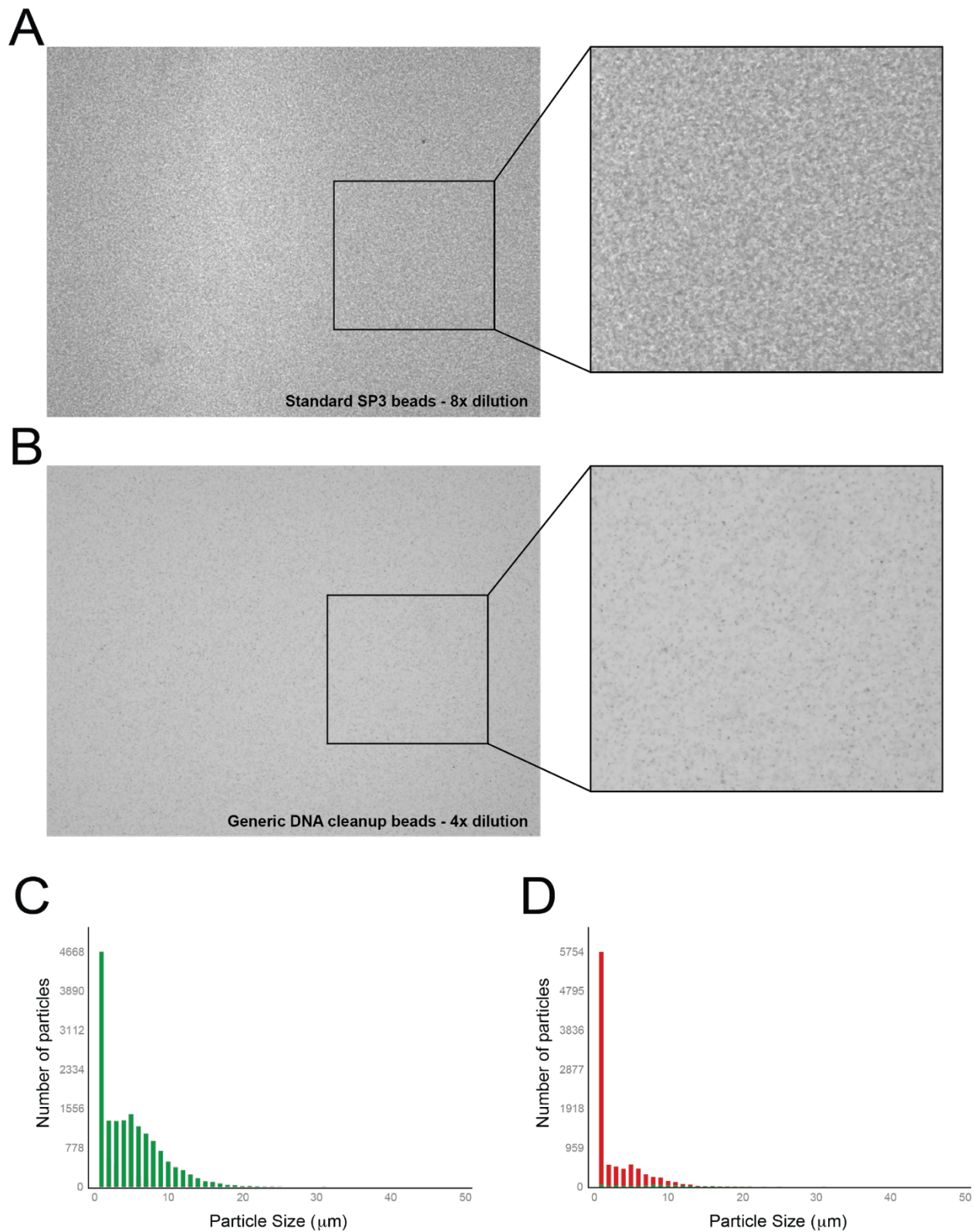


**Figure S19. Binding poses for the BEH-1 atropisomeric compound.** A) Binding pose showing the S atropisomer compound with an MOE predicted affinity value of -5.0090. BEH-1-(S<sub>a</sub>) is making a covalent bond with Cys231 in the active site (Glycine-rich loop) of CHEK2 along with various other hydrogen-bonding interactions in the hinge region and active site of the kinase. B) Binding pose showing the R atropisomer compound with an MOE predicted affinity value of -4.3410. BEH-1-(R<sub>a</sub>) is making a covalent bond with Cys231 in the active site of CHEK2 along with various other hydrogen-bonding interactions in the DFG motif and glycine-rich loop of the kinase. C) Binding pose showing the R atropisomer compound with an MOE predicted affinity value of -4.0134. BEH-1-(R<sub>a</sub>) is making a covalent bond with Cys231 in the active site of CHEK2 along with various other hydrogen-bonding and electrostatic interactions in the hinge region and active site of the kinase. D) Binding pose showing the R atropisomer compound with an MOE predicted affinity value of -3.8538. BEH-1-(R<sub>a</sub>) is making a covalent bond with Cys231 in the active site of CHEK2 along with various other hydrogen-bonding interactions in the DFG motif and active site of the kinase. All docking figures generated with CHEK2 PDB structure 2XBJ in MOE 2020.



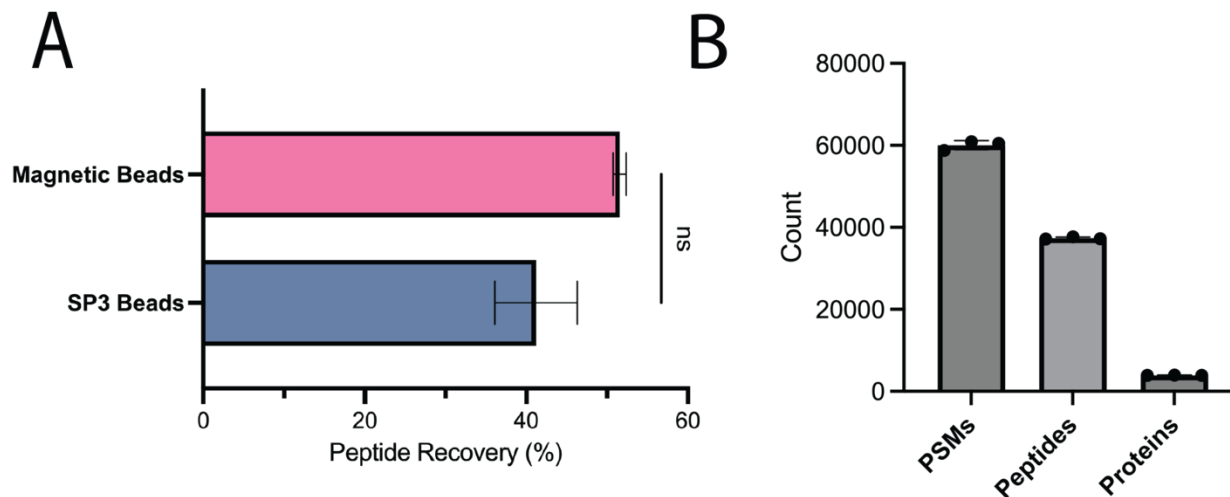


**Figure S20. Comparing cysteine chemoproteomic coverage obtained with standard SP3 resin versus generic DNA cleanup magnetic beads.** Chemoproteomic samples were prepared using Jurkat whole cell lysates capped with IAA (200 µM) clicked to biotin azide were subjected to chemoproteomic sample preparation in microcentrifuge tubes (n=2/condition) using the indicated volumes of resin, following the workflow v2.0, with digested peptides directly enriched on neutravidin beads after and resin protein cleanup. Statistical significance was calculated by performing Student's t-test.  $p < 0.05$  for significance cutoff, n=2. All MS data is available in **Table S7**.

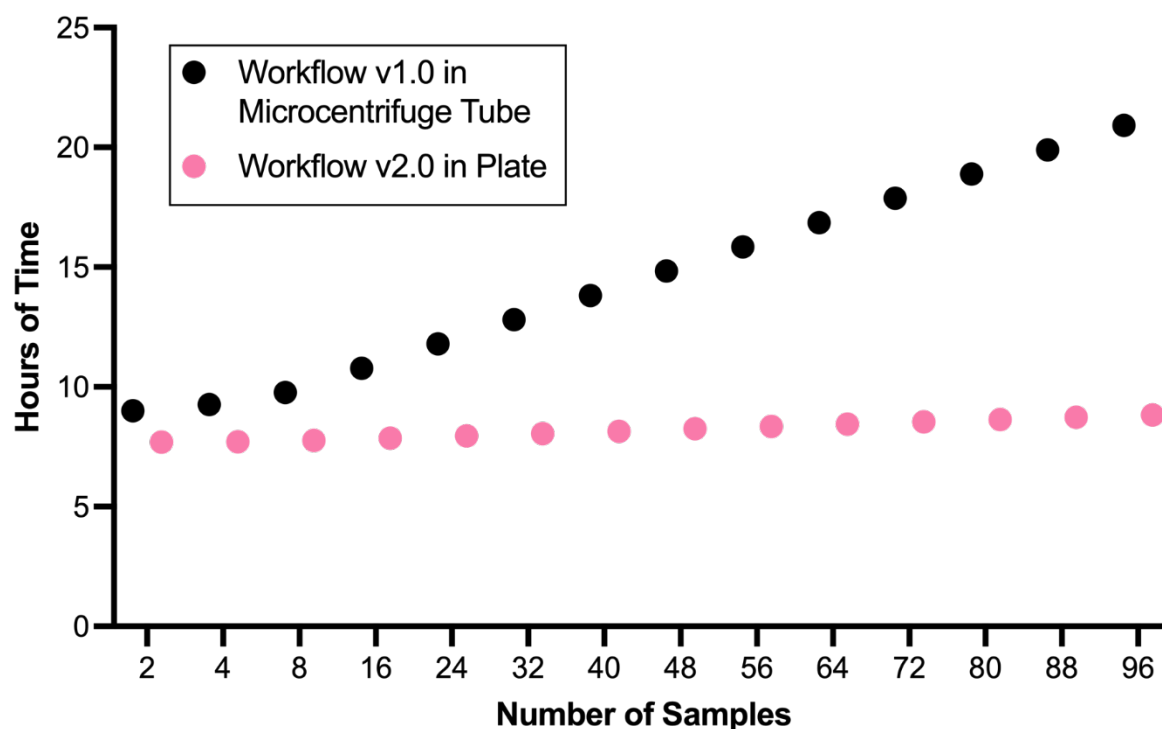


**Figure S21. Comparison of bead size and concentration using Countess.** A) Standard SP3 beads diluted 8x (10  $\mu$ L bead slurry in 80  $\mu$ L water) and imaged using Countess. B) Generic DNA cleanup beads diluted 4x (10  $\mu$ L bead slurry in 40  $\mu$ L water) and imaged using Countess C)

Particle size analysis for standard SP3 beads diluted 8x as generated by Countess. D) Particle size analysis for generic DNA cleanup beads diluted 4x as generated by Countess.



**Figure S22. Coverage for protein and peptide cleanup done with generic DNA cleanup magnetic beads.** A) Jurkat whole cell lysates (200  $\mu$ g) subjected to protein and peptide cleanup with standard SP3 magnetic beads and generic DNA cleanup magnetic beads. Percent recovery calculated based on total  $\mu$ g of peptides measured post-digest and pre-peptide cleanup. Peptide concentrations were measured pre- and post-cleanup for calculation of peptide recovery, using the Pierce Quantitative Peptide Concentration Assay. Statistical significance was calculated by performing Student's t-test.  $p < 0.05$  for significance cutoff,  $n=2$ . B) HEK293T whole cell lysates subjected to protein- and peptide-cleanup with on-resin digest, using generic DNA cleanup magnetic beads, followed by LC-MS/MS analysis. All MS data is available in **Table S7**.



**Figure S23. Estimated time saving with CySP-96 as compared to cysteine chemoproteomic preparation Workflow 1.0 in microcentrifuge tubes.** Data calculated based on the time taken to prepare 2-8 samples.

## Supplementary Tables

**Table S1.** Scaling of digest volumes. Table shows the relationship of digest volume adjustments as it relates to digest concentration and subsequent acetonitrile volumes to achieve  $\geq 95\%$  acetonitrile v/v for optimal peptide binding to SP3 beads.

| Digest Volume ( $\mu\text{L}$ ) | Digest Concentration (mg/mL) | Acetonitrile Volume to add ( $\mu\text{L}$ ) | Total Volume of Peptide Binding ( $\mu\text{L}$ ) |
|---------------------------------|------------------------------|----------------------------------------------|---------------------------------------------------|
| 75 $\mu\text{L}$                | 2.67                         | 1800 $\mu\text{L}$                           | 1895 $\mu\text{L}$                                |
| 50 $\mu\text{L}$                | 4.00                         | 1280 $\mu\text{L}$                           | 1350 $\mu\text{L}$                                |
| 25 $\mu\text{L}$                | 8.00                         | 640 $\mu\text{L}$                            | 685 $\mu\text{L}$                                 |
| 20 $\mu\text{L}$                | 10.00                        | 510 $\mu\text{L}$                            | 550 $\mu\text{L}$                                 |
| 15 $\mu\text{L}$                | 13.33                        | 380 $\mu\text{L}$                            | 415 $\mu\text{L}$                                 |

**Table S2.** Total volumes for each plate step.

\*Volumes should be maintained below 1200  $\mu$ L, as smaller volumes improve magnetic bead separation.

| Step                               | Volume Range ( $\mu$ L)                             | Plate        |
|------------------------------------|-----------------------------------------------------|--------------|
| Lysate treatment                   | 100-200<br>(compound + IAA ~4-6 $\mu$ L additional) | First plate  |
| CuAAC                              | ~130-230                                            |              |
| 1% SDS + benzonase                 | ~150-250                                            |              |
| Combine heavy + light              | ~300-500                                            |              |
| Add SP3/magnetic beads + >50% EtOH | ~700-1200*                                          |              |
| 80% EtOH washes                    | ~200-400                                            |              |
| DTT/IA                             | ~100                                                |              |
| Trypsin Digest                     | 100-150                                             |              |
| Avidin Enrichment                  | ~500                                                | Second plate |
| Washes                             | ~500                                                |              |
| Elution                            | 60/elution                                          |              |

**Table S3.** Conditions of Liquid-chromatography (LC)

| Experiment Type             | Parameter    | Condition                                                                                                                                                |
|-----------------------------|--------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|
| isoTOP-ABPP<br>Plate screen | Column       | 100 $\mu$ M ID fused silica capillary packed in-house with bulk C18 reversed phase resin (particle size, 1.9 $\mu$ m; pore size, 100 Å; Dr. Maisch GmbH) |
|                             | Mobile phase | Buffer A: water with 3% DMSO and 0.1% formic acid<br>Buffer B: 80% acetonitrile with 3% DMSO and 0.1% formic acid                                        |

|                                                      |                        |                                                                                                                                                |
|------------------------------------------------------|------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                      | Gradient and flow rate | 1 – 80 min, 6 – 40% B, 300 nL/min                                                                                                              |
|                                                      |                        | 80 – 89 min, 40 – 55% B, 300 nL/min                                                                                                            |
|                                                      |                        | 89 – 90 min, 55 – 95% B, 300 nL/min                                                                                                            |
|                                                      |                        | 90 – 99 min, 95% B, 300 nL/min                                                                                                                 |
|                                                      |                        | 99 – 100 min, 1% B, 300 nL/min                                                                                                                 |
|                                                      |                        | 99 – 120 min, 1% B, 300 nL/min                                                                                                                 |
|                                                      | Run time               | 120 minutes                                                                                                                                    |
|                                                      | Injection volume       | 5 uL                                                                                                                                           |
| isoTOP-ABPP<br>All workflow optimization experiments | Column                 | 100 µM ID fused silica capillary packed in-house with bulk C18 reversed phase resin (particle size, 1.9 µm; pore size, 100 Å; Dr. Maisch GmbH) |
|                                                      | Mobile phase           | Buffer A: water with 3% DMSO and 0.1% formic acid<br>Buffer B: 80% acetonitrile with 3% DMSO and 0.1% formic acid                              |
|                                                      | Gradient and flow rate | 0 – 5 min, 3 – 10% B, 300 nL/min                                                                                                               |
|                                                      |                        | 5 – 15 min, 10 - 20% B, 220 nL/min                                                                                                             |
|                                                      |                        | 15 – 64 min, 20 - 47% B, 220 nL/min                                                                                                            |
|                                                      |                        | 64 – 70 min, 47 – 95% B, 250 nL/min                                                                                                            |
|                                                      | Run time               | 70 minutes                                                                                                                                     |
| Sergi-bead cleanup                                   | Column                 | 100 µM ID fused silica capillary packed in-house with bulk C18 reversed phase resin (particle size, 1.9 µm; pore size, 100 Å; Dr. Maisch GmbH) |
|                                                      | Mobile phase           | Buffer A: water with 3% DMSO and 0.1% formic acid<br>Buffer B: 80% acetonitrile with 3% DMSO and 0.1% formic acid                              |
|                                                      |                        | 1 – 80 min, 3 – 40% B, 300 nL/min                                                                                                              |

|  |                        |                                     |
|--|------------------------|-------------------------------------|
|  | Gradient and flow rate | 80 – 89 min, 40 – 55% B, 300 nL/min |
|  |                        | 89 – 90 min, 55 – 95% B, 300 nL/min |
|  |                        | 90 – 99 min, 95% B, 300 nL/min      |
|  |                        | 99 – 100 min, 1% B, 300 nL/min      |
|  |                        | 99 – 120 min, 1% B, 300 nL/min      |
|  | Run time               | 120 minutes                         |

Table S4: Figure 1 MS data

Table S5: Figure 2 MS data

Table S6: Figures 3 + 4 MS data

Table S7: Figure 5 MS data

Table S8: MS raw file annotation

## Chemical Synthesis

**General Methods.** All solution-phase reactions were performed in oven-dried glassware under an atmosphere of dry N<sub>2</sub> except where water was used as a solvent. Reactions were monitored by thin-layer chromatography, and plates were visualized by fluorescence quenching under UV light or by staining with iodine. Concentration *in vacuo* refers to the removal of solvent on a Heidolph rotary evaporator under reduced pressure.

For compounds requiring purification, manual silica gel column chromatography was carried out using SiliCycle silica gel P60 as the stationary phase. Typical stationary phase column loading was a column height of 250 mm, a column diameter of 100 mm. Typical eluents used were hexanes/ethyl acetate, methanol/CH<sub>2</sub>Cl<sub>2</sub>, petroleum ether/ethyl acetate or hexanes/CH<sub>2</sub>Cl<sub>2</sub>.

Other reagents were purchased from Sigma-Aldrich (St. Louis, MO), Alfa Aesar (Ward Hill, MA), EMD Millipore (Billerica, MA), Fisher Scientific (Hampton, NH), Oakwood Chemical (West Columbia, SC), Combi-Blocks (San Diego, CA) and Cayman Chemical (Ann Arbor, MI) and used

without further purification.  $^1\text{H}$  NMR and  $^{13}\text{C}$  NMR spectra for characterization of new compounds and monitoring reactions were collected in  $\text{CDCl}_3$ ,  $\text{CD}_3\text{OD}$ , or  $\text{DMSO}-d_6$  (Cambridge Isotope Laboratories, Cambridge, MA) on a Bruker AV 400 MHz spectrometer or Bruker AV 500 MHz in the Department of Chemistry & Biochemistry at University of California, Los Angeles. All chemical shifts were reported in the standard notation of parts per million using the peak of residual proton signals of the deuterated solvent as an internal reference. Coupling constant ( $J$ ) units are in Hertz (Hz) to the nearest 0.1 Hz. Splitting patterns were indicated as follows: br, broad; s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; dd, doublet of doublets; dt, doublet of triplets. Low-resolution mass spectroscopy was performed on an Agilent Technologies InfinityLab LC/MSD single quadrupole LC/MS (ESI source).

Atropisomeric compounds BEH-1-( $S_a$ ) and BEH-1-( $R_a$ ) and fragments SO-05 to SO-253 used in this study were all synthesized as described below:

#### Materials and General Methods:

$^1\text{H}$  and  $^{13}\text{C}$  NMR spectra were recorded on the Varian VNMRS 400 MHz and 400 MHz Bruker Avance AV1 spectrometers at 25 °C. All chemical shifts were reported in parts per million (ppm) and were internally referenced to residual protio solvents unless otherwise noted. Fluorine spectra were referenced to an external TFA standard. Spectral data were reported as follows: chemical shift (multiplicity [singlet (s), doublet (d), triplet (t), quartet (q), pentet (p), and multiplet (m)], coupling constants [Hz], integration). Carbon spectra were recorded with complete proton decoupling. Conventional mass spectra were obtained using Advion Expression CMS (APCI and ASAP). All chemicals used in the synthesis of substrates were purchased from Sigma Aldrich, TCI, Frontier Scientific, Acros Organics, Combi Blocks, Oakwood, Matrix Scientific, or Fisher Scientific and were used as received without further purification, unless specifically noted. All normal phase flash column chromatography (FCC) was performed using Grade 60 Silica Gel (230-400 mesh) purchased from Fisher Scientific or performed on a Biotage Isolera One with a



Biotage SNAP cartridge (KP-SIL- 10-100g). CHIRALPAK ID and semipreparative CHIRALPAK ID columns were used from Daicel Technologies Corporation to separate the final compound.

Preparation of BEH-1:

**Synthesis of *N*-(3-cyano-7-ethoxy-4-((2-isopropylphenyl)amino)quinolin-6-yl)acetamide**

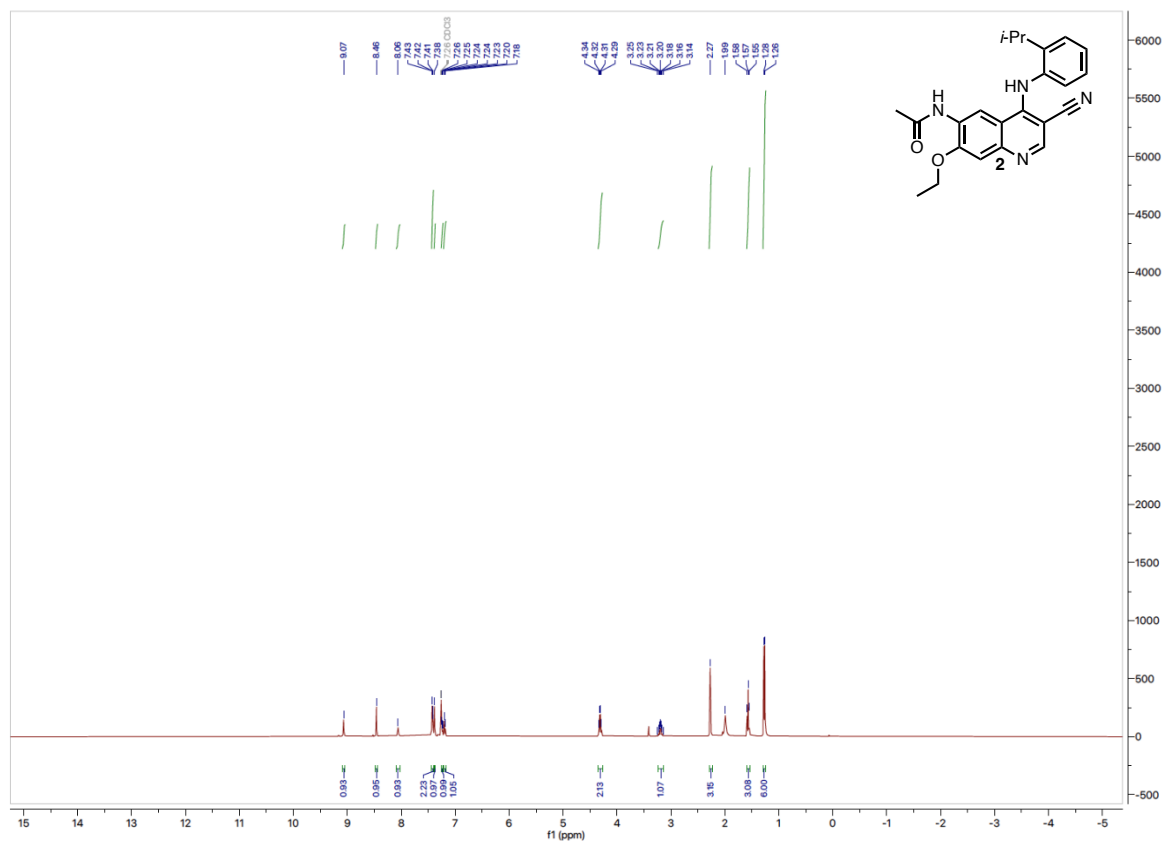
**(2):** *N*-(4-chloro-3-cyano-7-ethoxyquinolin-6-yl)acetamide (**1**, 3 g, 10.4 mmol, 1.0 eq) was added to a 75 mL pressure vessel, followed by 2-isopropylaniline (2.2 mL, 15.5 mmol, 1.5 eq), pyridine (1.7 mL, 20.7 mmol, 2.0 eq), and DMF (34.5 mL, 0.3M) at room temperature. The reaction mixture was stirred overnight at 130 °C. The reaction mixture was allowed to cool to room temperature and was diluted with ethyl acetate (30 mL) and washed with water (30x10 mL) and brine (30 mL). The organic layer was dried with anhydrous sodium sulfate, concentrated *in vacuo*, and purified using flash silica gel column chromatography (gradient up to 65% ethyl acetate in hexanes) to give *N*-(3-cyano-7-ethoxy-4-((2-isopropylphenyl)amino)quinolin-6-yl)acetamide (**2**, 2.8 g, 70% yield) as a beige solid.

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 9.07 (s, 1H), 8.46 (s, 1H), 8.06 (s, 1H), 7.42 (m, *J* = 4.9, 2.1 Hz, 2H), 7.38 (s, 1H), 7.24 (dd, *J* = 5.9, 2.7 Hz, 1H), 7.19 (d, *J* = 7.8 Hz, 1H), 4.31 (q, *J* = 7.0 Hz, 2H), 3.19 (m, *J* = 6.9 Hz, 1H), 2.27 (s, 3H), 1.57 (t, *J* = 6.9 Hz, 3H), 1.27 (d, *J* = 6.9 Hz, 6H).

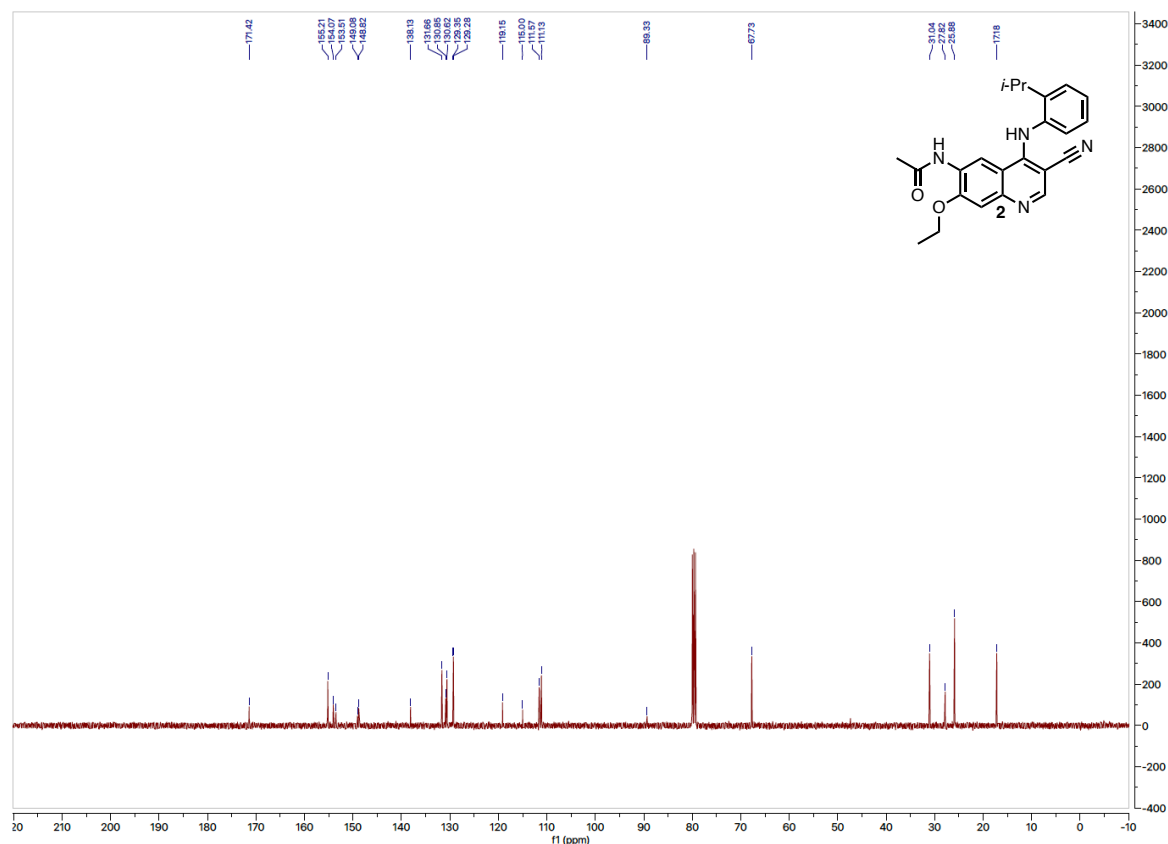
**<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 168.59, 152.38, 151.25, 150.68, 146.25, 146.00, 135.30, 128.83, 128.03, 127.79, 126.52, 126.45, 116.32, 112.18, 108.75, 108.30, 86.50, 64.90, 28.21, 24.99, 23.06, 14.35.

**MS (APCI):** Calculated for C<sub>23</sub>H<sub>24</sub>N<sub>4</sub>O<sub>2</sub> [M+H]<sup>+</sup>: 389.19; Found: 389.1982 m/z

<sup>1</sup>H NMR of *N*-(3-cyano-7-ethoxy-4-((2-isopropylphenyl)amino)quinolin-6-yl)acetamide (**2**) in CDCl<sub>3</sub>



$^{13}\text{C}$  NMR of **N**-(3-cyano-7-ethoxy-4-((2-isopropylphenyl)amino)quinolin-6-yl)acetamide (**2**) in  $\text{CDCl}_3$



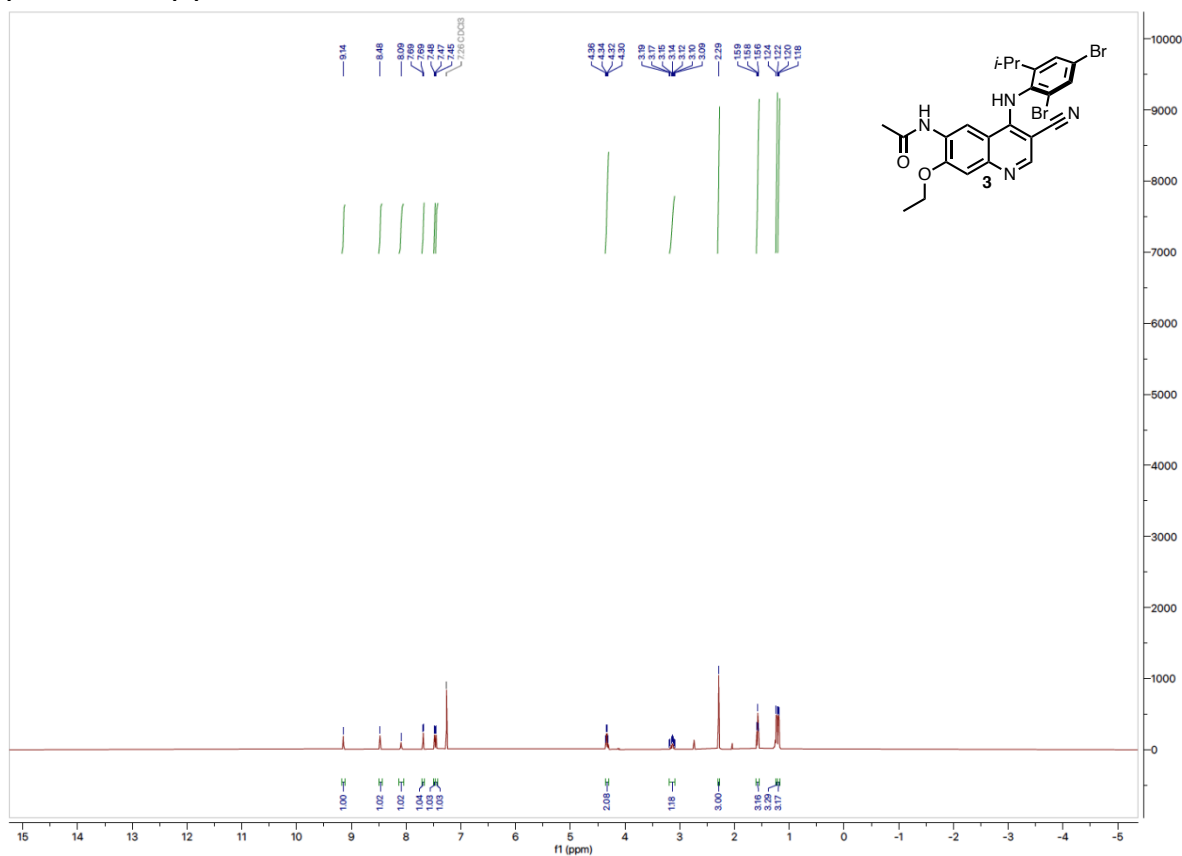
**Synthesis of N**-(3-cyano-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxyquinolin-6-yl)acetamide (**3**): **N**-(3-cyano-7-ethoxy-4-((2-isopropylphenyl)amino)quinolin-6-yl)acetamide (**2**, 800 mg, 2.1 mmol, 1.0 eq) was added to a 15 mL screw cap glass vial, followed by DCM (6.9 mL, 0.3 M) and *N*-Bromosuccinimide (366.6 mg, 4.3 mmol, 2.1 eq). The reaction was allowed to stir at room temperature for 30 minutes and was extracted with DCM (30 mL), washed with sat.  $\text{NaHCO}_3$  (3x30 mL), and brine (30 mL). The organic layer was dried with anhydrous sodium sulfate, concentrated *in vacuo*, and purified using flash silica gel column chromatography (gradient up to 60% ethyl acetate in hexanes) to give **N**-(3-cyano-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxyquinolin-6-yl)acetamide (**3**, 560 mg, 70% yield) as a beige solid.

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 9.14 (s, 1H), 8.48 (s, 1H), 8.09 (s, 1H), 7.69 (d, *J* = 2.3 Hz, 1H), 7.48 (d, *J* = 2.1 Hz, 1H), 7.45 (s, 1H), 4.33 (q, *J* = 7.0 Hz, 2H), 3.14 (m, *J* = 6.9 Hz, 1H), 2.29 (s, 3H), 1.58 (t, *J* = 7.0 Hz, 3H), 1.23 (d, *J* = 6.9 Hz, 3H), 1.19 (d, *J* = 6.9 Hz, 3H).

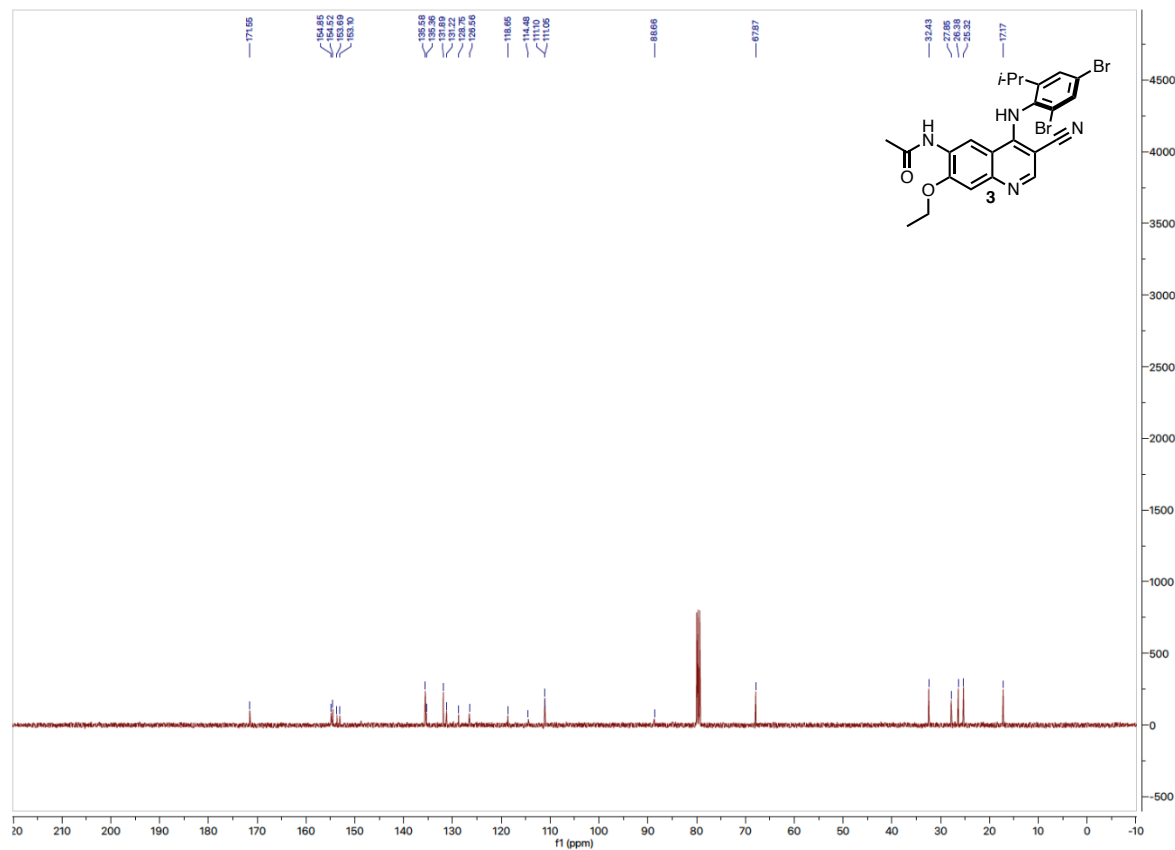
**<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 171.55, 154.85, 154.52, 153.69, 153.10, 135.58, 135.36, 131.89, 131.22, 128.75, 126.56, 118.65, 114.48, 111.05, 111.10, 88.66, 67.87, 32.43, 27.85, 26.38, 25.32, 17.17.

**MS (APCI):** Calculated for C<sub>23</sub>H<sub>22</sub>Br<sub>2</sub>N<sub>4</sub>O<sub>2</sub> [M+H]<sup>+</sup>: 547.26 m/z; Found: 547.0180 m/z

**<sup>1</sup>H NMR of *N*-(3-cyano-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxyquinolin-6-yl)acetamide (3) in CDCl<sub>3</sub>**



$^{13}\text{C}$  NMR of ***N*-(3-cyano-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxyquinolin-6-yl)acetamide (3)** in  $\text{CDCl}_3$



### Synthesis of 6-amino-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxyquinoline-3-carbonitrile

**(4).** *N*-(3-cyano-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxyquinolin-6-yl)acetamide (**3**, 1 g, 2.6 mmol, 1.0 eq) was added to a 100 mL round-bottom flask followed by methanol (25.7 mL, 0.1 M), and 6N HCl (4.4 mL, 0.58M). The reaction mixture was stirred overnight at 70 °C. The

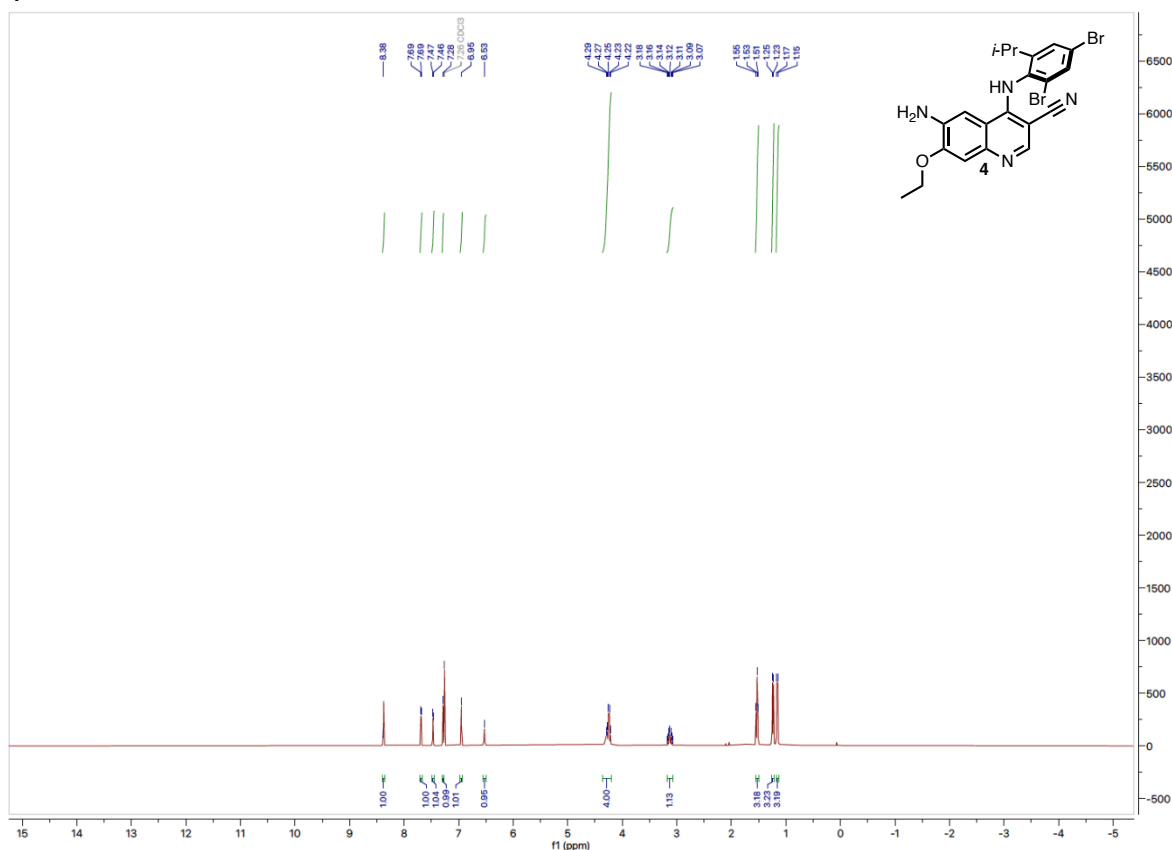
reaction mixture was then allowed to cool to room temperature and was extracted in ethyl acetate (30 mL), quenched with sat. NaHCO<sub>3</sub> (3x30 mL), and washed with brine (30 mL). The organic layer was dried with anhydrous sodium sulfate, concentrated *in vacuo*, and purified using flash silica gel column chromatography (gradient up to 50% ethyl acetate in hexanes) to give 6-amino-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxyquinoline-3-carbonitrile (**4**, 650 mg, 60% yield) as a bright yellow solid.

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 8.38 (s, 1H), 7.69 (d, *J* = 2.1 Hz, 1H), 7.47 (d, *J* = 2.3 Hz, 1H), 7.28 (s, 1H), 6.95 (s, 1H), 6.53 (s, 1H), 4.29 (s, 1H), 4.25 (q, *J* = 7.0 Hz, 2 H), 3.12 (m, *J* = 6.7 Hz, 1H), 1.53 (t, *J* = 7.0 Hz, 3H), 1.24 (d, *J* = 6.8 Hz, 3H), 1.16 (d, *J* = 6.8 Hz, 3H).

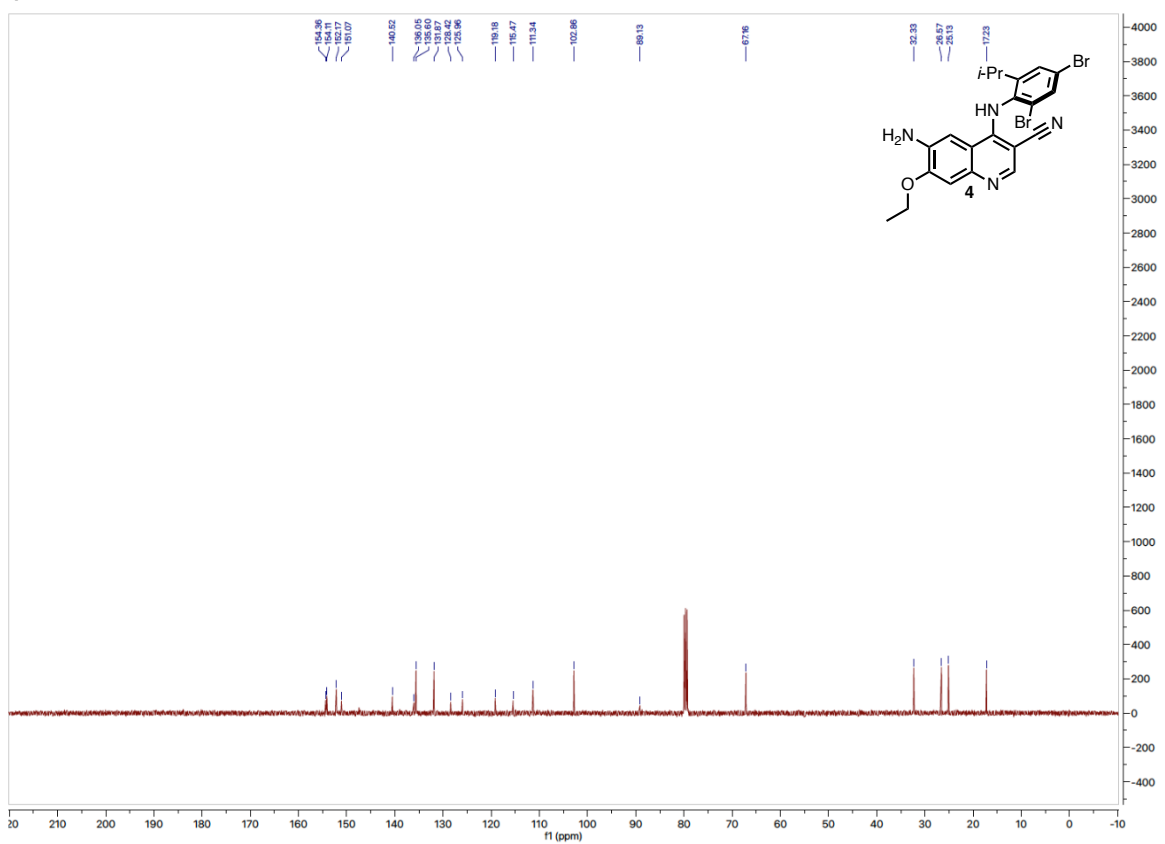
**<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 154.36, 154.11, 152.17, 151.07, 140.52, 136.05, 135.60, 131.87, 128.42, 125.96, 119.18, 115.47, 111.34, 102.86, 89.13, 67.16, 32.33, 26.57, 25.13, 17.23.

**MS (APCI):** Calculated for C<sub>21</sub>H<sub>20</sub>Br<sub>2</sub>N<sub>4</sub>O [M+H]<sup>+</sup>: 505.23 m/z; Found: 505.0071 m/z

**<sup>1</sup>H NMR of 6-amino-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxyquinoline-3-carbonitrile (**4**) in CDCl<sub>3</sub>**



$^{13}\text{C}$  NMR of **6-amino-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxyquinoline-3-carbonitrile (4)** in  $\text{CDCl}_3$



**Synthesis of 6-amino-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxy-5-fluoroquinoline-3-carbonitrile (5).** 6-amino-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxyquinoline-3-carbonitrile (**4**, 450 mg, 0.89 mmol, 1.0 eq) was added to a 50 mL round-bottom flask followed by selectfluor (316.3 mg, 0.89 mmol, 1.0 eq) and acetonitrile (17.8 mL, 0.05M). The reaction was allowed to stir at room temperature overnight. The reaction mixture was then diluted with ethyl acetate (30 mL), washed with water (30 mL), and brine (30 mL). The organic layer was dried with anhydrous sodium sulfate, concentrated *in vacuo*, and purified using flash silica gel column chromatography (gradient up to 30% ethyl acetate in hexanes) to give 6-amino-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxy-5-fluoroquinoline-3-carbonitrile (**5**, 115.6 mg, 25% yield) as a yellow solid.

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 8.32 (s, 1H), 7.82 (d, *J* = 27.8 Hz, 1H), 7.71 (d, *J* = 2.1 Hz, 1H), 7.48 (d, *J* = 2.1 Hz, 2H), 7.21 (s, 1H), 4.28 (q, *J* = 7.1 Hz, 2H), 4.19 (s, 2H), 3.19 (m, *J* = 6.7 Hz, 1H), 1.55 (t, *J* = 7.0 Hz, 3H), 1.28 (d, *J* = 6.9 Hz, 4H), 1.21 (d, *J* = 6.9 Hz, 4H).

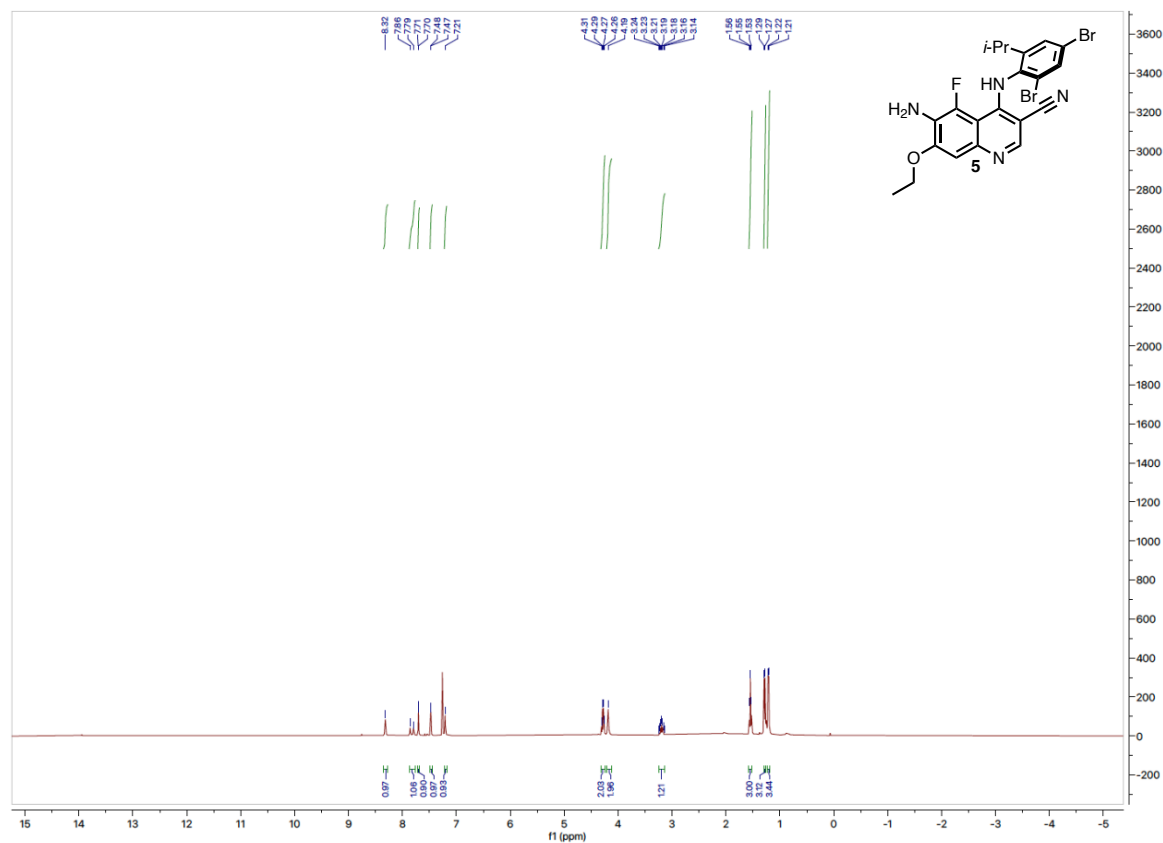
**<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 151.91, 151.66 (d, *J* = 11.3 Hz), 151.16, 149.10 (d, *J* = 4.9 Hz), 146.46, 144.11, 142.85, 132.87, 129.07, 126.28, 124.82 (d, *J* = 17.0 Hz), 123.87, 116.27, 105.19, 85.62, 64.97, 29.89, 23.96, 22.17, 14.52.

**<sup>19</sup>F NMR** (376 MHz, CDCl<sub>3</sub>) δ -139.99 (d, *J* = 25.2 Hz).

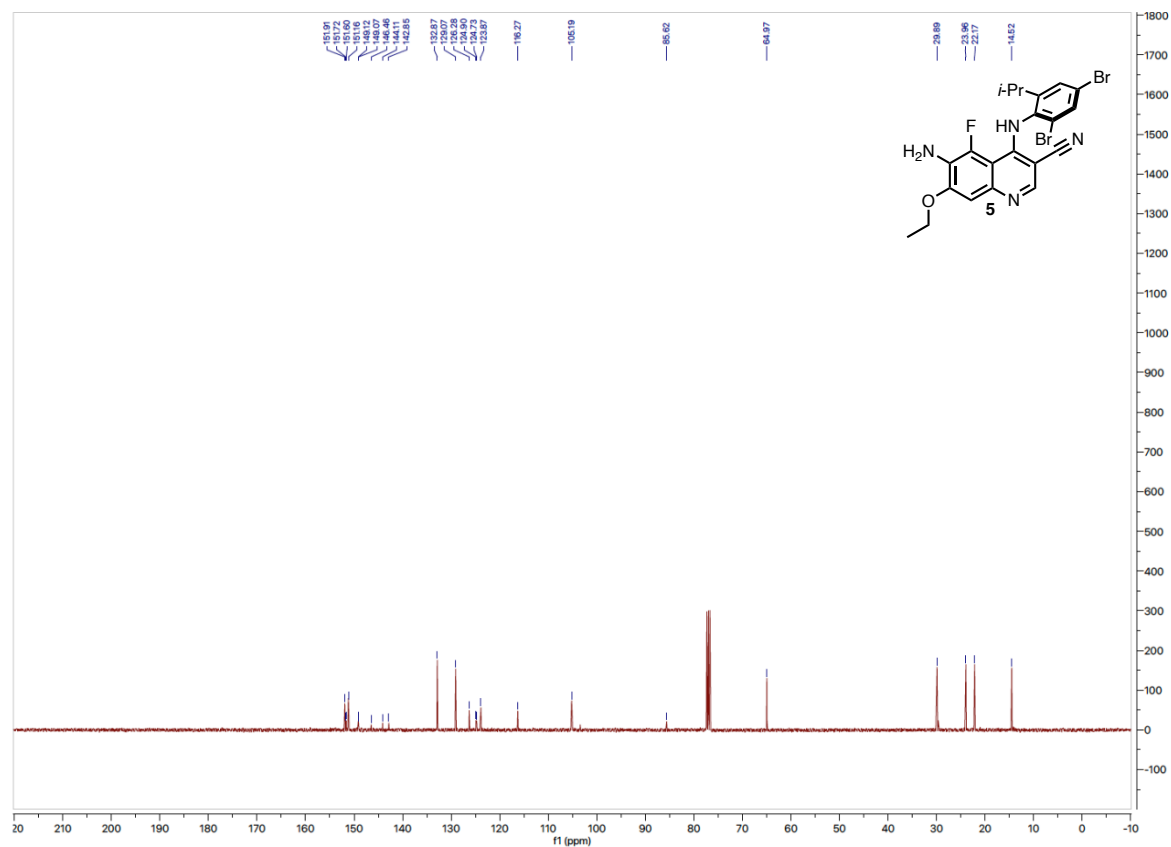
**MS (APCI):** Calculated for C<sub>21</sub>H<sub>20</sub>Br<sub>2</sub>FN<sub>4</sub>O [M+H]<sup>+</sup>: 523.22 m/z; Found: 522.9975 m/z

<sup>1</sup>H NMR of **6-amino-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxy-5-fluoroquinoline-3-carbonitrile (5)** in CDCl<sub>3</sub>

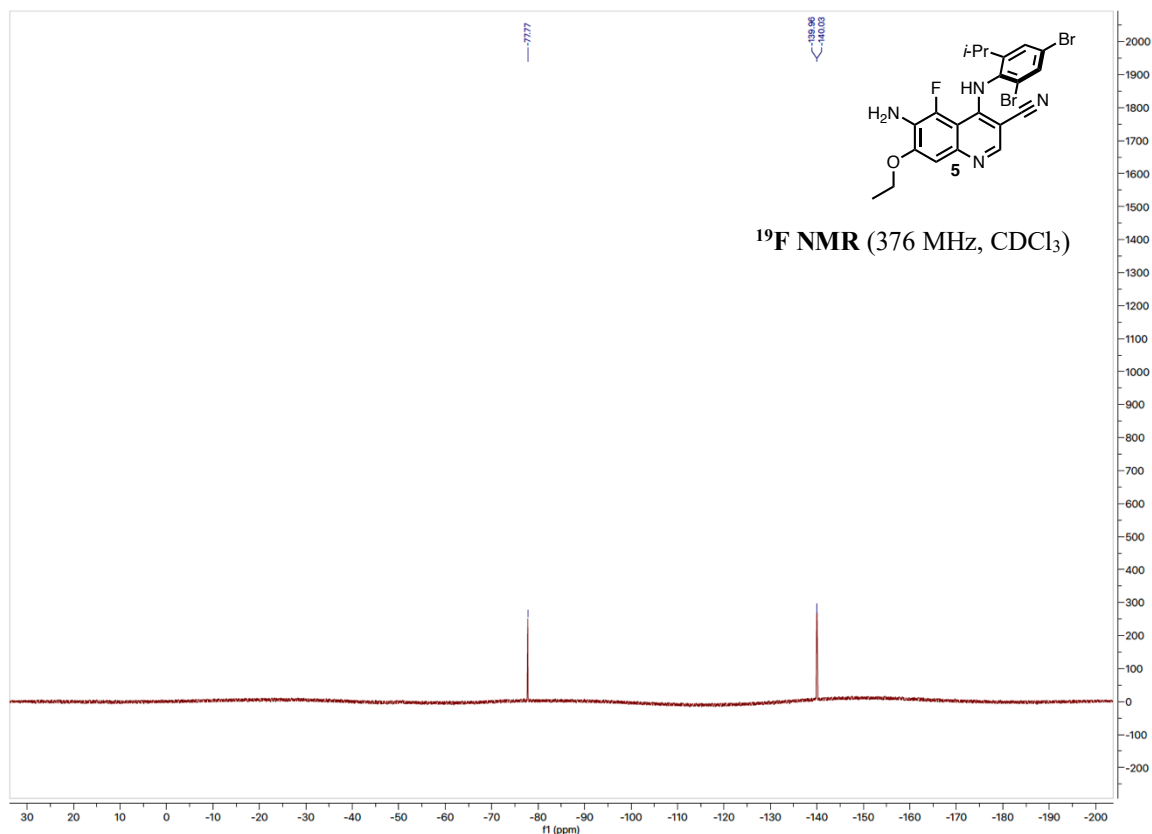




$^{13}\text{C}$  NMR of 6-amino-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxy-5-fluoroquinoline-3-carbonitrile (**5**) in  $\text{CDCl}_3$



$^{19}\text{F}$  NMR of **6-amino-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxy-5-fluoroquinoline-3-carbonitrile (5)** in  $\text{CDCl}_3$



**Synthesis of (E)-N-(3-cyano-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxy-5-fluoroquinolin-6-yl)-4-(dimethylamino)but-2-enamide (BeH-1).** First, (E)-4-(dimethylamino)but-2-enoic acid (43.7 mg, 0.26 mmol, 1.0 eq) was added to a 15 mL screw cap glass vial. The reaction vial was then purged with nitrogen three times, and DCM (0.4 mL, 0.66M), along with oxalyl chloride (45.2  $\mu$ L, 0.53 mmol, 2.0 eq), and 1-2 drops of DMF were subsequently added at 0 °C. The reaction vial was allowed to warm to room temperature, and was stirred for 2 hours. The reaction mixture was then concentrated *in vacuo*, dissolved in 0.4 mL of DCM, and kept aside. In another 15 mL screw cap glass vial, **5** (138.3 mg, 0.26 mmol, 1.0 eq) was added and the reaction vial was purged with nitrogen three times. THF (0.66 mL, 0.4M) was added, and the contents were added dropwise to the crude acid chloride product at 0 °C. Afterwards, *N,N*-Diisopropylethylamine (50.4  $\mu$ L, 0.29 mmol, 1.1 eq) was added dropwise and the reaction vial was allowed to warm to room temperature and was stirred overnight. The reaction mixture was then diluted with ethyl acetate (30 mL), and washed with brine (2x30 mL). The organic layer was

dried with anhydrous sodium sulfate, concentrated *in vacuo*, and purified using silica gel flash column chromatography (gradient up to 10% methanol in DCM) to give (E)-N-(3-cyano-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxy-5-fluoroquinolin-6-yl)-4-(dimethylamino)but-2-enamide (**6**, 41.7 mg, 25% yield) as a tan solid.

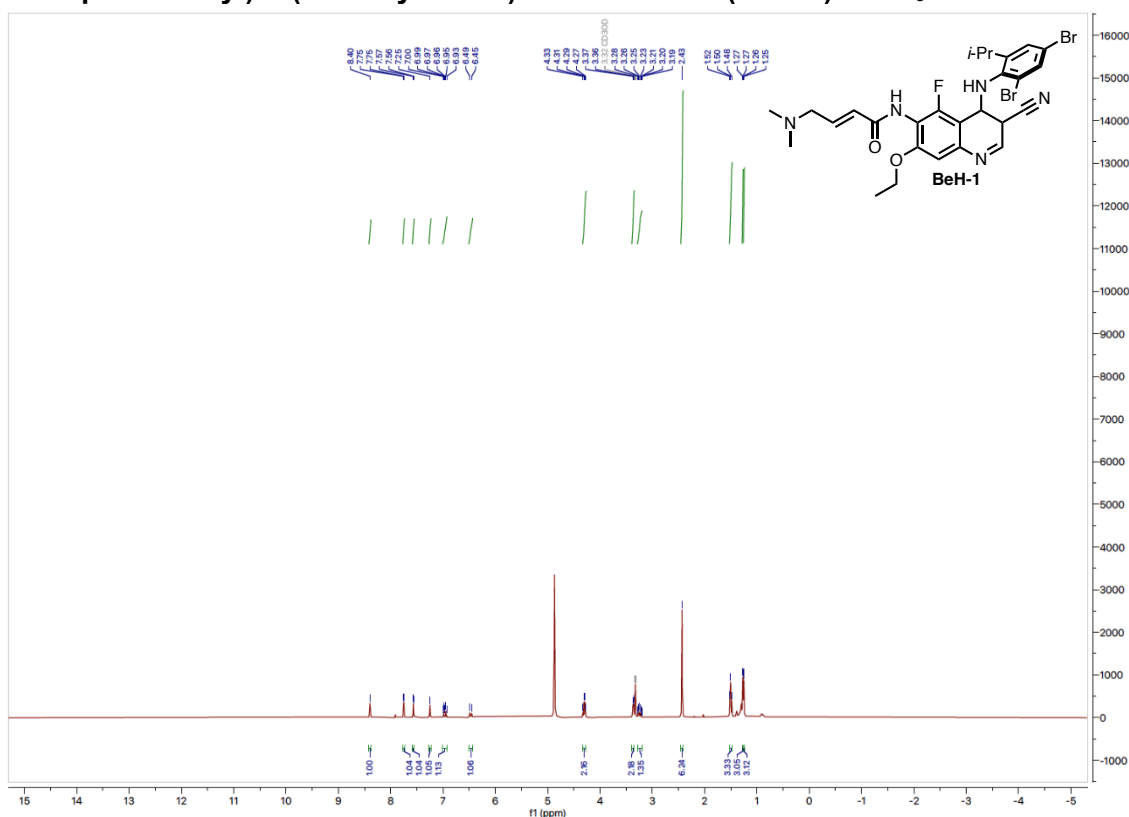
**<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD) δ 8.40 (s, 1H), 7.75 (d, *J* = 2.3 Hz, 1H), 7.57 (d, *J* = 2.3 Hz, 1H), 7.25 (s, 1H), 6.97 (dt, *J* = 15.4, 6.6 Hz, 1H), 6.47 (d, *J* = 15.4 Hz, 1H), 4.30 (q, *J* = 6.9 Hz, 2H), 3.37 (d, *J* = 7.4 Hz, 2H), 3.24 (m, *J* = 6.9 Hz, 1H), 2.43 (s, 6H), 1.50 (t, *J* = 7.0 Hz, 3H), 1.27 (d, *J* = 2.5 Hz, 3H), 1.25 (d, *J* = 2.6 Hz, 3H).

**<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD) δ 167.50, 160.10 (d, *J* = 7.4 Hz), 158.76, 157.25, 156.22, 155.10, 153.90 (d, *J* = 4.2 Hz), 151.62, 142.57, 136.10, 134.95, 131.40, 128.94 (d, *J* = 8.5 Hz), 126.20, 118.41, 116.37 (d, *J* = 17.3 Hz), 107.21, 87.24, 67.76, 61.96, 46.28, 32.27, 24.93, 24.29, 15.94.

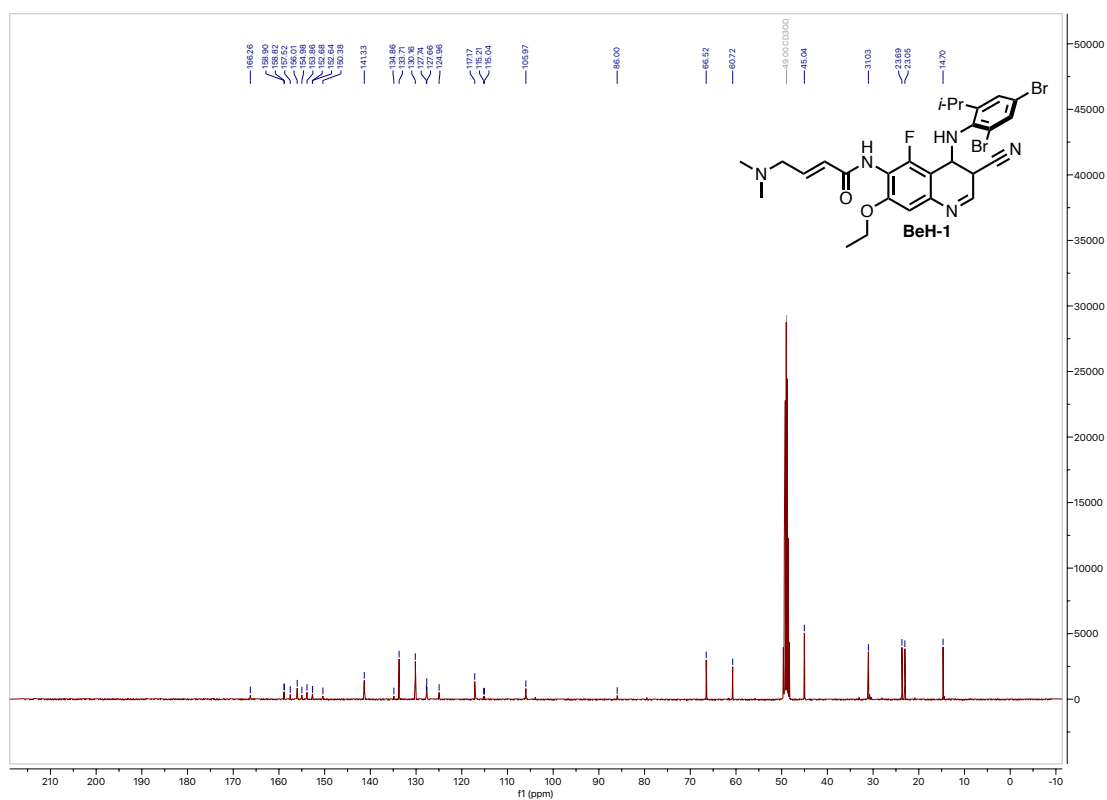
**<sup>19</sup>F NMR** (376 MHz, CD<sub>3</sub>OD) δ -118.23.

**MS (APCI):** Calculated for C<sub>21</sub>H<sub>20</sub>Br<sub>2</sub>FN<sub>4</sub>O [M+H]<sup>+</sup>: 634.36 m/z; Found: 634.0666 m/z

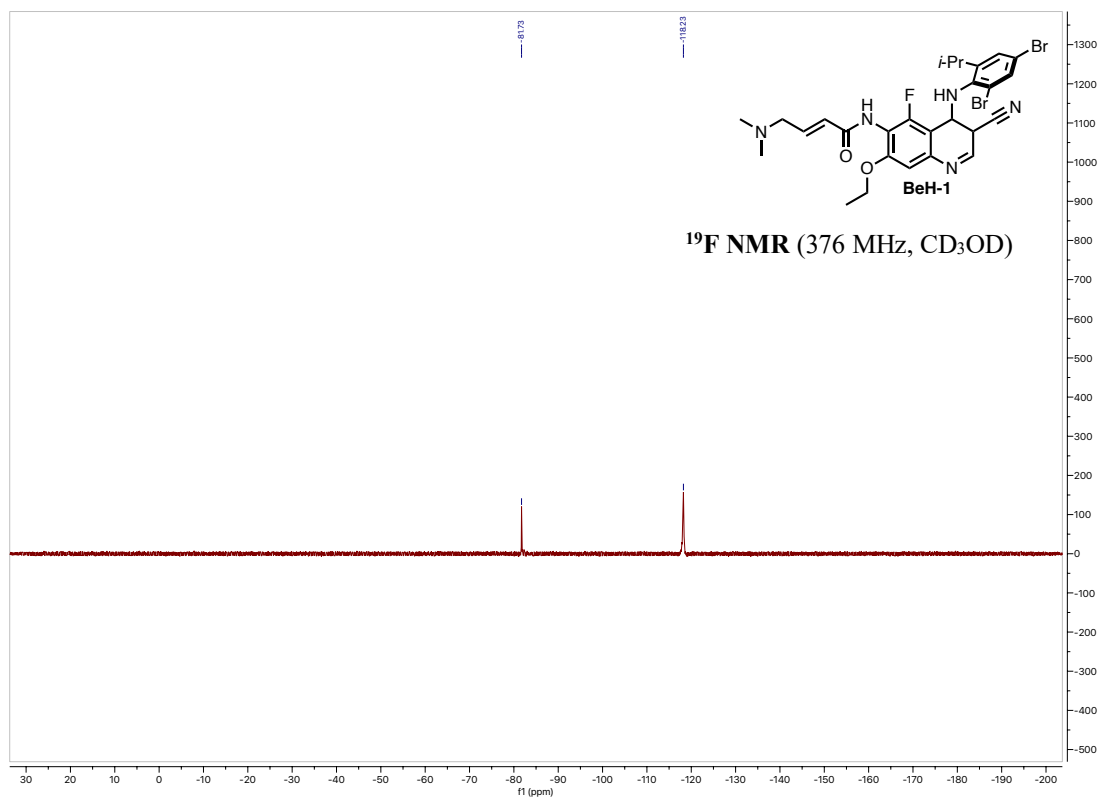
**<sup>1</sup>H NMR of (E)-N-(3-cyano-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxy-5-fluoroquinolin-6-yl)-4-(dimethylamino)but-2-enamide (BEH-1) in CD<sub>3</sub>OD**



$^{13}\text{C}$  NMR of (*E*)-*N*-(3-cyano-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxy-5-fluoroquinolin-6-yl)-4-(dimethylamino)but-2-enamide (BEH-1) in  $\text{CD}_3\text{OD}$

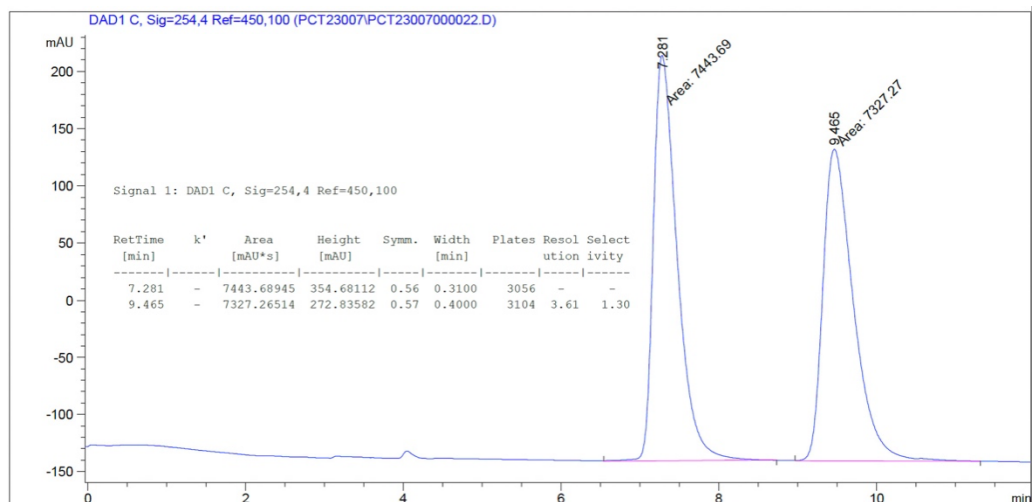


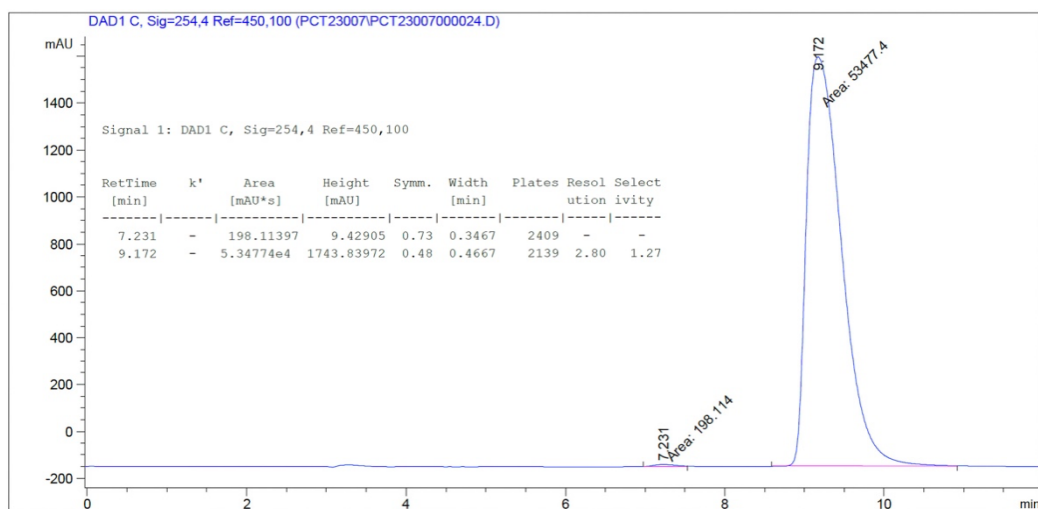
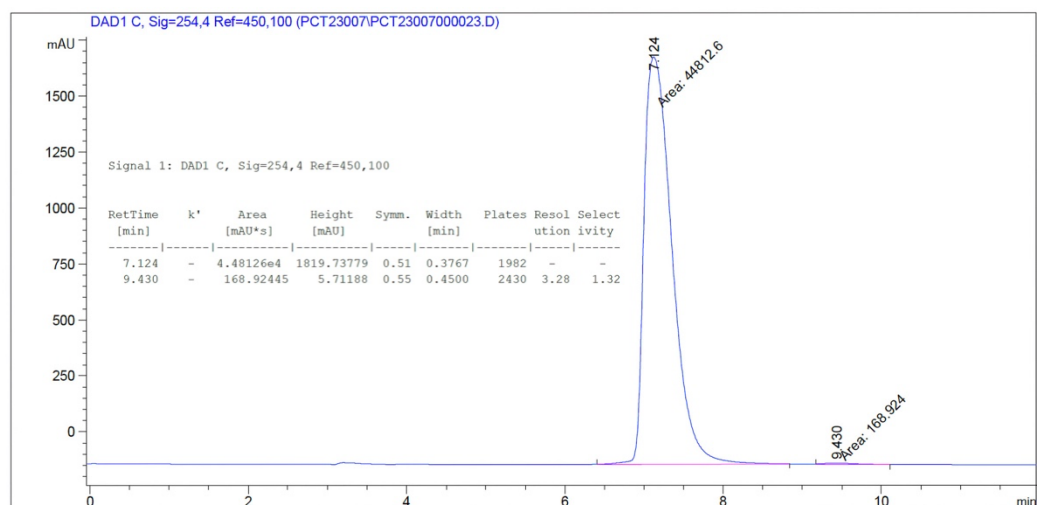
$^{19}\text{F}$  NMR of (*E*)-*N*-(3-cyano-4-((2,4-dibromo-6-isopropylphenyl)amino)-7-ethoxy-5-fluoroquinolin-6-yl)-4-(dimethylamino)but-2-enamide (BEH-1) in  $\text{CD}_3\text{OD}$



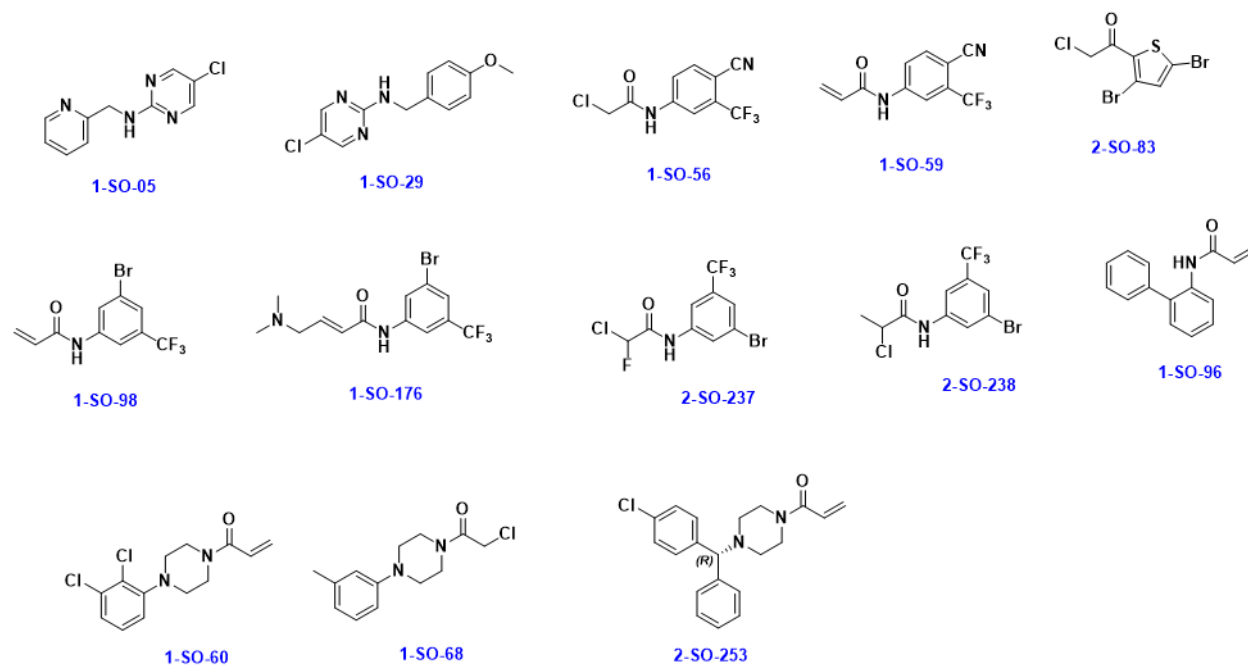
### Chiral HPLC Trace of **BEH-1**

**BEH-1** was measured with HPLC analysis using a ChiralPak ID Column in Hexanes/EtOH/DEA (70/30/0.1) with a flow rate of 1 mL/min and an injection volume of 20  $\mu$ L.





## Structures of compounds (SO-05 to SO-253):



## General procedures: 1-3

### General procedure 1:

To a solution of the amine (1.1 equiv.) and the aldehyde (1.0 equiv.) in anhydrous MeOH (1.0 mL) was added AcOH (1.0 equiv.). The reaction was stirred at 25 °C for 3 h. NaBH<sub>3</sub>CN (2.0 equiv.) was added at 0 °C. The reaction was stirred at 25 °C for another 1 h. LCMS showed the reaction was completed. The residue was poured into ice water. The aqueous phase was extracted three times with ethyl acetate. The combined organic phase was washed with brine, dried with anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated in vacuo, and purified by silica gel flash column chromatography to give the titled compound.

### General procedure 2:

The amine (1 equiv.) was added to a round bottom flask with a stir bar and dissolved in DCM with stirring. Triethylamine (2 equiv.) was added and the reaction was cooled to 0 °C. Chloroacetyl chloride (2 equiv.) was added dropwise to the flask, and the reaction was allowed to warm to room

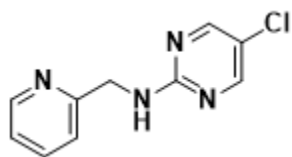


temperature as it stirred for 2 h. The reaction was washed with water, extracted with ethyl acetate, dried over Na<sub>2</sub>SO<sub>4</sub>, and purified by silica gel flash column chromatography to give the titled compound.

### General procedure 3:

The amine (1 equiv.) was added to a round bottom flask with a stir bar and dissolved in DCM. Triethylamine (1 equiv.) was added and the reaction was cooled to 0 °C. Acryloyl chloride (2 equiv.) was added dropwise to the flask, and the reaction was allowed to warm to room temperature as it stirred for 2 h. The reaction was washed with water, extracted with ethyl acetate, dried over Na<sub>2</sub>SO<sub>4</sub>, and purified silica gel flash column chromatography to give the titled compound.

### Synthesis of 5-chloro-N-(pyridin-2-ylmethyl)pyrimidin-2-amine (1-SO-05)

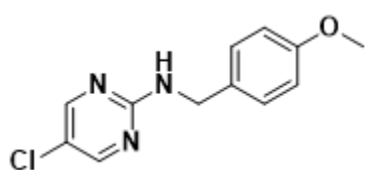


Prepared according to general procedure 1, using 5-chloropyrimidin-2-amine (266.1 mg, 2.05 mmol, 1.1 equiv. ) as the amine; picolinaldehyde (200 mg, 1.87 mmol, 1 equiv.) as the aldehyde. Product: off-white (356 mg, 86.4%)

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ 8.58 (s, 1H), 8.23 (s, 2H), 7.77 (s, 1H), 7.42 (d, J = 7.9 Hz, 1H), 7.33 – 7.27 (m, 1H), 6.45 (s, 1H), 4.78 (s, 2H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>): δ 160.11, 156.69, 156.13, 147.46, 138.00, 122.35, 119.35, 77.16, 45.88.

LRMS (ESI, m/z): calc'd for C<sub>10</sub>H<sub>10</sub>ClN<sub>4</sub><sup>+</sup> [M+H]<sup>+</sup> 221.05; found 221.0

### Synthesis of 5-chloro-N-(4-methoxybenzyl)pyrimidin-2-amine (1-SO-29)



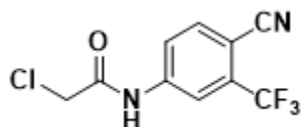
Prepared according to general procedure 1, using 5-chloropyrimidin-2-amine (266.1 mg, 2.42 mmol, 1.1 equiv.) as the amine; 4-methoxybenzaldehyde (300 mg, 2.20 mmol, 1 equiv.) as

the aldehyde. Product: white powder (450 mg, 81%)

**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):** δ 8.13 (s, 2H), 7.26 (d, J = 8.7 Hz, 2H), 6.87 (d, J = 8.6 Hz, 2H), 5.76 (s, 1H), 4.51 (d, J = 5.3 Hz, 2H), 3.80 (s, 3H). **<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>):** δ 159.05, 156.24, 130.64, 128.95, 119.02, 114.08, 55.31, 45.33.

**LRMS** (ESI, m/z): calc'd for C<sub>12</sub>H<sub>13</sub>ClN<sub>3</sub>O<sup>+</sup> [M+H]<sup>+</sup> 250.07; found 250.0

### Synthesis of 2-chloro-N-(4-cyano-3-(trifluoromethyl)phenyl)acetamide (1-SO-56)



Prepared according to general procedure 2, using 4-amino-2-(trifluoromethyl)benzonitrile (200 mg, 1.07 mmol, 1 equiv.) as the amine.

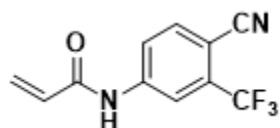
Product: off-white solid (250 mg, 90%).

**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):** δ 8.54 (s, 1H), 8.00 (d, J = 23.5 Hz, 2H), 7.84 (d, J = 8.5 Hz, 1H), 4.24 (s, 2H).

**<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>):** δ 164.53, 140.90, 135.99, 134.37, 134.05, 133.72, 123.36, 122.24, 120.64, 117.63, 115.27, 42.75.

**LRMS** (ESI, m/z): calc'd for C<sub>10</sub>H<sub>5</sub>ClF<sub>3</sub>N<sub>2</sub>O<sup>-</sup> [M-H]<sup>-</sup> 261.01; found 261.0

### Synthesis of N-(4-cyano-3-(trifluoromethyl)phenyl)acrylamide (1-SO-59)



Prepared according to general procedure 3, 4-amino-2-(trifluoromethyl)benzonitrile (200 mg, 1.07 mmol, 1 equiv.) as the amine.

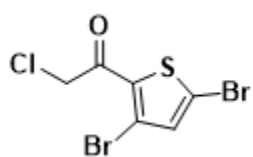
Product: off-white solid (200 mg, 79%).

**<sup>1</sup>H NMR (400 MHz, Methanol-d<sub>4</sub>):** δ 8.30 (s, 1H), 8.04 (d, J = 10.6 Hz, 1H), 7.92 (d, J = 8.5 Hz, 1H), 6.49 – 6.39 (m, 2H), 5.86 (d, J = 11.8 Hz, 1H).

**<sup>13</sup>C NMR (101 MHz, Methanol-d<sub>4</sub>):** δ 166.65, 145.02, 137.45, 134.77, 134.55, 134.36, 131.94, 129.87, 123.65, 118.57, 118.52, 104.75, 49.00.

**LRMS** (ESI, m/z): calc'd for C<sub>11</sub>H<sub>6</sub>F<sub>3</sub>N<sub>2</sub>O<sup>-</sup> [M-H]<sup>-</sup> 239.05; found 239.0

### Synthesis of 2-chloro-1-(3,5-dibromothiophen-2-yl)ethan-1-one(2-SO-83)



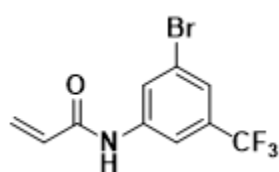
To a solution consisting of 1 g (4.133 mmol, 1.0 equiv.) 2,4-dibromothiophene and 560 mg (4.96 mmol, 1.2 equiv.) chloroacetyl chloride in 12 mL of absolute dichloromethane, 1.10 g (8.26 mmol, 2 equiv.) anhydrous

aluminum chloride was added portionwise, that was not exceeded 15 oC (ice bath) and stirred for 3 hrs at this temperature. When deemed complete via TLC, 2 N hydrochloric acid (10mL) was added, and stirred for 15 min until the complete destruction of the complex, the phases were separated and the aqueous phase was extracted twice with dichloromethane. The collected organic phases were washed successively with saturated sodium bicarbonate solution and water, dried with sodium sulfate, treated with activated carbon, filtered, and evaporated to yield the yellow solid product 1.2g (91 %).

**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)** : δ 7.12 (s, 1H), 4.77 (s, 2H).

**<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)**: δ 182.65, 138.80, 135.83, 123.71, 114.30, 47.67.

#### Synthesis of N-(3-bromo-5-(trifluoromethyl)phenyl)acrylamide (1-SO-98)



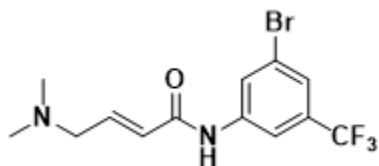
Prepared according to general procedure 3, 3-bromo-5-(trifluoromethyl)aniline (300 mg, 1.25 mmol, 1 equiv.) as the amine.  
Product: off-white solid (320 mg, 86%)

**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)**: <sup>1</sup>H NMR (400 MHz, Chloroform-d) δ 8.05 (s, 1H), 7.77 (s, 1H), 7.50 (s, 1H), 6.55 – 6.41 (m, 1H), 6.26 (dd, J = 16.9, 10.3 Hz, 1H), 5.85 (dd, J = 10.2, 1.0 Hz, 1H).

**<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)**: δ 163.91, 139.39, 132.99, 132.66, 132.33, 125.95, 124.25, 123.10, 121.52, 115.32.

**LRMS** (ESI, m/z): calc'd for C<sub>10</sub>H<sub>6</sub>BrF<sub>3</sub>NO<sup>-</sup> [M-H]<sup>-</sup> 291.97; found 291.9

#### Synthesis of (E)-N-(3-bromo-5-(trifluoromethyl)phenyl)-4-(dimethylamino)but-2-enamide (1-SO-176)



To a solution of (E)-4-(dimethylamino)but-2-enoic acid (194 mg, 1.5 mmol, 1.2 equiv.) in dichloromethane (CH<sub>2</sub>Cl<sub>2</sub>) (5 mL) at 0 °C, stirring under nitrogen, was added oxalyl chloride (317 μL, 2.50 mmol, 2 equiv.) via syringe followed by a catalytic amount of DMF (15 μL). The resulting

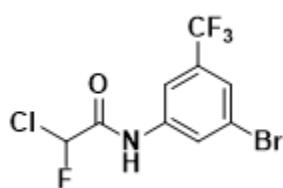
mixture was warmed to room temperature and stirred for 4 h. Next, the reaction mixture was

concentrated in vacuo to afford the crude (E)-4-(dimethylamino)but-2-enoyl chloride, which was used directly in the next step without further purification. To a stirring solution of 3-bromo-5-(trifluoromethyl)aniline (300  $\mu$ L, 1.25 mmol, 1 equiv.) and Et<sub>3</sub>N (253  $\mu$ L, 2.5 mmol, 2 equiv.) in CH<sub>2</sub>Cl<sub>2</sub> (5 mL) at 0 °C, was slowly added a cold solution of the crude (E)-4-(dimethylamino)but-2-enoyl chloride in CH<sub>2</sub>Cl<sub>2</sub> (5 mL). The reaction mixture was continually stirred at 25 °C, under nitrogen, for another 18 h. LCMS showed the reaction was completed. Next, the reaction was quenched with water. The aqueous phase was extracted three times with dichloromethane. The combined organic phase was washed with brine, dried with anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated in vacuo, and purified by silica gel flash column chromatography (7% methanol in DCM) to give (272 mg, 62%) of (E)-N-(3-bromo-5-(trifluoromethyl)phenyl)-4-(dimethylamino)but-2-enamide.

**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)** :  $\delta$  8.70 (s, 1H), 7.99 (s, 1H), 7.77 (s, 1H), 7.42 (s, 1H), 7.06 – 6.90 (m, 1H), 6.19 (d, J = 15.3 Hz, 1H), 3.07 (d, J = 5.8 Hz, 2H), 2.23 (s, 6H).

**<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)** :  $\delta$  164.81, 144.42, 140.24, 133.52, 133.19, 132.86, 132.53, 127.38, 126.47, 125.49, 124.66, 124.31, 123.33, 121.95, 115.95, 77.16, 60.63, 45.79.

**LRMS** (ESI, m/z): calc'd for C<sub>13</sub>H<sub>13</sub>BrF<sub>3</sub>N<sub>2</sub>O<sup>+</sup> [M-H]<sup>+</sup> 349.02; found 349.0.



#### Synthesis of N-(3-bromo-5-(trifluoromethyl)phenyl)-2-chloro-2-fluoroacetamide (2-SO-237)

To a stirred solution of sodium chlorofluoroacetate (336 mg, 2.50 mmol, 2 equiv.), HOBT•H<sub>2</sub>O (338 mg, 2.50 mmol, 2 equiv.), and propylphosphonic anhydride (T3P) (50 wt. % in CH<sub>2</sub>Cl<sub>2</sub>, 1.15 g, 2.04 mmol, 2 equiv.) in dry DMF (3.0 mL) was added DIPEA (323  $\mu$ L, 2.50 mmol, 2 equiv.) and 3-bromo-5-(trifluoromethyl)aniline (300 mg, 1.25 mmol, 1 equiv.) at 0 °C. After stirring overnight at ambient temperature, the reaction mixture was diluted with CH<sub>2</sub>Cl<sub>2</sub> and saturated sodium bicarbonate. The organic layer was separated and the aqueous phase was extracted with CH<sub>2</sub>Cl<sub>2</sub>. The combined organic layers were dried over anhydrous magnesium sulfate, filtered, and concentrated *in vacuo*. The resultant crude was purified by flash column chromatography on

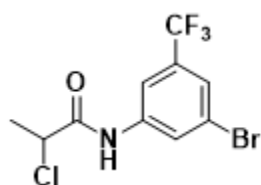
silica gel using 7% methanol/CH<sub>2</sub>Cl<sub>2</sub> to give the target compound (260 mg, 63% yield) as pale-yellowish oil.

**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)** : δ 8.05 (s, 1H), 7.79 (s, 1H), 7.60 (s, 1H), 6.44 (d, J = 50.9 Hz, 1H).

**<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)** : δ 162.12, 161.90, 137.63, 133.66, 132.99, 126.20, 125.49, 123.38, 115.68, 95.23, 92.67.

**LRMS** (ESI, m/z): calc'd for C<sub>9</sub>H<sub>6</sub>BrClF<sub>4</sub>NO<sup>+</sup> [M+H]<sup>+</sup> 333.92; found 333.9

#### Synthesis of N-(3-bromo-5-(trifluoromethyl)phenyl)-2-chloropropanamide (2-SO-238)



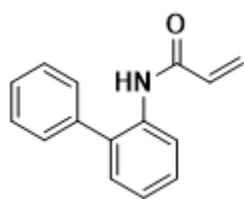
To a stirred solution of 3-bromo-5-(trifluoromethyl)aniline (500 μL, 2.08 mmol) in DCM (5ml), triethylamine (232 μL, 2.29 mmol, 1.1 equiv.), 2-chloropropanoyl chloride (317 mg, 2.5 mmol, 1.2 equiv.) at 0°C was added dropwise and maintained with an ice bath. The reaction mixture was stirred

at room temperature for an additional four hours. The reaction was washed with water, extracted with ethyl acetate, dried over Na<sub>2</sub>SO<sub>4</sub>, and purified by silica gel flash column chromatography (30% ethyl acetate in hexanes) to give the titled compound (560 mg, 81%).

**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)** : δ 8.39 (s, 1H), 8.02 (s, 1H), 7.77 (s, 1H), 7.55 (s, 1H), 4.56 (q, J = 7.1 Hz, 1H), 1.84 (s, 3H).

**<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)** : δ 168.21, 139.11, 133.91, 133.57, 133.24, 132.91, 126.32, 125.24, 124.63, 123.63, 115.85, 77.16, 56.39, 22.95.

**LRMS** (ESI, m/z): calcd for C<sub>10</sub>H<sub>9</sub>BrClF<sub>3</sub>NO<sup>+</sup> [M+H]<sup>+</sup> 329.94; found 329.9



#### Synthesis of N-([1,1'-biphenyl]-2-yl)acrylamide (1-SO-96)

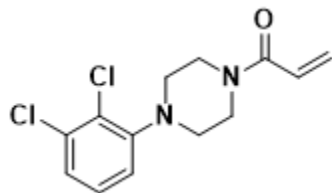
Prepared according to general procedure 3, [1,1'-biphenyl]-2-amine (250 mg, 1.48 mmol, 1 equiv.) as the amine. Product: white solid (240 mg, 74%)

**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)** : δ 8.44 (s, 1H), 7.49 (dd, J = 8.0, 6.5 Hz, 2H), 7.46 – 7.35 (m, 4H), 7.35 – 7.25 (m, 2H), 7.19 (t, J = 7.9 Hz, 1H), 6.27 (d, J = 17.6 Hz, 1H), 6.00 (d, J = 22.9 Hz, 1H), 5.66 (d, J = 11.4 Hz, 1H).

**<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)** : δ 138.42, 135.02, 131.85, 130.51, 129.74, 128.97, 128.54, 127.79, 124.91, 121.79, 77.16.

**LRMS** (ESI, m/z): calcd for  $C_{16}H_{13}NO^+$   $[M+H]^+$  224.10; found 224.1

**Synthesis of 1-(4-(2,3-dichlorophenyl)piperazin-1-yl)prop-2-en-1-one (1-SO-60)**



Prepared according to general procedure 2, 1-(2,3-dichlorophenyl)piperazine (200 mg, 0.865 mmol, 1 equiv.) as the amine. Product: off-white solid (151 mg, 61%)

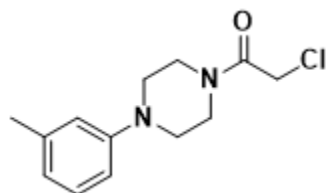
**$^1H$  NMR (400 MHz,  $CDCl_3$ )** :  $\delta$  5.73 (d,  $J$  = 12.5 Hz, 1H), 3.81 (d,  $J$  = 53.6 Hz, 4H), 3.03 (d,  $J$  = 4.9 Hz, 4H).

**$^{13}C$  NMR (101 MHz,  $CDCl_3$ )** :  $\delta$  165.58, 150.56, 134.24, 128.17, 127.80, 127.59, 127.41, 125.28, 118.82, 51.80, 51.06, 46.07, 42.16.

**LRMS** (ESI, m/z): calcd for  $C_{13}H_{15}Cl_2N_2O^+$   $[M+H]^+$  285.05; found 285.0

**Synthesis of 2-chloro-1-(4-(m-tolyl)piperazin-1-yl)ethan-1-one (SO-68)**

Prepared according to general procedure 2, using 1-(m-tolyl)piperazine (200 mg, 1.13 mmol, 1 equiv.) as the amine. Product: off-white solid (270 mg, 94%)

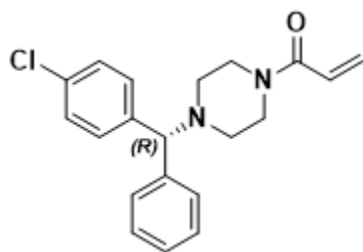


**$^1H$  NMR (400 MHz,  $CDCl_3$ )** :  $\delta$  7.18 (t,  $J$  = 7.8 Hz, 1H), 6.81 – 6.71 (m, 3H), 4.11 (s, 2H), 3.83 – 3.63 (m, 4H), 3.20 (dt,  $J$  = 20.1, 5.2 Hz, 4H), 2.33 (s, 3H).

**$^{13}C$  NMR (101 MHz,  $CDCl_3$ )** :  $\delta$  165.14, 150.80, 139.09, 129.15, 121.78, 117.71, 113.96, 49.83, 49.46, 46.28, 42.14, 40.84, 21.74.

**LRMS** (ESI, m/z): calcd for  $C_{13}H_{16}ClN_2O^-$   $[M-H]^-$  252.10; found 252.1

**Synthesis of (R)-1-(4-((4-chlorophenyl)(phenyl)methyl)piperazin-1-yl)prop-2-en-1-one (SO-253)**

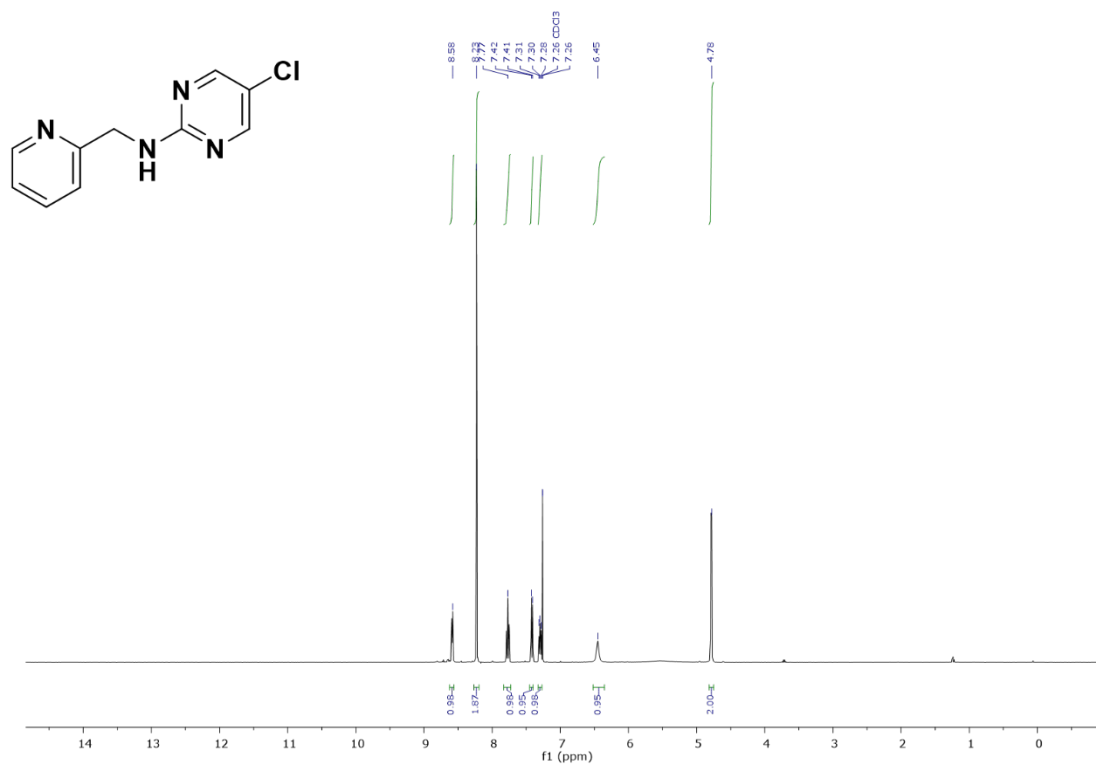


Prepared according to general procedure 3, (R)-1-((4-chlorophenyl)(phenyl)methyl)piperazine (200 mg, 0.697 mmol, 1 equiv.) as the amine. Product: pale-yellowish solid (164 mg, 69%).

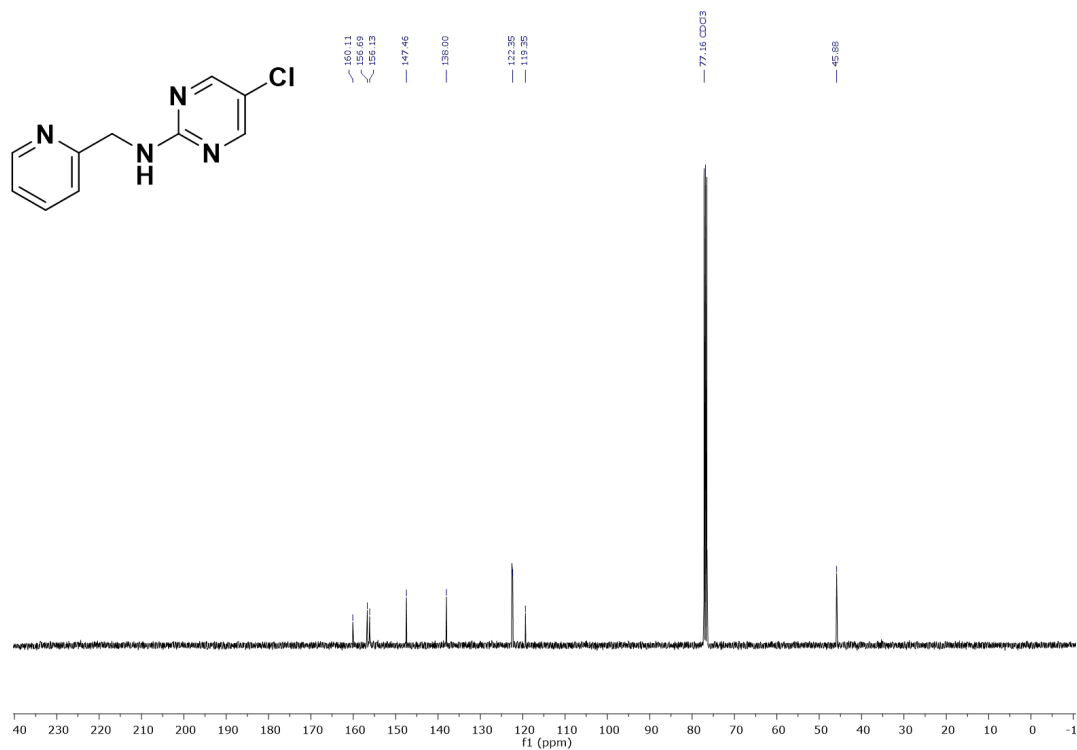
**$^1H$  NMR (400 MHz,  $CDCl_3$ )** :  $\delta$  7.38 – 7.33 (m, 4H), 7.29 (d,  $J$  = 7.3 Hz, 2H), 7.26 – 7.19 (m, 3H), 6.51 (dd,  $J$  = 16.8, 10.5 Hz, 1H), 6.26 (dd,  $J$  = 16.8, 1.9 Hz, 1H), 5.66 (dd,  $J$  = 10.5, 1.8 Hz, 1H), 4.22 (s, 1H), 3.61 (d,  $J$  = 54.6 Hz, 4H), 2.37 – 2.38 (m, 4H).

**$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )** :  $\delta$  165.32, 141.53, 140.74, 132.88, 129.14, 128.85, 128.78, 127.84, 127.80, 127.45, 127.41, 75.21, 52.06, 51.47, 45.91, 42.09.

**LRMS** (ESI, m/z): calcd for  $\text{C}_{20}\text{H}_{22}\text{ClN}_2\text{O}^-$   $[\text{M}-\text{H}]^-$  341.13; found 341.1

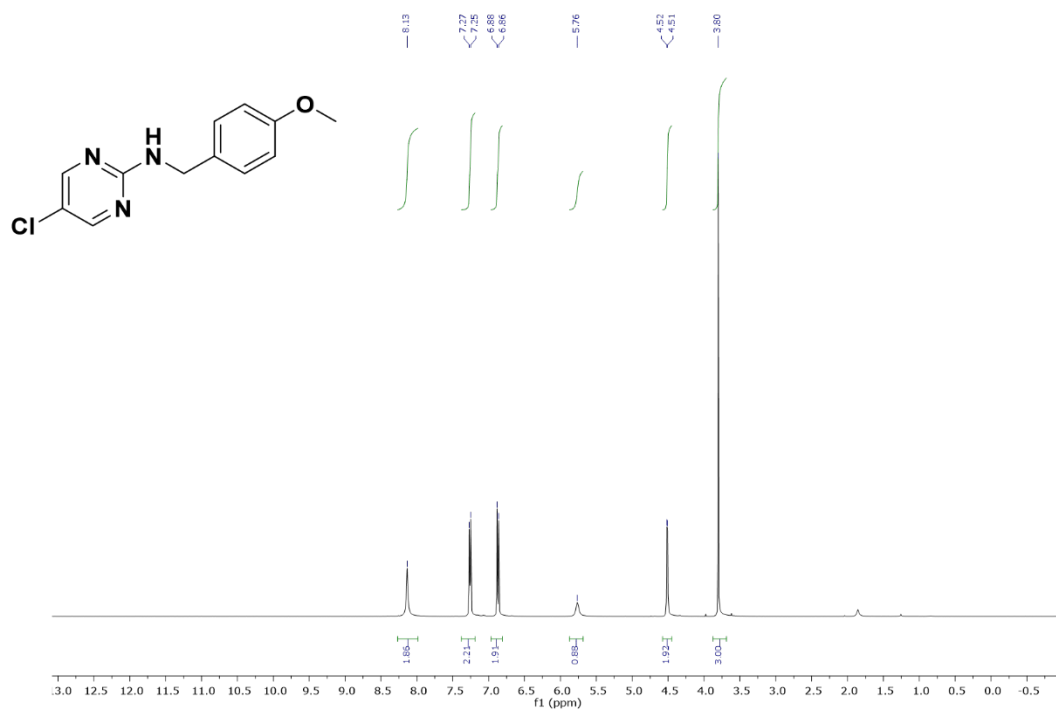


**<sup>1</sup>H NMR of 5-chloro-N-(pyridin-2-ylmethyl)pyrimidin-2-amine (1-SO-05) in CDCl<sub>3</sub>**

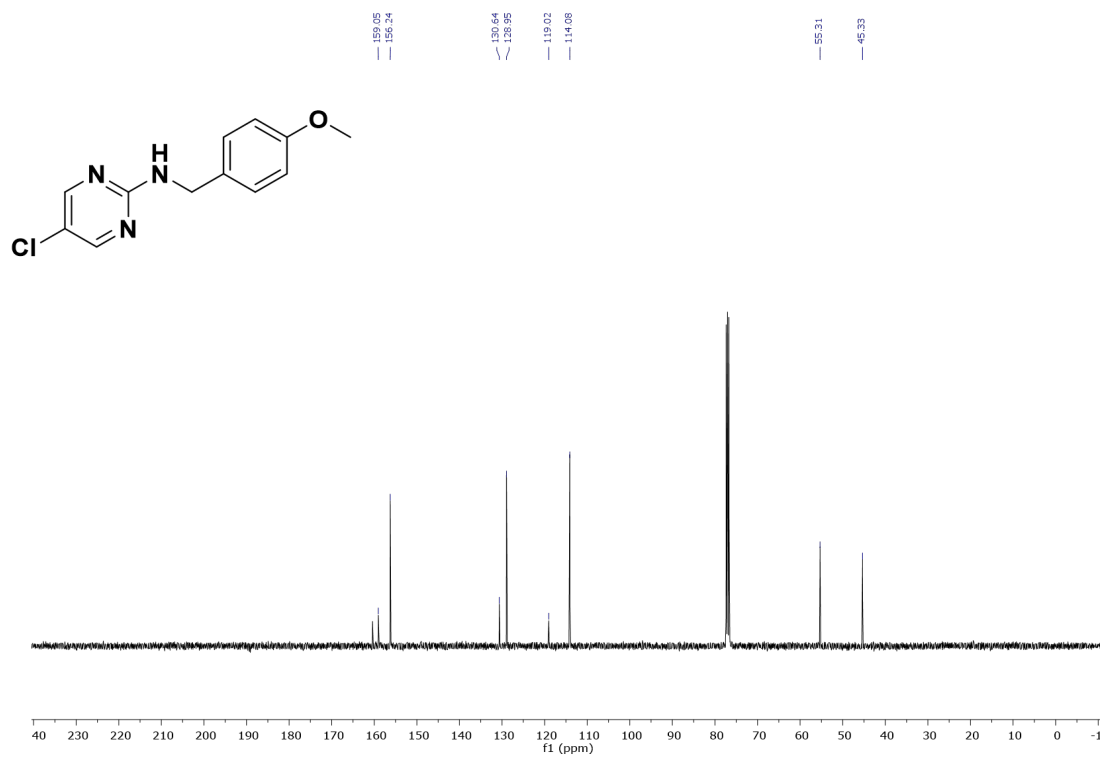


**<sup>13</sup>C NMR of 5-chloro-N-(pyridin-2-ylmethyl)pyrimidin-2-amine (1-SO-05) in CDCl<sub>3</sub>**

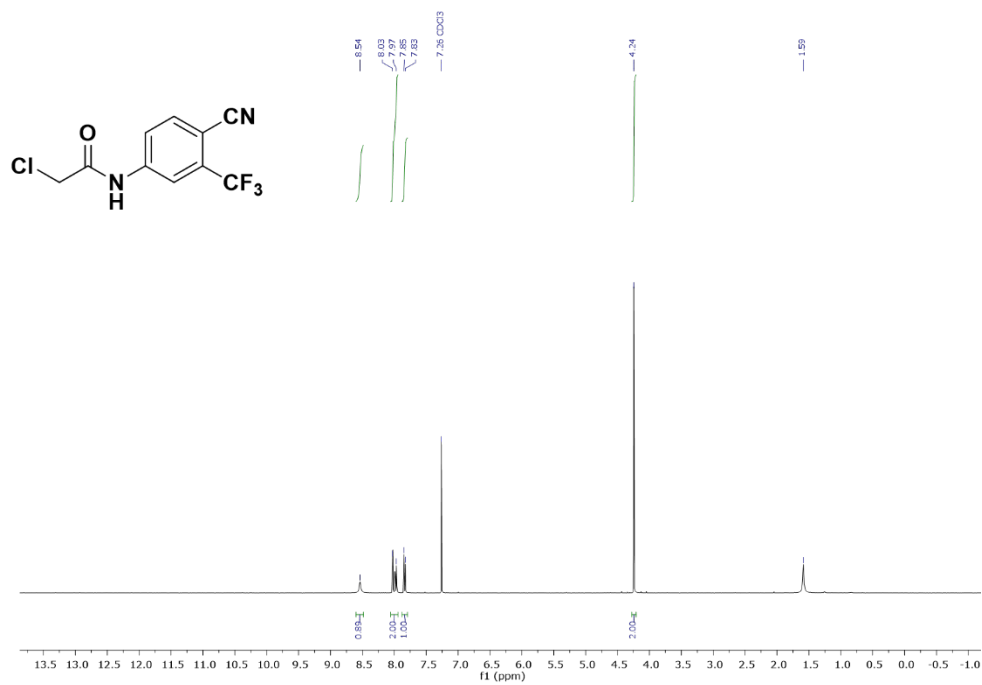




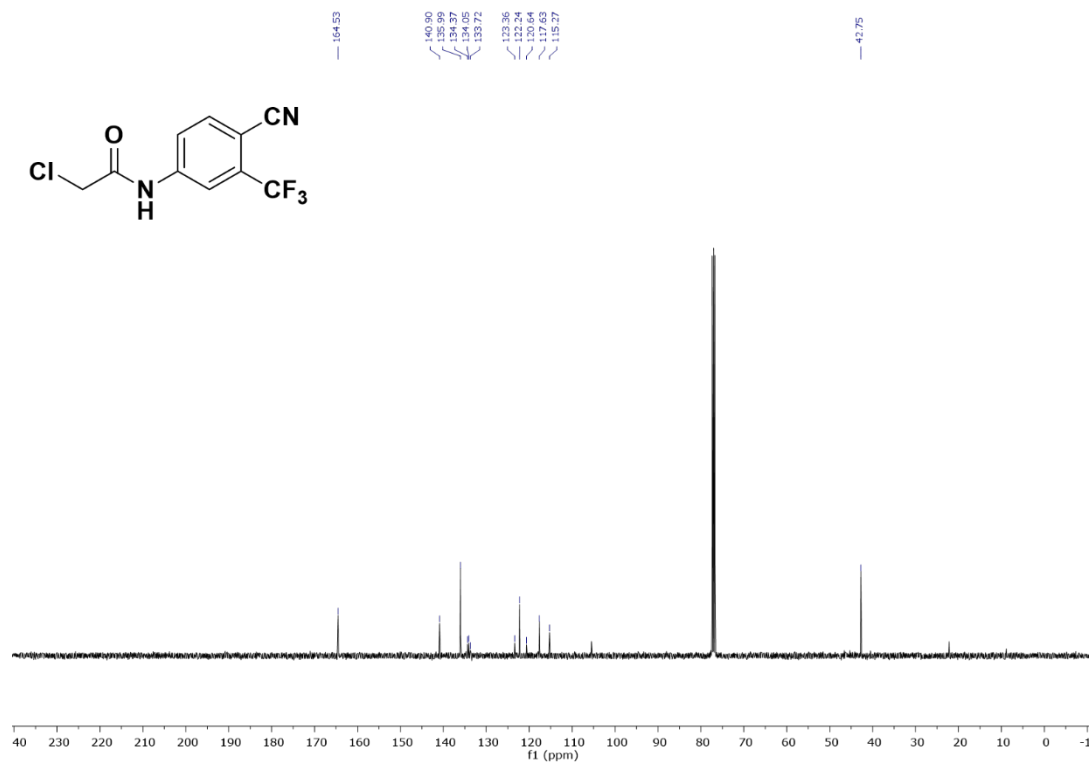
<sup>1</sup>H NMR of **Synthesis of 5-chloro-N-(4-methoxybenzyl)pyrimidin-2-amine (1-SO-29)** in CDCl<sub>3</sub>



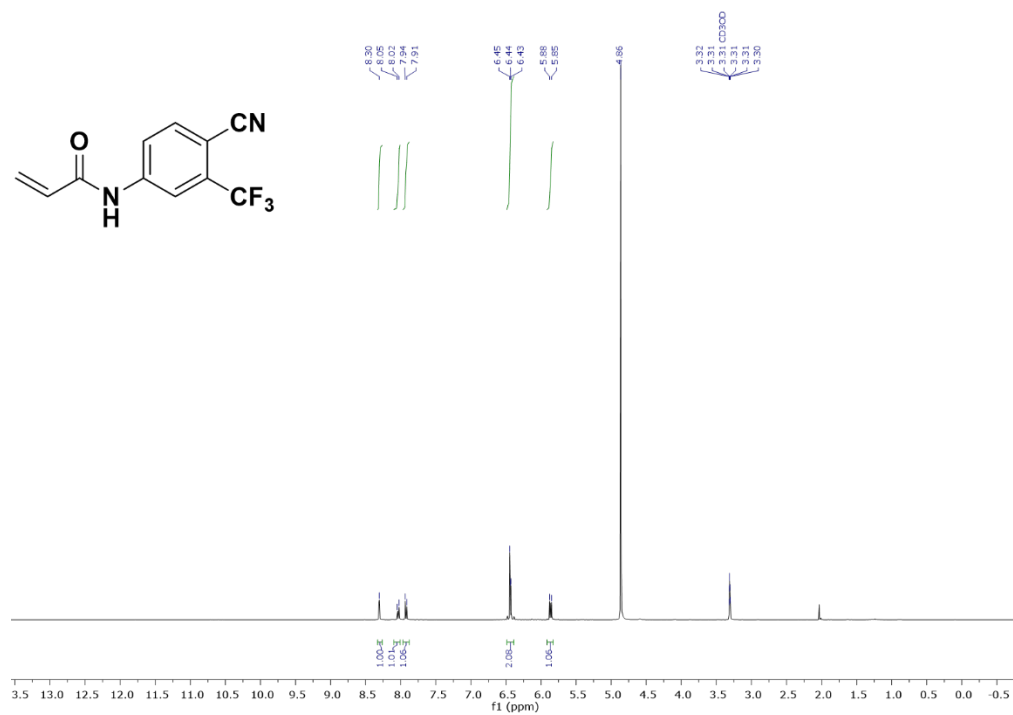
<sup>13</sup>C NMR of **Synthesis of 5-chloro-N-(4-methoxybenzyl)pyrimidin-2-amine (1-SO-29)** in CDCl<sub>3</sub>



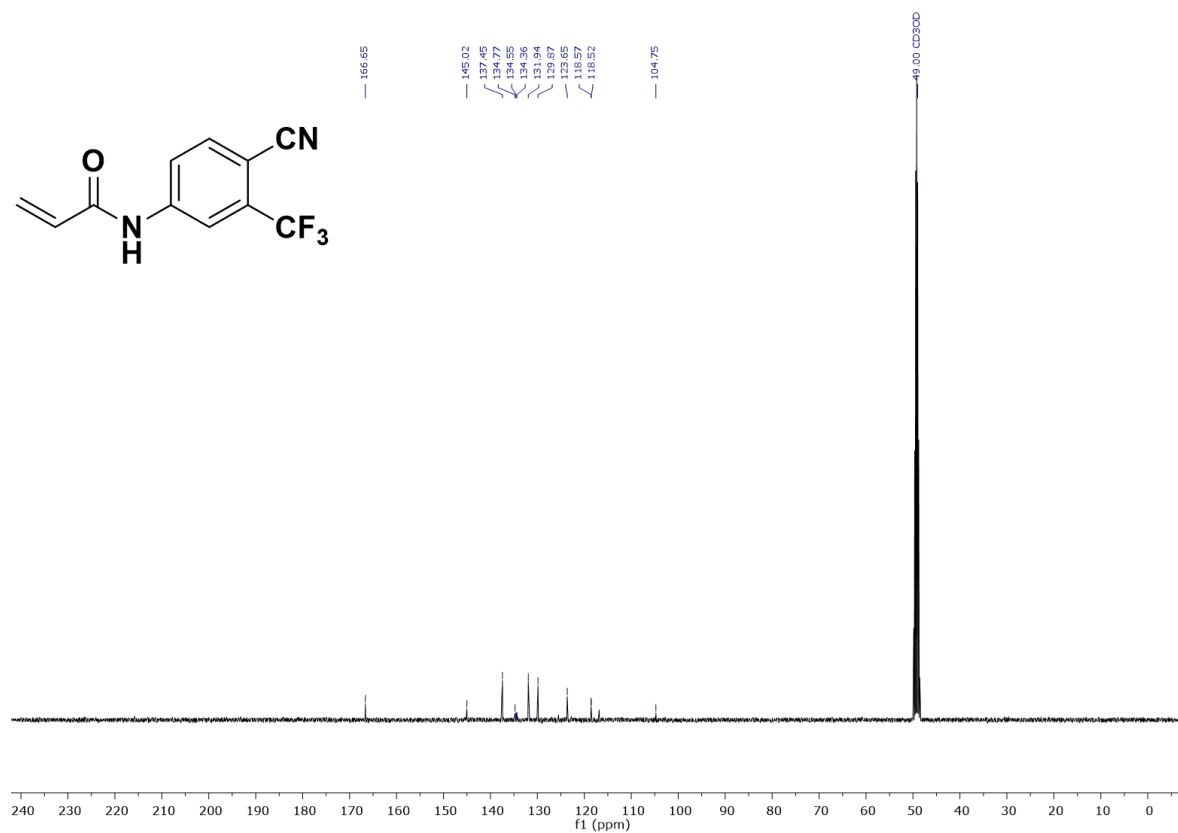
**<sup>1</sup>H NMR of 2-chloro-N-(4-cyano-3-(trifluoromethyl)phenyl)acetamide (1-SO-56) in CDCl<sub>3</sub>**



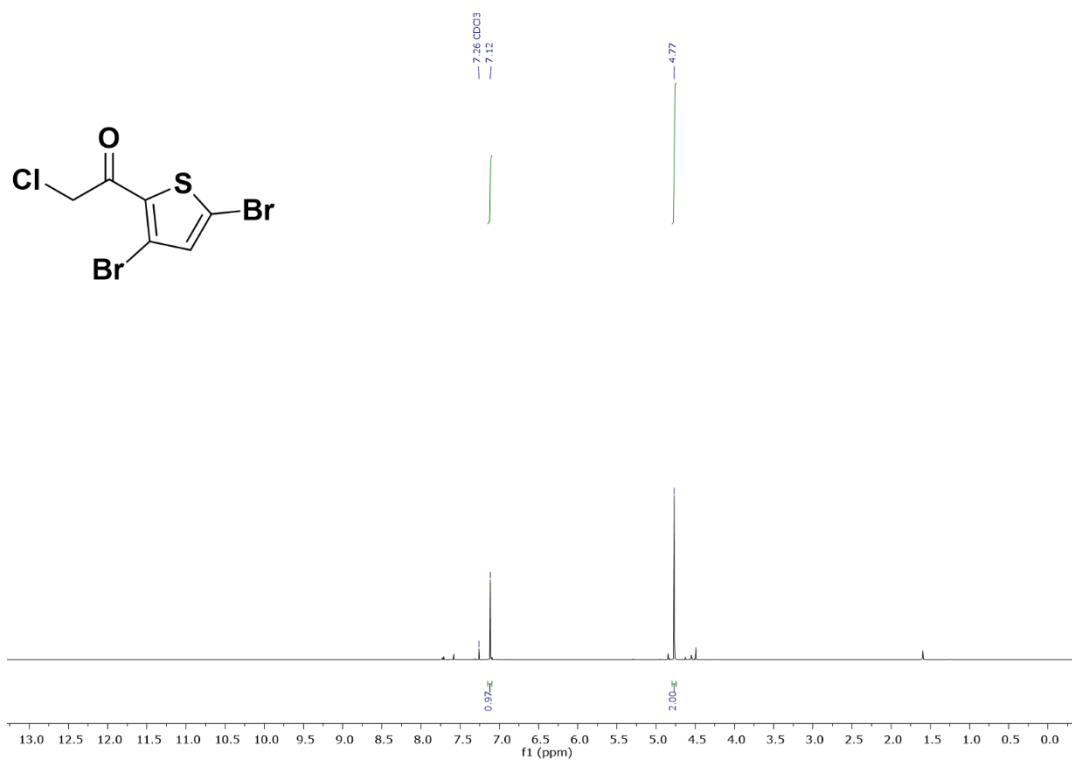
**<sup>13</sup>C NMR of 2-chloro-N-(4-cyano-3-(trifluoromethyl)phenyl)acetamide (1-SO-56) in CDCl<sub>3</sub>**



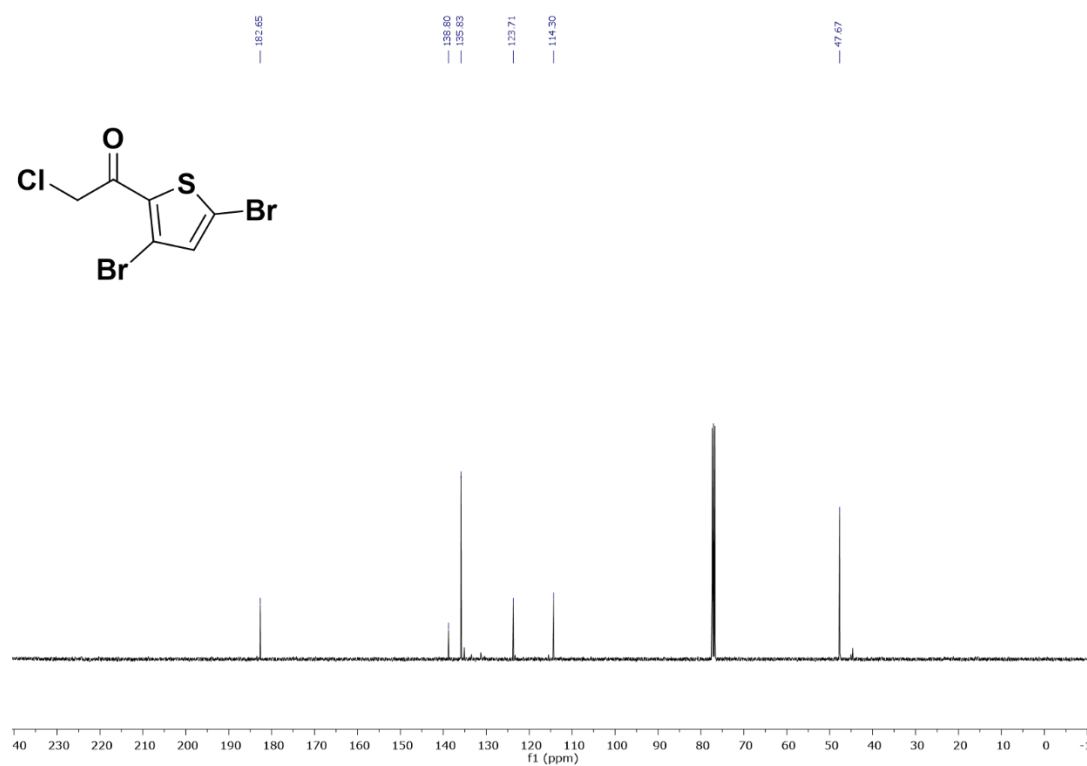
**<sup>1</sup>H NMR of N-(4-cyano-3-(trifluoromethyl)phenyl)acrylamide (1-SO-59) in CD<sub>3</sub>OD**



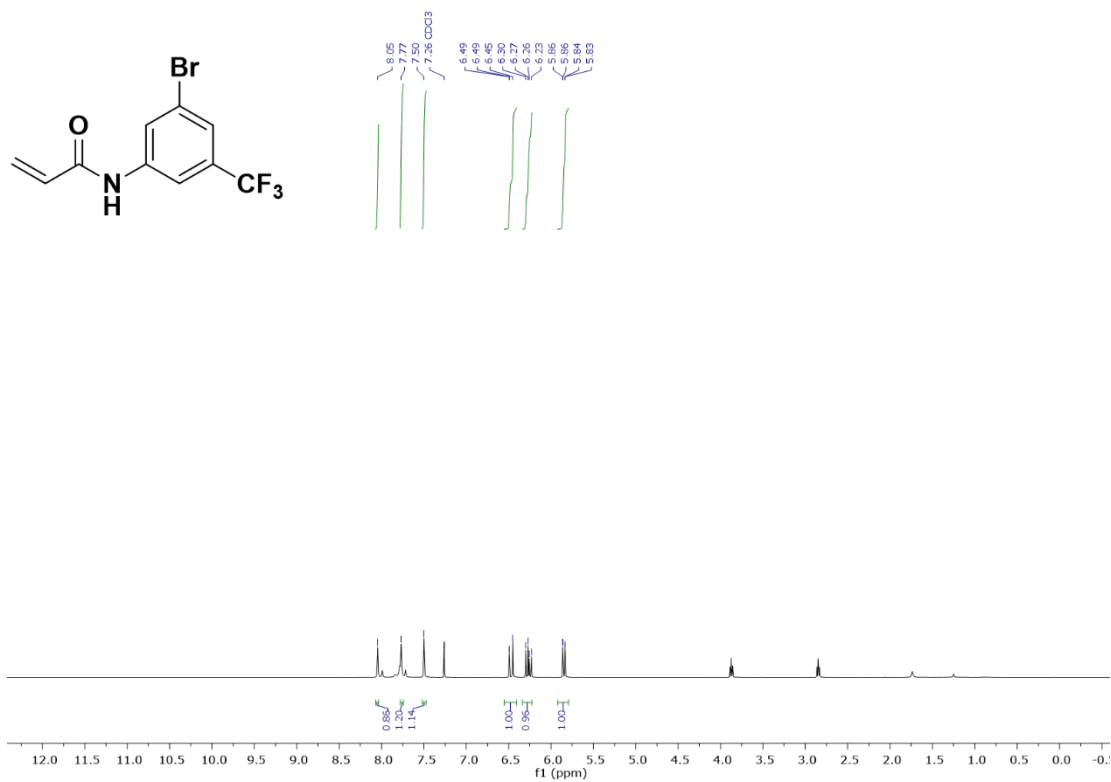
**<sup>13</sup>C NMR of N-(4-cyano-3-(trifluoromethyl)phenyl)acrylamide (1-SO-59) in CD<sub>3</sub>OD**



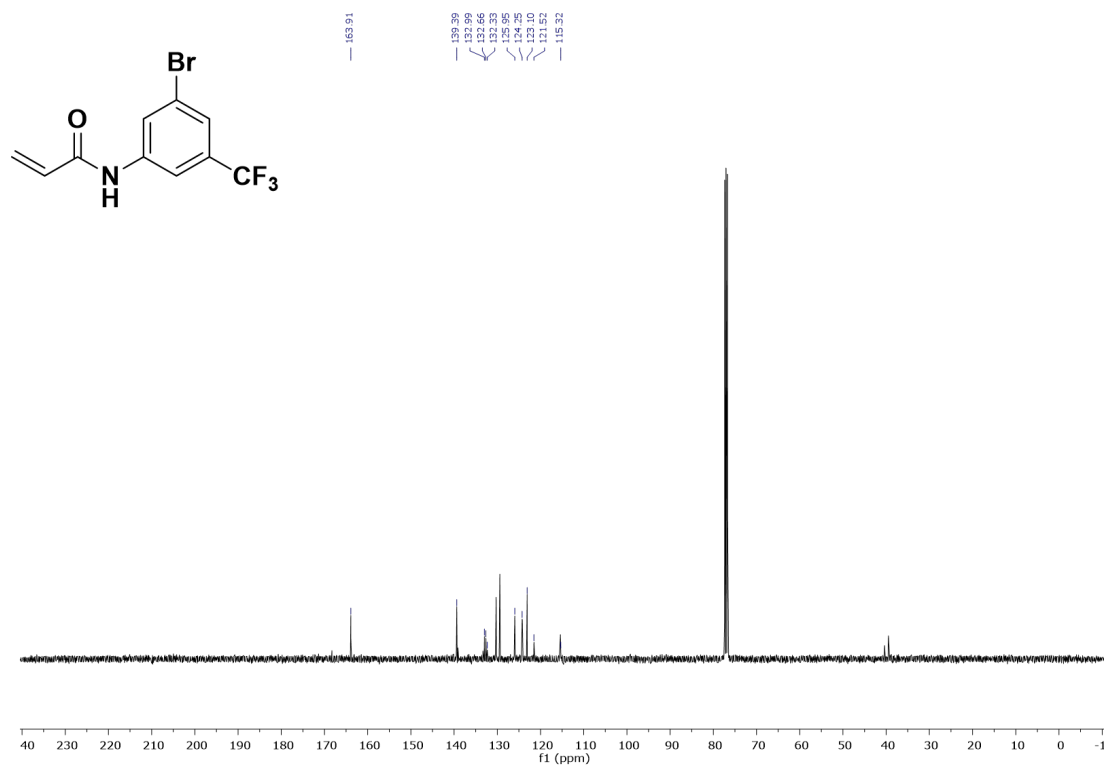
<sup>1</sup>H NMR of 2-chloro-1-(3,5-dibromothiophen-2-yl)ethan-1-one (1-SO-83) in CDCl<sub>3</sub>



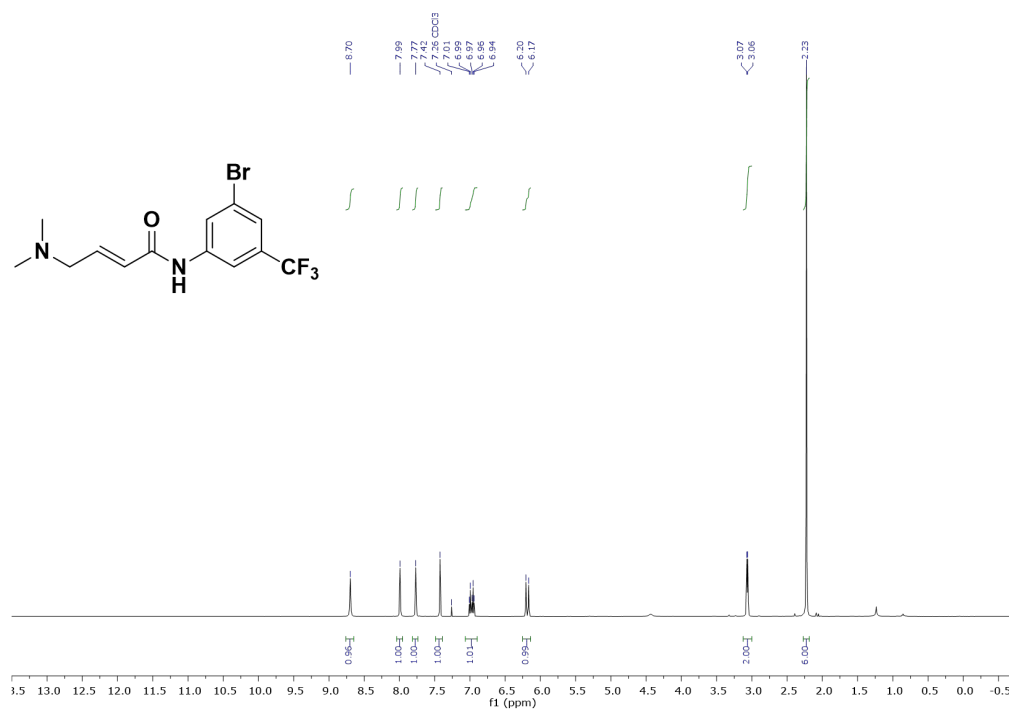
<sup>13</sup>C NMR of 2-chloro-1-(3,5-dibromothiophen-2-yl)ethan-1-one (1-SO-83) in CDCl<sub>3</sub>



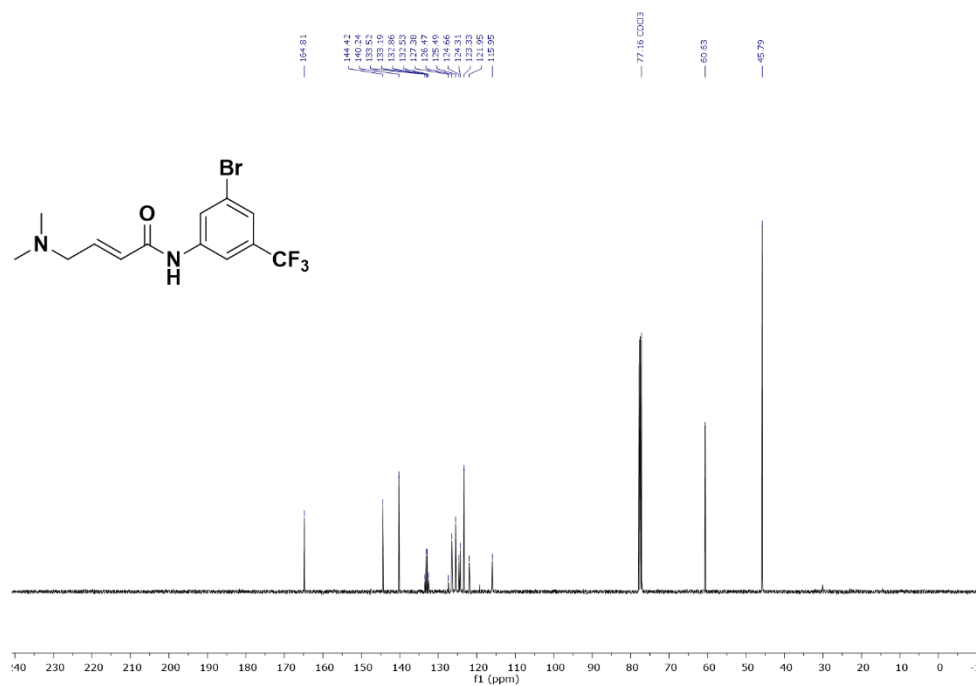
$^1\text{H}$  NMR of **N-(3-bromo-5-(trifluoromethyl)phenyl)acrylamide (1-SO-98)** in  $\text{CDCl}_3$



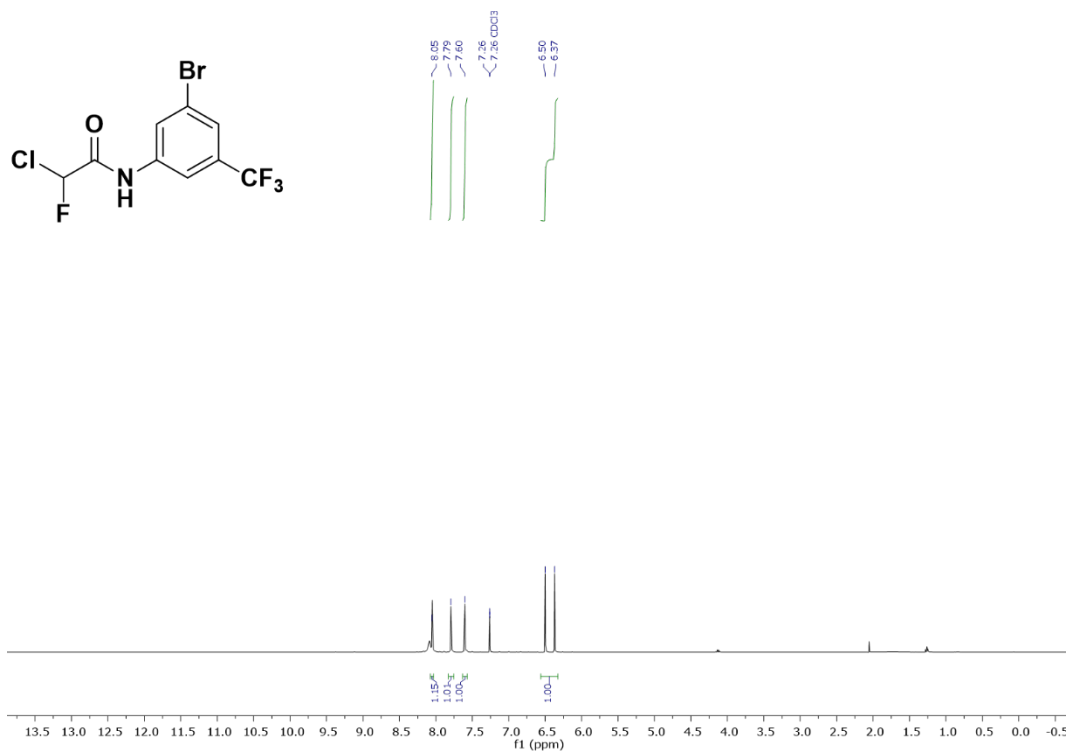
$^{13}\text{C}$  NMR of **N-(3-bromo-5-(trifluoromethyl)phenyl)acrylamide (1-SO-98)** in  $\text{CDCl}_3$



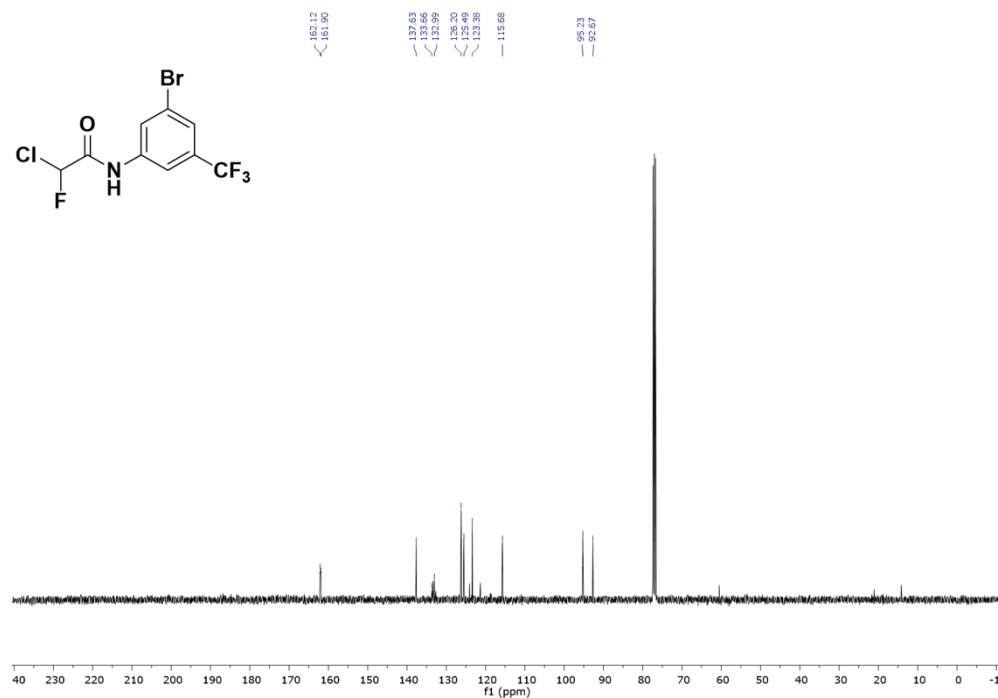
**<sup>1</sup>H NMR of (E)-N-(3-bromo-5-(trifluoromethyl)phenyl)-4-(dimethylamino)but-2-enamide (1-SO-176) in CDCl<sub>3</sub>**



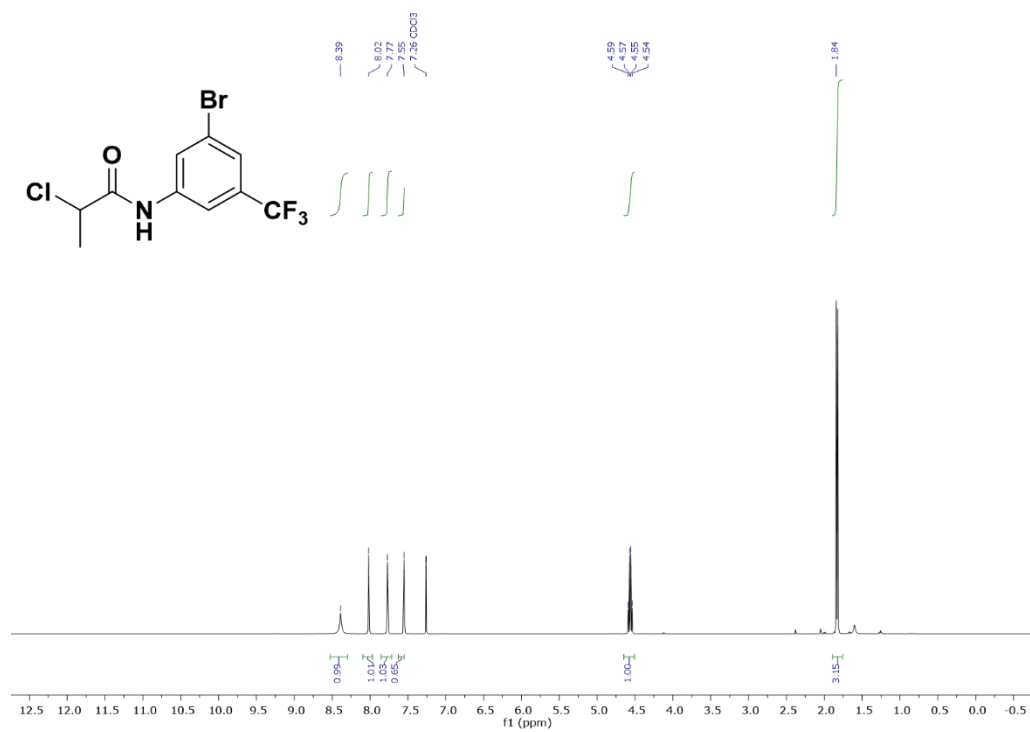
**<sup>13</sup>C NMR of (E)-N-(3-bromo-5-(trifluoromethyl)phenyl)-4-(dimethylamino)but-2-enamide (1-SO-176) in CDCl<sub>3</sub>**



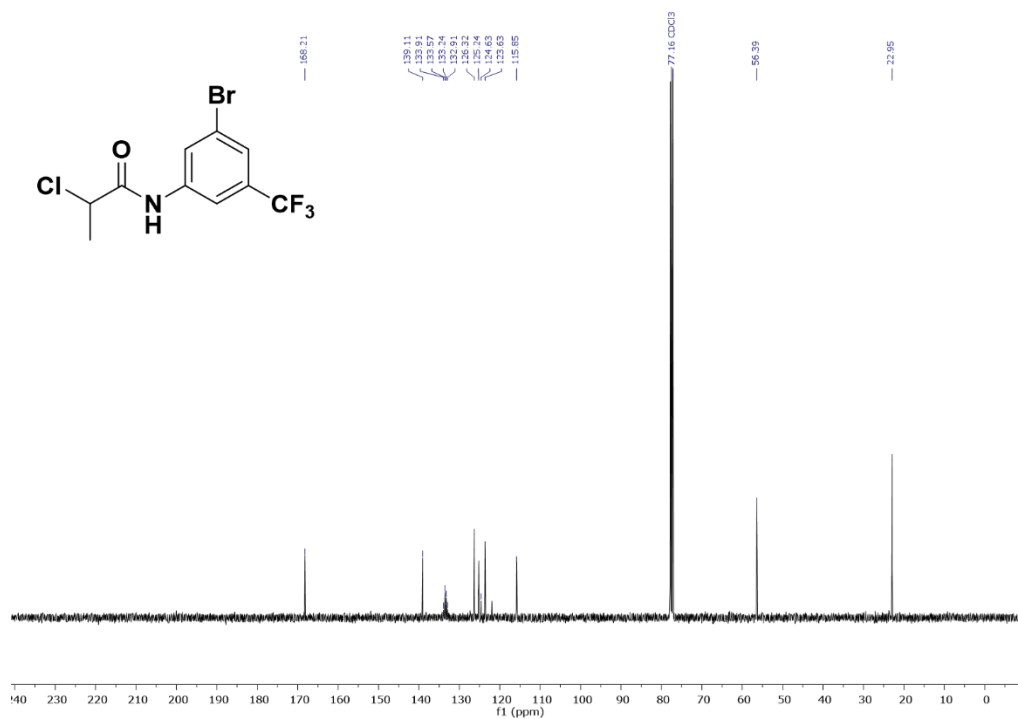
<sup>1</sup>H NMR of **N-(3-bromo-5-(trifluoromethyl)phenyl)-2-chloro-2-fluoroacetamide (2-SO-237)** in CDCl<sub>3</sub>



<sup>13</sup>C NMR of **N-(3-bromo-5-(trifluoromethyl)phenyl)-2-chloro-2-fluoroacetamide (2-SO-237)** in CDCl<sub>3</sub>

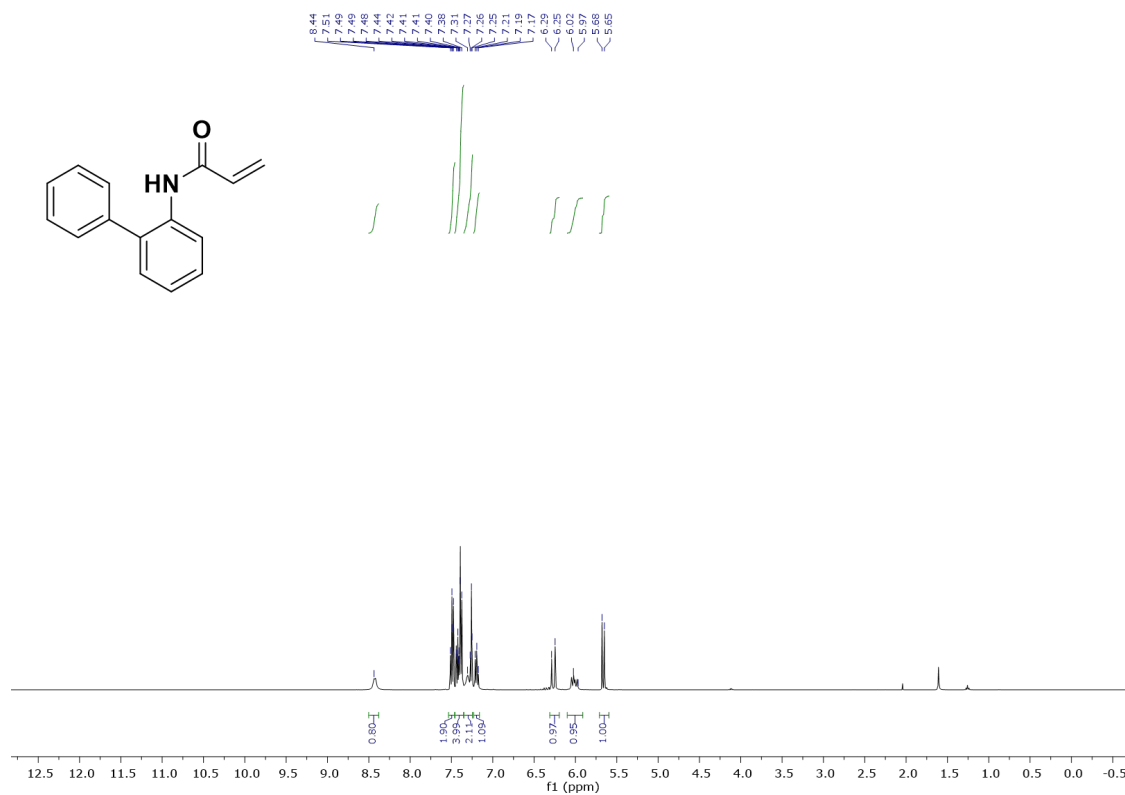


<sup>1</sup>H NMR of **N-(3-bromo-5-(trifluoromethyl)phenyl)-2-chloropropanamide (2-SO-238)** in CDCl<sub>3</sub>

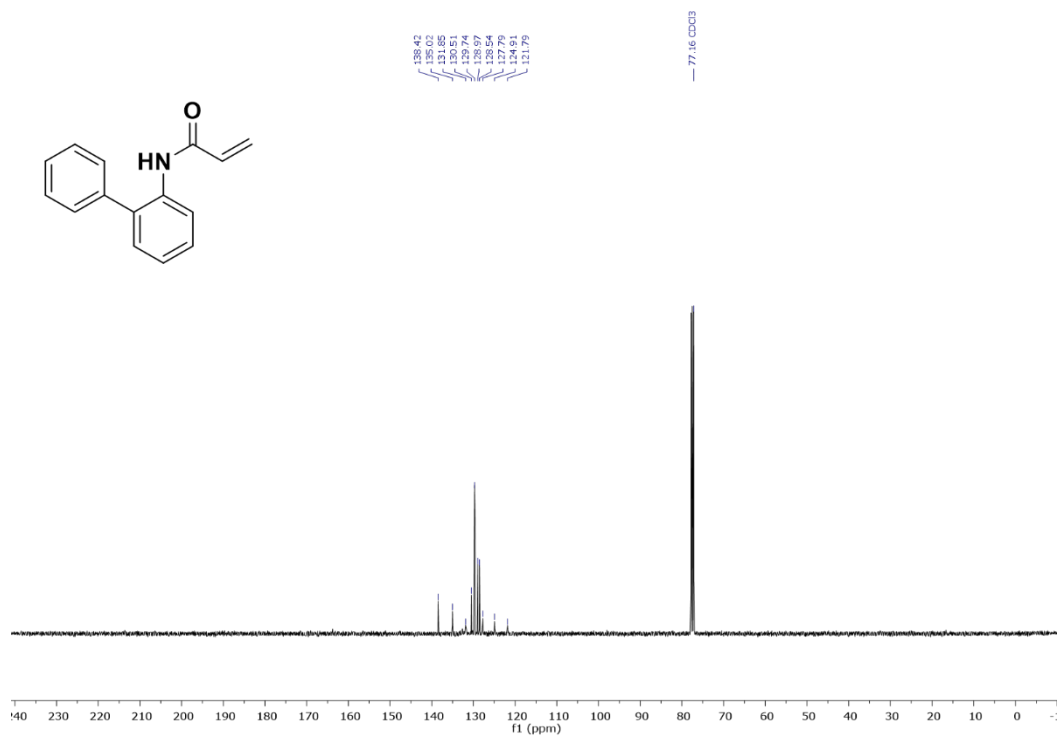


<sup>13</sup>C NMR of **N-(3-bromo-5-(trifluoromethyl)phenyl)-2-chloropropanamide (2-SO-238)** in CDCl<sub>3</sub>

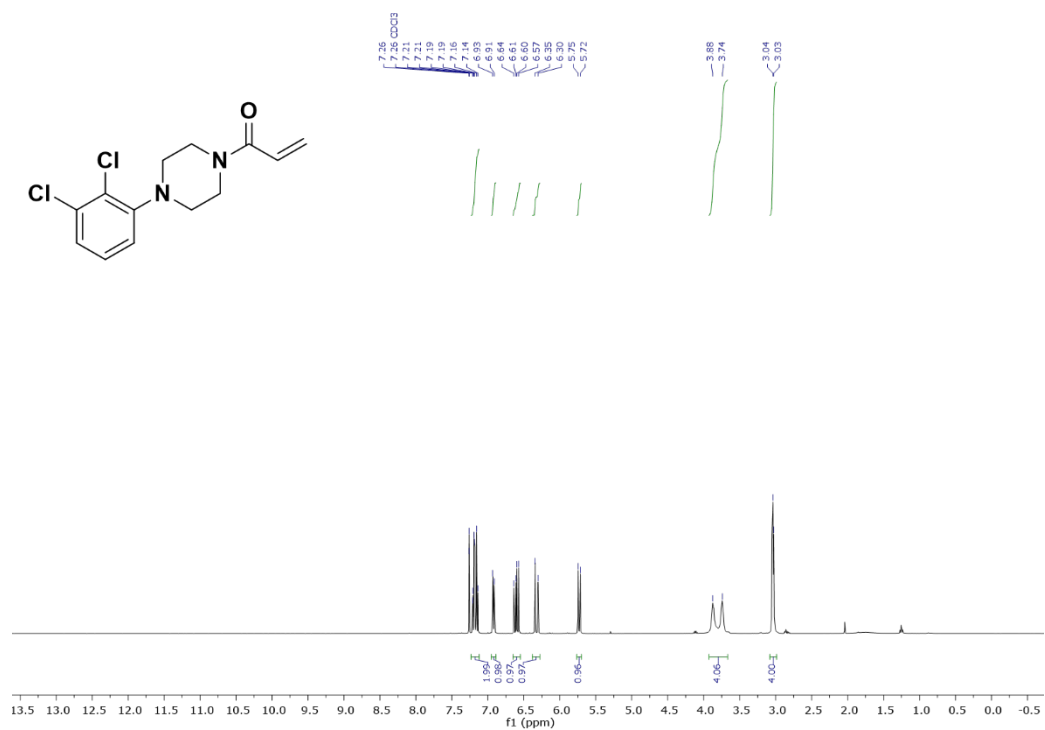




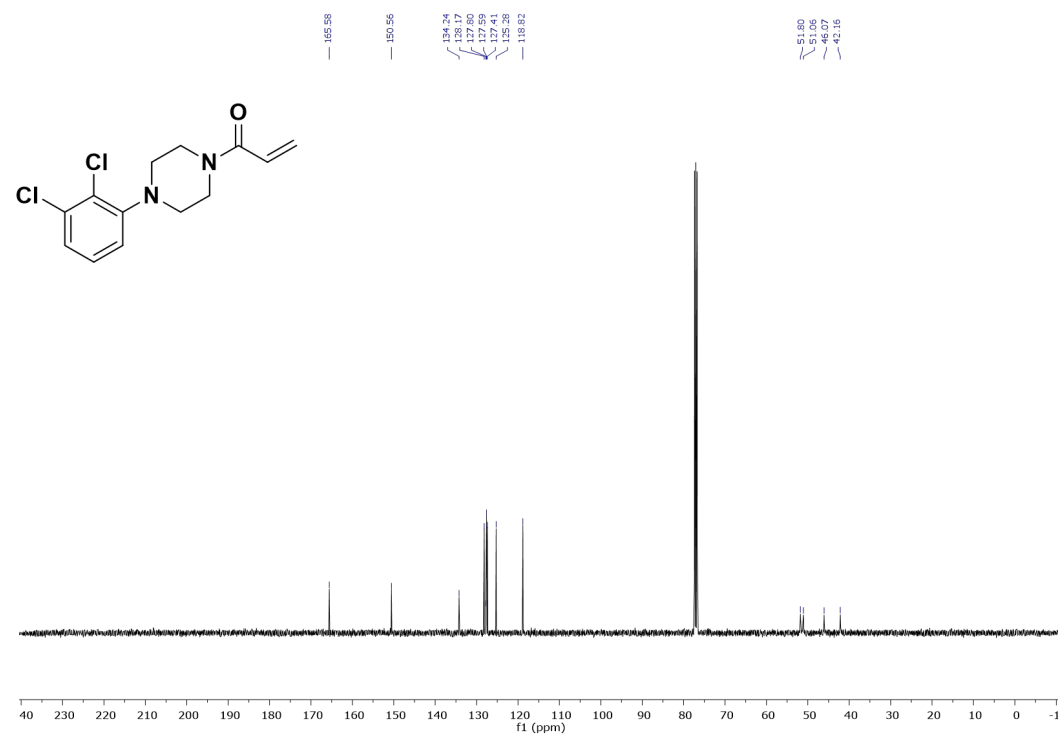
**<sup>1</sup>H NMR of 1N-([1,1'-biphenyl]-2-yl)acrylamide (1-SO-96) in CDCl<sub>3</sub>**



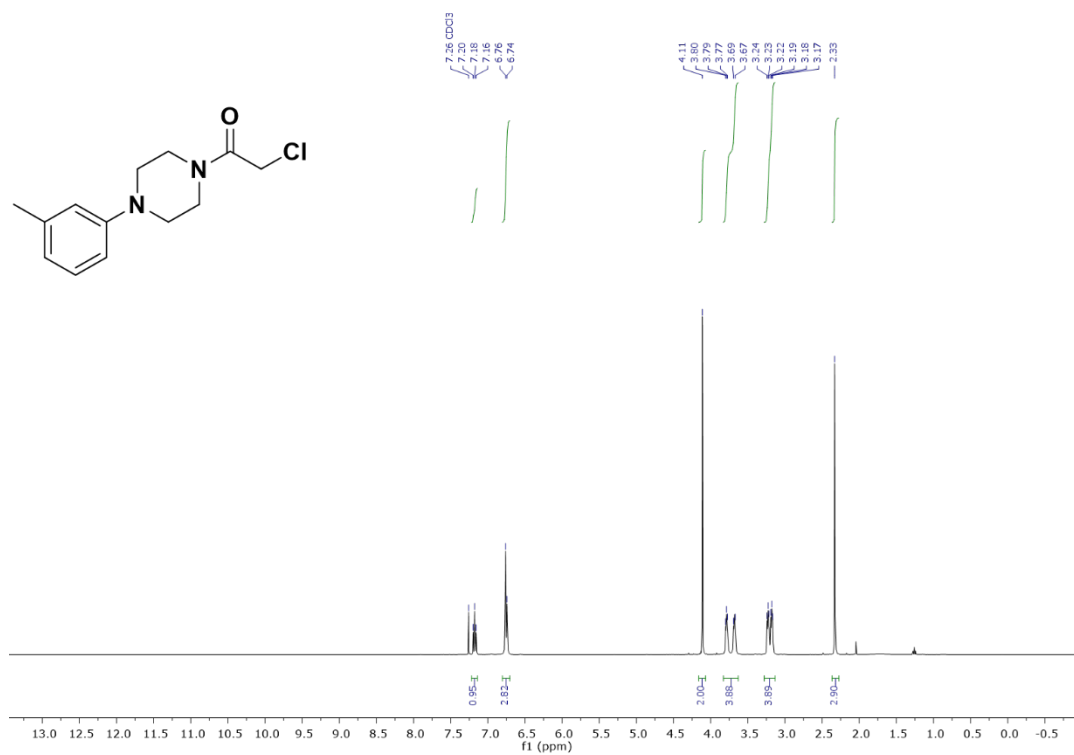
**<sup>13</sup>C NMR of N-([1,1'-biphenyl]-2-yl)acrylamide (1-SO-96) in CDCl<sub>3</sub>**



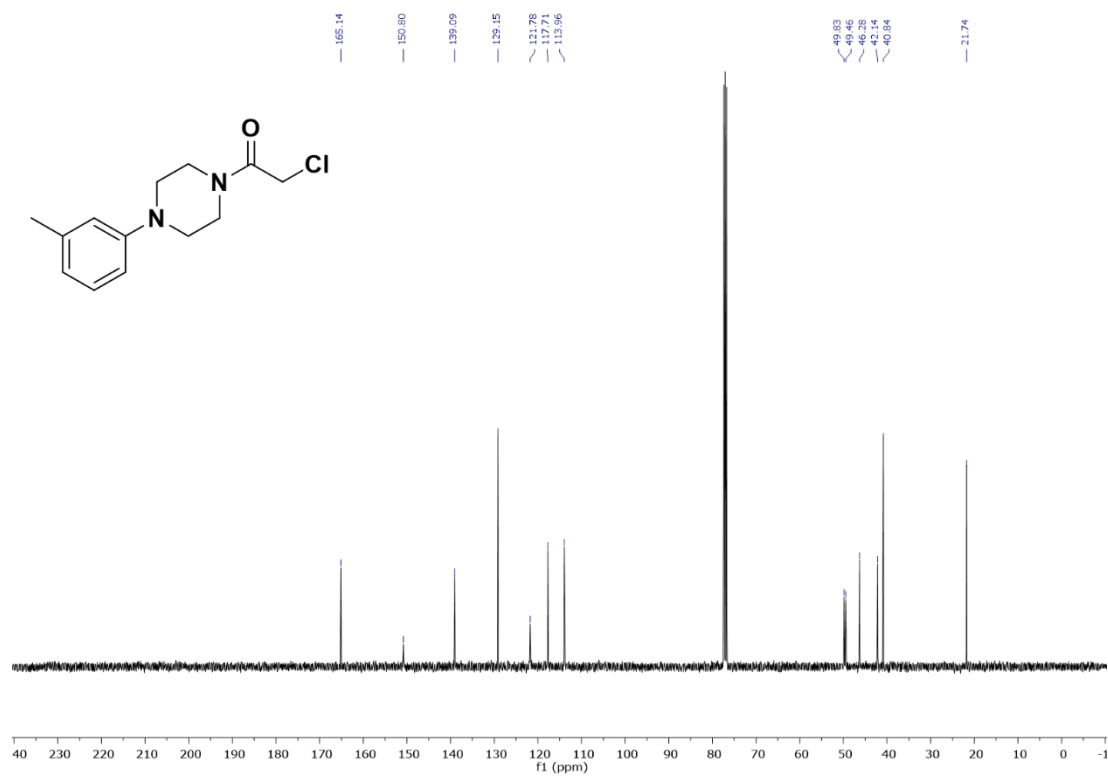
**<sup>1</sup>H NMR of 1-(4-(2,3-dichlorophenyl)piperazin-1-yl)prop-2-en-1-one (1-SO-60) in CDCl<sub>3</sub>**



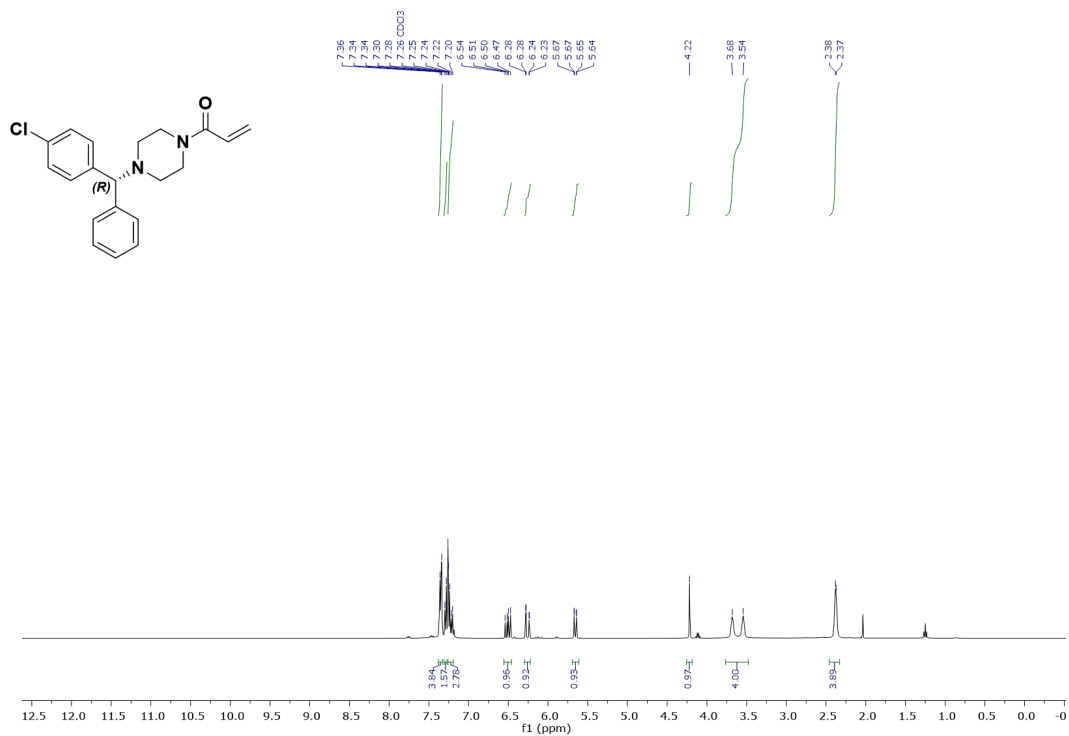
**<sup>13</sup>C NMR of 1-(4-(2,3-dichlorophenyl)piperazin-1-yl)prop-2-en-1-one (1-SO-60) in CDCl<sub>3</sub>**



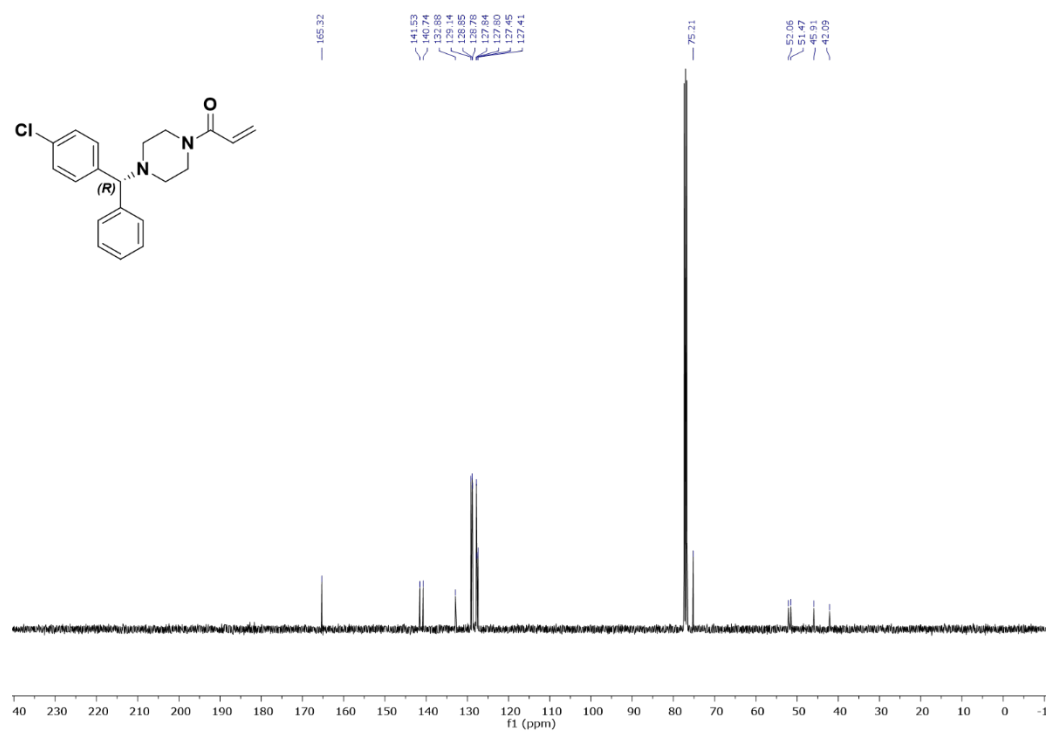
<sup>1</sup>H NMR of 2-chloro-1-(4-(m-tolyl)piperazin-1-yl)ethan-1-one (1-SO-68) in CDCl<sub>3</sub>



<sup>13</sup>C NMR of 2-chloro-1-(4-(m-tolyl)piperazin-1-yl)ethan-1-one (1-SO-68) in CDCl<sub>3</sub>



**<sup>1</sup>H NMR of (R)-1-(4-((4-chlorophenyl)(phenyl)methyl)piperazin-1-yl)prop-2-en-1-one (2-SO-253) in CDCl<sub>3</sub>**



**<sup>13</sup>C NMR of (R)-1-(4-((4-chlorophenyl)(phenyl)methyl)piperazin-1-yl)prop-2-en-1-one (2-SO-253) in CDCl<sub>3</sub>**

## References

1. Weerapana, E. *et al.* Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **468**, 790–795 (2010).
2. Yan, T. *et al.* Enhancing Cysteine Chemoproteomic Coverage through Systematic Assessment of Click Chemistry Product Fragmentation. *Anal. Chem.* **94**, 3800–3810 (2022).
3. Cao, J. *et al.* Multiplexed CuAAC Suzuki-Miyaura Labeling for Tandem Activity-Based Chemoproteomic Profiling. *Anal. Chem.* **93**, 2610–2618 (2021).
4. Boatner, L. M., Palafox, M. F., Schweppe, D. K. & Backus, K. M. CysDB: a human cysteine database based on experimental quantitative chemoproteomics. *Cell Chem. Biol.* **30**, 683-698.e3 (2023).
5. Yan, T. *et al.* SP3-FAIMS Chemoproteomics for High-Coverage Profiling of the Human Cysteinome\*. *Chembiochem* **22**, 1841–1851 (2021).
6. Burton, N. R. *et al.* Solid-Phase Compatible Silane-Based Cleavable Linker Enables Custom Isobaric Quantitative Chemoproteomics. *J. Am. Chem. Soc.* **145**, 21303–21318 (2023).
7. Collins, D. M. *et al.* Preclinical Characteristics of the Irreversible Pan-HER Kinase Inhibitor Neratinib Compared with Lapatinib: Implications for the Treatment of HER2-Positive and HER2-Mutated Breast Cancer. *Cancers (Basel)* **11**, (2019).

### Chapter 3: Delineating Cysteine-Reactive Compound Modulation of Cellular Proteostasis Processes

Reprinted with permission from

**Shikwana, F.\***, Julio, A. R.\* , Truong, C., Burton, N. R., Dominguez, E. R., Turmon, A. C., Cao, J. & Backus, K. M. Delineating cysteine-reactive compound modulation of cellular proteostasis processes. *Nat. Chem. Biol.* (2024). doi:10.1038/s41589-024-01760-9

Copyright © 2024 Springer Nature America, Inc

Ashley R. Julio<sup>1,2,6</sup>, Flowreen Shikwana<sup>1,2,6</sup>, Cindy Truong<sup>1</sup>, Nikolas R. Burton<sup>1,2</sup>, Emil R. Dominguez III<sup>1</sup>, Alexandra C. Turmon<sup>1,2</sup>, Jian Cao<sup>2</sup>, Keriann M. Backus<sup>1,2,3,4,5</sup>

<sup>1</sup>Department of Biological Chemistry, David Geffen School of Medicine, UCLA, Los Angeles, CA 90095 (USA); <sup>2</sup>Department of Chemistry and Biochemistry, UCLA, Los Angeles, CA 90095 (USA); <sup>3</sup>DOE Institute for Genomics and Proteomics, UCLA, Los Angeles, CA 90095 (USA); <sup>4</sup>Jonsson Comprehensive Cancer Center, UCLA, Los Angeles, CA 90095 (USA); <sup>5</sup>Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, UCLA, Los Angeles, CA 90095 (USA).

<sup>6</sup>These authors contributed equally: Ashley R. Julio, Flowreen Shikwana.

Corresponding Author: Keriann M. Backus, Biological Chemistry Department, David Geffen School of Medicine, UCLA, Los Angeles, CA, 90095, USA, E-mail: kbackus@mednet.ucla.edu

## Abstract

Covalent modulators and covalent degrader molecules have emerged as drug modalities with tremendous therapeutic potential. Towards realizing this potential, mass spectrometry-based chemoproteomic screens have generated proteome-wide maps of potential druggable cysteine residues. However, beyond these direct cysteine-target maps, the full scope of direct and indirect activities of these molecules on cellular processes and how such activities contribute to reported modes-of-action, such as degrader activity, remains to be fully understood. Using chemoproteomics, we identified a cysteine-reactive small molecule degrader of the SARS-CoV-2 non-structural protein 14 (nsp14), which effects degradation through direct modification of cysteines in both nsp14 and in host protein disulfide isomerases. This degrader activity was further potentiated by generalized electrophile-induced global protein ubiquitylation, proteasome activation, and widespread aggregation and depletion of host proteins, including formation of stress granules. Collectively, we delineate the wide-ranging impacts of cysteine-reactive electrophilic compounds on cellular proteostasis processes.

## Introduction

Mass spectrometry-based chemoproteomics have enabled the discovery of tens of thousands of potentially druggable cysteine residues<sup>1–8</sup>. These screening platforms and resulting datasets represent a potentially transformative resource for chemical probe and drug development efforts. In typical chemoproteomics screens, compounds are applied to cells at low to mid-micromolar concentrations, under the working model that initial screening hits may only have modest potency. Thus, initial scout molecules nearly always engage many cysteines, and the impact of this polypharmacology on target deconvolution and downstream mode-of-action and functional studies remains largely unclear.

Nearly all screening hits fail to conform to the stringent criteria set forth for bona fide chemical probes<sup>9–11</sup> and instead function as scout or pathfinder molecules. While progressing covalent-

fragment-like screening hits to lead-like molecules is feasible<sup>12–14</sup>, several cautionary examples of polypharmacology driving biological effects also exist, most notably RA190<sup>15</sup>, which was originally identified as an proteasomal ubiquitin receptor ADRM1 inhibitor<sup>16</sup> but subsequently found to increase global polyubiquitylation and alter cell cycle progression through a still cryptic mechanism not involving cysteine labeling in ADRM1<sup>15</sup>. Thus, defining how to delineate bona fide pathfinder compounds from more promiscuous acting molecules is an essential step to realizing the full potential of cysteine chemoproteomics in target hunting.

A handful of prior reports provide hints that cysteine-reactive small molecules, used at the concentrations applied in chemoproteomic screens, can cause cell stress and altered proteostasis response mechanisms<sup>15,17</sup>. The formation of stress granules (SGs) comprised of phase separated RNA and proteins is a general cellular response to stress<sup>18</sup>. The classic inducer of SGs is sodium arsenite<sup>18</sup>, which causes oxidative stress and also induces global increases in polyubiquitylation<sup>19</sup>. A second major response to cell stress is formation of Sequestosome 1 (SQSTM1/p62) aggresomes and activation of autophagy mediated degradation<sup>20</sup>. Proteasomes are recruited to SGs and aggresomes to facilitate clearance of aggregates<sup>21</sup>. Whether and to what extent cysteine-reactive druglike molecules impact these cell stress response processes remains to be seen.

The use of cysteine-reactive molecules to specifically modulate proteostasis mechanisms is a burgeoning area for both functional biology and drug development. Covalent PROTACs<sup>22–28</sup> and molecular glues<sup>29–33</sup> function by covalent modification of either an E3 ubiquitin ligase or a bait protein, primarily at cysteine residues, followed by ternary complex formation and subsequent ubiquitylation and proteasomal degradation of the target protein. Many of the reported covalent degraders show modest mid-micromolar potency<sup>29,30,34,35</sup>.

Here, we assess the impact of cysteine-reactive electrophilic compounds on cellular proteostasis processes. Using cell-based cysteine chemoproteomics we first identified a covalent degrader of the SARS-CoV-2 non-structural protein 14 (nsp14). Mechanistic studies reveal that degradation



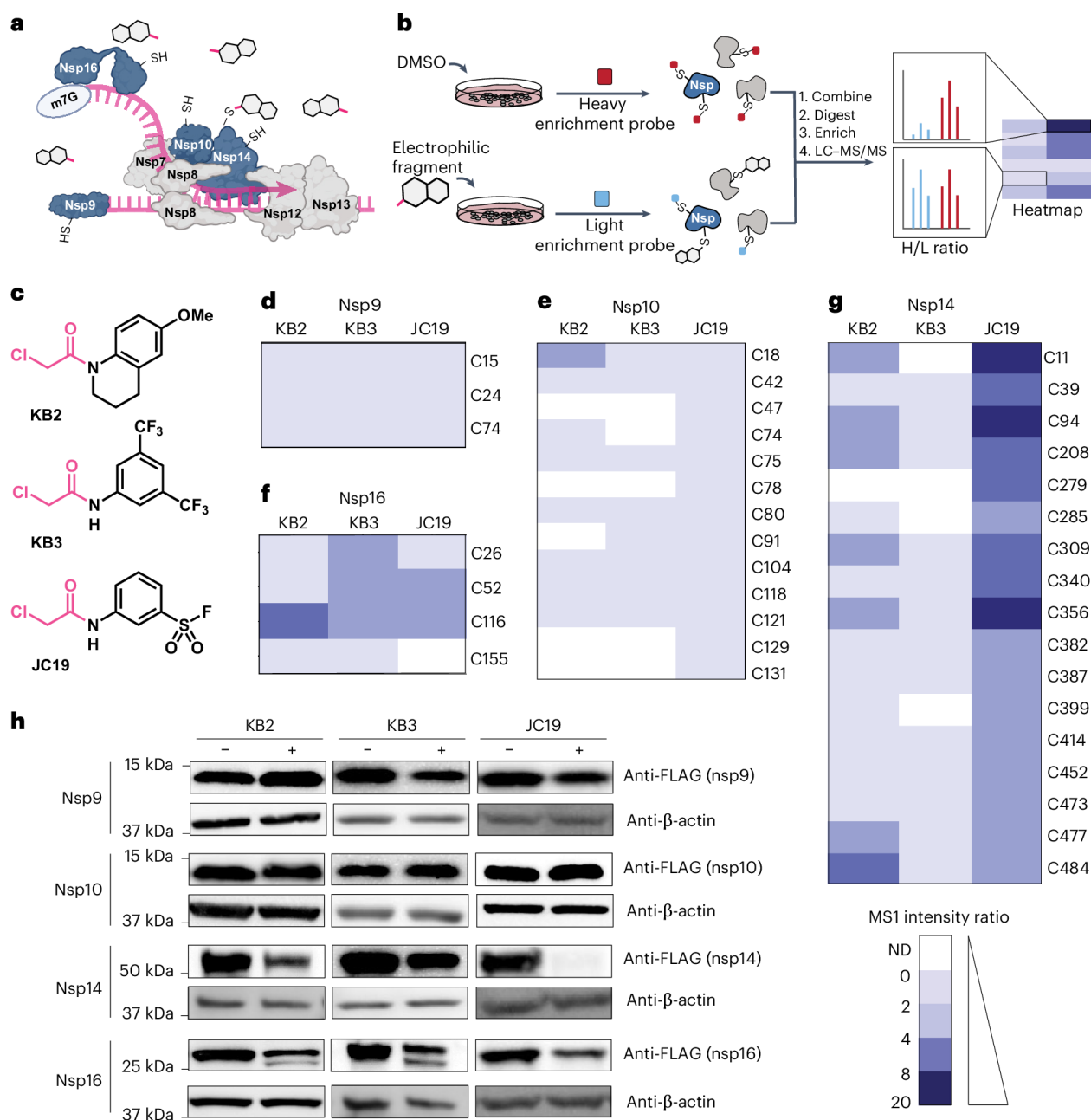
occurs via covalent modification of both nsp14 and host chaperones together with compound-induced association with the proteasome and recruitment into aggresomes. We then broaden our analysis to consider established degrader molecules and electrophilic compounds used widely for chemoproteomic screens and find that these compounds induce global changes in cellular proteostasis, including markedly elevated ubiquitylation, and proteasome activation. SG formation was observed to be a ubiquitous feature of cysteine-reactive compounds, rationalizing the observed pervasive aggregation and depletion of nucleoporins. Enabled by our rich multimodal proteomic datasets and disclosure of covalent-induced cell-stress, we lay the foundation for ensuring on-target mode-of-action for covalent inhibitors and degrader molecules.

## Results

**Chemoproteomic analysis of SARS-CoV-2 nonstructural proteins.** Due to their cysteine-rich sequences (**Supplementary Table 1, Supplementary Fig. 1**), we selected a panel of four individual non-structural proteins (nsps) from SARS-CoV-2 (**Fig. 1A**), nsp9<sup>36</sup>, nsp10<sup>37,38</sup>, nsp14<sup>37,39</sup>, and nsp16<sup>38</sup>) to serve as test case proteins for cysteine chemoproteomic mode-of-action studies. Using isotopic tandem orthogonal proteolysis- activity based protein profiling (isoTOP-ABPP)<sup>2,3</sup> (**Fig. 1B**) we screened cells overexpressing each nsp with a focused panel of three electrophilic fragments KB2 (**1**), KB3 (**2**), and JC19 (**3**) (**Fig. 1C, Extended Data Fig. 1**). Promiscuously reactive scout molecules KB2 and KB3<sup>2</sup> were selected to provide rapid insight into the ligandability of nsp cysteines. Our selection of the dual electrophile-containing compound JC19 was inspired by a number of established bioactive compounds with two electrophilic sites, including RA190<sup>15</sup>, triterpenoids such as bardoxolone<sup>40</sup>, and manumycin A<sup>29</sup>. We additionally hypothesized that JC19 could enable identification of additional “glueable” interactions, analogous to the recently reported Molecular COUPLrs<sup>41</sup>.

We find that nsp9 and nsp10 were generally insensitive to compound treatment (**Fig. 1D,E**). By contrast, nsp14 and, to a lesser extent, nsp16 showed elevated ratios across the majority of

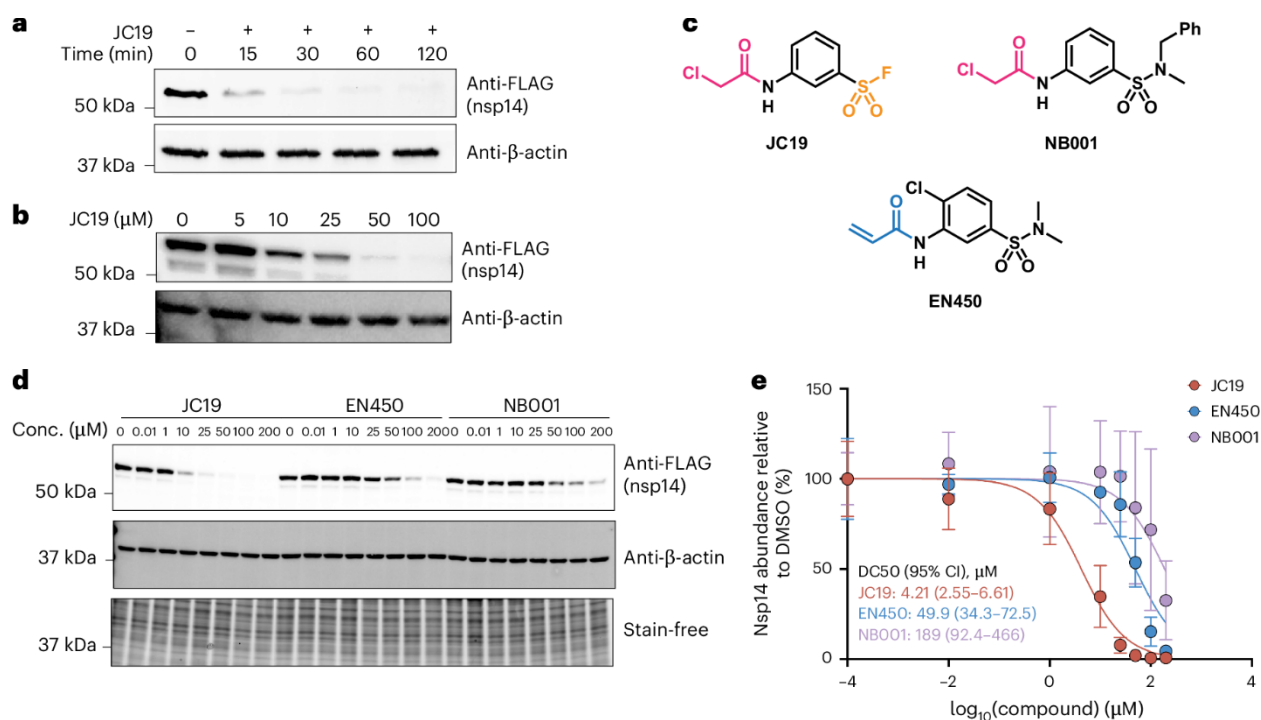
cysteines detected (**Fig. 1F,G**). While for nsp16 this effect was similar for all three compounds tested, for nsp14 the heightened ratios were specific to JC19. The presence of many putative ligandable cysteines in nsp14 was unexpected as prior studies had revealed that generally only a small fraction of proteinaceous cysteines in a given protein sequence are ligandable<sup>2,5,7</sup>. We speculated that the observed high ratios might reflect compound-induced depletion of nsp14 and nsp16. This hypothesis was corroborated by the pronounced compound-induced decrease in detectable nsp16 and nsp14, with nsp14 depletion specific to JC19 (**Fig. 1H**).



**Figure 1. Chemoproteomic profiling of SARS-CoV-2 nonstructural proteins with electrophilic scout fragments identifies JC19 as a putative nsp14 degrader.** (A) Schematic of SARS-CoV-2 nonstructural proteins (nsps). (B) IsoTOP-ABPP workflow used to identify ligandable cysteines within select nsps (colored blue in A). (C) Chemical structures of KB2, KB3, and JC19. Electrophilic chloroacetamide moiety is colored pink. (D-G) Heatmaps representing the unlogged MS1 isoTOP-ABPP ratio for each identified cysteine in nsp9 (D), nsp10 (E), nsp16 (F), and nsp14 (G) based on the median value from multiple biological replicates of treatment with 100  $\mu$ M of each compound for 1 h. ND = not detected. (H) Immunoblot of nsp abundance changes in response to treatment with each electrophilic fragment at 100  $\mu$ M for 1 h, relative to a DMSO-treated control. MS experiments shown in 'D-G' were conducted in  $\geq 3$  biological replicates in HEK293T cells. All MS data can be found in **Supplementary Table 1**.

**Both electrophilic moieties contribute to nsp14 depletion.** Given the numerous reports of electrophilic compounds functioning as degrader molecules<sup>22-31,33-35,42</sup>, we hypothesized that JC19 was acting as a covalent degrader of nsp14. We found that nsp14 depletion occurs rapidly, beginning as early as 15 minutes post treatment, with near complete loss of detectable protein after 1 hour (**Fig. 2A**). Nsp14 depletion was corroborated by immunocytochemistry (**Supplementary Fig. 2A,B**) and was found to generalize to HeLa cells (**Supplementary Fig. 2C**). The modest potency of JC19 (**Fig. 2B**), with an apparent degradation coefficient (DC50) of 4.2  $\mu$ M, aligns with many recent reports for covalent degraders<sup>29,30,34,35</sup>.

We next assembled a set of JC19-related analogues (**Fig. 2C, Extended Data Fig. 2A**). Our main goals were to establish the portions of the molecule required for activity, to increase potency, and to identify structurally matched inactive controls. We included structurally related compound EN450 (**4**), a reported molecular glue degrader of the oncogenic transcription factor NF-kappa-B p105 subunit (NFKB1)<sup>34</sup>, as well as JC17 (**5**) and JC18 (**6**)<sup>4</sup> and generated novel analogues (NB92 (**7**), NB177 (**8**), NB179 (**9**), and NB001 (**10**)). Unexpectedly, none of the analogues reached the potency of JC19 (**Fig. 2D-E, Extended Data Fig. 2B-G**). For most analogues, with the exception of JC18 (**Extended Data Fig. 2B**), which lacks a chloroacetamide, nsp14 depletion could still be observed at higher concentrations and longer treatment times. This focused SAR analysis indicated that the mechanism of JC19 degrader activity was likely at least partially distinct to that of all other compounds tested.



**Figure 2. Both the chloroacetamide and sulfonyl fluoride electrophiles are required for the rapid depletion of nsp14.** HEK293T cells transiently expressing nsp14-FLAG for A-E. (A-B) Immunoblot assessing (A) time-dependent (JC19, 100 μM) and (B) dose-dependent (JC19, 0.5 h) changes to nsp14 abundance. (C) Chemical structures of JC19 and analogues NB001 and EN450. Cysteine-reactive chloroacetamide and acrylamide moieties are colored pink and blue, respectively. Electrophilic sulfonyl fluoride moiety is colored orange. (D) Cells treated with the indicated concentrations of JC19 or NB001 (0.5 h), or EN450 (3 h), followed by immunoblot of the detergent soluble proteome. (E) Quantification of the detergent soluble fraction of nsp14 at each tested concentration to generate DC50 curves. Data are presented as mean ± s.d. ( $n = 3$  per data point). All western blot data are representative of three independent measurements.

**Nsp14 depletion involves the ubiquitin-proteasome system.** We next investigated whether JC19-induced nsp14 depletion was a result of protein degradation by cellular homeostasis machinery. We pretreated cells with proteasome inhibitors MG132 or bortezomib, which each conferred partial protection against JC19-induced nsp14 degradation (**Fig. 3A**). Analogous treatments using autophagy inhibitors revealed a similar partial rescue afforded by bafilomycin (**Supplementary Fig. 3**), which blocks lysosomal acidification by inhibiting V-type proton ATPase<sup>43</sup>.

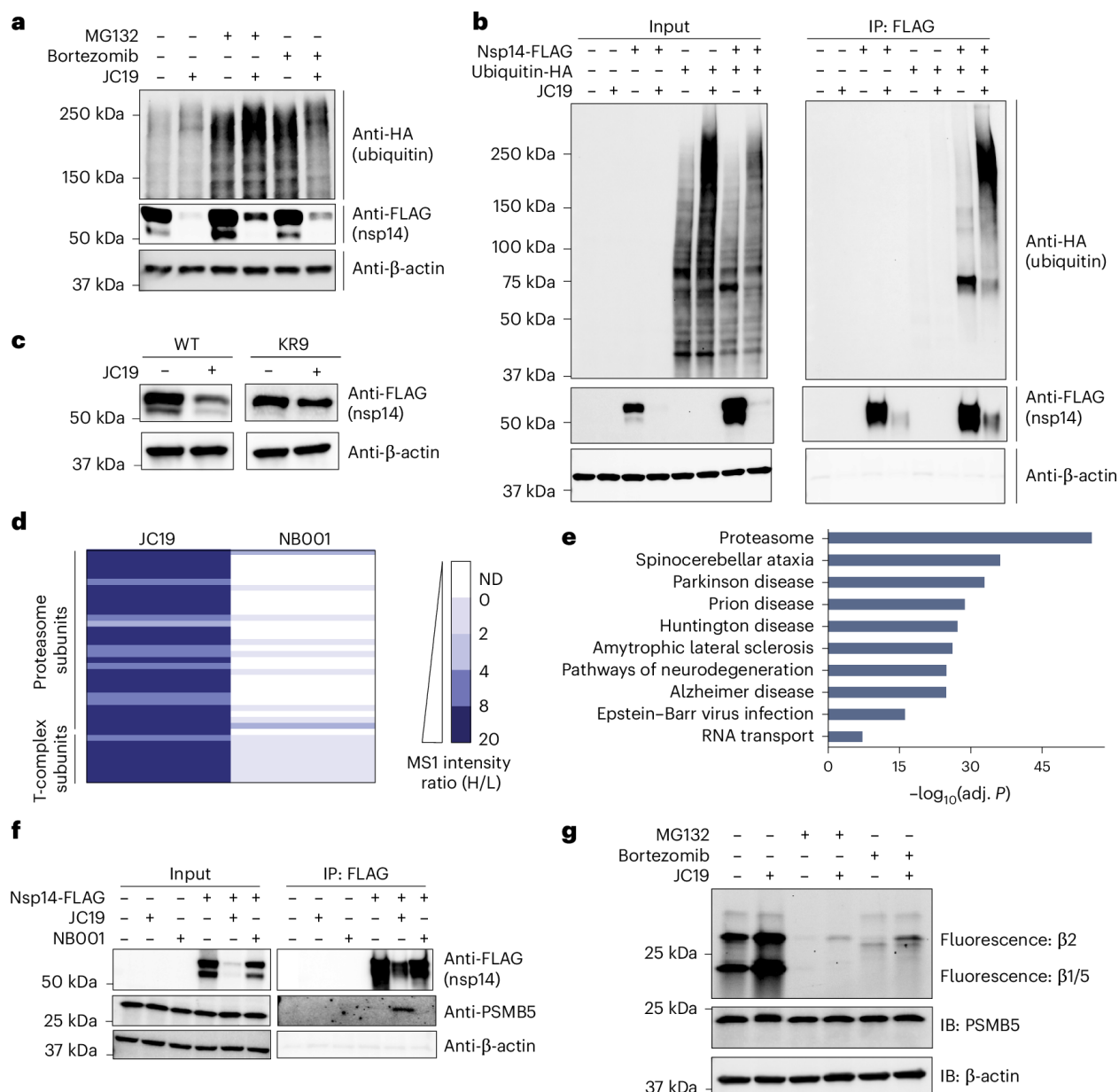
**JC19 induces nsp14 ubiquitylation and proteasome association.** As the mechanism of nsp14 depletion was observed to be, at least partially, proteasome-dependent (**Fig. 3A**), we tested the contributions of proteasome-mediated degradation to nsp14 depletion. JC19 treatment induced a marked increase in nsp14 polyubiquitylation, detectable both with co-expression of HA-tagged ubiquitin (**Fig. 3B**) and with antibodies that recognize endogenous ubiquitin (**Extended Data Fig. 3A**). The signal for heightened polyubiquitylation extended broadly across the host proteome (**Fig. 3B, Extended Data Fig. 3A**), pointing towards the likelihood of additional targets sensitive to JC19-induced degradation. In contrast, the comparatively inactive compound NB001 did not cause similar increased polyubiquitination, providing evidence of a JC19-specific effect (**Supplementary Fig. 4**).

To further delineate the isopeptide linkage sites and assess the contributions of K48 versus K63 ubiquitylation, we co-expressed nsp14 with mutated HA-Ub constructs (K48R, K63R, K48, and K63). We find that nsp14 is partially mono-ubiquitinated in the absence of electrophilic compounds (**Extended Data Fig. 3B**). A decrease in mono-ubiquitinated was observed for the K63 mutant ubiquitin constructs (K63R and K48), implicating a K63 linkage as predominantly responsible for the mono-ubiquitinated nsp14. In contrast, JC19-induced polyubiquitination of nsp14 was predominantly imparted by K48 linkages (**Extended Data Fig. 3C**). This mixture of isopeptide linkages with K63 monoubiquitylation and K48 polyubiquitylation is suggestive of a K48-K63 branched linkage<sup>44,45</sup>.

To further assess the contributions of nsp14 polyubiquitylation to the mechanism of nsp14 depletion, we mutated nine potential lysine ubiquitylation sites to arginine residues (hereafter referred to as nsp14 KR9), including those identified as ubiquitinated (GlyGly modified) by mass spectrometry (**Extended Data Fig. 3D**). Both wild-type (WT) and KR9 mutant proteins show similar subcellular localization (**Supplementary Fig. 5A,B**) and both interact comparably with the known nsp14 binding partner nsp10 (**Supplementary Fig. 5C**). Consistent with a requirement for

ubiquitylation in nsp14 degradation, the nsp14 KR9 was nearly completely insensitive to JC19 treatment (**Fig. 3C, Supplementary Fig. 5D**).

We next opted to use affinity purification-mass spectrometry (AP-MS) to capture proteostasis factors responsible for nsp14 polyubiquitylation (**Extended Data Fig. 3E**). Without compound treatment, we find that nsp14 enriches known interactors, including Inosine-5'-monophosphate dehydrogenase 2 (IMPDH2) and Alpha-galactosidase (GLA)<sup>46</sup>, and JC19 and NB001-treatments induced increased interactions with 133 and 21 proteins, respectively (mean SILAC H/L log<sub>2</sub> ratio >1.5) (**Supplementary Table 2**), despite compound-induced reduction in nsp14 abundance. Subunits of the proteasome and the T-complex protein ring complex (TriC, the cellular chaperonin<sup>47</sup>) together with proteins related to protein aggregation were specifically enriched with JC19 treatment (**Fig. 3D,E**). Providing further evidence of a JC19-dependent nsp14-proteasome interaction, the proteasome 20S subunit beta 5 (PSMB5) was co-immunoprecipitated by nsp14 upon JC19, but not NB001, treatment (**Fig. 3F**).



**Figure 3. JC19-induced nsp14 degradation is mediated by ubiquitylation and the proteasome.** (A) HEK293T cells transiently expressing nsp14-FLAG were pretreated with proteasome inhibitors MG132 (50  $\mu$ M, 14 h) or bortezomib (20  $\mu$ M, 14 h), then treated with DMSO or JC19 (50  $\mu$ M, 0.5 h). (B) HEK293T cells transiently expressing the indicated nsp14-FLAG and/or ubiquitin-HA constructs were treated with DMSO or JC19 (100  $\mu$ M, 1 h) and immunoprecipitated on FLAG resin. (C) HEK293T cells transiently expressing either WT or arginine mutant nsp14 KR9 (K13R, K34R, K288R, K318R, K339R, K423R, K433R, K440R, K469R) were treated with DMSO or JC19 (25  $\mu$ M, 0.5 h). (D) Mean unlogged MS1 intensity H/L ratio of proteasome and T-complex subunits in AP-MS experiments ( $n = 3$  for each compound). (E) KEGG pathway analysis of proteins with mean H/L  $\log_2$  ratio  $> 1.5$  in JC19 AP-MS dataset. KEGG 2021 Human database used. (F) Mock and nsp14-FLAG-expressing HEK293T cells were treated with JC19, NB001, or DMSO (100  $\mu$ M, 1 h) and immunoprecipitated on FLAG resin. (G) HEK293T cells were pretreated with DMSO, 10  $\mu$ M MG132, or 5  $\mu$ M bortezomib for 3 h, then treated with JC19 (100  $\mu$ M, 1 h), and finally with fluorescent proteasome activity probe



Me4BodipyFL-Ahx3Leu3VS (500 nM, 1 h). In-gel fluorescence used to visualize proteasome activity. For MS experiments shown in 'D-E'  $n = 3$  per treatment condition. All MS data can be found in **Supplementary Table 2**. Panels a,c,f,g are representative of two independent measurements, and panel b is representative of four independent measurements.

**Knockdown of E3 ligases does not block nsp14 depletion.** While prior reports have demonstrated direct recruitment of substrate proteins to the proteasome<sup>48,49</sup>, nearly all reported molecular glue and heterobifunctional degrader molecules require engagement of ubiquitin ligase machinery. As our AP-MS identified JC19-induced interactions for five E3 ubiquitin ligases, we next used siRNA knockdowns to test the possible involvement of each E3 in nsp14 degradation. No change in nsp14 abundance nor sensitivity to compound treatment was observed upon knockdown of E3 ubiquitin-protein ligase HUWE1, a known target of electrophilic compounds<sup>1</sup> which has been recently implicated in clearance of protein aggregates<sup>50</sup>, nor E3 ubiquitin-protein ligase UBR4, or E3 ubiquitin-protein ligase TRIM21 (**Supplementary Fig. 6A-C**). For E3 ubiquitin-protein ligase CHIP (STUB1), the possibility of a slight rescue was apparent for one siRNA (**Supplementary Fig. 6D**). While the E3 ubiquitin-protein ligase HECTD1 knockdown was not protective against JC19 treatment, in the absence of compound treatment, HECTD1 depletion afforded a modest decrease in nsp14 abundance, which may result from decreased cellular proliferation upon HECTD1 knockout<sup>51</sup> (**Supplementary Fig. 6E**).

**Cysteine-reactive electrophiles activate the proteasome.** Surprised by the general insensitivity of nsp14 to depletion of individual E3s, we chose to revisit two unusual aspects of this system, namely, the incomplete MG132 rescue (**Fig. 3A**) and the compound-induced increased global polyubiquitylation (**Fig. 3B**). This latter observation parallels that reported previously for cysteine-reactive compounds RA190 (**11**) and eeyarestatin (**12**)<sup>15,52</sup>, which both induced depletion of nsp14 (**Supplementary Fig. 7**). Given these global impacts on ubiquitylation, we next tested whether electrophilic compounds might impact proteasome function directly, using the activity-based probe, Me4BodipyFL-Ahx3Leu3VS<sup>53,54</sup>. For JC19 and RA190 treatments, we

observe a striking increase in proteasome activity (**Fig. 3G, Supplementary Fig. 8A**). Residual proteasome activity was additionally observed with proteasome inhibition, which implicates electrophilic compound treatment-induced formation of a pool of MG132/bortezomib refractory proteasomes that rationalize the only partial proteasome inhibitor rescue from nsp14 degradation (**Fig. 3A**). More broadly, we find that proteasome activation is a ubiquitous feature of all active and inactive nsp14 degraders analyzed (**Supplementary Fig. 8B,C**).

**Cell death pathways do not potentiate nsp14 depletion.** Puzzled by the lack of a clear E3 responsible for nsp14 degradation, we next expanded our mechanistic studies. We find that nsp14 has a half-life of ~8h and is not degraded upon translation arrest (**Supplementary Fig. 9**), ruling out this potential confounder of degrader discovery efforts<sup>55</sup>. Given the comparatively modest DC50, we reasoned that electrophile-induced depletion of intracellular glutathione, well known to induce ferroptosis<sup>56</sup>, could be involved in the depletion mechanism. Glutathione depletion did not impact nsp14 abundance (**Supplementary Fig. 10A**), and bulk cellular glutathione levels were insensitive to compound treatment (**Supplementary Fig. 10B**). Treatment with RSL3, an established phospholipid hydroperoxide glutathione peroxidase GPX4 inhibitor and ferroptosis inducer<sup>56</sup>, failed to impact nsp14 abundance (**Supplementary Fig. 10C**). Poly [ADP-ribose] polymerase 1 (PARP1) cleavage, which is a marker of programmed cell death, was not observed after JC19 treatment (**Supplementary Fig. 11A**) and membrane permeability was maintained during the treatment course for both JC19 and EN450 (**Supplementary Fig. 11B**).

Looking beyond cell death processes, we investigated the involvement of the Kelch-like ECH-associated protein 1 (KEAP1)- nuclear factor erythroid 2-related factor 2 (NRF2) system, which is central to cellular response to electrophiles<sup>57</sup>, and the unfolded protein response (UPR), the mechanism by which cells sense and eliminate misfolded proteins from the ER<sup>58</sup>. No NRF2 activation nor increased expression of UPR markers were observed in response to JC19 (**Extended Data Fig. 4A-F**). Nsp14 abundance was also insensitive to non-cysteine-reactive UPR

activators (**Extended Data Fig. 4B**), further supporting a cysteine-dependent mechanism of depletion.

**Covalent labeling of nsp14 cysteines potentiates depletion.** As a cysteine-reactive electrophile was essential for nsp14 degrader activity (**Extended Data Fig. 2**), we postulated that direct covalent alkylation of nsp14 was involved in the degradation mechanism. Following the workflow shown in **Fig. 4A**, using an alkynylated version of JC19 (JC19yne (**13**); **Extended Data Fig. 1**), we observed enrichment and competition of nsp14 in cell-based and lysate-based labeling experiments (**Fig. 4B**, **Supplementary Fig. 12**). These findings support direct covalent modification of nsp14, which we further confirmed by streptavidin-based affinity pulldown and immunoblot analysis (**Fig. 4C**). Nsp14 was also preferentially enriched by JC19yne over JC19 when treatment was performed in cells and lysates (**Supplementary Fig. 12B,C**), which mitigates the possible confounding effects of rapid nsp14 depletion during dual JC19 and JC19yne treatments.

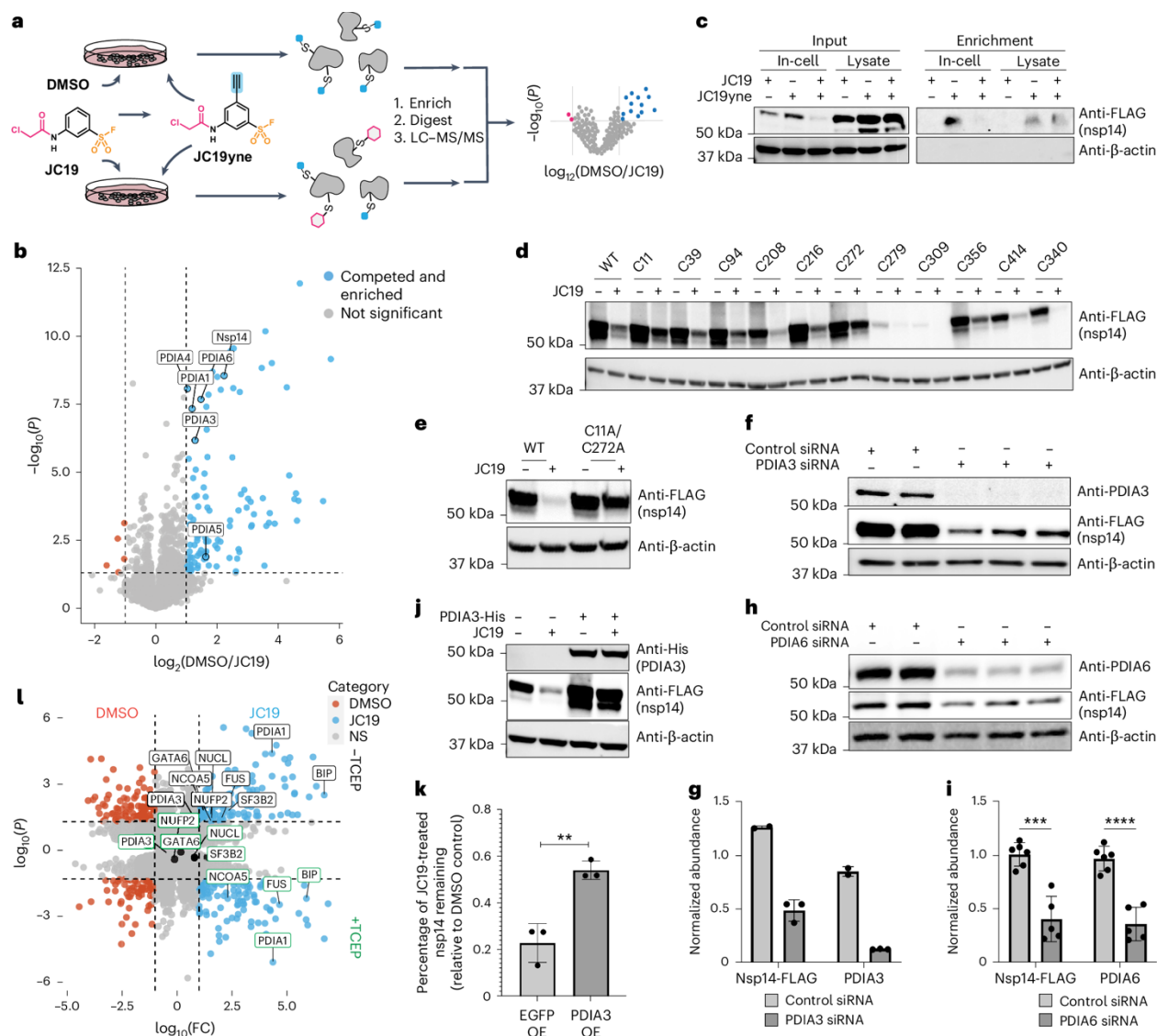
To pinpoint the likely site(s) of covalent alkylation, we evaluated a focused set of nsp14 cysteine mutants (**Supplementary Fig. 1**). We observed that the Cys11Ala and Cys272Ala mutants each afforded partial protection from compound-induced nsp14 depletion (**Fig. 4D**). Providing further evidence of these cysteines as the likely labeling sites, the double mutant (Cys11Ala and Cys272Ala) was largely insensitive to JC19 treatment (**Fig. 4E**).

**Covalent labeling of PDIA3 and PDIA6 potentiates depletion.** Given the comparatively modest potency of JC19 (**Fig. 2E**), we reasoned that a number of host protein cysteines would be subject to covalent modification alongside the nsp14 labeling sites. Therefore, we broadened our analysis of our JC19yne enrichment datasets to consider host proteins labeled by JC19(yne) compounds (**Fig. 4B**, **Supplementary Fig. 12**). Four protein disulfide-isomerases (PDIs; PDIA1, PDIA3, PDIA4, and PDIA6) were markedly enriched by JC19yne, with this enrichment competed by pre-

treatment with JC19 (**Fig. 4B**, **Supplementary Fig. 12A-C**). These data indicated that the PDI family of protein chaperones are bona fide targets of JC19.

We next asked if the JC19 analogues also engaged the PDIs. IsoTOP-ABPP analysis of our compound set identified 14,665 total unique cysteines from 4,839 proteins, including 20 PDI cysteines (**Supplementary Fig. 13**), multiple of which were labeled by JC19. Notably, no widespread PDI labeling was observed for the less active molecules EN450, JC17, and NB179 (**Extended Data Fig. 2**). As the PDI family are oxidoreductases that play essential roles in protein folding<sup>59</sup>, with covalent PDI inhibitors implicated in inducing the UPR<sup>60</sup>, we postulated that JC19 nsp14 depletion activity was potentiated by covalent PDI engagement.

To test this hypothesis, we pursued parallel chemical (**Supplementary Fig. 14**) and genetic approaches (**Fig. 4F-K**). The previously reported irreversible covalent PDI inhibitors AA147 (**14**)<sup>60</sup> and 16F16 (**15**)<sup>61</sup> afforded nsp14 depletion (**Supplementary Fig. 14A-C**). siRNA depletion of PDIA3 or PDIA6 also significantly decreased nsp14 abundance, as detected by immunoblot (**Fig. 4F-I**). Pointing towards mechanistic specificity for PDIA3 and PDIA6, nsp14 was insensitive to PDIA1 and PDIA4 gene expression knockdown (**Supplementary Fig. 15A-D**). Bulk proteomic analysis substantiated these findings revealing significant nsp14 depletion afforded by PDIA6 knockdown; intriguingly, Ras-related protein Rab-21 (RAB21) expression was also significantly increased upon PDIA6 depletion, which was further corroborated by immunoblot (**Supplementary Fig. 15E-G**). RAB21 expression is required for endosome and autophagosome-lysosome fusion<sup>62,63</sup>, and, consistent with PDIA3/6 depletion contributing to increased autophagosomal flux, knockdown induced an increase in the autophagy marker microtubule-associated proteins 1A/1B light chain 3A (LC3I/II) (**Supplementary Fig. 15G**). Further implicating PDI activity as central to the nsp14 depletion mechanism, we find that PDIA3 overexpression attenuated nsp14 sensitivity to JC19 (**Fig. 4J,K**). Collectively, our data support a mechanism by which rapid nsp14 depletion by JC19 is achieved by destabilization of nsp14 via both direct labeling of nsp14 cysteines and through PDI inhibition.



**Figure 4. Direct labeling of nsp14 and the protein disulfide isomerases (PDIs) potentiates nsp14 degradation.** (A) Workflow of JC19 competition experiment. HEK293T cells transiently expressing nsp14-FLAG were treated with either DMSO or 100  $\mu$ M JC19 for 20 minutes, and all replicates then treated with 100  $\mu$ M JC19yne for 20 minutes. (B) Differentially enriched proteins from the experiment outlined in (A), DMSO ( $n = 6$ ) and JC19 ( $n = 6$ ). (C) Cells and lysates were treated as described in (A) and immunoblot used to detect nsp14 enrichment. (D) HEK293T cells transiently expressing WT nsp14-FLAG or the indicated cysteine mutants (all cysteine to alanine mutations, except C208V) were treated with either DMSO or JC19 (50  $\mu$ M, 1 h). (E) HEK293T cells transiently expressing WT nsp14-FLAG or the C11A/C272A double mutant were treated with DMSO or JC19 (50  $\mu$ M, 1 h). (F) Immunoblot of nsp14-FLAG abundance upon siRNA knockdown of PDIA3. (G) Quantification of immunoblot in (F). Data are presented as mean  $\pm$  s.d. (control  $n = 2$ , PDIA3 targeting siRNA  $n = 3$ ). (H) Immunoblot of nsp14-FLAG abundance upon siRNA knockdown of PDIA6. (I) Quantification of nsp14 ( $p = 0.000173$ ) and PDIA6 ( $p = 0.000039$ ) abundance following siRNA knockdown of PDIA6. Data are presented as mean  $\pm$  s.d. (control  $n = 5$ , PDIA6 targeting siRNA  $n = 6$ ). (J) HEK293T cells transiently coexpressing nsp14-FLAG and EGFP or nsp14-FLAG and PDIA3-His were treated with either DMSO or JC19 (25  $\mu$ M, 0.5 h). (K) Quantification of immunoblot in (J) ( $p = 0.0041$ ). Data are presented as mean  $\pm$  s.d. ( $n = 3$ ). (L)

HEK293T cells transiently expressing PDIA6-His8x were affinity purified under denaturing conditions either in the absence (black border labels) or presence of TCEP reductant (green border labels). Volcano plot displays comparison between enriched proteins for groups treated with DMSO versus JC19 (50  $\mu$ M, 1 h) with black colored dots for proteins showing sensitivity to TCEP ( $n = 3$  for each group). P-values calculated with unpaired Student's t-test where applicable. All MS data can be found in **Supplementary Table 3**. All immunoblots are representative of two independent measurements.

**PDI crosslinking contributes to nsp14 depletion.** While PDI inhibition was sufficient to decrease nsp14 abundance, we postulated that JC19 crosslinking activity potentiated degradation, rationalizing the rapid degradation kinetics (**Fig. 2A**). Indeed, both JC19 and the structurally related bis-chloroacetamide derivative NB82 (**16**) induced formation of high molecular weight forms of PDIA6 (**Supplementary Fig. 16**), with the later also a potent nsp14 degrader. No such gel shifts were observed for the non-degrader and non-crosslinking analogue NB92.

PDIs interact with substrates reversibly via mixed disulfides. Thus, we surmised that JC19 may drive protein degradation by irreversibly trapping PDI substrates, which could lead to widespread protein misfolding. Therefore, using AP-MS under denaturing conditions, we next delineated the disulfide-mediated and compound-induced PDIA6 interactome (**Extended Data Fig. 5A**), comparing treatment with either JC19, the non-crosslinkable analogue NB92, or TCEP to delineate mixed disulfide interactions from JC19 crosslinked proteins. Notable crosslinked proteins include PDIA1, Endoplasmic reticulum chaperone BiP (BiP), RNA-binding protein FUS (FUS), and Nuclear receptor coactivator 5 (NCOA5) (**Fig. 4L, Extended Data Fig. 5B**), known protein regulators of ER folding, autophagy and the UPR, which suggest that PDI crosslinking may cause substantial protein folding stress within the ER, although the specific mechanism remains somewhat opaque, as illustrated by the absence of substantial JC19-induced activation of known UPR markers (**Extended Data Fig. 4**).

**IsoTOP-ABPP identifies widespread changes to the proteome.** Thus far, our data are consistent with direct alkylation of both nsp14 and PDI crosslinking as central to the mechanism of JC19-induced proteasome-mediated nsp14 depletion. However, the modest potency of JC19, compound-induced global polyubiquitylation, and proteasome activation, together hinted at the likelihood of polypharmacology-driven modulation of general proteostasis processes. We hypothesized that these activities are a generalized consequence of cell-based treatments with scout electrophilic compounds, at doses exceeding those recommended for potent, selective, and optimized chemical probes, and that these activities were likely contributing to degrader activity. To test these hypotheses, we broadened our analysis to consider electrophile-sensitive host proteins. We harnessed our rich isoTOP-ABPP datasets (**Supplementary Table 1,4**) to pinpoint compound-induced changes to protein abundance, inferred from identification of proteins that harbor multiple high-ratio cysteines, which we termed multi-high-ratio proteins (MHRPs). For the JC19, EN450, KB2, and JC36 (**17**) datasets, 15-25% of all proteins were categorized as MHRPs (**Fig. 5A**). Given the outsized impact of this subset of compounds on the proteome (**Extended Data Fig. 6**), we next investigated whether specific fractions of the proteome, as annotated by UniProtKB cellular components, were disproportionately sensitive to compound treatment. While MHRPs were present in all cellular components (**Extended Data Fig. 6**), we observed that a marked increase in the frequency of nuclear pore complex (NPC) protein subunits within the MHRP subset (18/33 nucleoporins; **Fig. 5B**). This observation was striking, given the important functions of the NPC in shuttling protein and RNA cargo in and out of the nucleus.

**Direct labeling and indirect mechanisms induce degradation.** To further confirm protein abundance changes, we subjected JC19- and EN450-treated cells to bulk proteomic analysis (**Supplementary Fig. 17A**). JC19 and EN450 were selected given their shared impact on the NPC together with JC19's unique PDI labeling activity. In aggregate, JC19 and EN450 treatment decreased abundance of 245 total proteins, including nsp14 (**Fig. 5C, Supplementary Fig. 17B-**

**C).** The compound-depleted protein subset was enriched for nucleoporins, including E3 SUMO-protein ligase RanBP2 (RANBP2) and Nuclear pore complex protein Nup160 (NUP160) (**Fig. 5D, Supplementary Fig. 18 and 19**). The electrophile-sensitive nucleoporins were observed to be located throughout the NPC (**Supplementary Fig. 20**). Immunoblot analysis substantiated these findings and showed that the effect extended to additional cysteine-reactive electrophiles (**Fig. 5E, Supplementary Fig. 21**). Our data implicate nucleoporin depletion as a general consequence of multiple promiscuous cysteine-reactive compounds.

We hypothesized that for many of the host proteins, similar to nsp14, cysteine modifications should drive protein depletion. We compared the protein targets statistically enriched by clickable EN450yne (**18**) or JC19yne (**Fig. 5F**) with the MHRPs and the proteins statistically depleted by the corresponding parent compounds (**Fig. 5A,C**). While these groups did have substantial overlap indicative of covalent modification-induced protein depletion, many depleted proteins were not enriched by either alkyne analogue (**Fig. 5G**). Exemplifying this process, while nsp14 was robustly enriched by JC19yne (**Fig. 4B, Supplementary Fig. 12B**) and depleted by JC19 (**Supplementary Fig. 17B**), EN450yne failed to capture nsp14 (**Supplementary Fig. 22A**), despite EN450 also affording marked protein depletion (**Supplementary Fig. 17C**). Similarly, protein SEC13 homolog (SEC13), while not depleted by JC19, was the only nucleoporin significantly enriched by JC19yne (**Supplementary Fig. 22B**), despite the dramatic impact of JC19 on nucleoporin abundance. None of the five nucleoporins enriched by EN450yne were significantly depleted in the bulk dataset (**Supplementary Fig. 22A, Supplementary Fig. 18**), which provides further evidence that direct modification is not required for depletion of most nucleoporins. Taken together, these data hint at the likelihood of multiple mechanisms of depletion occurring in response to electrophilic compound treatment, with direct alkylation of proteins synergizing with indirect activity.



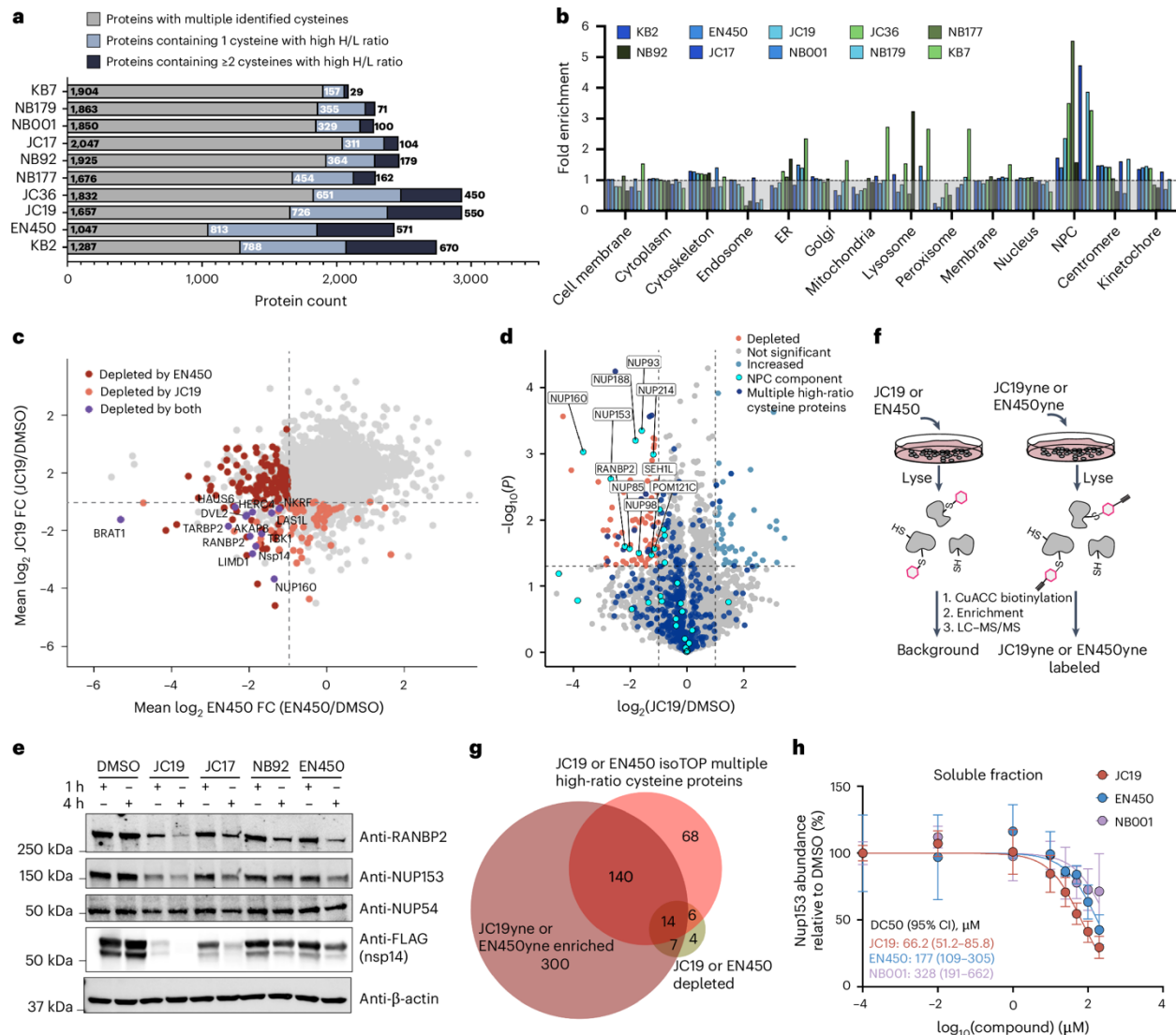
**Electrophiles induce nucleoporin aggregation.** Given the involvement of PDIs and the ubiquitin-proteasome system (UPS) in JC19-mediated nsp14 depletion (**Fig. 3A,B,D,E and Fig. 4F-K**), we next asked whether these systems also contributed to host protein depletion, using the nucleoporins as a case-study. We find that nucleoporins are generally insensitive to PDI depletion (**Supplementary Fig. 23**). Knockdowns of key stress-modulating E3 ubiquitin-protein ligases HUWE1, UBR4, STUB1, and TRIM21 did not impart a substantial protective effect against nucleoporin depletion (**Supplementary Fig. 24A-D**). HECTD1 knockdown revealed protection against JC19-mediated depletion of Nuclear pore complex protein Nup153 (NUP153), although this effect was not rescued by HECTD1 overexpression (**Supplementary Fig. 24E,F**). For RANBP2 or NUP153, proteasome inhibition also failed to prevent compound-induced depletion (**Supplementary Fig. 25A**).

Guided by the proteasome-independent mechanism of depletion and prior reports of nucleoporin aggregation<sup>64</sup>, we hypothesized that electrophilic compounds were inducing protein aggregation. Indeed, we observed marked compound-induced accumulation of NPC subunits in the detergent insoluble cellular fraction by immunoblot and bulk proteomic analysis (**Figure 25A,B**), consistent with nucleoporin aggregation. Looking more broadly, of the proteins identified in both soluble and insoluble datasets for JC19 and EN450, a handful of proteins were shown to be significantly depleted in both soluble and insoluble (bone fide degraded), while a greater number of proteins were significantly depleted from the soluble and significantly accumulated in the insoluble (aggregated) (**Supplementary Fig. 25C-F**).

We next assessed the dose dependence of the aggregator activity and whether this effect extended to nsp14. We find that at lower doses (< 50  $\mu$ M) JC19 primarily induces nsp14 aggregation, while it imparts degradation at higher concentrations (**Fig. 2D-E, Extended Data Fig. 7A-D**). In contrast, EN450 and NB001 primarily caused nsp14 aggregation (**Fig. 2D-E, Extended Data Fig. 7B-C,E**). Among the three compounds tested, JC19 treatment was also observed to most substantially impact NUP153 abundance both in the soluble and insoluble

fractions (**Fig. 5H, Extended Data Fig. 7F-G**). While aggregation was primarily observed in the insoluble fraction (**Extended Data Fig. 7F-G**), a suggestive modest decrease in NUP153 abundance was observed at higher concentrations, hinting at the possibility of slight aggregation-induced degradation (**Fig. 5H, Extended Data Fig. 7F**).

Guided by these findings, we also opted to revisit whether one of our previously identified E3 ligases might be responsible for degrading nsp14 from the insoluble fraction. Suggestive of such activity, knockdown of either STUB1 or HECTD1 afforded modest and specific protection from JC19-mediated nsp14 depletion from the insoluble fraction (**Supplementary Fig. 26**), potentially implicating these E3 ligases in the clearance of nsp14 aggregates.



**Figure 5. Multimodal proteomics implicates global rewiring of the proteome by electrophilic compounds.** (A) The number of proteins in indicated isoTOP-ABPP experiments with a single high-ratio cysteine (light blue) and multiple high-ratio cysteines (navy blue). (B) Cellular compartment (UniprotKB annotated) enrichment analysis of proteins with multiple ( $\geq 2$ ) cysteines with a  $\log_2$  H/L ratio  $> 1$  in the indicated isoTOP-ABPP experiments, relative to all proteins with multiple identified cysteines. Fold-change enrichment calculated by dividing the frequency in multiple high-ratio cysteine proteins by frequency in all identified multi-cysteine proteins. (C) Analysis of proteins identified in both JC19 (100  $\mu$ M, 1 h,  $n = 3$ ) and EN450 (100  $\mu$ M, 4 h,  $n = 3$ ) bulk proteome experiments. Proteins with a  $\log_2$ (fold-change)  $< -1$  and p-value  $< 0.05$  are considered significantly depleted. (D) LFQ analysis of JC19-induced changes to protein abundance (comparing DMSO versus JC19 (100  $\mu$ M, 1 h),  $n = 3$  per group). (E) HEK293T cells transiently expressing nsp14-FLAG were treated with DMSO, JC19, JC17, NB92, or EN450 (100  $\mu$ M, 1 or 4 h). (F) Workflow for JC19yne and EN450yne enrichment experiments. (G) Overlap of proteins identified as statistically depleted by JC19 or EN450 in bulk proteome datasets ( $n = 3$  for each compound), proteins identified as statistically enriched by JC19yne ( $n = 6$ ) or EN450yne ( $n = 4$ ), and proteins identified as having multiple ( $\geq 2$ ) cysteines with elevated ratios for JC19 or EN450 by isoTOP-ABPP ( $n \geq 3$  for each compound). Proteins included in the Venn diagram include only those identified in all representative datasets. (H) Quantification of nup153 intensity in the detergent soluble fraction from replicate western blots as in **Extended Data Fig. 7F**. Data are presented as mean  $\pm$  s.d. ( $n = 3$  per data point). An unpaired Student's t-test was performed to generate p-values where applicable. All MS data can be found in **Supplementary Table 4**. Western blot data is representative of two independent measurements.

**Cysteine-reactive compounds induce stress granule formation.** Many features of our system were suggestive of cellular stress and dysregulated proteostasis, including (1) the accumulation of proteins into the detergent insoluble fraction, consistent with protein aggregation; (2) enrichment of aggregation-related processes in our AP-MS dataset (**Fig. 3E**); (3) compound-induced increases in global polyubiquitylation (**Fig. 3B**); (4) increased proteasome activity (**Supplementary Fig. 8**); (5) involvement of protein folding machinery (**Fig. 4F-K**); and (6) rapid kinetics (**Fig. 2A**). Therefore, hypothesized that cell stress mechanisms were involved in the activity of JC19 and potentially also previously reported modestly potent covalent degraders.

Guided by the presence of NPC subunits<sup>18</sup> in stress granules (SGs) and the shared cysteine reactivity of electrophilic small molecules and stress-granule-inducing oxidative stressors (e.g. H<sub>2</sub>O<sub>2</sub> and sodium arsenite), we next asked whether JC19 or EN450 could induce stress granules—we hypothesized that stress-induced protein aggregation was contributing to the depletion of nsp14, nsp16, and the compound-sensitive NPC subunits. Using Ras GTPase-activating protein-binding protein 1 (G3BP1), a core component and well-known nucleator of

stress granules<sup>65</sup>, reporter cells, we find that characteristic G3BP1 SG puncta were formed in response to JC19, EN450, and NB001 alongside sodium arsenite, as positive control (**Fig. 6A**). Formation of SGs was observed at modest compound concentrations (25-50  $\mu$ M) (**Supplementary Fig. 27**) that have been used for molecular glue degrader analyses<sup>34</sup>. To test how ubiquitous SG-formation is in response to electrophilic compounds, we next broadened the scope of compounds evaluated in our G3BP1 reporter cell line to include additional scout fragments KB2, KB5 (**19**), KB7 (**20**), and covalent-reversible compound KB23 (**21**) that features a cyanoacrylamide<sup>2</sup> (**Extended Data Fig. 1**). All scout fragments robustly induced formation of G3BP1 puncta (**Fig. 6A**). We also tested additional bioactive compounds, including RA190, bardoxolone (**22**), a widely utilized NRF2 activator<sup>66</sup>, sulforaphane (**23**), a covalent modulator of the ubiquitin-proteasome and autophagic pathways with anticancer activity<sup>67</sup> and FDA-approved anti-cancer drugs ibrutinib (**24**) and afatinib (**25**), which are covalent kinase inhibitors, auranofin (**26**), which targets Thioredoxin reductase 1/2 (TXNRD1/2)<sup>68</sup> and was recently reported to increase the activity of the DNA damage checkpoint Serine/threonine-protein kinase Chk1 (CHK1)<sup>69</sup>. We subjected cells to compound dosing consistent with prior reports—for afatinib and ibrutinib we chose to treat with 10  $\mu$ M compound, which, while a dosage used widely<sup>70-72</sup> is in substantial excess of that required to achieve full target engagement<sup>73</sup>. With the exception of ibrutinib, all cysteine-reactive compounds induced G3BP1 puncta consistent with SG formation (**Fig. 6A, Extended Data Fig. 8A**). Notably, 10  $\mu$ M afatinib was the minimum concentration required to induce SG formation (**Extended Data Fig. 8B**), which suggests that more potent and less cysteine-reactive compounds (e.g. ibrutinib<sup>73</sup>) and clinically relevant submicromolar compound concentrations (e.g. afatinib low dose) are not expected to form SGs.

**RANBP2 and PSMB2 are recruited to stress granules.** In addition to G3BP1 accumulation into puncta, stress granules are defined by sequestration of RNA and RNA-binding proteins (RBPs). Therefore, we next asked whether compound treatment would induce protein-RNA interactions<sup>74</sup>.

XRNAX<sup>75</sup> analysis, which reports RBPs bounds to RNA, revealed marked increases in RNA association in response to JC19 treatment for many SG proteins, including FUS, Far upstream element-binding protein 1 (FUBP1), and Eukaryotic translation initiation factor 4B (EIF4B) (**Fig. 6B**). Additionally, nearly all (13 of 16) nucleoporins identified showed increased association with RNA upon treatment with JC19 (**Supplementary Table 5**). In total, 5 nucleoporins that were depleted in our JC19 bulk dataset (NUP93, NUP214, NUP153, NUP98, RANBP2) displayed increased RNA association upon JC19 treatment, suggestive of sequestration into stress granules (**Fig. 6B**). Immunofluorescence analysis corroborated the compound-induced accumulation of RANBP2 into the G3BP1 puncta upon both JC19 and EN450 treatment (**Extended Data Fig. 8C**). Additional proteins that were depleted by JC19 also showed increased association with RNA by XRNAX, including A-kinase anchor protein 8 (AKAP8), A-kinase anchor protein 8-like (AKAP8L), and cellular tumor antigen p53 (TP53) (**Fig. 6B**). We expect that the increased RNA-AKAP8L interaction likely rationalizes the observed abundance changes in response to both EN450 and JC19. The accumulation of TP53 is particularly noteworthy given its tumor suppressor activity and the prior report of naturally occurring isothiocyanate induced selective depletion of TP53 in a manner insensitive to proteasome inhibition<sup>76</sup>.

As prior reports have shown sequestration of proteasome subunits to SGs to modulate SG dynamics<sup>77,78</sup>, we also asked if our compounds induced sequestration of the proteasome into SGs. We observe robust association of the proteasome subunit beta type-2 (PSMB2) with G3BP1 puncta in a compound-dependent manner (**Extended Data Fig. 8D**). We expect that this local accumulation of proteasome is likely a major contributor to electrophilic compound-induced protein degradation, particularly for less potent degraders.

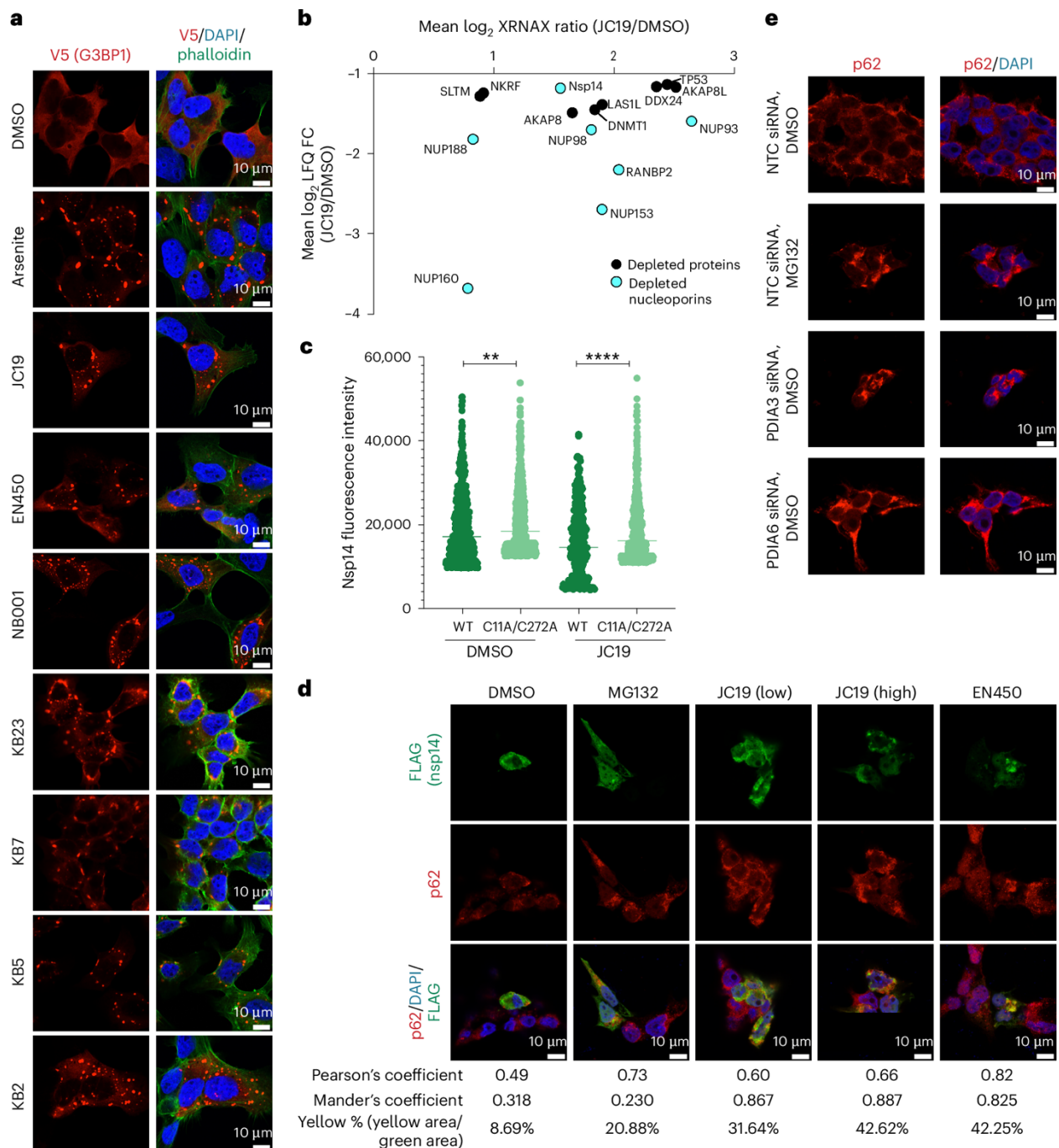
**Nsp14 and nsp16 accumulate into the aggresome.** Given the widespread nature of the observed protein accumulation into stress granules, we next asked whether compound treatment similarly led to accumulation of either nsp14 or nsp16 into SGs. Gratifyingly, and consistent with

our immunoblot analysis (**Fig. 4E**), we observed robust compound-induced and cysteine-dependent depletion of nsp14 by immunofluorescence (**Fig. 6C**). However, no substantial colocalization was observed between nsp14 nor nsp16 and G3BP1 (**Supplementary Fig. 28**). We hypothesized that nsp14 and nsp16 were accumulating in the aggresome, which is a perinuclear-localized body of misfolded and aggregated proteins that forms when cellular degradation machinery is overwhelmed<sup>79</sup>. Gene Ontology (GO) analysis of our bulk insoluble proteomic dataset (**Supplementary Table 5**) revealed that the aggresome was the only enriched cellular component among the aggregated protein subset (**Supplementary Fig. 29**). Both nsp14 and nsp16 show compound-induced increased colocalization with p62, a well-known component and modulator of the aggresome<sup>80</sup>, upon JC19 and EN450 treatment (**Fig. 6D and Supplementary Fig. 30**). These data support the hypothesis that cell-based treatment with cysteine-reactive electrophilic compounds can induce aggresome formation and further implicate compound-dependent protein destabilization in the mechanism of nsp14 depletion.

Given the observed association PDI labeling and nsp14 depletion, together with the well-established chaperone activity of the PDIs, we further hypothesized that decreased PDI chaperone activity would potentiate protein aggregation. Genetic depletion of either PDIA3 or PDIA6 induced marked accumulation of p62 into aggresome-like bodies (**Fig. 6E, Extended Data Fig. 9**) with similar appearance to those afforded by compound treatment (**Fig. 6D**). This causal relationship between PDI inhibition and aggresome formation has not been described previously, to our knowledge, and it provides key mechanistic rationale for how PDI crosslinking leads to nsp14 aggregation. Therefore, we put forth the model whereby nsp14 labeling by JC19 causes initial protein destabilization, which is further potentiated by PDI-inactivation-induced proteome-wide protein aggregation and aggresome formation. The nsp14 and host aggregates are subsequently degraded via the UPS and autophagy. Thus, by revealing that protein aggregation is a ubiquitous feature of cysteine-reactive small molecules, which is intimately intertwined with



protein degradation, our work highlights both the complexities and opportunities for ongoing and future efforts to achieve precise pharmacological control over protein abundance.



**Figure 6. Cysteine-reactive electrophilic fragments induce global proteome remodeling through protein aggregation and stress granule formation.** (A) V5-tagged G3BP1 recruitment to stress granules (SGs) was assayed in response to DMSO or sodium arsenite (400  $\mu$ M, 1 h), with all other compounds (100  $\mu$ M, 1 h). (B) Analysis of the 15 proteins shared between the JC19 XRNAX dataset and the proteins significantly depleted by JC19 in the bulk proteome dataset ( $\log_2(\text{fold-change}) < -1$ ,  $p\text{-value} < 0.05$  calculated from unpaired Student's t-test). Proteins are

plotted based on the mean XRNAX  $\log_2(\text{H/L})$  SILAC ratio ( $n = 3$ ) and mean bulk proteome  $\log_2(\text{fold-change})$  ( $n = 3$ ). (C) HEK293T cells transiently expressing WT or C11A/C272A nsp14-FLAG were treated with DMSO or JC19 (100  $\mu\text{M}$ , 1 h). Nsp14-FLAG fluorescence intensity for each cell was quantified by ImageJ (WT, DMSO:  $n = 637$ ; C11A/C272A, DMSO:  $n = 587$ ; WT, JC19:  $n = 317$ ; C11A/C272A, JC19:  $n = 491$ ). All data points are represented with median value shown. Statistical significance calculated with unpaired Student's t-tests \*\*  $p=0.0034$ , \*\*\*\*  $p=0.00000006$ . (D) HEK293T cells transiently expressing nsp14-FLAG were treated with DMSO, MG132 (10  $\mu\text{M}$ , 5 h), a low dose of JC19 (25  $\mu\text{M}$ , 0.5 h), a high dose of JC19 (100  $\mu\text{M}$ , 1 h), or EN450 (100  $\mu\text{M}$ , 1 h). (E) HEK293T cells transfected with either a non-targeting control (NTC) siRNA, NTC siRNA and treated with MG132 (10  $\mu\text{M}$ , 8 h), PDIA3 siRNA, or PDIA6 siRNA. p62 used as a marker for aggresomes/aggregated proteins. Additional images shown in **Extended Data Fig. 9**. All scale bars = 10  $\mu\text{m}$ . All MS data can be found in **Supplementary Table 4 and Supplementary Table 5**. All microscopy is representative of three independent measurements.

## Discussion

Our study implicates the confluence of covalent protein alkylation and activation of cellular stress response mechanisms in causing widespread proteome remodeling in response to electrophilic compounds. We observe formation of stress granules in response to nearly every cysteine-reactive compound evaluated, including the drugs afatinib and auranofin. Of note Afatinib-induced stress granules were only observed to occur at 10  $\mu\text{M}$ , which, while a dosage that is reported in a number of prior studies<sup>70,72,81</sup>, is >50-fold above<sup>82</sup> the dose required to effect full covalent modification of epidermal growth factor receptor (EGFR)<sup>73</sup>, supporting that SG formation is likely not a clinically relevant feature of afatinib and highlighting the importance of not dosing at concentrations far in excess of those required for on-target activity.

Cell-based chemoproteomic screening using low-mid-micromolar concentrations (e.g. 10-50  $\mu\text{M}$ ) of covalent compounds, such as the molecules reported by our study, is an increasingly common strategy for target hunting<sup>2,5,7,28,30,69,83–89</sup>. Given the ubiquitous nature of the cell stress response mechanisms observed here, we expect that many of the molecules used in prior studies at the indicated concentrations could also have stress-causing activity that may, in some cases, complicate target identification and subsequent delineation of on- versus off-target activity. “Scout” molecules with limited recognition features typically engage most targets with only modest potency (i.e. ~10-20  $\mu\text{M}$  apparent  $\text{IC}_{50}$ ). Thus these compounds fail to fulfill the important and



more stringent criteria of bona fide chemical probes, such as those put forth by the Chemical Probes portal and related resources<sup>9–11,90</sup>.

Despite this limitation, we do still believe that, with suitably stringent controls, covalent fragments can serve as useful pathfinder compounds<sup>9</sup> for target discovery and remain meritorious of their new-found position in the target-discovery toolbox. We look forward to continuing to build upon established best practices to ensure on-target mechanisms, including screening “smarter” compound libraries that feature more tempered cysteine reactive electrophiles and enantiomeric compound pairs, selection of targets that show clear SAR across a compound library analyzed, and careful use of covalent compound-resistant cysteine mutants, that together help delineate on-target activity from more general electrophilic stress<sup>2,5,12,88,91–94</sup>.

Together with these valuable practices, we additionally suggest strategies to help delineate on-and off-target activities of covalent fragments (**Extended Data Fig. 10**). We would additionally encourage caution when no clear SAR is apparent, when relative degradation correlates with relative compound reactivity towards thiols, with more reactive compounds resulting in more degradation, and when high compound concentrations are required to effect degradation. Looking to the future, we are excited to further delineate additional protein biomarkers and mechanisms of cell stress activation. We are particularly optimistic that our work will help to bolster the chemoproteomic discovery of actionable target-compound pairs that can serve as starting points and additions to the existing toolbox of high quality covalent chemical probes and drugs.

## Acknowledgements

This study was supported by a Beckman Young Investigator Award (K.M.B.), DOD-Advanced Research Projects Agency (DARPA) D19AP00041 (K.M.B.), National Institutes of Health DP2 GM146246-02 (K.M.B.), Packard Fellowship (K.M.B.), and NIGMS UCLA Chemistry Biology Interface T32GM136614 (A.R.J). We thank Dr. Irene Zohn for providing the HECTD1 plasmid.

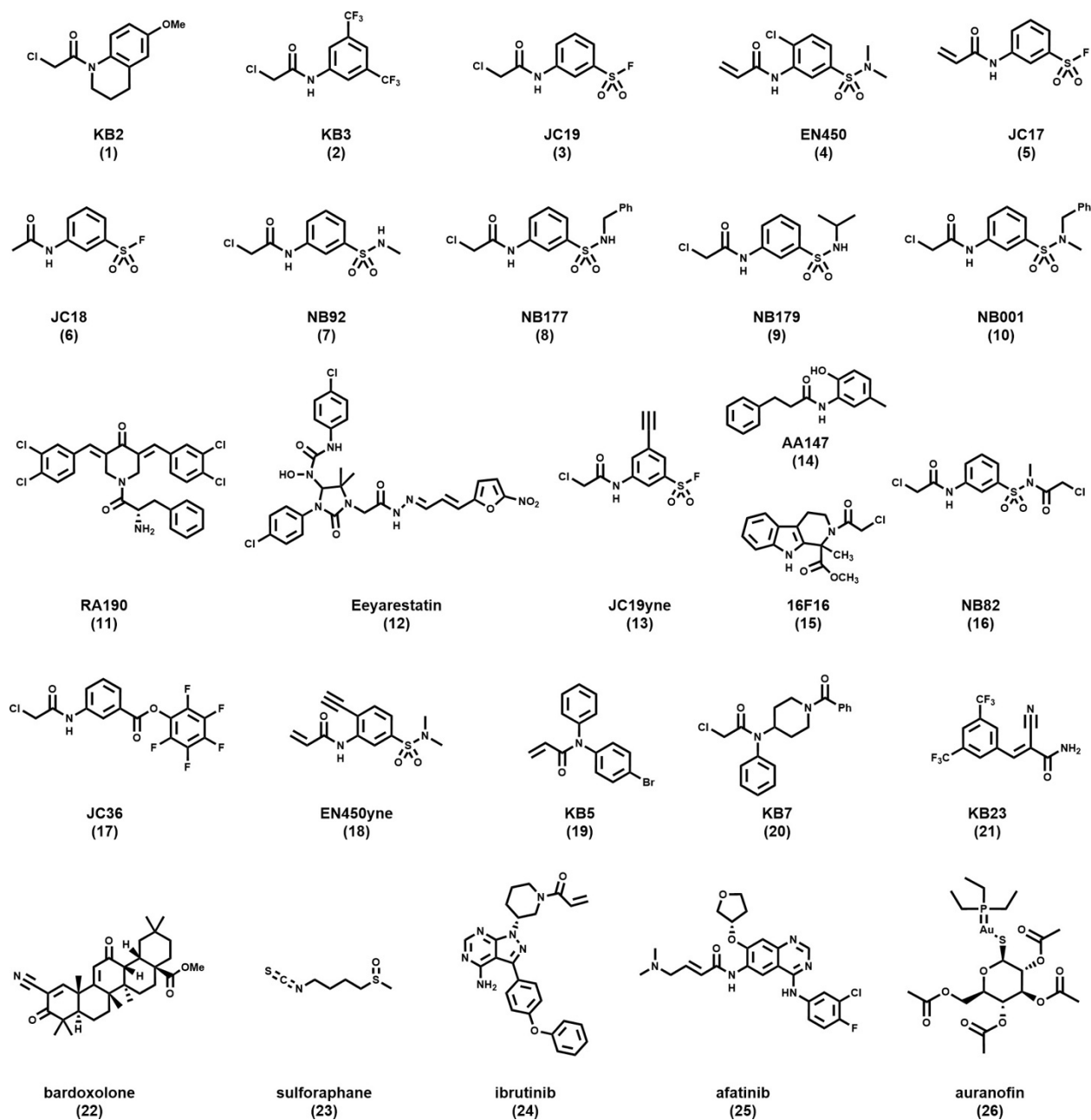
We additionally thank all members of the Backus labs for helpful suggestions and the UCLA Broad Stem Cell Resource Center Microscopy Core for assistance with microscopy. Figure 1A was made using Biorender.com.

### **Author Contributions**

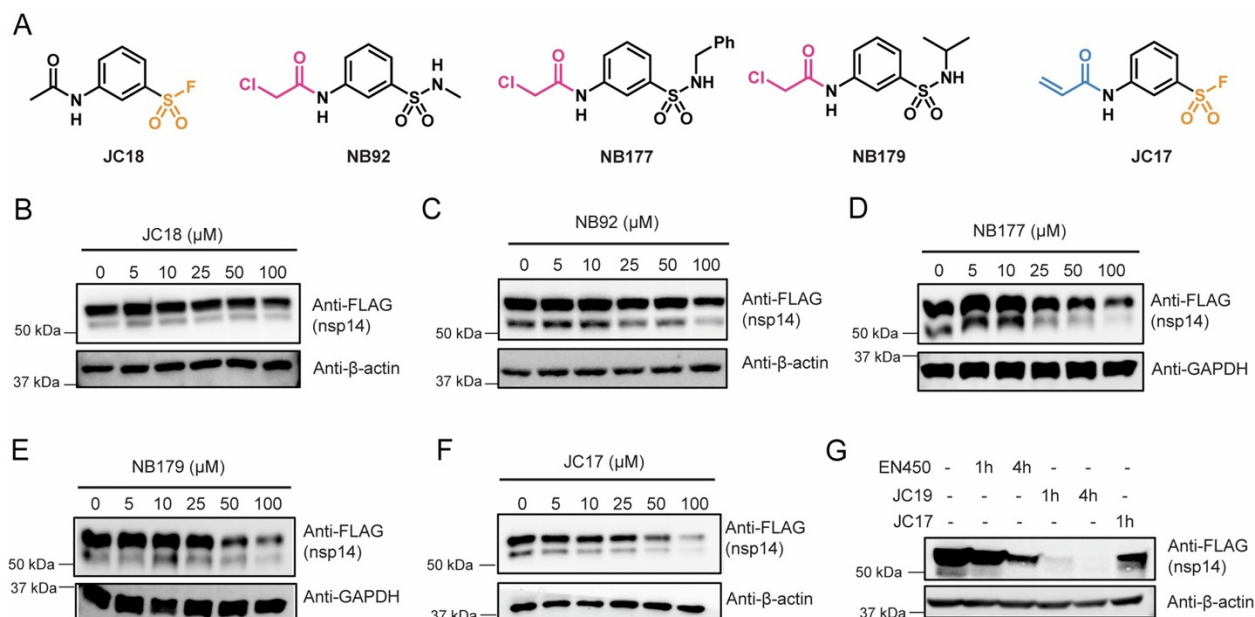
A.R.J., F.S, and K.M.B. conceived of the project. A.R.J., F.S., and K.M.B designed experiments. A.R.J. and F.S. performed, collected, and analyzed data for all biochemical experiments, including bulk proteomics and protein-directed ABPP, western blot analyses, AP-MS proteomics, and cloning of plasmids. F.S. performed, collected, and analyzed data for all isoTOP-ABPP experiments with the help of C.T. and performed all knockdown experiments. A.R.J. performed, collected, and analyzed all data for all imaging experiments. A.C.T performed, collected, and analyzed data for IncuCyte Live-Cell Imaging. N.R.B, E.D, and J.C. performed syntheses of JC19 and all analogues. C.T. wrote software. A.R.J. and F.S contributed to the figures. A.R.J., F.S, and K.M.B. wrote the manuscript with assistance from all authors.

### **Competing interests**

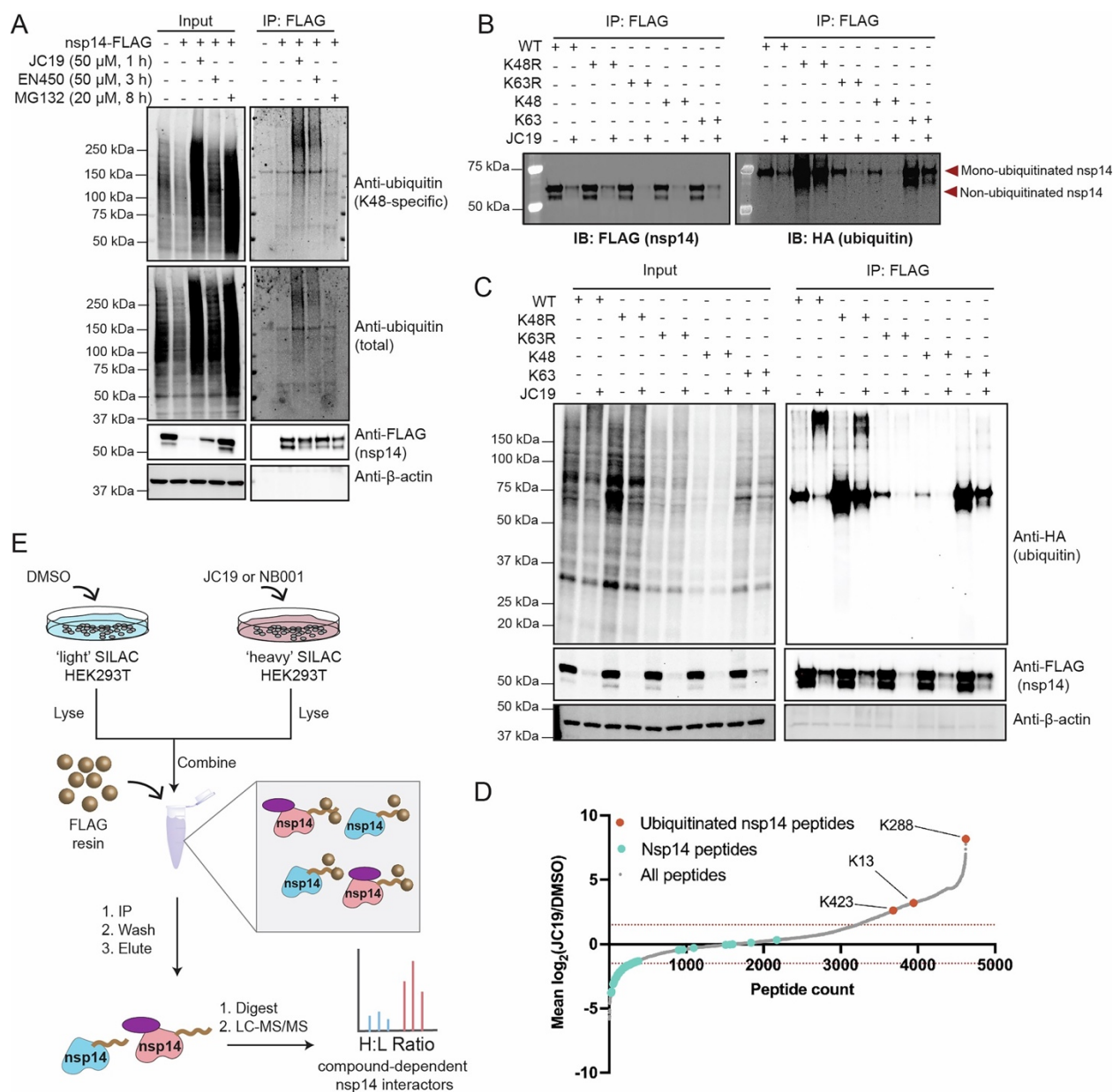
K.M.B. is a member of the advisory board at Matchpoint Therapeutics. The remaining authors declare no competing interests.



**Extended Data Fig. 1 Compound structures.** Shown are the structures of all compounds used for the study, ordered in the order of appearance in the text.

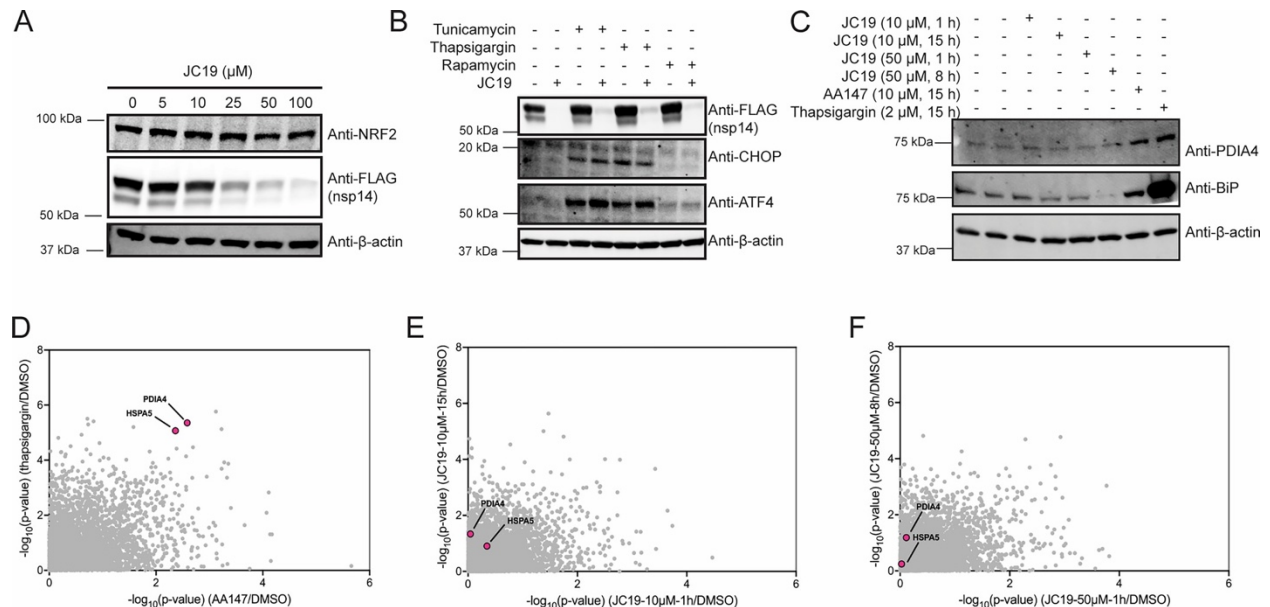


**Extended Data Fig. 2 JC19 analogs have varying degrees of nsp14 depletion efficiency. (A)** Chemical structures of JC19 analogs. Cysteine-reactive chloroacetamide warhead is colored pink, cysteine-reactive acrylamide warhead is colored blue and sulfonyl fluoride warhead is colored orange. **(B–F)** HEK293T cells transiently expressing nsp14-FLAG were treated with the indicated concentrations of JC18 **(B)**, NB92 **(C)**, NB177 **(D)**, NB179 **(E)** or JC17 **(F)** for 0.5 h, and the soluble lysate fraction assayed by immunoblot. **(G)** HEK293T cells transiently expressing nsp14-FLAG were treated with 100 μm of the indicated compounds for the indicated times, and the soluble lysate fraction was assayed by immunoblot. All western blot data are representative of 2 independent experiments.

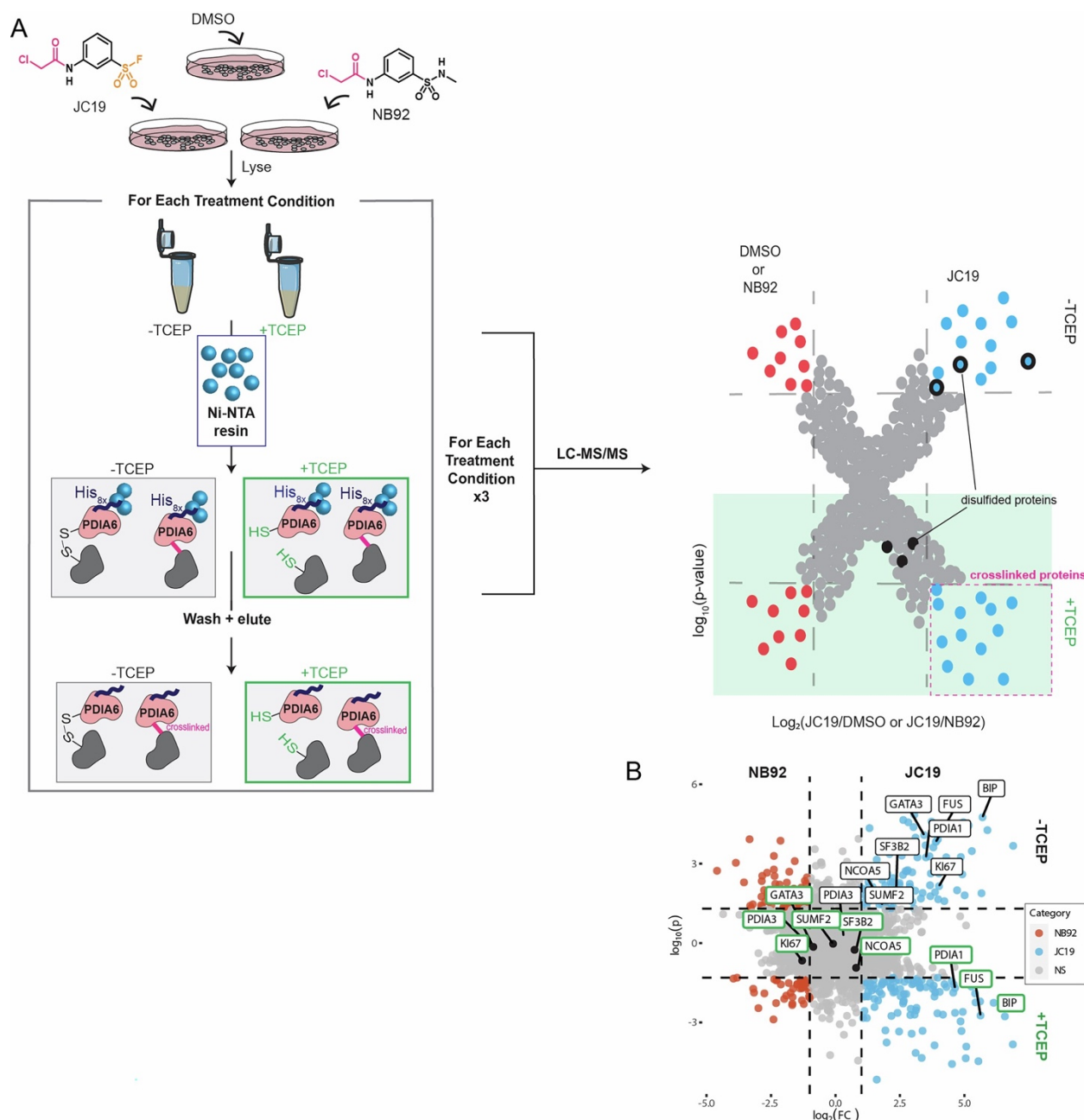


**Extended Data Fig. 3 JC19 induces linkage of predominantly K48-linked polyubiquitin chains on nsp14.** (A) HEK293T cells transiently expressing nsp14-FLAG were treated with the indicated compounds, lysed and immunoprecipitated on FLAG resin. (B) HEK293T cells transiently expressing nsp14-FLAG or the indicated ubiquitin construct were treated with DMSO or JC19 (50  $\mu$ M, 1 h), lysed and immunoprecipitated on FLAG resin. Immunoblot analysis was used to probe FLAG and ubiquitin expression for the immunoprecipitated fraction. Ubiquitin constructs used include WT (wild-type), K48R, K63R, K48 (all ubiquitin lysines mutated to arginine except K48) and K63 (all ubiquitin lysines mutated to arginine except K63). (C) Same immunoblot as shown in b, including inputs, loading control and full-length ubiquitin blot to depict polyubiquitin smears. Polyubiquitylation is primarily due to K48-linked ubiquitin, as indicated by the attenuation (K48R) or abrogation (K63) of the high-molecular-weight polyubiquitin smear when using ubiquitin constructs containing a mutant K48. (D) 'Heavy' SILAC HEK293T cells transiently expressing nsp14-FLAG were treated with 100  $\mu$ M JC19 for 1 h ( $n = 3$ ), while 'light' SILAC HEK293T cells transiently expressing nsp14-FLAG were treated with an equal volume of DMSO for 1 h ( $n = 3$ ).

Lysates were combined, immunoprecipitated on FLAG resin, proteolyzed and subjected to LC–MS/MS analysis. Mean  $\log_2(H/L)$  ratios of detected peptides are depicted, highlighting all nsp14 peptides and ubiquitinated nsp14 peptides (GlyGly modified). (E) Workflow for AP–MS experiments. All MS data can be found in Supplementary Table 2. Western blot data are representative of 2 independent experiments.

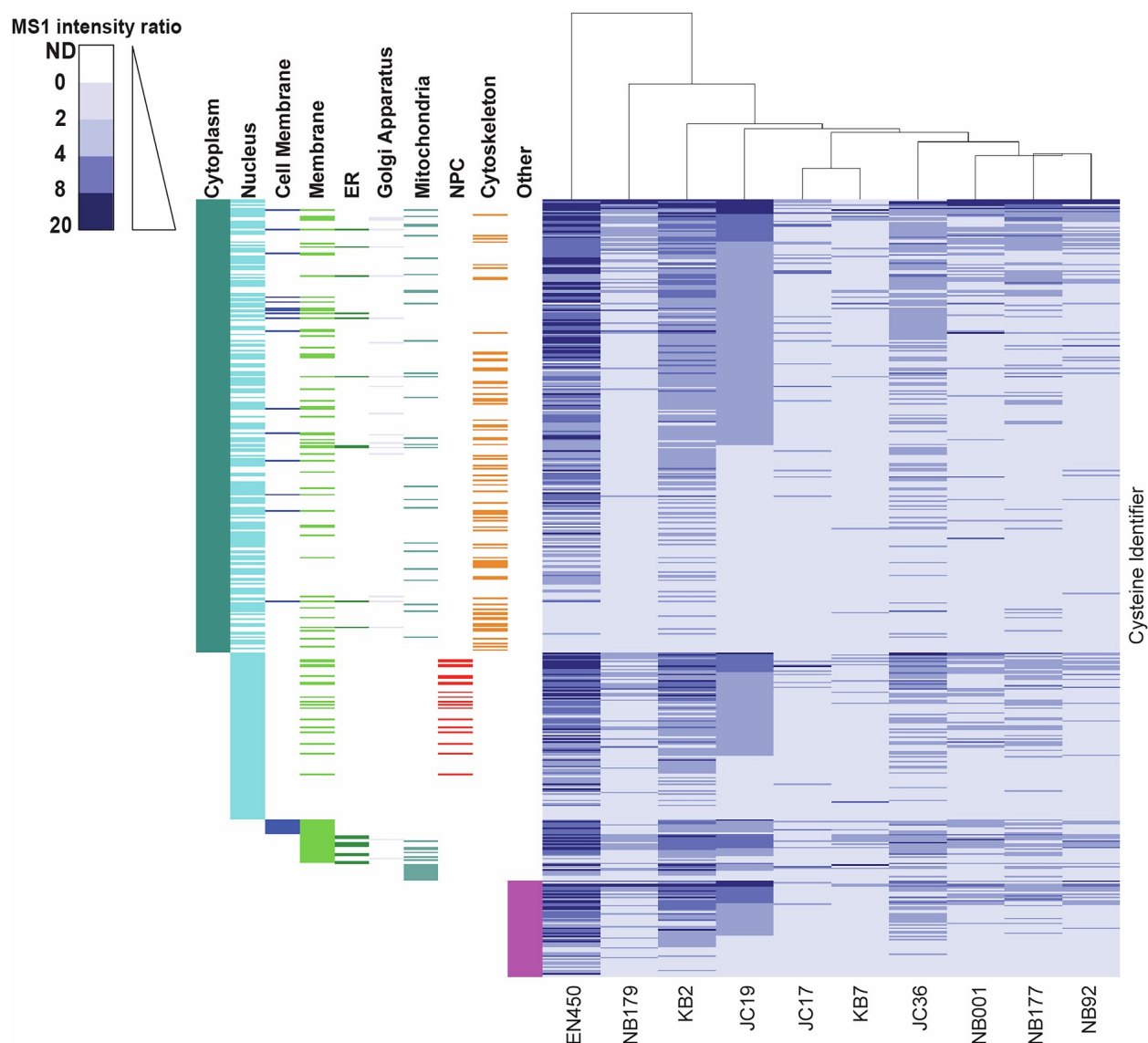


**Extended Data Fig. 4 Common cell stress-sensing pathways are not responsible for nsp14 depletion.** (A) HEK293T cells transiently expressing nsp14-FLAG were treated with the indicated concentrations of JC19 (1 h). Immunoblot analysis was used to visualize abundance of nsp14 and induction of NRF2 expression in the soluble lysate fraction (0.3% CHAPS in PBS) for each condition. (B) HEK293T cells transiently expressing nsp14-FLAG were pretreated with either DMSO, tunicamycin (12 μg/mL, 8 h), thapsigargin (2 μM, 8 h) or rapamycin (1 μM, 8 h), and then either treated with DMSO or JC19 (50 μM, 1 h). Immunoblot analysis was used to visualize abundance of nsp14 and induction of UPR markers in the soluble lysate fraction (0.3% CHAPS in PBS) for each condition. (C) HEK293T cells were treated with the indicated compounds, and immunoblot analysis was used to visualize induction of UPR markers (lysed in RIPA). (D–F) HEK293T cells were treated with vehicle DMSO, thapsigargin (2 μM, 15 h) or AA147 (10 μM, 15 h;  $n = 3$  per group; (D) or DMSO, 10 μM JC19 for 1 h or 10 μM JC19 for 15 h ( $n = 3$  per group; (E) or DMSO, 50 μM JC19 for 1 h or 50 μM JC19 for 15 h ( $n = 3$  per group; (F) and lysed in RIPA lysis buffer. Bulk proteomics sample preparation and label-free quantification proteomics were used to measure intensities of proteins and p-values generated for each treatment compared to vehicle using a Student's unpaired t-test.  $-\log_{10}(p\text{-values})$  were plotted and significance of BiP (HSPA5) and PDIA4 highlighted. All MS data can be found in Supplementary Table 2. Panel A is representative of one independent measurement, and panels B and C are representative of three independent measurements.



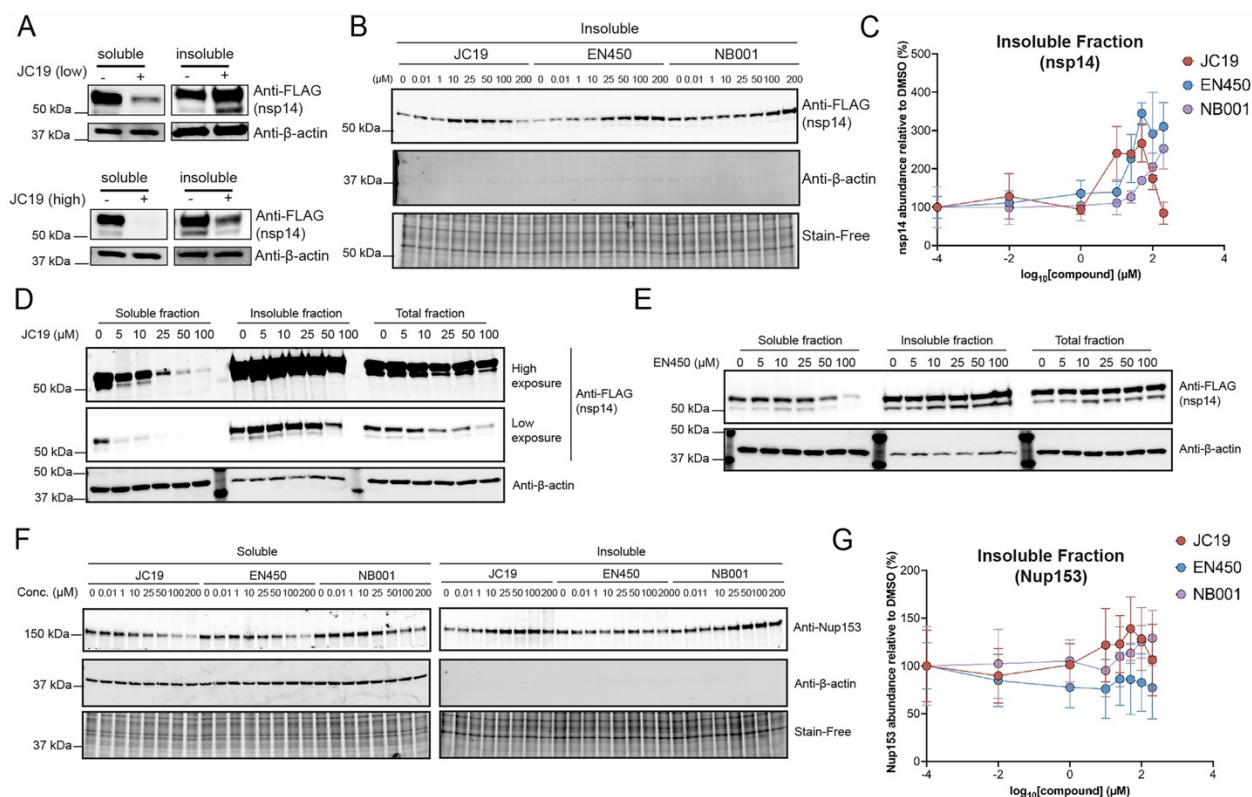
**Extended Data Fig. 5 PDIA6-His pulldown.** (A) HEK293T cells transiently expressing PDIA6-His protein were treated with DMSO, NB92 or JC19 for 1 h at 50  $\mu$ m. Cells were lysed, and each condition was split into 2 tubes, one treated with TCEP and one without TCEP. Samples were then incubated with Ni-NTA resin, washed, eluted and prepared for LC-MS/MS analysis. Proteins significantly enriched by JC19 in both the +TCEP and -TCEP conditions were identified to be true cross-linked proteins to PDIA6. Proteins enriched by JC19 only in the -TCEP condition but not +TCEP condition were identified to be linked to PDIA6 by disulfide bond. (B) HEK293T cells transiently expressing PDIA6-His8x were affinity purified under denaturing conditions either in the absence (black border labels) or presence of TCEP reductant (green border labels). Volcano plot displays comparison between enriched proteins for groups treated with DMSO versus NB92 (50  $\mu$ m, 1 h) with black colored dots for proteins showing sensitivity to TCEP ( $n=3$  for each group). An unpaired Student's t-test was performed to calculate p-values. All MS data can be found in Supplementary Table 3.



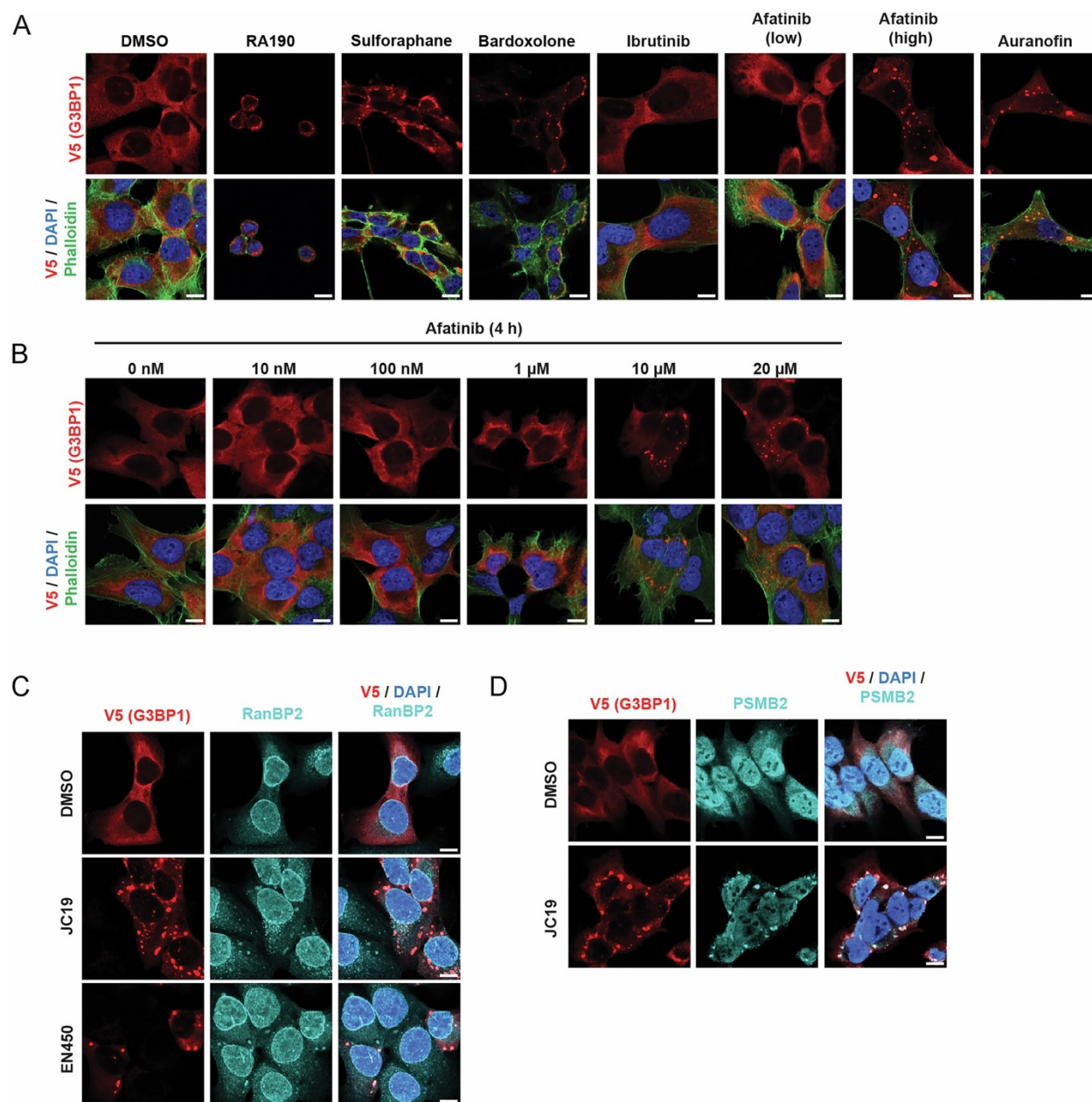


**Extended Data Fig. 6 High-ratio cysteines as identified by isoTOP-ABPP belong to proteins spanning various subcellular compartments.** Heatmap depicting unlogged MS1 isoTOP-ABPP cysteine ratios for high coverage cysteines belonging to proteins with multiple cysteines identified. Cell compartment annotations from UniProt have been provided for each cysteine identifier. For generation of data, each compound was used in triplicate ( $n = 3$ ) at  $100 \mu\text{M}$  for 1 h (exception: 4 h treatment for EN450) in nsp14-expressing HEK293T cells. MS data can be found in Supplementary Table 4.

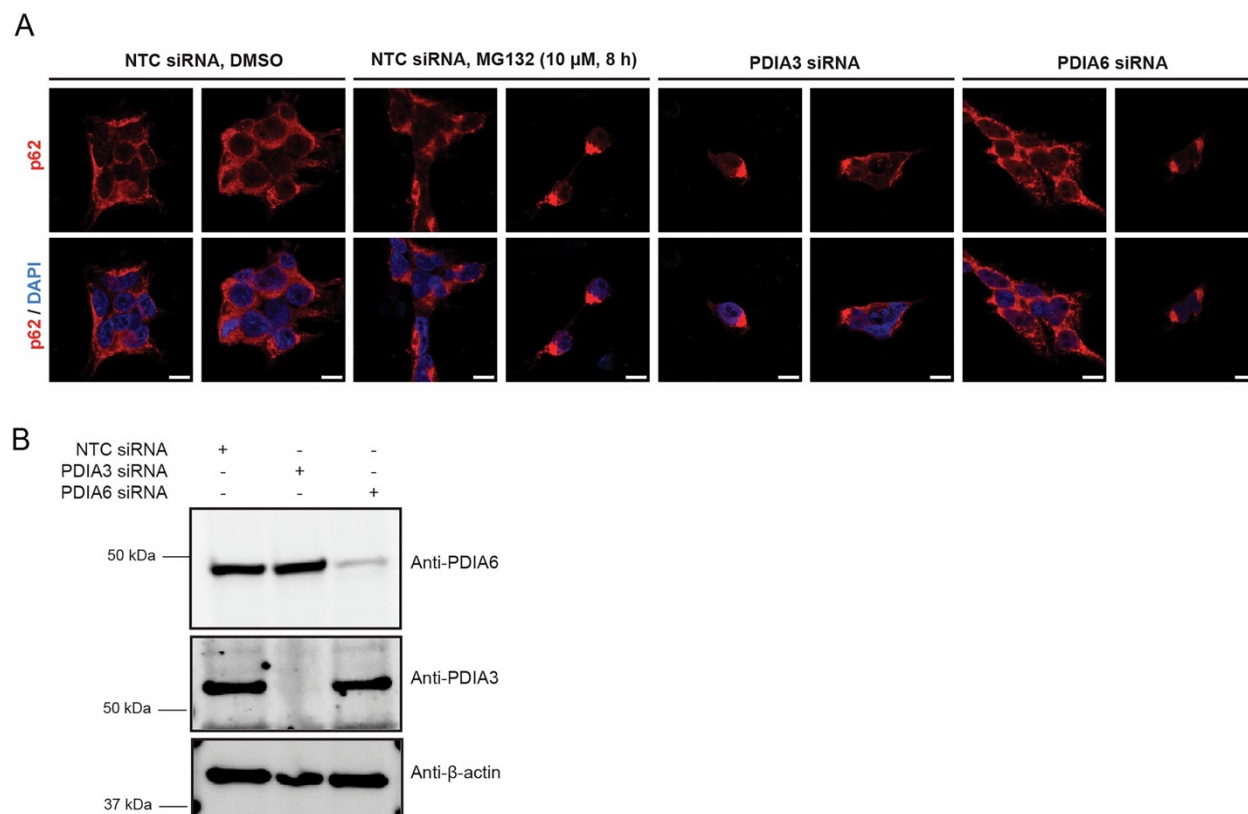




**Extended Data Fig. 7 Delineation of aggregation and degradation properties of JC19, EN450 and NB001.** (A) HEK293T cells transiently expressing nsp14-FLAG were treated with DMSO, a low dose of JC19 (25 μM for 0.5 h) or a high dose of JC19 (100 μM for 1 h) and subjected to immunoblot analysis. Cells were lysed in 0.3% CHAPS in PBS to generate the soluble lysate, and after clearance by centrifugation, the insoluble debris was solubilized in 8 M urea in PBS to generate 'insoluble' lysate. (B) HEK293T cells transiently expressing nsp14-FLAG were treated with the indicated concentrations of JC19 or NB001 for 0.5 h or EN450 for 3 h, and the insoluble lysate was analyzed by immunoblot. (C) Quantification of nsp14 intensity from replicate western blots as in B. Data are presented as mean ± s.d. (n = 3 per data point). D,E, Nsp14-expressing HEK293T cells were treated with the indicated concentrations of JC19 for 0.5 h (D) or EN450 for 3 h (E), and were then separated into 3 fractions: 'soluble' fraction (cells lysed in 0.3% CHAPS and cleared by centrifugation), 'insoluble' fraction (insoluble debris from CHAPS lysis re-solubilized using 8 M urea) and 'total' fraction (cells lysed in 8 M urea with 3× freeze/thaw). (F) HEK293T cells transiently expressing nsp14-FLAG were treated with the indicated concentrations of JC19 or NB001 for 0.5 h or EN450 for 3 h, and the soluble and insoluble lysates were analyzed by immunoblot. (G) Quantification of nsp14 intensity from replicate western blots as in F. Data are presented as mean ± s.d. (n = 3 per data point). All western blot data is representative of three independent measurements.



**Extended Data Fig. 8 Various cysteine-reactive electrophiles induce stress granule formation.** (A) U2OS cells stably expressing V5-tagged G3BP1 were treated with DMSO, RA190 (10  $\mu$ M, 1 h), sulforaphane (50  $\mu$ M, 1 h), bardoxolone (10  $\mu$ M, 1 h), ibrutinib (10  $\mu$ M, 14 h), low afatinib (100 nM for 14 h), high afatinib (10  $\mu$ M, 1 h) or auranofin (10  $\mu$ M, 1 h). Cells were then fixed, permeabilized and subjected to immunofluorescence microscopy. (B) U2OS cells stably expressing V5-tagged G3BP1 were treated with the indicated concentration of afatinib for 4 h. Cells were then fixed, permeabilized and subjected to immunofluorescence microscopy. (C) Immunofluorescence microscopy of G3BP1 and RANBP2 in response to DMSO, JC19 (100  $\mu$ M, 1 h) and EN450 (100  $\mu$ M, 1 h). (D) Immunofluorescence microscopy of G3BP1 and PSMB2 in response to DMSO and JC19 (100  $\mu$ M, 1 h). Images were acquired on an LSM880 confocal microscope at  $\times 63$  objective and 2 $\times$  manual zoom. All scale bars = 10  $\mu$ m. All images are representative of two independent measurements.



**Extended Data Fig. 9 Genetic knockdown of PDIA3 and PDIA6 lead to increases in aggregated proteins (aggresomes).** (A) Immunofluorescence microscopy of HEK293T cells transfected with either a non-targeting control (NTC) siRNA for 48 h, NTC siRNA for 48 h and treated with MG132 (10  $\mu$ M, 8 h), PDIA3 siRNA for 48 h or PDIA6 siRNA for 48 h, using p62 as a marker for aggresomes/aggregated proteins. Images were acquired on an LSM880 confocal microscope at  $\times 63$  objective and  $2\times$  manual zoom. All scale bars = 10  $\mu$ m. (B) Immunoblot analysis PDIA3 and PDIA6 knockdown efficiency for the cells used in a and Fig. 6e. All data are representative of three independent measurements.

| Challenge | Recommendation                                                                                                                                                                                                                          |
|-----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1         | Pair cell- and lysate-based chemoproteomic studies to delineate cell-specific activity—additional care should be taken in progressing hits that fail to reconfirm in vitro.                                                             |
| 2         | Assess the number of cysteines per protein identified with high ratios together with measures of bulk changes in protein abundance to delineate direct labeling from changes in abundance or protein stability.                         |
| 3         | Use protein-directed ABPP analysis with tailored click probes to confirm target engagement and occupancy.                                                                                                                               |
| 4         | Restrict follow-up studies to cysteines labeled at near complete (~100%) occupancy by prioritizing lead compounds that show reproducible high fold-change competition values for the cysteine of interest across replicate experiments. |
| 5         | Confirm clear SAR using multiple active and inactive control compounds, including enantiomeric pairs of compounds when possible.                                                                                                        |
| 6         | Assess generalized cell stress (e.g. stress granule formation, global polyubiquitylation, and aggresome formation), and relative compound reactivity towards thiols such as glutathione.                                                |
| 7         | Community-wide publication-associated data deposition into centralized repositories, such as CysDB, to facilitate identification of pan-electrophile sensitive proteins for exclusion from future studies.                              |
| 8         | For covalent compounds targeting non-essential cysteines, use compound-resistant cysteine mutants (e.g., cysteine-to-alanine/serine mutants) as controls.                                                                               |
| 9         | For covalent degraders, ternary complex formation should be assessed in vitro, and the POI and implicated ligase assessed via stringent genetic and biochemical characterization.                                                       |
| 10        | Assess electrophile reactivity (e.g. cysteine or glutathione reactivity) of test compounds as a broad metric of compound quality to use as context in chemoproteomic studies.                                                           |

**Extended Data Fig. 10 Recommendations for best practices.** Suggested steps to facilitate delineation of on- versus off-target activities of covalent fragments.

## Experimental Methods

**Cell lines, culture conditions.** Cell culture reagents including Dulbecco's phosphate-buffered saline (DPBS), Dulbecco's Modified Eagle Medium (DMEM) media, and penicillin/streptomycin (Pen/Strep) were purchased from Fisher Scientific. Fetal Bovine Serum (FBS) was purchased



from Avantor Seradigm (lot #214B17). All cell lines were obtained from ATCC and were maintained at a low passage number (<20 passages). U2OS cells were provided by Dr. Melody Li and used with author's permission. HEK293T (ATCC: CRL-3216), HeLa (ATCC: CCL-2), and U2OS (ATCC: HTB-96) cells were cultured in DMEM supplemented with 10 % FBS and 1 % antibiotics (Penn/Strep, 100 U/mL). Stable-Isotope Labeling by Amino Acids in Cell Culture (SILAC) cells were maintained in DMEM for SILAC supplemented with 10% dialyzed FBS, 1 % antibiotics (Penn/Strep, 100 U/mL), and isotopically labeled arginine and lysine. All media was filtered (0.22 µm) prior to use. Cells were maintained in a humidified incubator at 37 °C with 5 % CO<sub>2</sub>. Cell lines were tested monthly for mycoplasma using the Mycoplasma Detection Kit (InvivoGen).

**Cloning of plasmids.** List of plasmids with detailed information used in this study can be found in **Supplementary Table 9**. pDONR223 SARS-CoV-2 vectors containing sequences of nsp9, nsp10, nsp14, and nsp16 were gifts from Fritz Roth (Addgene plasmid #149312, 149313, 149316, 149318, respectively) and subcloned using GateWay cloning into C-terminal FLAG destination vector, which was a kind gift from T Wucherpfennig. For GateWay Cloning, 300 ng of both the donor and destination vector were combined in a microcentrifuge tube, following addition of TE buffer up to 8 µL. 2 µL of Gateway LR Clonase II enzyme mix (Invitrogen; 11791020) was added to the mixture, vortexed briefly, and allowed to incubate for 2 hours at room temperature before being transferred to 4°C for 12-16 hours. The enzymes were then inactivated at 65°C for 15 minutes, and the reaction mixture transformed into competent TOP10 cells, colonies grown up in LB media containing 100 µg/mL ampicillin, and successful cloning confirmed by sequencing. Plasmid for stable overexpression of the G3BP1-TurboID-V5 construct was generated via two rounds of restriction cloning. First, the G3BP1 sequence was PCR amplified from pENTR4\_G3BP1 (a gift from Thomas Tuschl, Addgene plasmid #127104) using Phusion polymerase (Berkeley) and restriction cloned via HindIII and NotI restriction sites into C1(1-29)-

TurboID-V5\_pCDNA3 (a gift from Alice Ting, Addgene plasmid #107173)<sup>95</sup> using T4 DNA ligase. Then, the G3BP1-TurboID-V5 sequence was PCR amplified from this construct using Phusion polymerase (Berkeley) and restriction cloned into both the empty Piggybac transposon vector (a kind gift from Dr. Melody Li's lab) and the pLenti CMV GFP Hygro vector (a gift from Eric Campeau and Paul Kaufman) (Addgene plasmid #17446) for generation of constructs that can be used for doxycycline-inducible stable expression or constitutive stable expression, respectively. HA-HECTD1 in pCMV plasmid was gifted by Irene Zohn at Children's National. The PDIA6-His8x plasmid was generated through Gateway cloning of entry plasmid human PDIA6 in pDONR221 (DNASU HsCD00451137) into pEGc1-DEST (DNASU EvNO00534626).

#### **Generation of U2OS cell line stably expressing doxycycline-inducible G3BP1-TurboID-V5.**

U2OS cells were transfected at 70-80% confluency. For transfection of a single well of a 6-well plate, plasmid (0.5 ug of G3BP1-TurboID-V5\_Piggybac transposon vector + 0.5ug of the ePiggybac transposase vector), reduced-serum OptiMEM (up to 100uL) and Lipofexin (Lambda Biotech, TS310) (3 µL) were mixed and incubated for 15 min at room temperature, followed by adding dropwise to the cells. 6-8 hours after transfection, media was replaced with fresh culture media and cells were incubated for 48 hours. After 48 hours, media was replaced with fresh media containing 1 µg/mL of puromycin (Fisher, AAJ672368EQ) for selection of stably transfected cells. Selection media was replaced every 24-48 hours until colonies of transfected cells were visible. Cells were expanded to larger plates and cryogenically frozen in FBS containing 10% DMSO for future use.

#### **Generation of HEK293T cell line stably expressing constitutively expressed G3BP1-**

**TurboID-V5.** For preparation of lentiviruses, HEK293T cells in 10 cm plates were transfected at ~70-80% confluency with lentiviral vector pLenti\_CMV\_GFP\_Hygro (656-4) containing G3BP1-TurboID-V5 (10 µg; Addgene #17446) with the lentiviral packaging plasmids pVSVG (4 µg;

Addgene #8454) and  $\Delta$ 8.9 (8  $\mu$ g; Addgene #2221) and 66  $\mu$ L of Turbo DNAfectin3000 (Lamda Biotech, TS310) in antibiotic-free media for 6 hours. After 6 hours, the DNAfectin-containing media was replaced with fresh antibiotic-free media and the cells were left to incubate for 48 hours for lentiviral generation. The media was collected (containing virus) and stored at 4°C and cells were incubated for another 24 hours in fresh media. After 24 hours, the lentivirus-containing media was collected, and added to the previously harvested media. 1/3 volume of Lenti-X concentrator (Takara Bio, Cat# 631231) was added to the total harvested media and incubated 16 hours at 4°C. The lentivirus was pelleted at 1500 *g* for 45 minutes at 4°C and resuspended in 500  $\mu$ L plain DMEM and stored in 100  $\mu$ L aliquots at -80°C. 300  $\mu$ L of reconstituted virus was added dropwise to HEK293T cells at 70-80% confluency in antibiotic-free media containing 8  $\mu$ g/mL polybrene and allowed to incubate for 48 hours. Media was then replaced with DMEM containing 400  $\mu$ g/mL hygromycin and allowed to incubate for 48 hours for selection of transduced cells. Selection media was replaced every 24-48 hours until the appearance of visible colonies of transduced cells. Cells were expanded to larger plates and cryogenically frozen in FBS containing 10% DMSO for future use.

**Mutagenesis.** Point mutations (C11A, C39A, C94A, C208A, C208V, C216A, C272A, C279A, C309A, C356A, C414A, C340A) were created by PCR-based site-directed mutagenesis as previously described<sup>96</sup>. All primers used to generate mutant constructs can be found in **Supplementary Table 8**. Mutant plasmids were then transformed into competent TOP10 cells, colonies selected and grown in SOC media at 37°C for >2 hours, and successful mutagenesis confirmed by sequencing.

**Transfections.** Cells were transfected at 70-80% confluency. For a 10-cm plate, plasmid (5  $\mu$ g), serum-free DMEM (350  $\mu$ L) and PEI MAX- Transfection Grade Linear Polyethylenimine Hydrochloride (MW 40,000) (Polysciences, Inc., 24765-1) (35  $\mu$ L of 1 mg/mL, pH 7.4) were mixed

and incubated for 15 minutes at room temperature. Transfection cocktail was then added dropwise to the cells and incubated for ~36 hours. Transfection reagents were scaled according to volume of media in plates of various sizes.

**Cell treatments.** All compounds were made up as 50 mM stock solutions in DMSO. For cells transiently expressing a protein of interest, treatments were conducted 24 hours post-transfection. Otherwise, cells were grown to ~80% confluency prior to treatment. Cells were treated with either DMSO (vehicle) or the compound concentration indicated in the respective figure legend, and incubated at 37°C for the indicated time. For isoTOP-ABPP analyses, all fragments were screened at 100  $\mu$ M *in situ* for 1 hour, except for EN450, which was screened at 100  $\mu$ M for 4 hours.

**Cell harvesting and lysis.** Cells were washed with cold PBS, scraped into cold PBS and harvested by centrifugation at 1800 *g* for 3 minutes. Cell pellets were then washed twice with cold PBS. Unless otherwise indicated, cells were lysed in 0.3% 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) in PBS (pH 7.4) with 1X protetriethylaminase inhibitor cocktail (Sigma, 11836170001) and incubated on ice for 30 minutes (protetriethylaminase inhibitor cocktail was excluded from any lysates that required CuAAC 'click' conjugation). After lysis, cellular debris was clarified by centrifugation at 21,100 *g* for 5 minutes and the soluble fraction transferred to a fresh microcentrifuge tube. If applicable, the insoluble fractions were generated by re-solubilizing the pelleted debris with 8M urea in PBS, shaking at 600 rpm for 30 minutes at 37°C. Protein concentrations were determined using a BioRad DC protein assay kit from BioRad Life Science (Hercules, CA) and the lysate diluted to 2-5 mg/mL, with all samples of the same experiment normalized to the same concentration.



**Western Blots.** Lysate was normalized to 2-5 mg/mL, 25  $\mu$ L of each lysate was added to 8  $\mu$ L of 4X SDS loading dye (BioRad), and the samples were denatured at 95°C for 5 minutes. At least 60  $\mu$ g of protein was resolved on a 4-20% SDS-PAGE gel (BioRad). Gels were transferred to nitrocellulose or activated PVDF membranes (BioRad) using a semi-dry transfer system (BioRad) and membranes blocked in 5% (w/v) milk in Tris-buffered saline (TBS) for 30 minutes by rocking at room temperature. Membranes were rocked with primary antibodies (**Supplementary Table 6**) in 5% (w/v) milk in TBS overnight (14-16 hours) at 4°C, then washed 3 times with TBS for 5 minutes. Membranes were then rocked with secondary antibodies (**Supplementary Table 6**) in 5% (w/v) milk in TBST (Tris-buffered saline with 0.1% Tween20) for 1 hour at room temperature, then washed 3 times with TBS. Membrane was imaged on a BioRad ChemiDoc Imaging System. Antibodies used were listed in **Supplementary Table 6**. ImageJ<sup>97</sup> was used to normalize and quantify band intensities.

**Calculation of apparent DC<sub>50</sub>.** HEK293T cells were reverse transfected with nsp14-FLAG for 24 hours (seeded at  $3.0 \times 10^6$  cells per well in a 6-well plate with 1.5  $\mu$ g of DNA per well). The cells were then treated with JC19, EN450, and NB001 at final concentrations of 0  $\mu$ M (vehicle DMSO), 0.01  $\mu$ M, 1  $\mu$ M, 10  $\mu$ M, 25  $\mu$ M, 50  $\mu$ M, 100  $\mu$ M, and 200  $\mu$ M. Cells were treated with JC19 and NB001 for a duration of 30 minutes at 37°C and EN450 for a duration of 3 hours at 37°C. Cells were harvested, lysed, and soluble fractions were collected. The insoluble portion was resolubilized with 8M Urea in PBS at 37°C for 30 minutes with shaking. All lysates were normalized to 2-3 mg/mL. 25  $\mu$ L of each lysate was added to 8  $\mu$ L of 4X SDS loading dye (BioRad) and the samples were denatured at 95°C for 5 minutes. Samples were resolved on 4-20% SDS-PAGE gels (BioRad) and transferred to a nitrocellulose membrane. The nsp14-FLAG signal was detected by an anti-FLAG primary antibody (**Supplementary Table 6**). Membranes were imaged using a BioRad ChemiDoc Imaging System. Nsp14 bands of each sample were quantified using ImageJ<sup>97</sup> and plotted in GraphPad Prism 7 to obtain an apparent DC<sub>50</sub> value. Briefly, the intensity

of the nsp14 band for each replicate of compound treatment was quantified by ImageJ<sup>97</sup>, and intensity values were normalized to the DMSO control in GraphPad Prism. Normalized values were then transformed to logarithms in GraphPad Prism and the IC50 calculated using a nonlinear regression curve fit.

**Immunoprecipitation for Western Blot.** HEK293T cells were transfected with the specified constructs for 36 hours and treated with the specified compounds indicated on the respective figures (**Figure 3B and Figure 3G**). Cells were harvested, lysed, and normalized as described in 'Cell harvesting and lysis.' To prepare anti-FLAG EZView resin (Sigma, F2426) for immunoprecipitation, 10  $\mu$ L of anti-FLAG EZView resin per sample was washed in 600  $\mu$ L TBS 2 times, then resuspended in 10  $\mu$ L TBS/sample. Lysates were immunoprecipitated by rotating 500  $\mu$ L of lysates with 10  $\mu$ L of anti-FLAG EZView resin per sample at 4°C for 2 hours. At least 50  $\mu$ L of the original lysates was kept for detection of input samples. Resin was spun down at 8200 *g* for 1.5 minutes and supernatant aspirated, followed by 3 washes with 500  $\mu$ L of ice cold TBS (pH 7.4). Resin was then boiled in 30  $\mu$ L of 4X SDS loading dye (absent of Beta-mercaptoethanol) for 10 minutes to release immunoprecipitated proteins. 80  $\mu$ g of each input lysate and 15  $\mu$ L of each immunoprecipitated sample was loaded and separated on 4-20% SDS-PAGE gels (BioRad). Gels were transferred to nitrocellulose membranes and taken through the standard western blotting procedure as described in the 'Western Blots' section.

**siRNA Knockdowns.** 10  $\mu$ M siRNA stocks (**Supplementary Table 7**) resuspended in nuclease-free duplex buffer (IDT, Cat. No. 11-01-03-01) were combined with 140  $\mu$ L of DharmaFECT Transfection Reagent #1 (Horizon Discovery, Cat. No. T-2001-01) at a dilution of 1:100 with HBSS (Gibco) for a final concentration of 25 nM of siRNA per well for a 6-well cell culture plate. The siRNA and Transfection Reagent #1 mixture in HBSS was incubated at ambient temperature for 20 mins. HEK293T cells at a density of  $\sim 4.5 \times 10^5$  cells/well were then reverse transfected with

the incubated siRNA/Transfection Reagent #1 mix. Cells were left in the 37°C incubator for 48 hours before either harvesting or transfecting. For transfections, cells were transfected with nsp14-FLAG plasmid (as described above) for 24 hours, treated with compounds (if stated), and then harvested as described in 'Cell harvesting and lysis.' All samples were normalized and prepared for analysis by SDS-PAGE and western blot analysis as described in 'Western blots.' For analysis by mass spectrometry, samples were prepared as stated below in the "Protein abundance sample preparation" section. All siRNAs used in this study are listed in **Supplementary Table 7.**

**IsoTOP-ABPP sample preparation.** HEK293T cells were transfected with nsp14-FLAG for 36 hours (unless otherwise indicated) and treated with either DMSO or electrophilic fragment in at least biological duplicate. Cells were then harvested, lysed, and lysate concentration normalized to 2 mg/mL as described in 'Cell harvesting and lysis'. 200 µL of normalized lysates were then treated with 200 µM iodoacetamide alkyne (IAA) for 1 hour in the dark at ambient temperature. Samples were then subjected to bioorthogonal copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC) 'click' conjugation with either heavy (DMSO treated samples) or light (compound treated samples) biotin-azide tags containing a TEV cleavable linker<sup>3</sup> for 1 hour at ambient temperature. Click reaction was performed by adding 24 µL of a premixed cocktail of click reagents (TEV tags (4 µL of 5 mM stock per sample, final concentration = 100 µM), TCEP (4 µL of fresh 50 mM stock in water per sample, final concentration 1 mM), TBTA (12 µL of 1.7 mM stock in DMSO/t-butanol 1:4 per sample, final concentration = 100 µM), and CuSO<sub>4</sub> (4 µL of 50 mM stock in water per sample, final concentration = 1 mM) to each 200 µL lysate sample. Protein was then solubilized by addition of SDS (final concentration = 1%), and treated with benzonase (0.5 µL) at 37°C for 30 minutes. Equal volumes (200 µL) of heavy and light samples were then combined and subject to Single-Pot Solid-Phase-enhanced sample-preparation (SP3) clean-up. For preparation of SP3 beads, 40 µL Sera-Mag SpeedBeads Carboxyl Magnetic Beads per sample, hydrophobic (GE

Healthcare, 65152105050250, 50 µg/µL, total 1 mg) and 40 µL Sera-Mag SpeedBeads Carboxyl Magnetic Beads per sample, hydrophilic (GE Healthcare, 45152105050250, 50 µg/µL, total 1 mg) were mixed and washed with water three times, then resuspended in 80 µL water per sample. 80 µL of the bead slurry was then transferred to the lysate samples (800 µg protein input) and incubated for 5 minutes at room temperature with shaking (1000 rpm). 600 µL of absolute ethanol was added to each sample and incubated for 10 minutes at room temperature with shaking (1000 rpm). The beads were then washed (3x, 500 µL 80% ethanol) using a magnetic rack (Sergi Lab Supplies, Cat. No. 1005a). Beads were then resuspended in 200 µL 2M urea in 0.5% SDS/PBS. DTT was added to a final concentration of 10 mM to each sample and incubated at 65°C for 15 minutes. Iodoacetamide was added to each sample to a final concentration of 20 mM and incubated at 37°C for 30 minutes. Next, absolute ethanol (500 µL) was added to each sample, and the samples were incubated for 5 min at room temperature with shaking (1000 rpm). The beads were then washed (3x, 500 µL 80% ethanol) using a magnetic rack. Next, beads were resuspended in 150 µL 2M urea in PBS and 3 µL of 1 mg/mL trypsin solution was added. Lysates were digested for 12-16 hours at 37°C with shaking (200 rpm). After digestion, 3.8 mL acetonitrile (> 95% of the final volume) was added to each sample and the mixtures were incubated for 10 min at room temperature with shaking (1000 rpm). The beads were then washed (3x, 1 mL acetonitrile) with a magnetic rack. Peptides were eluted from SP3 beads with 100 µL of 2% DMSO in molecular biology-grade water for 30 min at 37°C with shaking (1000 rpm). The elution was repeated again with 100 µL of 2% DMSO in molecular biology-grade water and the elution fractions from the same sample combined. The samples were then enriched on Pierce streptavidin agarose beads in PBS (Thermo Scientific, Cat. No. 20353). To prepare streptavidin resin, 50 mL of resin per sample was washed in 10 mL PBS and collected by centrifugation at 1800 g for 3 minutes. PBS was then aspirated and the resin resuspended in 1 mL PBS/per sample. 1 mL of the resin slurry was then added to each sample and samples incubated for 2 hours with rotation at ambient temperature. Samples were then spun down at 1800 g for 3 minutes

and supernatant aspirated. Streptavidin resin was then washed 2x with 1 mL PBS, then washed 2x with 1 mL water, collecting resin by centrifuging at 1800 g for 3 minutes in between each wash. Final wash was aspirated and the beads were resuspended in 75  $\mu$ L of 1X TEV solution (50 mM Tris, pH 8, 0.5 mM EDTA, 1 mM DTT). 1.5  $\mu$ L TEV protease (2 mg/mL; MacroLab, UC Berkeley) was added. Samples were then incubated with rotation at 30°C for 7 hours. Samples were then spun at 1800 g for 3 minutes and the supernatant was collected. The beads were then washed with 20  $\mu$ L of water, spun, and the wash was collected and combined with the supernatant. Trifluoroacetic acid (TFA) was added to the samples to a final concentration of 0.5%. Samples were then cleaned up with Pierce C18 spin tips (Thermo Fisher, Cat. No. 87784) according to the manufacturer's protocol. The samples were dried (SpeedVac) and reconstituted with mass spectrometry buffer (5% acetonitrile and 1% formic acid in molecular biology-grade water) and analyzed by LC-MS/MS.

**AP-MS sample preparation.** Stable-Isotope Labeling by Amino Acids in Cell Culture (SILAC) HEK293T cells were used for affinity purification-mass spectrometry (AP-MS) sample preparation (**Figure 3D, 3E, 3F**). 6 10-cm plates of 'heavy' labeled HEK293T cells and 6 10-cm plates of 'light' labeled HEK293T cells were transfected with nsp14-FLAG for 36 hours. All 6 'light' labeled plates were then treated with DMSO for 1 hour, while 3 'heavy' labeled plates were treated with 100  $\mu$ M JC19 for 1 hour and 3 'heavy' labeled plates were treated with 100  $\mu$ M NB001 for 1 hour. All cells were harvested, lysed, and lysate concentrations normalized to 2.5 mg/mL. 120  $\mu$ L of anti-FLAG EZView resin (Sigma, F2426) (10  $\mu$ L/sample) was washed 2 times with 1X TBS and then resuspended in 120  $\mu$ L of 1X TBS. 10  $\mu$ L of washed EZView resin was added to each 350  $\mu$ L lysate sample and rotated at 4°C for 2 hours. Resin was then spun down at 8200 g for 1.5 minutes and supernatant aspirated, followed by 3 washes with 500  $\mu$ L of ice-cold TBS (pH 7.4). Bound proteins were eluted in 40  $\mu$ L 8M urea in PBS at 37°C and 400 rpm for 30 minutes. Resin was spun down and eluant transferred to a new microcentrifuge tube. Resin was washed in an

additional 20  $\mu$ L of 8M urea in PBS and combined with previous eluant fraction. 50  $\mu$ L of each 'heavy' lysate was then combined with a corresponding 50  $\mu$ L of 'light' lysate. DTT was added to each sample to a final concentration of 10 mM and incubated at 65°C for 15 minutes. Iodoacetamide was added to each sample to a final concentration of 20 mM and incubated at 37°C for 30 minutes. Samples were diluted with PBS such that final urea concentration was 2M. 7  $\mu$ L of 1 mg/mL trypsin was added to each sample and incubated at 37°C and 200 rpm for 12-16 hours to digest proteins. Trifluoroacetic acid (TFA) was added to the samples to a final concentration of 0.5%. Samples were then cleaned up with Pierce C18 spin tips (Thermo Fisher, Cat. No. 87784) according to the manufacturer's protocol. The samples were dried (SpeedVac) and reconstituted with mass spectrometry buffer (5% acetonitrile and 1% formic acid in molecular biology-grade water) and analyzed by LC-MS/MS.

**AP-MS for His-tagged PDIA6 sample preparation.** HEK293T cells in 6 cm plates were transfected with a PDIA6-His<sub>8x</sub> plasmid for 24 hours followed by treatment with either vehicle (DMSO), NB092 (50  $\mu$ M), or JC19 (50  $\mu$ M) for 1 hour. Each condition was done in triplicate. Cells were then harvested and lysed by freeze thaw lysis (3x) using 0.3% CHAPS in 8M urea dissolved in PBS with the addition of protease inhibitor cocktail. Lysates were centrifuged at 10,000 g for 5 minutes at 4°C to remove debris. Protein concentrations were determined using a BioRad DC protein assay kit from BioRad Life Science (Hercules, CA) and the lysate diluted to 5.5 mg/mL, with all samples of the same experiment normalized to the same concentration. Lysates from each condition were split in half so that the input for each condition was 165  $\mu$ g. One set of the conditions were diluted down to 1.1 mg/mL and incubated with 50  $\mu$ L of washed Ni-NTA agarose resin (Thermo Scientific™ No. PI88222) per condition. The other set was first treated with 5 mM TCEP for 40 minutes then diluted down to 1 mM TCEP and 1.1 mg/mL final protein concentration before incubating with 50  $\mu$ L of washed Ni-NTA resin per condition.

Prior to incubating samples on Ni-NTA resin, each 50  $\mu$ L aliquot of resin was washed with 400  $\mu$ L of 6M urea in TBS with 10 mM imidazole for a total of 5 times. Resin was spun down at 1300 g for 3 minutes to separate resin from supernatant. Imidazole was added to the diluted lysates for a final concentration of 10 mM and let incubate with Ni-NTA resin for 2 hours at ambient temperature. Resin was spun down for each condition and the supernatant was removed. Resin was then washed 5 times with 400  $\mu$ L of 6M urea in TBS with 10 mM imidazole. Protein from resin was eluted with 290 mM imidazole in 6M urea in TBS. Elutions were then subjected to SP3 sample cleanup as outlined in the *Bulk proteomics sample preparation* section below.

**Proteasome activation assay.** HEK293T cells were seeded in a 6-well plate and grown to 90% confluence. Cells were pretreated as indicated in figure legends, then treated with 500 nM fluorescent proteasome probe Me4BodipyFL-Ahx3Leu3VS for 1 hour. Cells were then harvested and lysed as described in 'Cell harvesting and lysis.' Normalized cell lysates were then run on SDS-PAGE gel (BioRad) and in-gel fluorescence used to visualize proteasome activity.

**Click probe non-competition and competition enrichment sample preparation.** HEK293T cells were transfected with nsp14-FLAG for 36 hours. For non-competition enrichment experiments, cells or lysates were either treated with 100  $\mu$ M electrophilic compound ( $n = 6$  for in-cell labeling,  $n = 3$  for lysate labeling) or 100  $\mu$ M alkynylated electrophilic probe ( $n = 6$  for in-cell labeling,  $n = 3$  for lysate labeling) for 20 minutes. For competition-based enrichment experiments, cells were either treated with DMSO ( $n = 6$  for in-cell labeling,  $n = 3$  for lysate labeling) or 100  $\mu$ M electrophilic compound ( $n = 6$  for in-cell labeling,  $n = 2$  for lysate labeling) for 20 minutes, then all conditions treated with 100  $\mu$ M alkynylated electrophilic compound for 20 minutes. Cells were then harvested, lysed, and lysate concentration normalized to 3 mg/mL. 200  $\mu$ L each lysate was then subject to CuAAC 'click' biotinylation by treatment with 24  $\mu$ L of premixed cocktail of click reagents (12  $\mu$ L 1.7mM TBTA per sample, 4  $\mu$ L 50mM copper sulfate per sample, 4  $\mu$ L 50mM biotin-

azide per sample, 4  $\mu$ L 50mM TCEP per sample). Samples were then cleaned via SP3 clean-up. 20  $\mu$ L Sera-Mag SpeedBeads Carboxyl Magnetic Beads per sample, hydrophobic (GE Healthcare, 65152105050250, 50  $\mu$ g/ $\mu$ L, total 1 mg) and 20  $\mu$ L Sera-Mag SpeedBeads Carboxyl Magnetic Beads per sample, hydrophilic (GE Healthcare, 45152105050250, 50  $\mu$ g/ $\mu$ L, total 1 mg) were mixed and washed with water three times and resuspended in 40  $\mu$ L of water per sample. 40  $\mu$ L of the bead slurry was then transferred to the lysate samples (500  $\mu$ g protein input) and incubated for 5 minutes at room temperature with shaking (1000 rpm). 500  $\mu$ L of absolute ethanol was added to each sample and incubated for 10 minutes at room temperature with shaking (1000 rpm). The beads were then washed (3x, 400  $\mu$ L 80% ethanol) with a magnetic rack. Proteins were eluted from SP3 beads with 100  $\mu$ L of 0.2% SDS in PBS for 30 min at 37°C with shaking (1000 rpm). The elution was repeated again with 100  $\mu$ L of 0.2% SDS in PBS and the elution fractions from the same sample combined. Streptavidin resin (Thermo Scientific, Cat. No. 20353) was then prepared by washing 50  $\mu$ L of streptavidin resin per sample in 10mL PBS and collecting resin by centrifugation at 1800 *g* for 5 minutes, PBS aspirated, and resin resuspended in 1 mL PBS per sample. 1mL of streptavidin slurry was added to each sample and samples rotated at room temperature for 1.5 hours. Streptavidin resin was collected via centrifugation at 1500 *g* for 2 minutes and washed 1 time with 1 mL 0.2% SDS in PBS, 3 times with 1 mL PBS, and 3 times with 1 mL molecular biology-grade water. Resin was then resuspended in 200  $\mu$ L 6M urea. DTT was then added to each sample to a final concentration of 10 mM and incubated at 65°C for 15 minutes. Iodoacetamide was then added to each sample to a final concentration of 20 mM and incubated at 37°C for 30 minutes. Urea was diluted to ~2M, resin collected via centrifugation at 1500 *g* for 2 minutes, supernatant aspirated, and resin resuspended in 150  $\mu$ L 2M urea in PBS. Then, 3  $\mu$ L of 1 mg/mL trypsin was added to each sample and samples were subject to digestion for 12-16 hours at 37°C and 200 rpm. Resin was then spun down and supernatant transferred to a new tube (containing digested peptides). Resin was washed with an additional 50  $\mu$ L of molecular biology-grade water, spun down, and supernatant added to the previous peptide



fraction. Trifluoroacetic acid (TFA) was added to the samples to a final concentration of 0.5%. Samples were then cleaned up with Pierce C18 spin tips (Thermo Fisher, Cat. No. 87784) according to the manufacturer's protocol. The samples were dried (SpeedVac) and reconstituted with mass spectrometry buffer (5% acetonitrile and 1% formic acid in molecular biology-grade water) and analyzed by LC-MS/MS.

**Bulk proteomics sample preparation.** HEK293T cells were transfected with nsp14-FLAG for 36 hours and treated with either DMSO or electrophilic compound in biological triplicate. Cells were then harvested, lysed, and lysate concentration normalized to 3 mg/mL. Samples were then cleaned using SP3 clean-up. 20  $\mu$ L Sera-Mag SpeedBeads Carboxyl Magnetic Beads per sample, hydrophobic (GE Healthcare, 65152105050250, 50  $\mu$ g/ $\mu$ L, total 1 mg) and 20  $\mu$ L Sera-Mag SpeedBeads Carboxyl Magnetic Beads per sample, hydrophilic (GE Healthcare, 45152105050250, 50  $\mu$ g/ $\mu$ L, total 1 mg) were mixed and washed with water three times and resuspended in 40  $\mu$ L water per sample. 40  $\mu$ L of the bead slurry was then transferred to the lysate samples (500ug protein input) and incubated for 5 minutes at room temperature with shaking (1000 rpm). 500  $\mu$ L of absolute ethanol was added to each sample and incubated for 10 minutes at room temperature with shaking (1000 rpm). The beads were then washed (3x, 400uL 80% ethanol) with a magnetic rack. Beads were then resuspended in 200  $\mu$ L 2M urea in 0.5% SDS/PBS. DTT was added to each sample to a final concentration of 10 mM to each sample and incubated at 65°C for 15 minutes. Iodoacetamide was added to each sample to a final concentration of 20 mM and incubated at 37°C for 30 minutes. Absolute ethanol (400  $\mu$ L) was then added to each sample, and the samples were incubated for 5 minutes at room temperature with shaking (1000 rpm). Beads were then washed (3x, 400uL 80% ethanol) with a magnetic rack. Next, beads were resuspended in 200  $\mu$ L 2M urea in PBS and 3  $\mu$ L of 1 mg/mL trypsin solution was added. Proteins were digested for 12-16 hours at 37°C with shaking (200 rpm). After digestion, 3.8 mL acetonitrile (> 95% of the final volume) was added to each sample and the

mixtures were incubated for 10 minutes at room temperature with shaking (1000 rpm). The beads were then washed (3x, 1mL acetonitrile) with a magnetic rack. Peptides were eluted from SP3 beads with 100  $\mu$ L of 2% DMSO in molecular biology-grade water for 30 min at 37°C with shaking (1000 rpm). The elution was repeated again with 100  $\mu$ L of 2% DMSO in molecular biology-grade water and the elution fractions from the same sample combined. The samples were dried (SpeedVac), then reconstituted with 5% acetonitrile and 1% formic acid in molecular biology-grade water and analyzed by LC-MS/MS.

**Pathway analysis.** For KEGG and Gene Ontology analyses, the subset of proteins to be analyzed was searched using the Enrichr search algorithm<sup>98</sup>, and plotted using the GSEAPy package<sup>99</sup>. The entirety of the proteins in the dataset used as the ‘background’ dataset for statistical comparison. Pathway groups with an adjusted p-value <0.05 were considered significantly enriched over background.

**Cellular Compartment analysis.** For the cell compartment analysis done on the isoTOP-ABPP samples, the isoTOP data was first filtered based on proteins containing multiple identified cysteines across the aggregate dataset. The bulk multi-cysteine containing protein dataset was crossed with UniprotKB keyword annotations and the proteins were grouped based on each keyword annotations for cellular compartment. Then, for each compound, proteins were grouped as having more than 2 cysteines with high ratios or as having less than 2 cysteines with high ratios. The frequency at which proteins from each cellular compartment were identified from the whole dataset versus the dataset with more than 2 high ratio cysteines was calculated and, finally, a fold change ratio was calculated to assess enrichment of proteins from a particular subcellular compartment.

**Immunocytochemistry.** Sterilized coverslips were placed in each well of a 24 well plate and coated with poly-d-lysine for 30 minutes at 37°C. Cells were seeded on coverslips overnight, then treated with the indicated compounds. For nsp14 and nsp16 experiments, viral proteins were transfected into cells for at least 24 hours prior to compound treatment. For doxycycline-inducible U2OS cells, cells were cultured with media containing 1 µg/mL doxycycline for at least 24 hours prior to fixation for expression of G3BP1-TurboID-V5. Following treatment, media was aspirated and each well washed 1 time with 500 µL DPBS, followed by fixation with 3.7% formaldehyde in DPBS for 15 minutes at room temperature. Each well was then washed 2 times with 500 µL DPBS and permeabilized with 500 µL 0.1% TritonX in DPBS for 6 minutes at room temperature. Cells were then washed 2 times with 500 µL of DPBS and blocked in 1% FBS in DPBS for 30 minutes at room temperature. Cells were then incubated with 300 µL primary antibodies (**Supplementary Table 6**) in 0.1% FBS in PBS overnight at 4°C. Cells were then washed 2 times with 500 µL 0.1% Tween20 in PBS and 1 time with 500 µL PBS. Cells were then incubated with 300 µL secondary antibodies (**Supplementary Table 6**) for 1 hour at room temperature. Cells were then incubated with 300 µL 1:400 phalloidin-488 stain (Invitrogen, A12379) in PBS for 30 minutes at room temperature, followed by staining with 500 µL 1 µg/mL DAPI stain in PBS for 5 minutes. Cells were then washed 2 times with 500 µL PBS and mounted onto glass slides using Aqua-Poly/Mount mounting media (Polysciences, Inc., 18606-20). Samples were left to sit in the dark for 24 hours. Slides were then imaged on a Zeiss LSM880 confocal microscope at 63X oil objective with 2X manual zoom. Quantification of fluorescence intensity was done on ImageJ<sup>97</sup> software using particle picker. Unpaired student's t-test was performed to generate p-value between control and treatment groups. Colocalization analyses were done on ImageJ using the Coloc2 plugin.

**Glutathione assay.** Cells were grown to 80% confluence and treated with either DMSO (vehicle), 100 µM JC19 for 1 hour, or 500 µM Buthionine sulfoximine (BSO) for 24 hours in duplicate. Total

glutathione content in cells was measured using Calorimetric Glutathione Assay (G Biosciences, 786-075) according to the manufacturer's Cell Lysate Preparation protocol using the end-point method. Only deviation from protocol is that cell lysates were prepared at 4X the concentration suggested in the protocol.

**IncuCyte Live Cell Imaging.** HeLa cells were seeded in a 96-well plate at a density of 100,000 cells/well and allowed to adhere for 6 hours in a tissue culture incubator at 37° C with 5% CO<sub>2</sub>. 30 mins before first live cell time point, addition of media containing Cytotox Green reagent (Cat. No. 4633) (final dilution 1:4000) and compound was performed. An IncuCyte SX5 Live-Cell Imaging and Analysis System (Sartorius) was used for data acquisition. Over the course of 13 hours, at each 1 hour time point, four images per well were taken using both the 10X phase contrast and green lenses with the adherent cell-by-cell option selected. Analysis was performed using the IncuCyte 2019B Rev2 software. In brief, the green integrated intensity per well was normalized to 0d0h0m per condition and each condition was plated in triplicate.

**XRNAX.** Stable-Isotope Labeling by Amino Acids in Cell Culture (SILAC) HEK293T cells were used for XRNAX experiments. 'Light' labeled cells were treated with DMSO for 1 hour in duplicate and 'heavy' labeled cells were treated with 100 µM JC19 for 1 hour in duplicate. Cells were then exposed to 254 nm light at 200 mJ/cm<sup>2</sup> on ice to initiate crosslinking of RNA to RNA-binding proteins. Cells were then harvested and cells counted on a cell counter. Equal number of heavy and light labeled cells (~6.0x10<sup>6</sup> cells) were then combined and collected by centrifugation at 1800 g for 5 minutes. Cells were lysed in 1 mL TRIzol (Invitrogen, 15596026) until homogenous and let to sit for 10 minutes at room temp. 200 µL of chloroform was added to each sample and samples shook vigorously to mix, then let to incubate at room temperature for 5 minutes. Samples were centrifuged at 7000 g for 10 minutes at 4°C. Samples separate into three phases; the top layer was discarded and the middle interphase layer (light pink/white jelly-like) transferred to a new

tube. Wash interphase layer 2x with 1 mL TE buffer + 0.1% SDS. Disintegrate pellet with 1 mL TE buffer + 0.1% SDS and spin down at 5000 *g* for 5 minutes and transfer supernatant to new tube. Repeated this disintegration step 3 additional times to get 4 fractions for each sample. To each fraction, added 60  $\mu$ L 5M NaCl, 1  $\mu$ L glycoblu (Fisher, AM9515), and 1 mL isopropanol, mix, and incubate on ice for 15 minutes to encourage precipitation. Spun each tube at 18,000 *g* and -9°C for 15 minutes and removed supernatant from all fractions. Used 1 mL 80% ethanol to combine pellets from the 4 fractions from each sample and spin down pellets at 18,000 *g* for 1 minute. Removed ethanol supernatant and added 900  $\mu$ L RNase-free water to reconstitute pellets. Incubated each sample at 65°C for 5 minutes and let cool to room temperature. Added 100  $\mu$ L 10X DNase buffer (100 mM Tris-HCl, 1 mM CaCl<sub>2</sub>, 25 mM MgCl<sub>2</sub>, pH 7.5) to each sample followed by addition of 1  $\mu$ L DNase to each sample and incubated at 37°C for 1.5 hours with shaking (600 rpm). Transferred 1 mL samples to 2 mL microcentrifuge tubes and added 60  $\mu$ L 5M NaCl and 1 mL isopropanol to each sample, then incubated samples on ice for 30 minutes. Spun down samples at 18,000 *g* for 15 minutes at -5°C, then removed supernatant and resuspended pellet in 1 mL 80% ethanol and transferred back to microcentrifuge tubes. Spun down samples at 21.1k *g* for 5 minutes to pellet RNA and removed ethanol supernatant carefully so as not to disturb pellets and let pellets dry for ~5-10 minutes. Resuspended pellets in 100  $\mu$ L PBS and let sit at room temp for about an hour to allow full resuspension. Samples were then cleaned using SP3 clean-up. 20  $\mu$ L Sera-Mag SpeedBeads Carboxyl Magnetic Beads per sample, hydrophobic (GE Healthcare, 65152105050250, 50  $\mu$ g/ $\mu$ L, total 1 mg) and 20  $\mu$ L Sera-Mag SpeedBeads Carboxyl Magnetic Beads per sample, hydrophilic (GE Healthcare, 45152105050250, 50  $\mu$ g/ $\mu$ L, total 1 mg) were mixed and washed with water three times and resuspended in 40  $\mu$ L water per sample. 40  $\mu$ L of the bead slurry was then transferred to the lysate samples (500ug protein input) and incubated for 5 minutes at room temperature with shaking (1000 rpm). 500  $\mu$ L of absolute ethanol was added to each sample and incubated for 10 minutes at room temperature with shaking (1000 rpm). The beads were then washed (3x, 400uL

80% ethanol) with a magnetic rack. Beads were then resuspended in 200  $\mu$ L 2M urea in 0.5% SDS/PBS. DTT was added to each sample to a final concentration of 10 mM to each sample and incubated at 65°C for 15 minutes. Iodoacetamide was added to each sample to a final concentration of 20 mM and incubated at 37°C for 30 minutes. Absolute ethanol (400  $\mu$ L) was then added to each sample, and the samples were incubated for 5 minutes at room temperature with shaking (1000 rpm). Beads were then washed (3x, 400  $\mu$ L 80% ethanol) with a magnetic rack. Next, beads were resuspended in 200  $\mu$ L 2M urea in PBS and 3  $\mu$ L of 1 mg/mL trypsin solution was added. Proteins were digested for 12-16 hours at 37°C with shaking (200 rpm). After digestion, 3.8 mL acetonitrile (> 95% of the final volume) was added to each sample and the mixtures were incubated for 10 minutes at room temperature with shaking (1000 rpm). The beads were then washed (3x, 1mL acetonitrile) with a magnetic rack. Peptides were eluted from SP3 beads with 100  $\mu$ L of 2% DMSO in molecular biology-grade water for 30 min at 37°C with shaking (1000 rpm). The elution was repeated again with 100  $\mu$ L of 2% DMSO in molecular biology-grade water and the elution fractions from the same sample combined. The samples were dried (SpeedVac), then reconstituted with 5% acetonitrile and 1% formic acid in molecular biology-grade water and analyzed by LC-MS/MS.

**Liquid-chromatography tandem mass-spectrometry (LC-MS/MS) acquisition.** The samples were analyzed by liquid chromatography tandem mass spectrometry using a Thermo Scientific™ Orbitrap Eclipse™ Tribrid™ mass spectrometer. Peptides were fractionated online using a 18 cm long, 100  $\mu$ m inner diameter (ID) fused silica capillary packed in-house with bulk C18 reversed phase resin (particle size, 1.9  $\mu$ m; pore size, 100 Å; Dr. Maisch GmbH). The 70-minute water/acetonitrile gradient was delivered using a Thermo Scientific™ EASY-nLC™ 1200 system at different flow rates (Buffer A: water with 3% DMSO and 0.1% formic acid and Buffer B: 80% acetonitrile with 3% DMSO and 0.1% formic acid). The detailed gradient for each experiment is outlined in **Supplementary Table 10**. Data was collected with charge exclusion (1, 8,>8). Data

was acquired using a Data-Dependent Acquisition (DDA) method consisting of a full MS1 scan (Resolution = 120,000) followed by sequential MS2 scans (Resolution = 15,000) to utilize the remainder of the 1 second cycle time. Precursor isolation window was set as 1.6 and normalized collision energy was set as 30%. Some datasets were collected with a high-field asymmetric waveform ion mobility (FAIMS) device with current voltages of -35, -45, and -55. Files of MS data with corresponding experiments can be found in **Supplementary Table 11**.

**Data Compilation and Statistics.** Raw data collected by LC-MS/MS were searched with MSFragger and FragPipe (version 17-19)<sup>100,101</sup>. The proteomic workflow and its collection of tools was set as default. Precursor and fragment mass tolerance was set as 20 ppm. Missed cleavages were allowed up to 1. Peptide length was set 7 - 50 and peptide mass range was set 500 - 5000. For SILAC based experiments, MS1 labeling quant was enabled with Light set as K;R+0 and Heavy set as K+8.0142 and R+10.0083. MS1 intensity ratio of heavy and light labeled cysteine peptides and compiled for individual proteins were reported. For isoTOP-ABPP experiments, MS1 labeling quant was enabled with Light set as C+521.3074 and Heavy set as C+527.3213. MS1 intensity ratio of heavy and light labeled cysteine peptides were reported. For Glygly searches, a variable modification of +114.0429 was included on lysine. Calibrated and deisotoped spectrum files produced by FragPipe were retained and reused for this analysis. For protein-level analyses quantified using label-free quantification, identified proteins were filtered by Perseus<sup>102</sup> to retain only proteins identified in at least 2 replicates (for experiments with 3 replicates/condition) or 3 replicates (for experiments with 6 replicates/condition) of at least 1 experimental condition per experiment. Missing values were then imputed by Perseus based on the normal distribution. For experiments utilizing MS1 quantitation (e.g. isoTOP-ABPP and SILAC), custom python scripts were implemented to compile labeled peptide and protein datasets. Unique proteins, unique cysteines, and unique peptides were quantified for each dataset. Unique proteins were established based on UniProtKB protein IDs. Unique peptides were found based on sequences

containing a modified cysteine residue. Unique cysteines were classified by an identifier consisting of a UniProtKB protein ID and the amino acid number of the modified cysteine (ProteinID\_C#); residue numbers were found by aligning the peptide sequence to the corresponding UniProtKB protein sequence. When there are multiple cysteines in one peptide, all the cysteine residue numbers will be reported as ProteinID\_C#\_C# though it is undistinguishable on which cysteine the modification is. Statistical values including the exact n, statistical test, and significance are reported in the Figure Legends. Statistical significance was defined as p-value < 0.05 and unless indicated otherwise determined by an unpaired 2-tailed Student's t-test.

**Statistics.** Statistical analysis was performed using GraphPad Prism (v9.4.1) for Mac (GraphPad Software). Statistical values including the exact n, statistical test, and significance are reported in the Figure Legends. Statistical significance was defined as p-value < 0.05 and unless indicated otherwise determined by an unpaired 2-tailed Student's t-test.

**Reporting Summary.** Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### **Data Availability**

All supporting data for this study can be found within the article and Supplementary Information documents. The MS data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE<sup>103</sup> partner repository with the dataset identifiers PXD046278 (isoTOP-ABPP data in Figures 1 and 5), PXD046393 (all other proteomic data corresponding to Figures 3-6), and PXD053865 (PDIA6-His AP-MS and Extended Data Fig. 4 proteomics). Publicly available databases used are the The UniProtKB Consortium (<https://www.uniprot.org/>); 2021 release of the Kyoto Encyclopedia of Genes and Genomes



PATHWAY database (<https://www.genome.jp/kegg/pathway.html>); the 2021 release of the Gene Ontology Resource (<https://geneontology.org/>) Molecular Function, Cellular Component, and Biological Process classes. Source data is provided as source data files for Figures and Extended Data Figures, and within the Supplementary Information file for Supplementary Figures.

### Code Availability

All code used for this work is available at <https://github.com/BackusLab>.

### References

1. Boatner, L. M., Palafox, M. F., Schweppe, D. K. & Backus, K. M. CysDB: a human cysteine database based on experimental quantitative chemoproteomics. *Cell Chem. Biol.* **30**, 683-698.e3 (2023).
2. Backus, K. M. et al. Proteome-wide covalent ligand discovery in native biological systems. *Nature* **534**, 570–574 (2016).
3. Weerapana, E. et al. Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **468**, 790–795 (2010).
4. Cao, J. et al. Multiplexed CuAAC Suzuki-Miyaura Labeling for Tandem Activity-Based Chemoproteomic Profiling. *Anal. Chem.* **93**, 2610–2618 (2021).
5. Vinogradova, E. V. et al. An Activity-Guided Map of Electrophile-Cysteine Interactions in Primary Human T Cells. *Cell* **182**, 1009-1026.e29 (2020).
6. Yan, T. et al. SP3-FAIMS Chemoproteomics for High-Coverage Profiling of the Human Cysteinome\*. *Chembiochem* **22**, 1841–1851 (2021).
7. Kuljanin, M. et al. Reimagining high-throughput profiling of reactive cysteines for cell-based screening of large electrophile libraries. *Nat. Biotechnol.* **39**, 630–641 (2021).
8. Burton, N. R. et al. Solid-Phase Compatible Silane-Based Cleavable Linker Enables Custom Isobaric Quantitative Chemoproteomics. *J. Am. Chem. Soc.* **145**, 21303–21318 (2023).
9. Mader, M. M. et al. Which small molecule? selecting chemical probes for use in cancer research and target validation. *Cancer Discov.* **13**, 2150–2165 (2023).

10. Antolin, A. A. et al. The Chemical Probes Portal: an expert review-based public resource to empower chemical probe assessment, selection and use. *Nucleic Acids Res.* **51**, D1492–D1502 (2023).
11. Hartung, I. V., Rudolph, J., Mader, M. M., Mulder, M. P. C. & Workman, P. Expanding chemical probe space: quality criteria for covalent and degrader probes. *J. Med. Chem.* **66**, 9297–9312 (2023).
12. Kavanagh, M. E. et al. Selective inhibitors of JAK1 targeting an isoform-restricted allosteric cysteine. *Nat. Chem. Biol.* **18**, 1388–1398 (2022).
13. Ostrem, J. M., Peters, U., Sos, M. L., Wells, J. A. & Shokat, K. M. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* **503**, 548–551 (2013).
14. Janes, M. R. et al. Targeting KRAS Mutant Cancers with a Covalent G12C-Specific Inhibitor. *Cell* **172**, 578–589.e17 (2018).
15. Dickson, P. et al. Physical and functional analysis of the putative rpn13 inhibitor RA190. *Cell Chem. Biol.* **27**, 1371–1382.e6 (2020).
16. Anchoori, R. K. et al. A bis-benzylidene piperidone targeting proteasome ubiquitin receptor RPN13/ADRM1 as a therapy for cancer. *Cancer Cell* **24**, 791–805 (2013).
17. Gamayun, I. et al. Eeyarestatin Compounds Selectively Enhance Sec61-Mediated Ca<sup>2+</sup> Leakage from the Endoplasmic Reticulum. *Cell Chem. Biol.* **26**, 571–583.e6 (2019).
18. Zhang, K. et al. Stress granule assembly disrupts nucleocytoplasmic transport. *Cell* **173**, 958–971.e17 (2018).
19. Othumpangat, S., Kashon, M. & Joseph, P. Sodium arsenite-induced inhibition of eukaryotic translation initiation factor 4E (eIF4E) results in cytotoxicity and cell death. *Mol. Cell. Biochem.* **279**, 123–131 (2005).
20. Kopito, R. R. Aggresomes, inclusion bodies and protein aggregation. *Trends Cell Biol.* **10**, 524–530 (2000).
21. García-Mata, R., Bebök, Z., Sorscher, E. J. & Sztul, E. S. Characterization and dynamics of aggresome formation by a cytosolic GFP-chimera. *J. Cell Biol.* **146**, 1239–1254 (1999).
22. Tao, Y. et al. Targeted Protein Degradation by Electrophilic PROTACs that Stereoselectively and Site-Specifically Engage DCAF1. *J. Am. Chem. Soc.* **144**, 18688–18699 (2022).
23. Forte, N. et al. Targeted Protein Degradation through E2 Recruitment. *ACS Chem. Biol.* **18**, 897–904 (2023).
24. Zhang, X., Crowley, V. M., Wucherpennig, T. G., Dix, M. M. & Cravatt, B. F. Electrophilic PROTACs that degrade nuclear proteins by engaging DCAF16. *Nat. Chem. Biol.* **15**, 737–746 (2019).

25. Zhang, X. et al. DCAF11 Supports Targeted Protein Degradation by Electrophilic Proteolysis-Targeting Chimeras. *J. Am. Chem. Soc.* **143**, 5141–5149 (2021).
26. Spradlin, J. N. et al. Harnessing the anti-cancer natural product nimbolide for targeted protein degradation. *Nat. Chem. Biol.* **15**, 747–755 (2019).
27. Ward, C. C. et al. Covalent ligand screening uncovers a RNF4 E3 ligase recruiter for targeted protein degradation applications. *ACS Chem. Biol.* **14**, 2430–2440 (2019).
28. Henning, N. J. et al. Discovery of a covalent FEM1B recruiter for targeted protein degradation applications. *J. Am. Chem. Soc.* **144**, 701–708 (2022).
29. Isobe, Y. et al. Manumycin polyketides act as molecular glues between UBR7 and P53. *Nat. Chem. Biol.* **16**, 1189–1198 (2020).
30. Toriki, E. S. et al. Rational chemical design of molecular glue degraders. *ACS Cent. Sci.* **9**, 915–926 (2023).
31. Sarott, R. C. et al. Chemical Specification of E3 Ubiquitin Ligase Engagement by Cysteine-Reactive Chemistry. *J. Am. Chem. Soc.* **145**, 21937–21944 (2023).
32. Li, Y.-D. et al. Template-assisted covalent modification underlies activity of covalent molecular glues. *Nat. Chem. Biol.* (2024) doi:10.1038/s41589-024-01668-4.
33. Hassan, M. M. et al. Exploration of the Tunability of BRD4 Degradation by DCAF16 Trans-labelling Covalent Glues. *BioRxiv* (2023) doi:10.1101/2023.10.07.561308.
34. King, E. A. et al. Chemoproteomics-enabled discovery of a covalent molecular glue degrader targeting NF- $\kappa$ B. *Cell Chem. Biol.* **30**, 394-402.e9 (2023).
35. Hong, S. H., Osa, A., Wertz, I. E. & Nomura, D. K. Exploiting the cullin E3 ligase adaptor protein SKP1 for targeted protein degradation. *BioRxiv* (2023) doi:10.1101/2023.10.20.563371.
36. Littler, D. R., Gully, B. S., Colson, R. N. & Rossjohn, J. Crystal Structure of the SARS-CoV-2 Non-structural Protein 9, Nsp9. *iScience* **23**, 101258 (2020).
37. Ma, Y. et al. Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. *Proc Natl Acad Sci USA* **112**, 9436–9441 (2015).
38. Decroly, E. et al. Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2'-O-methyltransferase nsp10/nsp16 complex. *PLoS Pathog.* **7**, e1002059 (2011).
39. Imprachim, N., Yosaatmadja, Y. & Newman, J. A. Crystal structures and fragment screening of SARS-CoV-2 NSP14 reveal details of exoribonuclease activation and mRNA capping and provide starting points for antiviral drug development. *Nucleic Acids Res.* **51**, 475–487 (2023).

40. Liby, K. et al. The synthetic triterpenoids, CDDO and CDDO-imidazolide, are potent inducers of heme oxygenase-1 and Nrf2/ARE signaling. *Cancer Res.* **65**, 4789–4798 (2005).
41. Yang, D. et al. Systematic Targeting of Protein Complexes with Molecular COUPLRs. *BioRxiv* (2024) doi:10.1101/2024.07.16.603666.
42. Li, Y.-D. et al. Template-assisted covalent modification of DCAF16 underlies activity of BRD4 molecular glue degraders. *BioRxiv* (2023) doi:10.1101/2023.02.14.528208.
43. Mauvezin, C. & Neufeld, T. P. Bafilomycin A1 disrupts autophagic flux by inhibiting both V-ATPase-dependent acidification and Ca-P60A/SERCA-dependent autophagosome-lysosome fusion. *Autophagy* **11**, 1437–1438 (2015).
44. Ohtake, F., Tsuchiya, H., Saeki, Y. & Tanaka, K. K63 ubiquitylation triggers proteasomal degradation by seeding branched ubiquitin chains. *Proc Natl Acad Sci USA* **115**, E1401–E1408 (2018).
45. Ohtake, F., Saeki, Y., Ishido, S., Kanno, J. & Tanaka, K. The K48-K63 Branched Ubiquitin Chain Regulates NF- $\kappa$ B Signaling. *Mol. Cell* **64**, 251–266 (2016).
46. Gordon, D. E. et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020).
47. Grantham, J. The molecular chaperone cct/tric: an essential component of proteostasis and a potential modulator of protein aggregation. *Front. Genet.* **11**, 172 (2020).
48. Bashore, C. et al. Targeted degradation via direct 26S proteasome recruitment. *Nat. Chem. Biol.* **19**, 55–63 (2023).
49. Tofaris, G. K., Layfield, R. & Spillantini, M. G. alpha-synuclein metabolism and aggregation is linked to ubiquitin-independent degradation by the proteasome. *FEBS Lett.* **509**, 22–26 (2001).
50. Zhou, M. et al. HUWE1 amplifies ubiquitin modifications to broadly stimulate clearance of proteins and aggregates. *BioRxiv* (2023) doi:10.1101/2023.05.30.542866.
51. Sarkar, A. A. & Zohn, I. E. Hectd1 regulates intracellular localization and secretion of Hsp90 to control cellular behavior of the cranial mesenchyme. *J. Cell Biol.* **196**, 789–800 (2012).
52. Wang, Q., Li, L. & Ye, Y. Inhibition of p97-dependent protein degradation by Eeyarestatin I. *J. Biol. Chem.* **283**, 7445–7454 (2008).
53. Berkers, C. R. et al. Profiling proteasome activity in tissue with fluorescent probes. *Mol. Pharm.* **4**, 739–748 (2007).
54. Leestemaker, Y. et al. Proteasome activation by small molecules. *Cell Chem. Biol.* **24**, 725-736.e7 (2017).

55. Vetma, V. et al. Confounding Factors in Targeted Degradation of Short-Lived Proteins. *ACS Chem. Biol.* (2024) doi:10.1021/acscchembio.4c00152.
56. Dixon, S. J. et al. Ferroptosis: an iron-dependent form of nonapoptotic cell death. *Cell* **149**, 1060–1072 (2012).
57. Higgins, L. G. et al. Transcription factor Nrf2 mediates an adaptive response to sulforaphane that protects fibroblasts in vitro against the cytotoxic effects of electrophiles, peroxides and redox-cycling agents. *Toxicol. Appl. Pharmacol.* **237**, 267–280 (2009).
58. Hetz, C. The unfolded protein response: controlling cell fate decisions under ER stress and beyond. *Nat. Rev. Mol. Cell Biol.* **13**, 89–102 (2012).
59. Edman, J. C., Ellis, L., Blacher, R. W., Roth, R. A. & Rutter, W. J. Sequence of protein disulphide isomerase and implications of its relationship to thioredoxin. *Nature* **317**, 267–270 (1985).
60. Plate, L. et al. Small molecule proteostasis regulators that reprogram the ER to reduce extracellular protein aggregation. *eLife* **5**, (2016).
61. Hoffstrom, B. G. et al. Inhibitors of protein disulfide isomerase suppress apoptosis induced by misfolded proteins. *Nat. Chem. Biol.* **6**, 900–906 (2010).
62. Pei, Y. et al. RAB21 controls autophagy and cellular energy homeostasis by regulating retromer-mediated recycling of SLC2A1/GLUT1. *Autophagy* **19**, 1070–1086 (2023).
63. Jean, S., Cox, S., Nassari, S. & Kiger, A. A. Starvation-induced MTMR13 and RAB21 activity regulates VAMP8 to promote autophagosome-lysosome fusion. *EMBO Rep.* **16**, 297–311 (2015).
64. Liu, Y. et al. A Novel Mechanism For NF- $\kappa$ B Activation Via I $\kappa$ B Aggregation: Implications For Hepatic Mallory-Denk-Body Induced Inflammation. *Mol. Cell. Proteomics* (2020) doi:10.1074/mcp.RA120.002316.
65. Yang, P. et al. G3BP1 Is a Tunable Switch that Triggers Phase Separation to Assemble Stress Granules. *Cell* **181**, 325-345.e28 (2020).
66. Weiss-Sadan, T. et al. NRF2 activation induces NADH-reductive stress, providing a metabolic vulnerability in lung cancer. *Cell Metab.* **35**, 487-503.e7 (2023).
67. Liu, Y. et al. Sulforaphane enhances proteasomal and autophagic activities in mice and is a potential therapeutic reagent for Huntington's disease. *J. Neurochem.* **129**, 539–547 (2014).
68. Gromer, S., Arscott, L. D., Williams, C. H., Schirmer, R. H. & Becker, K. Human placenta thioredoxin reductase. Isolation of the selenoenzyme, steady state kinetics, and inhibition by therapeutic gold compounds. *J. Biol. Chem.* **273**, 20096–20101 (1998).

69. Zhang, J. et al. Systematic identification of anticancer drug targets reveals a nucleus-to-mitochondria ROS-sensing pathway. *Cell* **186**, 2361-2379.e25 (2023).
70. Mao, X. et al. Novel multi-targeted ErbB family inhibitor afatinib blocks EGF-induced signaling and induces apoptosis in neuroblastoma. *Oncotarget* **8**, 1555–1568 (2017).
71. Huang, F. et al. Repurposing of Ibrutinib and Quizartinib as potent inhibitors of necroptosis. *Commun. Biol.* **6**, 972 (2023).
72. Kapoor, I. et al. Resistance to BTK inhibition by ibrutinib can be overcome by preventing FOXO3a nuclear export and PI3K/AKT activation in B-cell lymphoid malignancies. *Cell Death Dis.* **10**, 924 (2019).
73. Lanning, B. R. et al. A road map to evaluate the proteome-wide selectivity of covalent kinase inhibitors. *Nat. Chem. Biol.* **10**, 760–767 (2014).
74. Seo, K. W. & Kleiner, R. E. Profiling dynamic RNA-protein interactions using small-molecule-induced RNA editing. *Nat. Chem. Biol.* (2023) doi:10.1038/s41589-023-01372-9.
75. Trendel, J. et al. The Human RNA-Binding Proteome and Its Dynamics during Translational Arrest. *Cell* **176**, 391-403.e19 (2019).
76. Wang, X. et al. Selective depletion of mutant p53 by cancer chemopreventive isothiocyanates and their structure-activity relationships. *J. Med. Chem.* **54**, 809–816 (2011).
77. Yasuda, S. et al. Stress- and ubiquitylation-dependent phase separation of the proteasome. *Nature* **578**, 296–300 (2020).
78. Gwon, Y. et al. Ubiquitination of G3BP1 mediates stress granule disassembly in a context-specific manner. *Science* **372**, eabf6548 (2021).
79. Johnston, J. A., Ward, C. L. & Kopito, R. R. Aggresomes: a cellular response to misfolded proteins. *J. Cell Biol.* **143**, 1883–1898 (1998).
80. Zhang, C. et al. Autophagic sequestration of SQSTM1 disrupts the aggresome formation of ubiquitinated proteins during proteasome inhibition. *Cell Death Dis.* **13**, 615 (2022).
81. Yu, C. et al. Afatinib combined with anti-PD1 enhances immunotherapy of hepatocellular carcinoma via ERBB2/STAT3/PD-L1 signaling. *Front. Oncol.* **13**, 1198118 (2023).
82. Li, D. et al. BIBW2992, an irreversible EGFR/HER2 inhibitor highly effective in preclinical lung cancer models. *Oncogene* **27**, 4702–4711 (2008).
83. Bar-Peled, L. et al. Chemical proteomics identifies druggable vulnerabilities in a genetically defined cancer. *Cell* **171**, 696-709.e23 (2017).
84. Abegg, D. et al. Proteome-Wide Profiling of Targets of Cysteine reactive Small Molecules by Using Ethynyl Benziiodoxolone Reagents. *Angew. Chem.* **127**, 11002–11007 (2015).

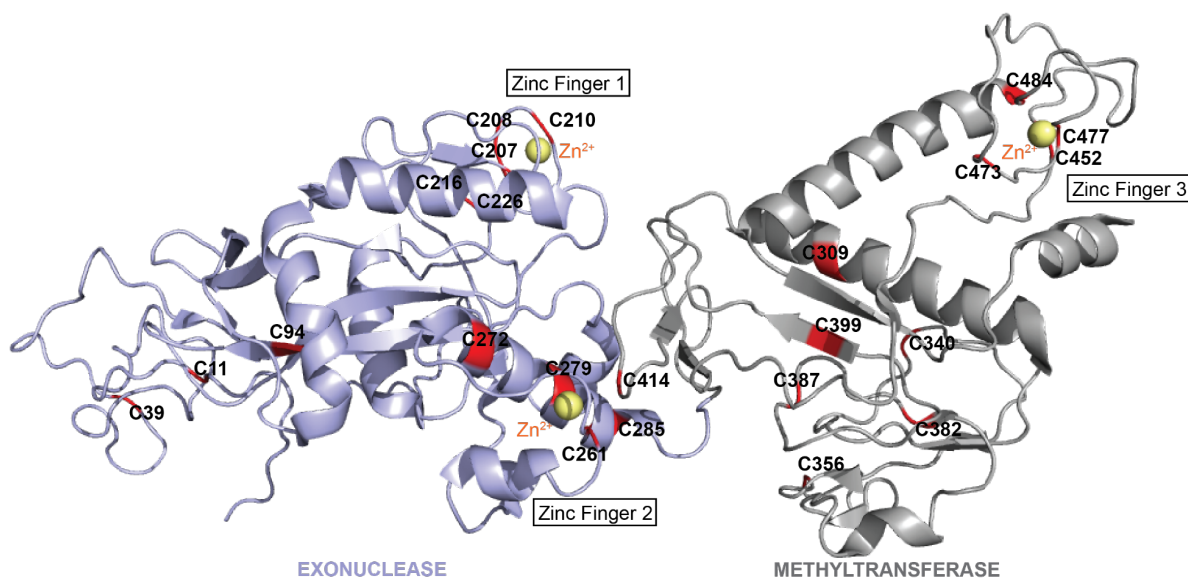
85. Grossman, E. A. et al. Covalent Ligand Discovery against Druggable Hotspots Targeted by Anti-cancer Natural Products. *Cell Chem. Biol.* **24**, 1368-1376.e4 (2017).
86. Wang, C., Weerapana, E., Blewett, M. M. & Cravatt, B. F. A chemoproteomic platform to quantitatively map targets of lipid-derived electrophiles. *Nat. Methods* **11**, 79–85 (2014).
87. Henning, N. J. et al. Deubiquitinase-targeting chimeras for targeted protein stabilization. *Nat. Chem. Biol.* **18**, 412–421 (2022).
88. Lazear, M. R. et al. Proteomic discovery of chemical probes that perturb protein complexes in human cells. *Mol. Cell* **83**, 1725-1742.e12 (2023).
89. Kathman, S. G. et al. Remodeling oncogenic transcriptomes by small molecules targeting NONO. *Nat. Chem. Biol.* **19**, 825–836 (2023).
90. Sterling, J., Baker, J. R., McCluskey, A. & Munoz, L. Systematic literature review reveals suboptimal use of chemical probes in cell-based biomedical research. *Nat. Commun.* **14**, 3228 (2023).
91. Li, H. et al. Assigning functionality to cysteines by base editing of cancer dependency genes. *Nat. Chem. Biol.* **19**, 1320–1330 (2023).
92. Chen, Y. et al. Direct mapping of ligandable tyrosines and lysines in cells with chiral sulfonyl fluoride probes. *Nat. Chem.* **15**, 1616–1625 (2023).
93. Resnick, E. et al. Rapid Covalent-Probe Discovery by Electrophile-Fragment Screening. *J. Am. Chem. Soc.* **141**, 8951–8968 (2019).
94. Dubiella, C. et al. Sulfopin is a covalent inhibitor of Pin1 that blocks Myc-driven tumors in vivo. *Nat. Chem. Biol.* **17**, 954–963 (2021).
95. Branon, T. C. et al. Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **36**, 880–887 (2018).
96. Liu, H. & Naismith, J. H. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.* **8**, 91 (2008).
97. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
98. Chen, E. Y. et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).
99. Xie, Z. et al. Gene Set Knowledge Discovery with Enrichr. *Curr. Protoc.* **1**, e90 (2021).

100. Teo, G. C., Polasky, D. A., Yu, F. & Nesvizhskii, A. I. Fast deisotoping algorithm and its implementation in the msfragger search engine. *J. Proteome Res.* 20, 498–505 (2021).
101. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D. & Nesvizhskii, A. I. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat. Methods* 14, 513–520 (2017).
102. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* 13, 731–740 (2016).
103. Perez-Riverol, Y. et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* 50, D543–D552 (2022).

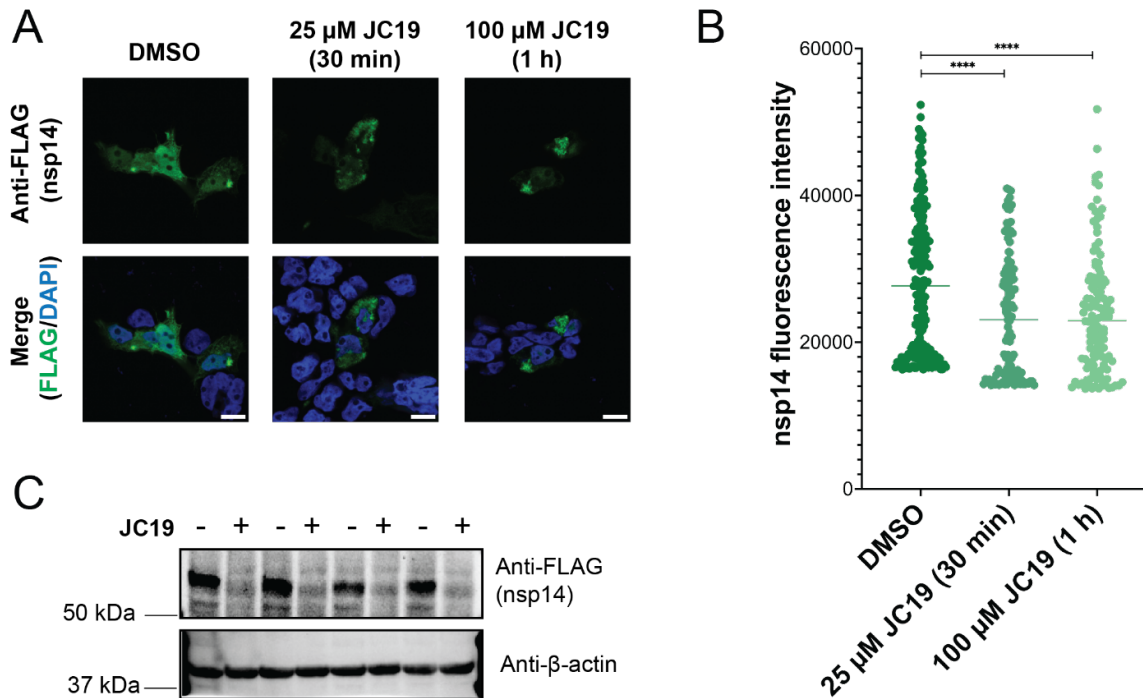


## Chapter 3

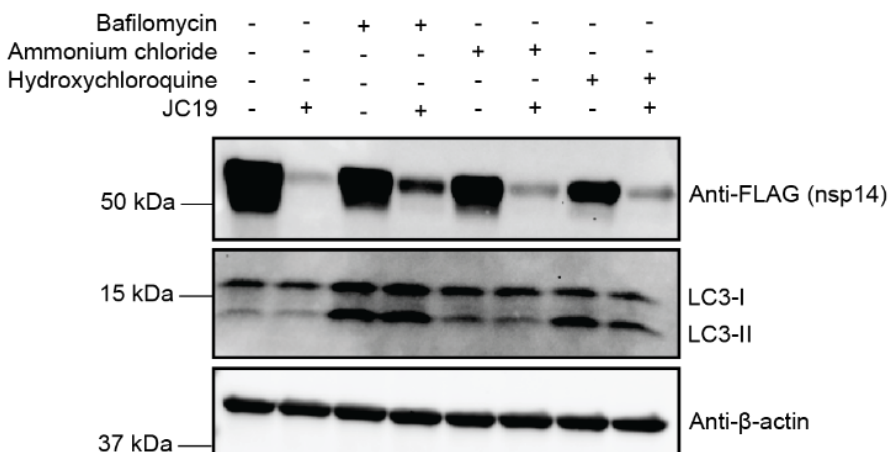
### Supporting Information



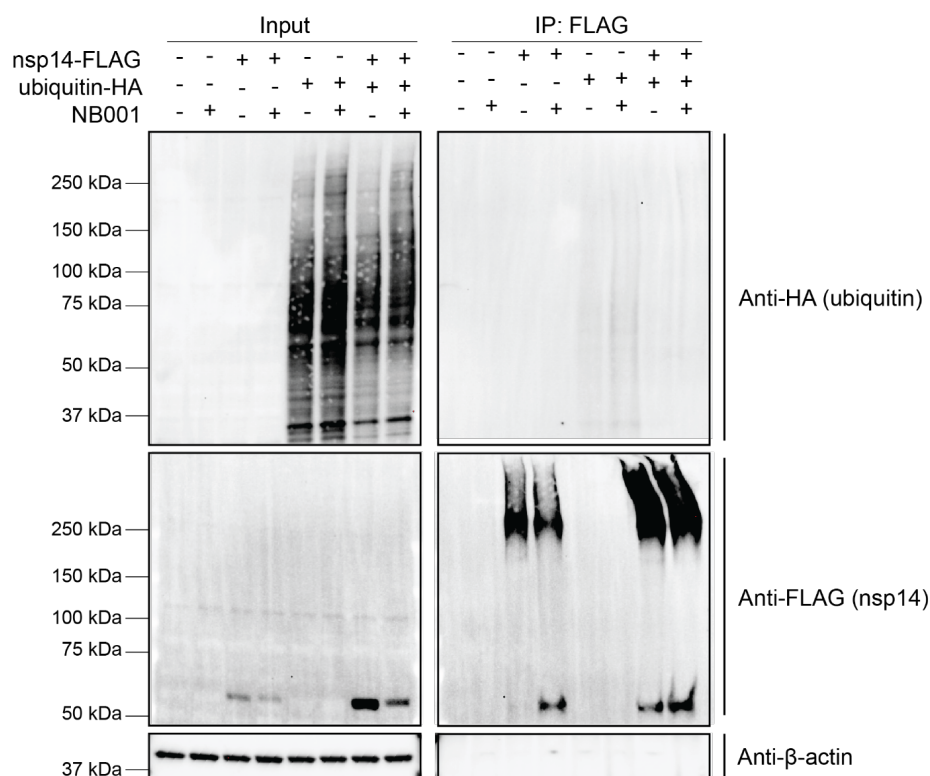
**Supplementary Figure 1. Structure of SARS-CoV-2 nonstructural protein 14 (nsp14) used in isoTOP-ABPP experiments.** Nsp14 (PDB: 7QGI)<sup>1</sup> with cysteine residues highlighted in red and labeled with residue number. The exonuclease domain is colored light purple and the methyltransferase domain is colored gray. Zinc ions of the zinc fingers are colored yellow with an orange Zn label. Protein structure generated using The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.



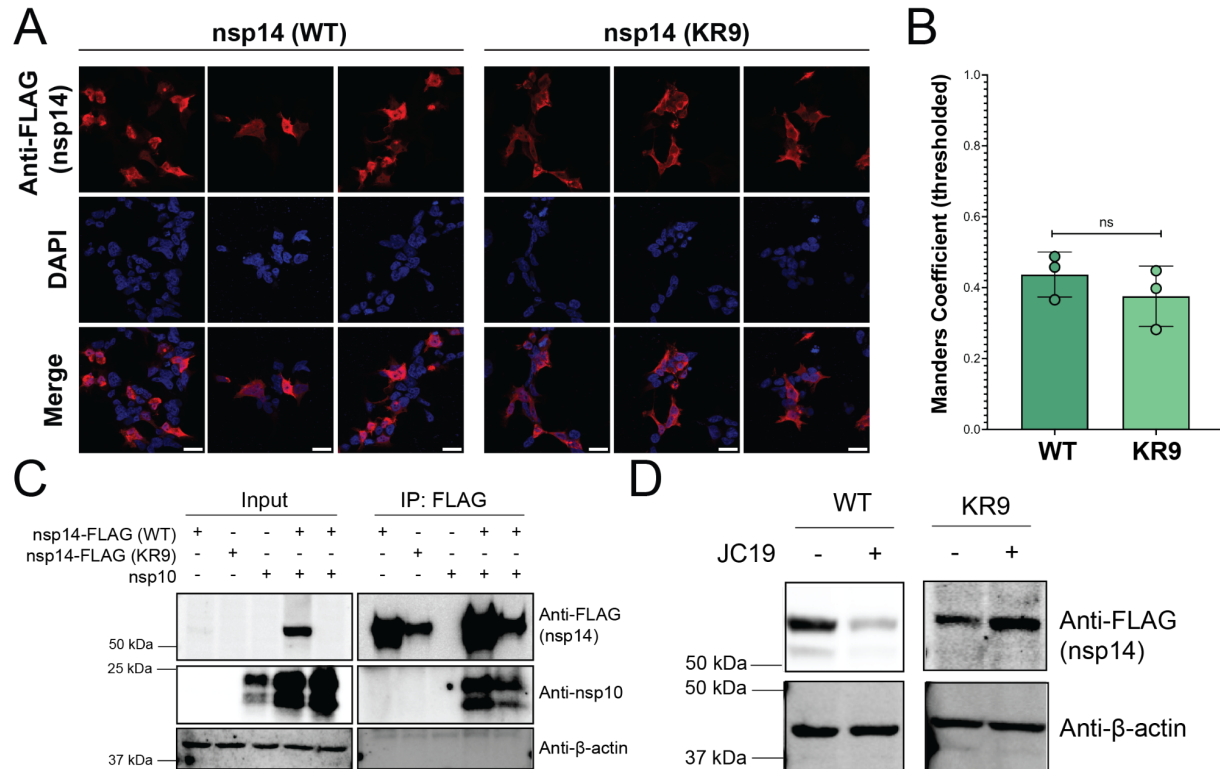
**Supplementary Figure 2. Nsp14 is depleted in HEK293T and HeLa cells.** (A) HEK293T cells transiently expressing nsp14-FLAG were treated with vehicle control (DMSO) or the indicated concentrations of JC19 for the indicated times. Cells were then fixed, permeabilized, and stained for immunofluorescence imaging. Cells were imaged on an LSM880 confocal microscope at 63X objective with 2X manual zoom. Scale bar = 10  $\mu$ m. (B) Quantification of nsp14-FLAG fluorescence intensity for each cell (ImageJ<sup>2</sup>) taken from 4 independent images prepared as described in (A) and acquired at 10X objective (DMSO:  $n = 200$ ; 25  $\mu$ M JC19:  $n = 115$ ; 100  $\mu$ M JC19:  $n = 125$ ). All data points are presented with median value denoted by line. Statistical significance calculated with unpaired Student's t-tests (\*\*\*\*p-value 25  $\mu$ M JC19 = 0.000013482, \*\*\*\* p-value 100  $\mu$ M JC19 = 0.000043023). (C) HeLa cells transiently expressing nsp14-FLAG were treated with either vehicle control (DMSO) or 100  $\mu$ M JC19 for 1 h ( $n = 4$ ) and immunoblot analysis was used to visualize abundance of nsp14 in the soluble lysate fraction. Immunoblot is representative of 2 independent measurements.



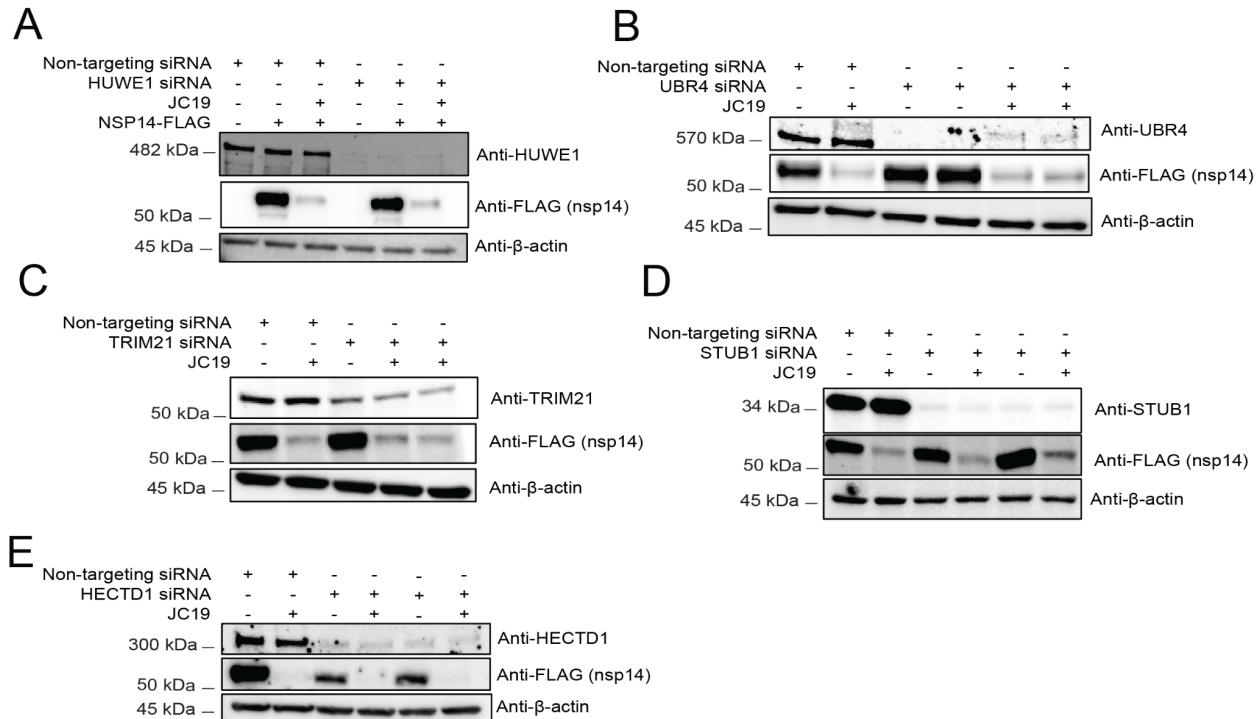
**Supplementary Figure 3. Autophagy and lysosomal inhibitors impart partial protection from nsp14 depletion.** HEK293T cells transiently expressing nsp14-FLAG were pretreated with either DMSO, bafilomycin (1  $\mu$ M, 6 h), ammonium chloride (10 mM, 6 h), or hydroxychloroquine (100  $\mu$ M, 6 h), followed by treatment with either DMSO or 50  $\mu$ M JC19 for 0.5 h. Immunoblot analysis was used to visualize abundance of nsp14 in the soluble lysate. LC3-II used as a marker of autophagy inhibition. Experiment is representative of 2 independent experiments.



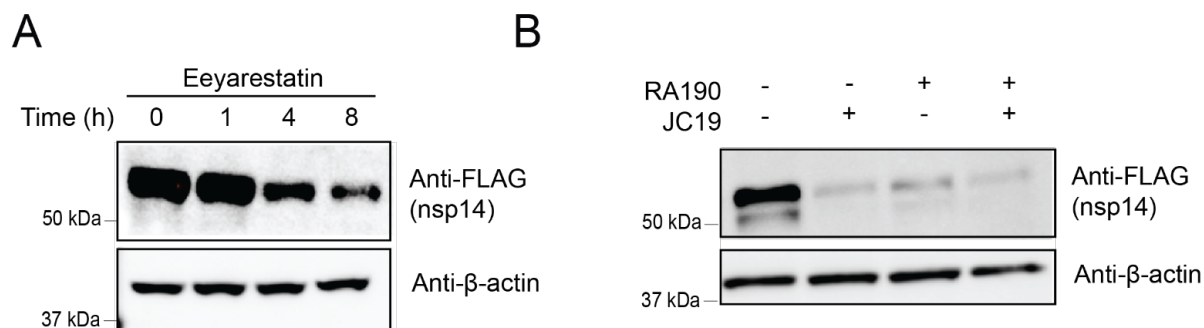
**Supplementary Figure 4. Nsp14 is not ubiquitinated by NB001.** HEK293T cells transiently expressing nsp14-FLAG, ubiquitin-HA, neither, or both were treated with either DMSO or 100  $\mu$ M NB001 for 1 h, immunoprecipitated on FLAG resin, and ran on immunoblot to assay ubiquitination of nsp14. Experiment is representative of one independent experiment.



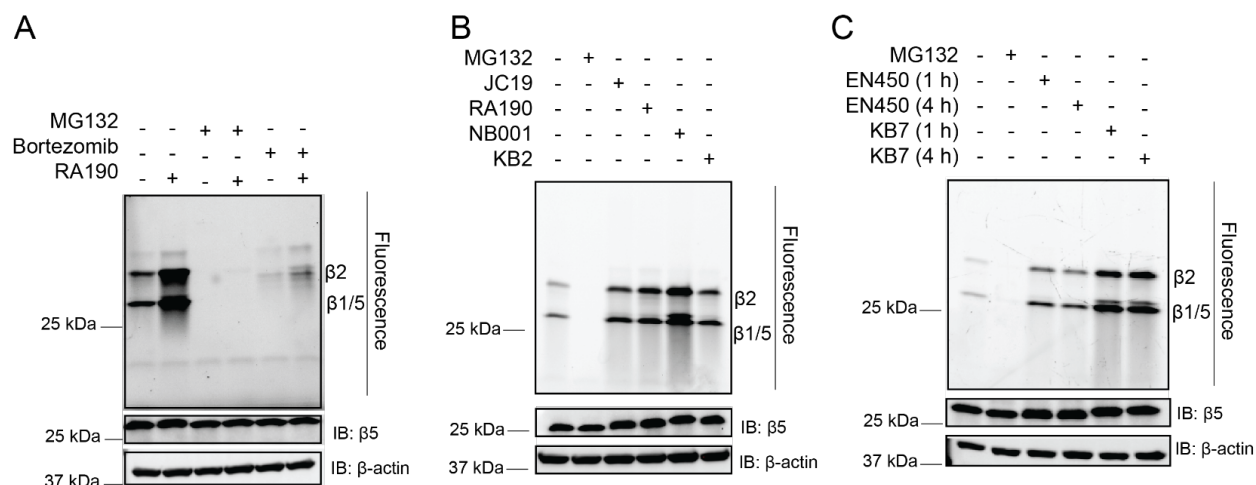
**Supplementary Figure 5. Nsp14 WT and KR9 have similar localization and interaction partners.** (A) Immunofluorescence microscopy of HEK293T cells transiently expressing both WT nsp14-FLAG and KR9 nsp14-FLAG (K13R, K34R, K288R, K318R, K339R, K423R, K433R, K440R, K469R). Mutated residues were selected based on those residues identified as ubiquitylated (GlyGly modified) in a JC19-dependent manner by LC-MS/MS (**Extended Data Fig. 3D**) as well as other solvent accessible lysines within nsp14 (PDB: 7QGI). Three representative images are shown. Scale bars represent 20  $\mu$ m. (B) Quantification of thresholded Manders coefficient (M2) for nsp14 and DAPI signal colocalization. Values represent the percentage of the nsp14 signal that overlaps with the DAPI signal. Colocalization analysis performed on ImageJ using the JACoP plugin. Data presented as mean  $\pm$  s.d. ( $n = 3$ ) and an unpaired t-test was used to calculate significance (ns,  $p$ -value  $> 0.05$ ). (C) Western blot depicting immunoprecipitation of nsp10 by nsp14 WT and nsp14 KR9. Cells expressing the indicated proteins were lysed and the soluble lysate fraction immunoprecipitated on FLAG resin. (D) Western blot analysis of nsp14 WT and nsp14 KR9 abundance after treatment with JC19 (25  $\mu$ M, 0.5 h). Panel c is representative of one independent experiment and panel d is representative of two independent experiments.



**Supplementary Figure 6. Knockdown of AP-MS identified E3 ubiquitin ligases does not protect nsp14 from JC19-mediated degradation.** (A) Immunoblot of siRNA knockdown of HUWE1 in HEK293T cells and the effect on nsp14-FLAG abundance in the presence and absence of JC19 treatment (25  $\mu$ M for 0.5 h). (B) Immunoblot of siRNA knockdown of UBR4 in HEK293T cells and the effect on nsp14-FLAG abundance in the presence and absence of JC19 treatment (25  $\mu$ M for 0.5 h). (C) Immunoblot of siRNA knockdown of TRIM1 in HEK293T cells and the effect on nsp14-FLAG abundance in the presence and absence of JC19 treatment (25  $\mu$ M for 0.5 h). (D) Immunoblot of siRNA knockdown of STUB1 in HEK293T cells and the effect on nsp14-FLAG abundance in the presence and absence of JC19 treatment (25  $\mu$ M for 0.5 h). (E) Immunoblot of siRNA knockdown of HECTD1 in HEK293T cells and the effect on nsp14-FLAG abundance in the presence and absence of JC19 treatment (100  $\mu$ M for 1 h). Detergent soluble fraction (0.3% CHAPS) was analyzed for all compounds. All western blots are representative of two independent measurements.

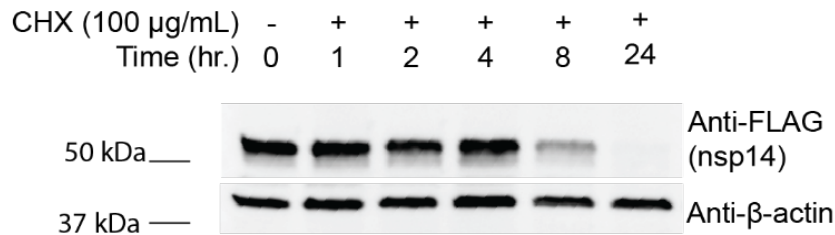


**Supplementary Figure 7. Both eeyarestatin and RA190 induce depletion of nsp14.** (A) HEK293T cells transiently expressing nsp14-FLAG were treated with 50  $\mu$ M eeyarestatin for the indicated time points and immunoblot used to assay nsp14 abundance. (B) HEK293T cells transiently expressing nsp14-FLAG were treated with either 2  $\mu$ M RA190 for 2 h, 50  $\mu$ M JC19 for 0.5 h, or pretreated with either DMSO or 2  $\mu$ M RA190 for 2 h and then treated with either DMSO or 50  $\mu$ M JC19 for 0.5 h. Immunoblot used to assay nsp14 abundance. Both panels are representative of two independent measurements.

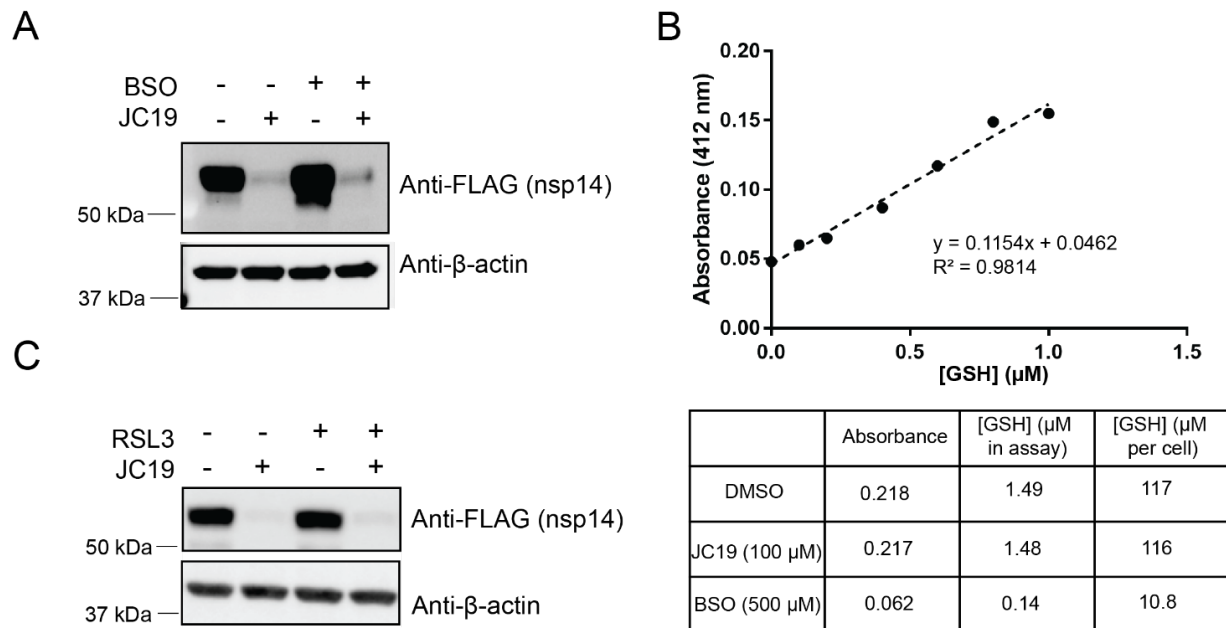


**Supplementary Figure 8. Cysteine-reactive small molecules activate the proteasome.** (A) HEK293T cells were pretreated with DMSO, MG132 (10  $\mu$ M for 5 h), or bortezomib (5  $\mu$ M for 5 h), followed by treatment with either DMSO or RA190 (5  $\mu$ M for 1 h). Cells were then treated with a fluorescent proteasome activity probe (500 nM for 1 h) and in-gel fluorescence used to assay proteasome activity. Immunoblot analysis was used to assay the abundance of PSMB5. (B) HEK293T cells were treated with DMSO, MG132 (10  $\mu$ M for 5 h), JC19 (100  $\mu$ M for 1 h), RA190 (5  $\mu$ M for 1 h), NB001 (100  $\mu$ M for 1 h), or KB2 (100  $\mu$ M for 1 h). Cells were then treated with a fluorescent proteasome activity probe (500 nM for 1 h) and in-gel fluorescence used to assay proteasome activity. Immunoblot analysis was used to assay the abundance of PSMB5. (C) HEK293T cells were treated with DMSO, MG132 (10  $\mu$ M for 5 h), EN450 (100  $\mu$ M for 1 and 4 h), or KB7 (100  $\mu$ M for 1 or 4 h). Cells were then treated with a fluorescent proteasome activity probe (500 nM for 1 h) and in-gel fluorescence used to assay proteasome activity. Immunoblot analysis was used to assay the abundance of PSMB5. Panels a and b are representative of two independent measurements. Panel c is representative of one independent measurement.

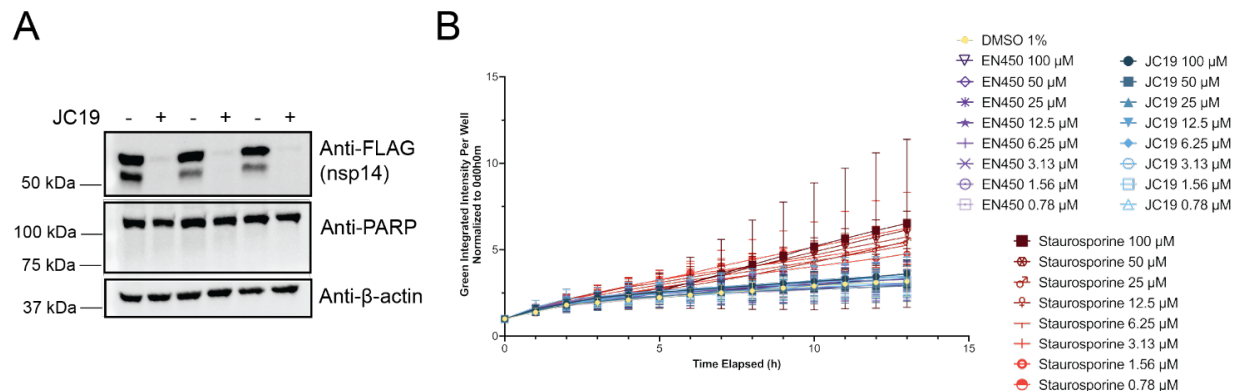




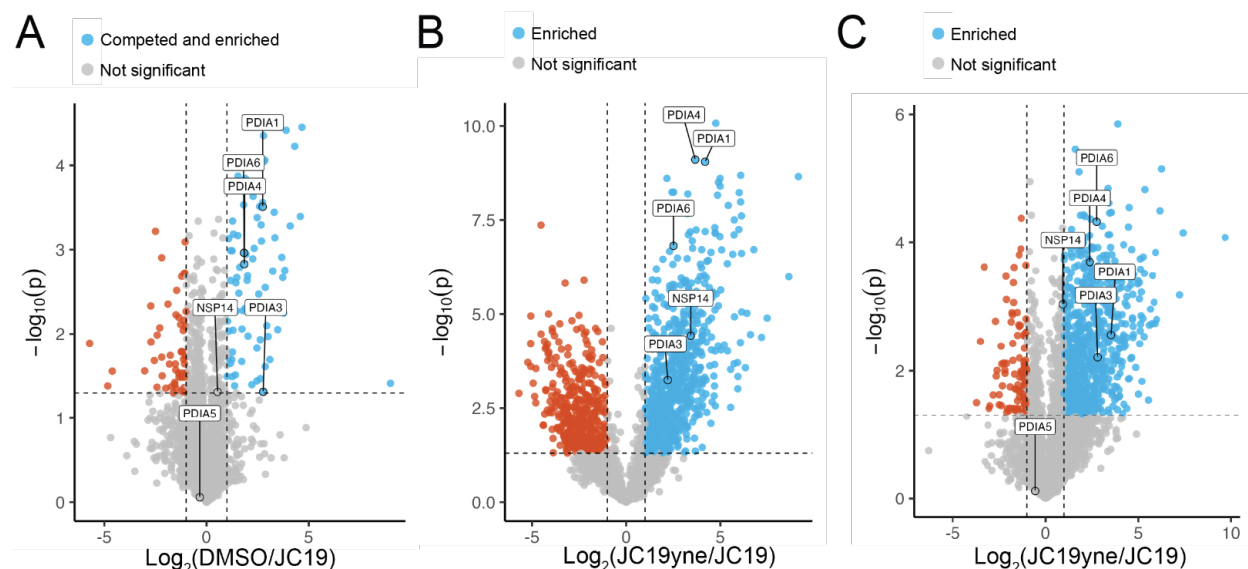
**Supplementary Figure 9. Nsp14 is turned over after 8+ h.** HEK293T cells transiently expressing nsp14-FLAG were treated with cycloheximide (CHX) for the indicated times. Cells were then lysed and soluble lysate subject to immunoblot to assay nsp14 abundance. Representative of one independent measurement.



**Supplementary Figure 10. JC19-mediated nsp14 degradation is not a result of altered glutathione levels.** (A) HEK293T cells transiently expressing nsp14-FLAG were pretreated with either DMSO or 500 µM buthionine sulfoximine (BSO) for 24 h to reduce glutathione levels, and then treated with either DMSO or 100 µM JC19 for 1 h. Immunoblot analysis was used to visualize abundance of nsp14 in the soluble lysate fraction for each condition. (B) Glutathione detection kit (G Biosciences, Cat#786-075) was used to detect glutathione (GSH) levels in cells treated with DMSO or 100 µM JC19 for 1 h (500 µM BSO treatment for 24 h used as positive control). Table depicts the standard curve of GSH measured in the assay, and the table below summarizes measurements of experimental samples (vehicle, JC19, and BSO). (C) HEK293T cells transiently expressing nsp14-FLAG were pretreated with 5 µM RSL3 for 5 h to induce ferroptosis, and then treated with either DMSO or 50 µM JC19 for 1 h. Immunoblot analysis was used to visualize abundance of nsp14 in the soluble lysate fraction for each condition. Western blot data is representative of two independent measurements.



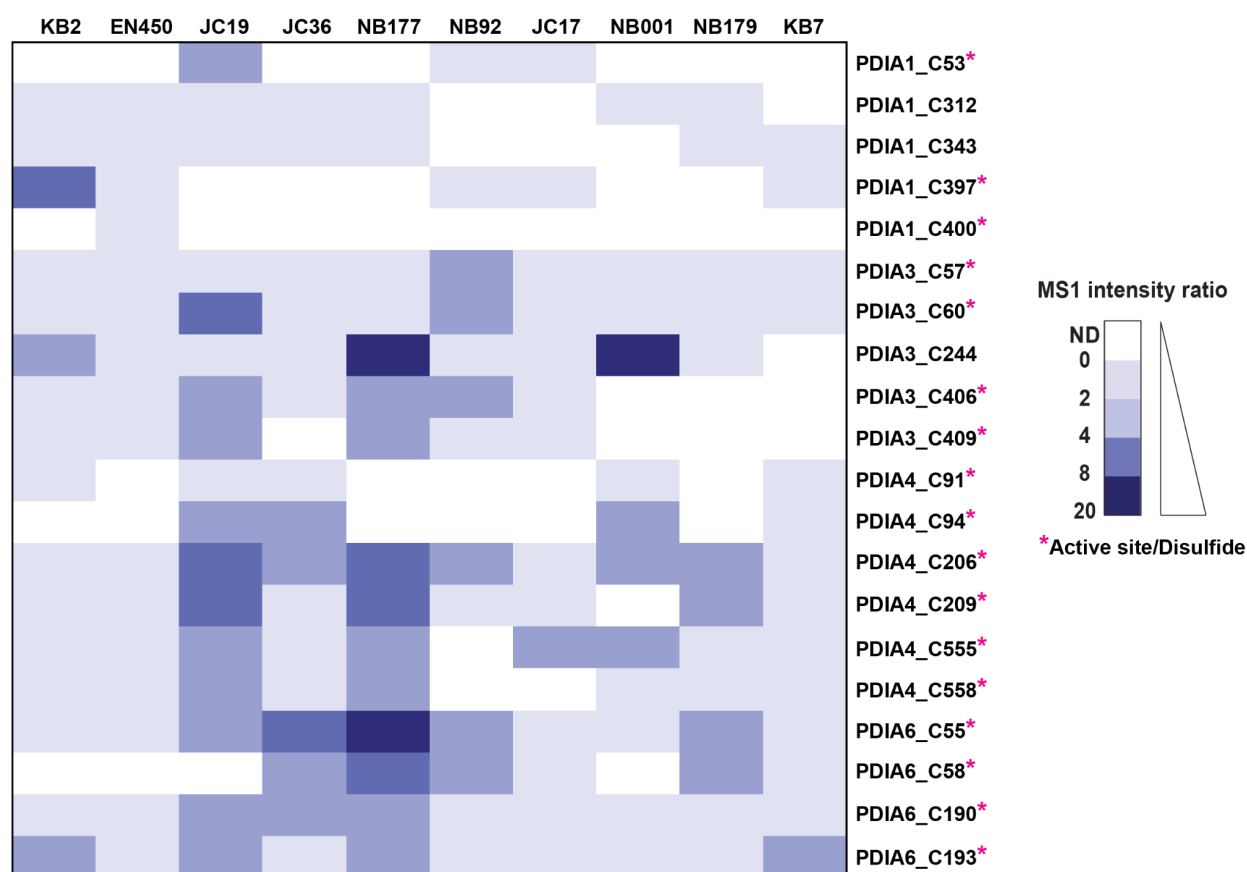
**Supplementary Figure 11. JC19 does not induce cytotoxicity to afford nsp14 depletion.** (A) HEK293T cells transiently expressing nsp14-FLAG were treated with either 100 μM JC19 or equal volume of DMSO for 1 h (n = 3). Immunoblot analysis was used to visualize abundance of nsp14 and PARP cleavage in the soluble lysate fraction for each condition. (B) HeLa cells were treated with the indicated concentrations of compounds (n = 3 per conditions) and imaged on a IncuCyte SX5 Live-Cell Imaging and Analysis System (Sartorius) using Cytotox Green reagent (Cat. No. 4633) (final dilution 1:4000). Measurements were taken every hour over the course of 13 h. Staurosporine was used as a positive control. Data are presented as mean ± standard deviation (n=3, 4 images per well).



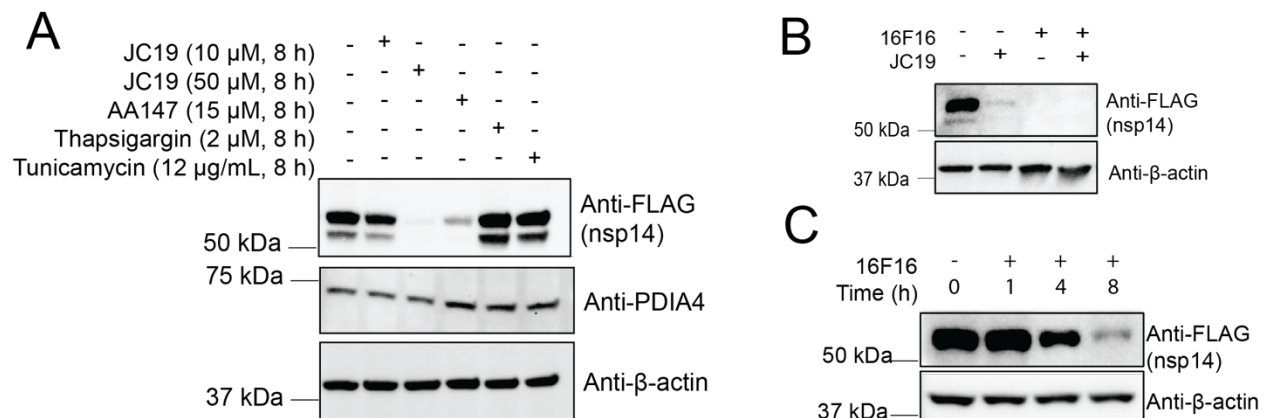
**Supplementary Figure 12. Nsp14 and the protein disulfide isomerase family are directly labeled by JC19.** (A) HEK293T cells transiently expressing nsp14-FLAG were lysed, cleared by centrifugation, and the soluble lysate treated with either DMSO (n = 3) or 100 μM JC19 (n = 2) for 20 minutes. All replicates were then treated with 100 μM JC19yne for 20 minutes, conjugated to biotin-azide via CuAAC, cleaned using SP3 clean-up, enriched on streptavidin resin, and proteolytically digested prior to LC-MS/MS analysis. Label-free quantitation was used to generate protein intensities for the volcano plot and an unpaired Student's t-test performed to generate p-values. (B) HEK293T cells transiently expressing nsp14-FLAG were treated with either 100 μM JC19 (n = 6) for 20 minutes or JC19yne (n = 6) for 20 minutes. Cells were then lysed and cleared by centrifugation. The soluble lysate was then conjugated to biotin-azide via CuAAC, cleaned



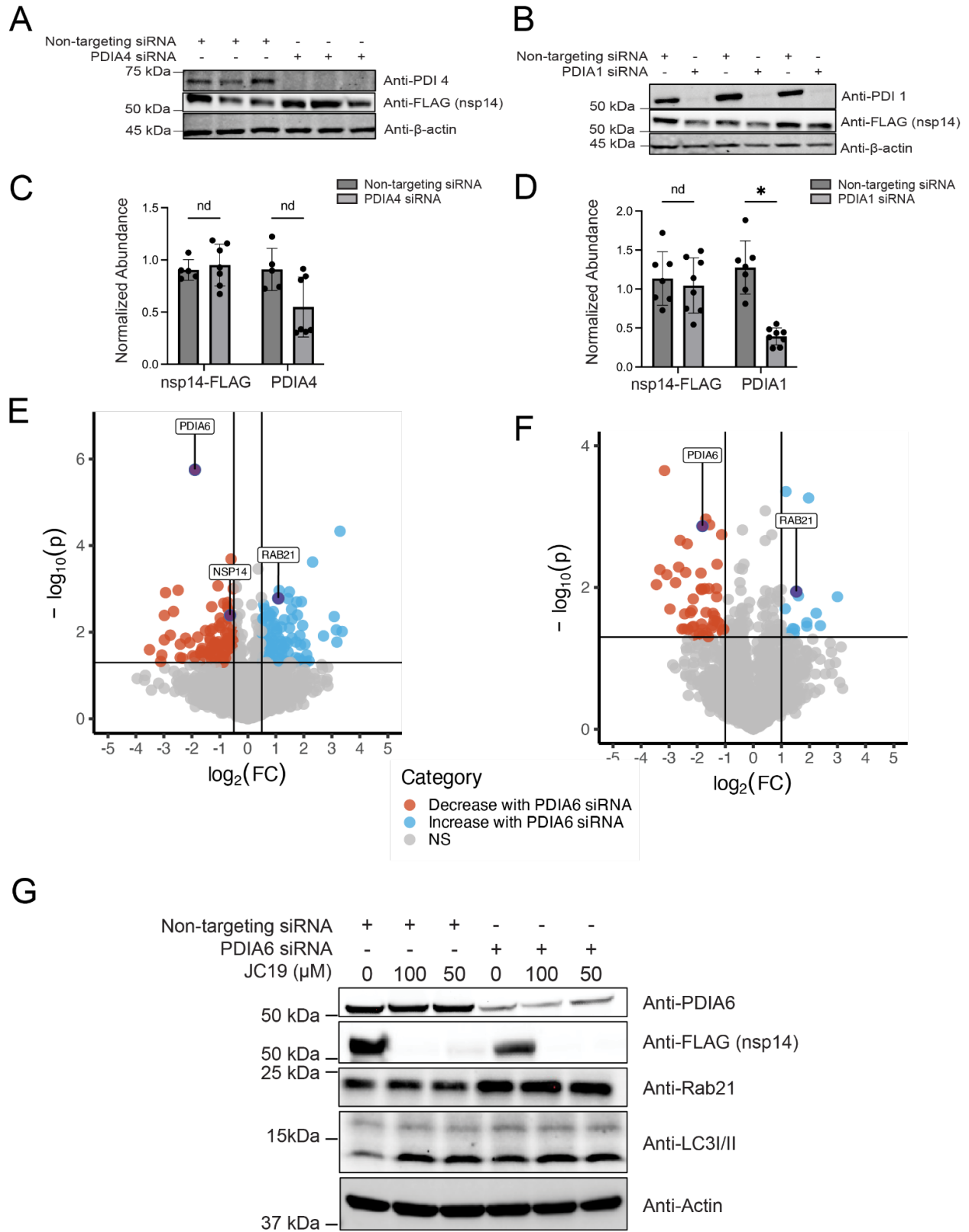
using SP3 clean-up, enriched on streptavidin resin, and proteolytically digested prior to LC-MS/MS analysis. Label-free quantitation was used to generate protein intensities for the volcano plot and an unpaired Student's t-test performed to generate p-values. (C) HEK293T cells transiently expressing nsp14-FLAG were lysed, cleared by centrifugation and treated with either 100  $\mu$ M JC19 (n = 3) for 20 minutes or JC19yne (n = 3) for 20 minutes. The soluble lysate was then conjugated to biotin-azide via CuAAC, cleaned using SP3 clean-up, enriched on streptavidin resin, and proteolytically digested prior to LC-MS/MS analysis. Label-free quantitation was used to generate protein intensities for the volcano plot and an unpaired Student's t-test performed to generate p-values. For all analyses, a student's t-test was performed between control and treatment groups to generate p-values, and proteins with a  $\log_2(\text{fold-change}) > 1$  and p-value  $< 0.05$  were considered significantly enriched. All MS data can be found in **Supplementary Table 3**.



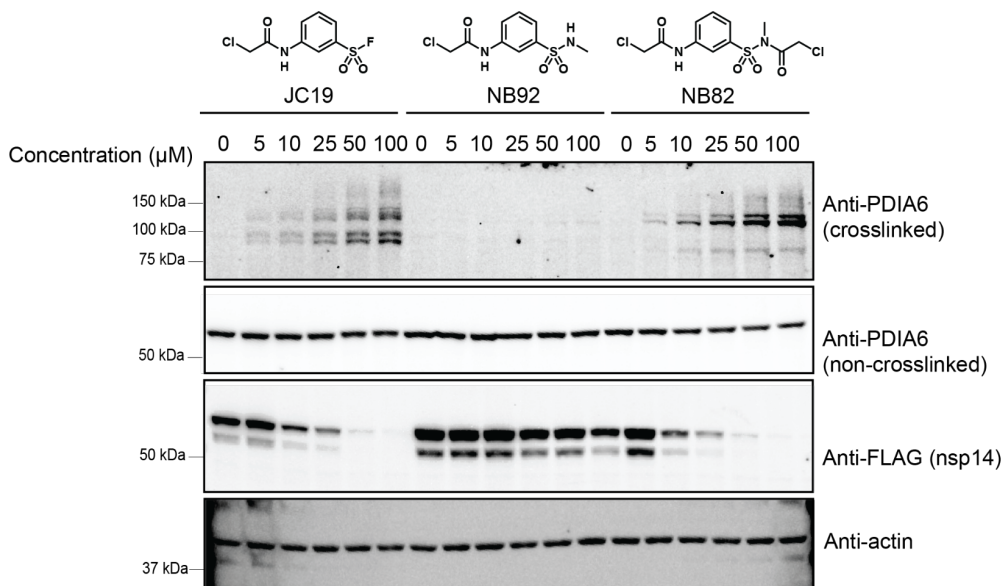
**Supplementary Figure 13. IsoTOP-ABPP ratios for protein disulfide isomerase (PDI) cysteines.** Heatmap representing the unlogged MS1 isoTOP-ABPP ratios for each identified cysteine in PDIA1, PDIA3, PDIA4, and PDIA6 based on the median value from treatment with the indicated compounds. All compound treatments were performed in biological triplicate (n=3) at 100  $\mu$ M for 1 h, except for EN450, which was treated at 100  $\mu$ M for 4 h. All MS data can be found in **Supplementary Table 4**.



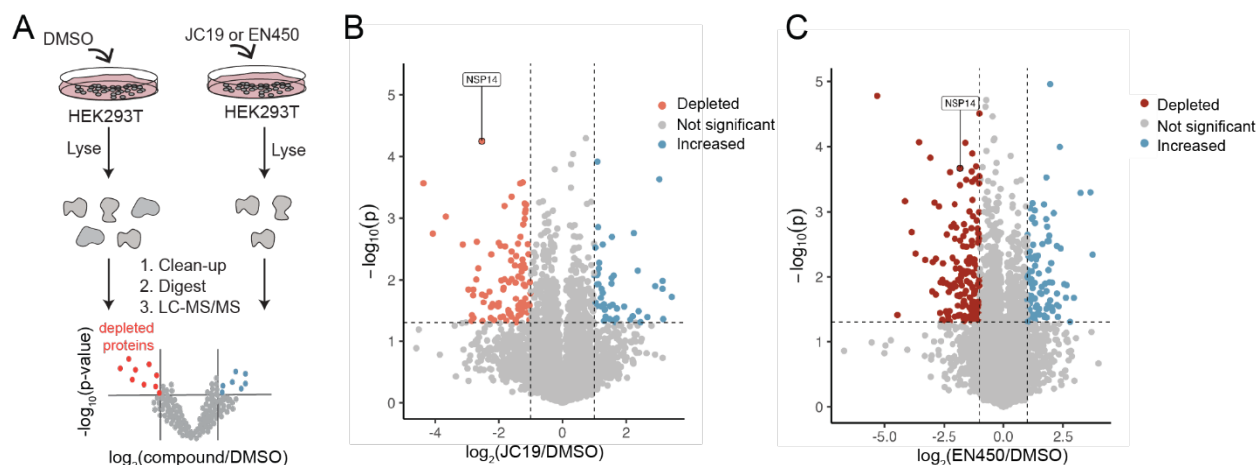
**Supplementary Figure 14. PDI inhibitors induce nsp14 depletion.** (A) HEK293T cells transiently expressing nsp14-FLAG were treated with the indicated compounds and the soluble lysate fraction assayed by immunoblot. (B) HEK293T cells transiently expressing nsp14-FLAG were treated with either DMSO or 10  $\mu$ M 16F16 for 14 h, then treated with either DMSO or 50  $\mu$ M JC19 for 1 h and immunoblot used to assay nsp14 abundance. (C) HEK293T cells transiently expressing nsp14-FLAG were treated with 10  $\mu$ M 16F16 for the indicated times and nsp14 abundance assayed by immunoblot. Western blot data is representative of three independent measurements.



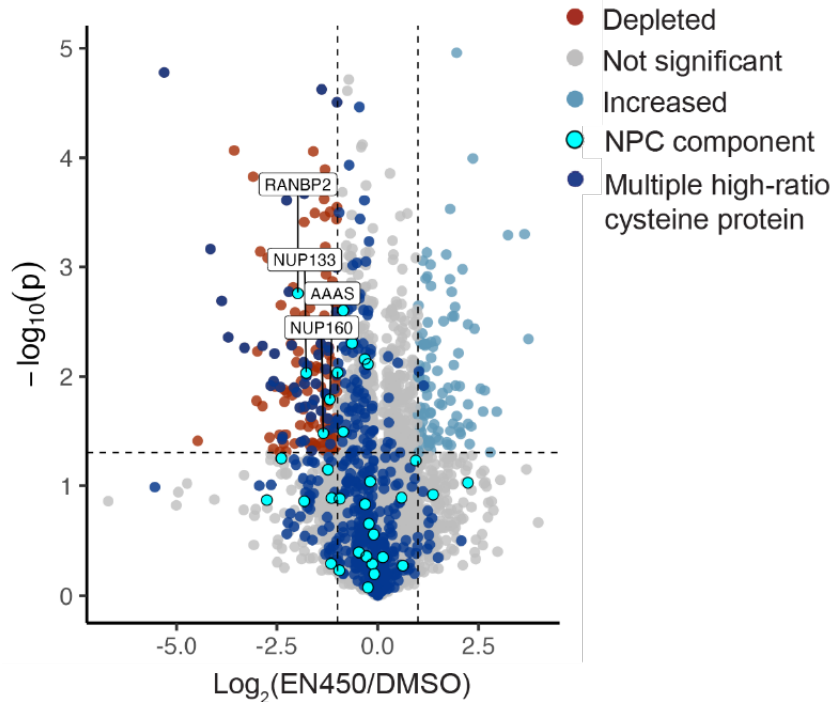
**Supplementary Figure 15. Effect of PDI knockdown on nsp14 and host proteins.** (A) Immunoblot analysis of the impact of siRNA-mediated knockdown of PDIA4 on nsp14-FLAG abundance (n = 3). (B) Immunoblot analysis the impact of siRNA-mediated knockdown of PDIA1 on nsp14-FLAG abundance (n = 3). (C) Quantification of both PDIA4 and nsp14 abundance (ImageJ<sup>2</sup>) upon siRNA-mediated knockdown of PDIA4. Data are presented as mean  $\pm$  s.d. (n=5 for non-targeting siRNA, n=7 for PDIA4 targeted siRNA). Statistical significance calculated with an unpaired Student's t-tests (nd = not significant). (D) Quantification of both PDIA1 and nsp14 abundance (ImageJ<sup>2</sup>) upon siRNA-mediated knockdown of PDIA1. Data are presented as mean  $\pm$  s.d. (n=7 for non-targeting siRNA, n=8 for PDIA1 targeted siRNA). Statistical significance calculated with an unpaired Student's t-tests \* p<0.05, PDIA1 p=0.00001. (E) Label-free quantification used to measure protein abundance in cells in which siRNA was used to knockdown PDIA6 as compared to scramble siRNA in HEK293T nsp14-overexpressing cells (n = 3 per group) and an unpaired Student's t-test performed to generate p-values. FC cutoff at log<sub>2</sub>(0.5). (F) Label-free quantification used to measure protein abundance in cells in which siRNA was used to knockdown PDIA6 as compared to scramble siRNA in HEK293T cells (n = 3 per group) and an unpaired Student's t-test used to generate p-values. log<sub>2</sub>(FC) cutoff set at 1. (G) Immunoblot of a PDIA6 siRNA knockdown in HEK293T cells transiently expressing NSP14. For all analyses, an unpaired Student's t-test was performed between control and treatment groups to generate p-values, and proteins with the indicated fold change (FC) cutoff and p-value < 0.05 were considered significantly enriched. All MS data can be found in **Supplementary Table 3**. Panel g is representative of two independent measurements.



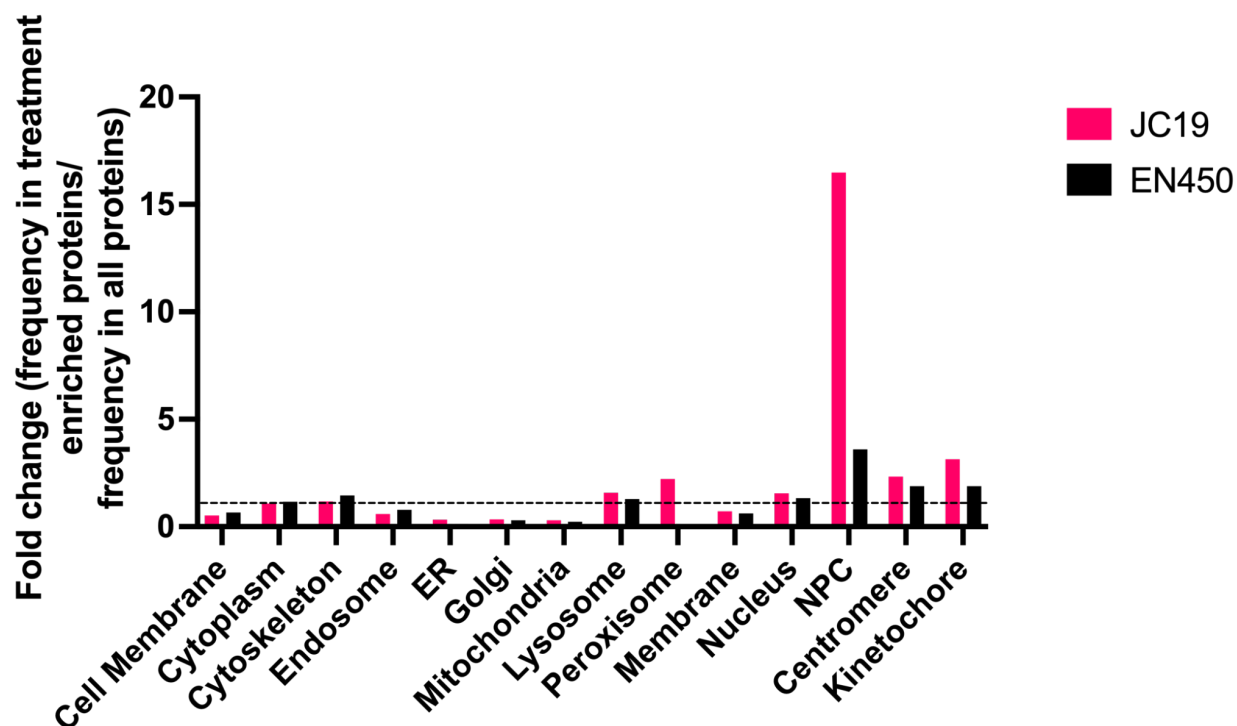
**Supplementary Figure 16. Compounds with PDIA6 cross linking activity are more active nsp14 depleters than non-crosslinking compounds.** HEK293T cells expressing nsp14-FLAG were treated with the indicated concentrations of compound for 0.5 h. Cells were then lysed and soluble lysate subject to immunoblot analysis. Western blot data is representative of two independent measurements.



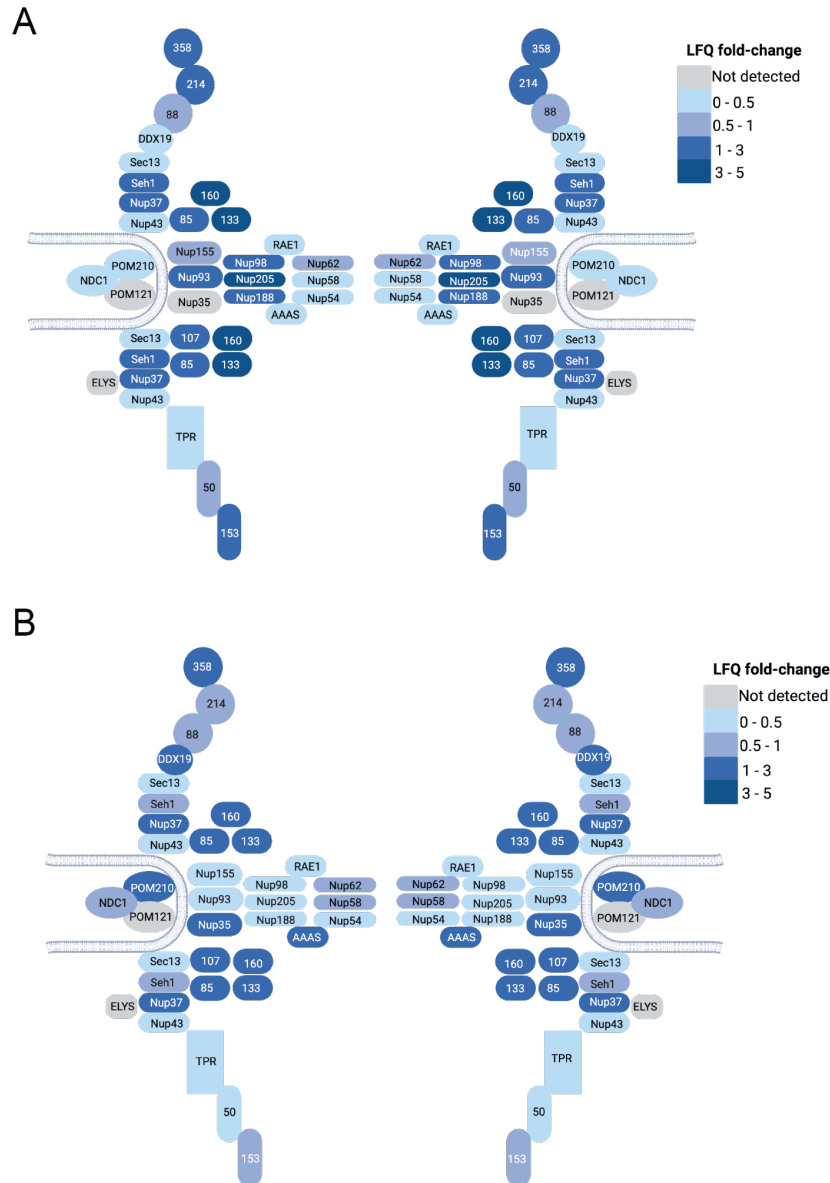
**Supplementary Figure 17. Nsp14 is depleted from cells when treated with JC19 and EN450.** (A) Workflow for bulk proteomics. (B-C) HEK293T cells transiently expressing nsp14-FLAG were treated with DMSO ( $n = 3$ ) or 100  $\mu\text{M}$  JC19 ( $n = 3$ ) for 1 h (B) or DMSO ( $n = 3$ ) or 100 $\mu\text{M}$  EN450 ( $n = 3$ ) for 3 h (C). Cells were then lysed and prepared for LC-MS/MS analysis via SP3 clean-up. Label-free quantification was used to quantify protein intensities in each replicate. For both plots, an unpaired Student's t-test was performed between control and treatment groups, and the average fold-change of the label-free intensity for each protein plotted against its p-value. Proteins with a  $\log_2$  fold-change  $< -1$  and p-value  $< 0.05$  are considered significantly depleted. All MS data can be found in **Supplementary Table 4**.



**Supplementary Figure 18. EN450 induces protein depletion.** HEK293T cells transiently expressing nsp14-FLAG were treated with DMSO (n = 3) or 100  $\mu$ M EN450 (n = 3) for 4 h. Cells were then lysed and prepared for LC-MS/MS analysis via SP3 clean-up. Red points represent proteins that were depleted in response to treatment. Dark blue points represent proteins with multiple high-ratio cysteines in the EN450 isoTOP-ABPP dataset. Cyan points represent components of the nuclear pore complex (NPC). The labeled proteins are the NPC proteins that were significantly depleted in response to EN450 treatment. Label-free quantification was used to measure protein intensities in each replicate and the fold-change of the average intensity for each protein plotted against its respective p-value calculated from an unpaired Student's t-test. Log<sub>2</sub>(fold-change) < -1 and p-values < 0.05 are considered significantly depleted. All MS data can be found in **Supplementary Table 4**.

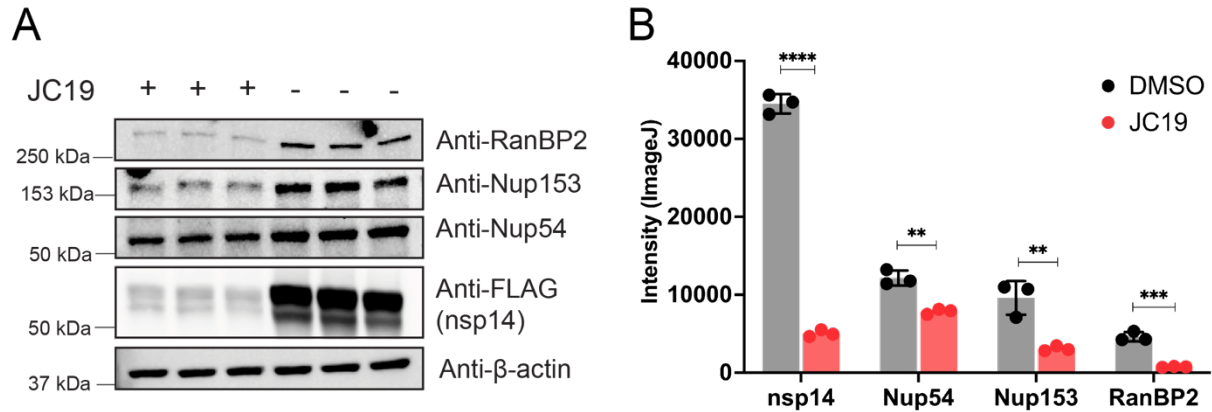


**Supplementary Figure 19. The nuclear pore complex is an enriched subcellular compartment within proteins depleted by JC19 and EN450.** All proteins identified from JC19 and EN450 soluble label-free bulk proteome MS analysis annotated by subcellular location using Uniprot annotations. The total representation of subcellular location for all proteins that were identified was compared to representation of subcellular locations for the significantly depleted proteins (significance calculated as stated in the soluble label-free bulk proteome MS analysis). Fold-change of those two values were plotted. MS data can be found in **Supplementary Table 4**.

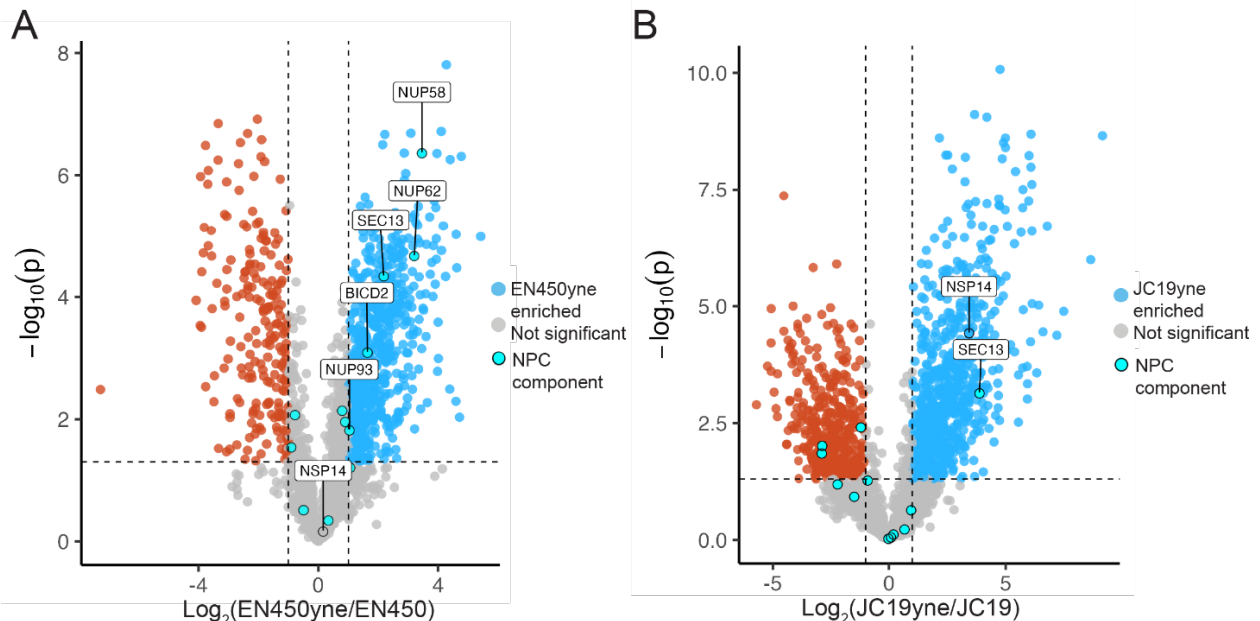


**Supplementary Figure 20. Depletion of nucleoporins by JC19 and EN450 is not biased to a specific region of the NPC.** Heatmaps of the nuclear pore complex, with each nucleoporin colored based on the LFQ fold-change observed in the bulk proteome analyses in response to JC19 (A) or EN450 (B). MS data can be found in **Supplementary Table 4**.



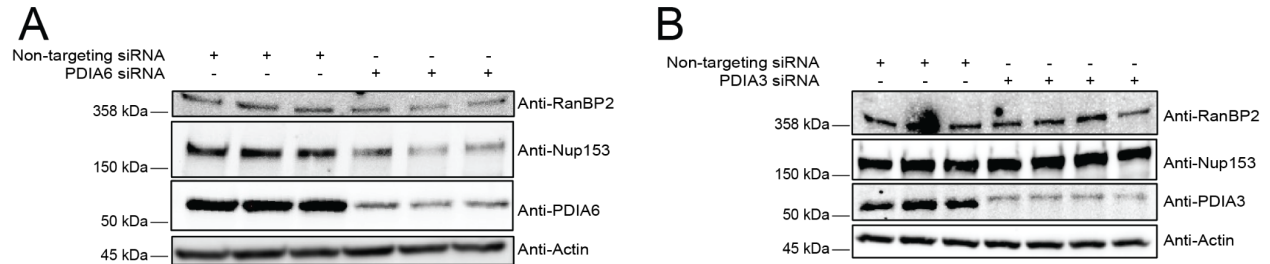


**Supplementary Figure 21. Nucleoporin proteins, in addition to nsp14, are sensitive to JC19 treatment.** (A) HEK293T cells transiently expressing nsp14-FLAG were treated with either DMSO or 100  $\mu$ M JC19 for 1 h ( $n = 3$ ). Immunoblot analysis was used to assay the abundance of RANBP2, NUP153, NUP54, and nsp14. Lysates solubilized with 0.3% CHAPS. (B) Quantification of the band intensities in (A) by ImageJ<sup>2</sup>. Data are presented as mean  $\pm$  standard deviation. ( $n=3$ ). An unpaired Student's t-test was performed to obtain the p-values (nsp14 p-value = 0.000002648, Nup54 p-value = 0.001862, Nup153 p-value = 0.006801, and RanBP2 p-value = 0.0003851).

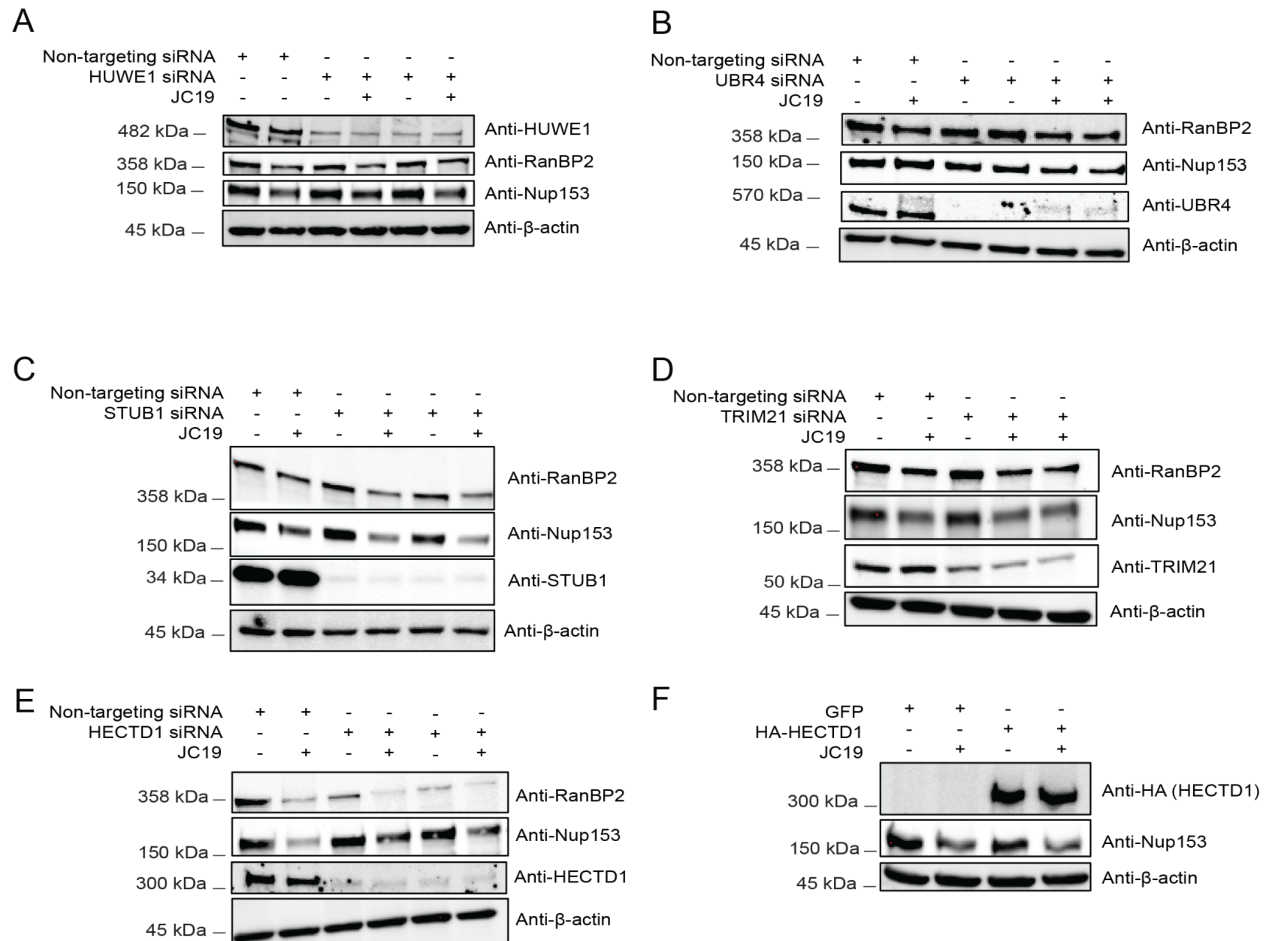


**Supplementary Figure 22. Nucleoporins are largely not directly labeled by JC19 or EN450 to afford depletion.** (A) HEK293T cells transiently expressing nsp14-FLAG were treated with either 100  $\mu$ M EN450 ( $n = 4$ ) or 100  $\mu$ M EN450yne ( $n = 4$ ) for 4 h. Cells were then lysed, proteins conjugated to biotin-azide via CuAAC, enriched on streptavidin, and on-resin trypsin digest performed to release peptides for LC-MS/MS. Label-free quantification was used to measure the intensity of each protein in all replicates. An unpaired Student's t-test was performed between control and treatment groups to generate a p-value, which was plotted against the average fold-change of each protein. Proteins with a  $\log_2(\text{fold-change}) > 1$  and p-value  $< 0.05$  are considered significantly enriched by EN450yne. (B) Data duplicated from Supplementary Figure 12, highlighting enriched nucleoporins. HEK293T cells transiently expressing nsp14-FLAG were

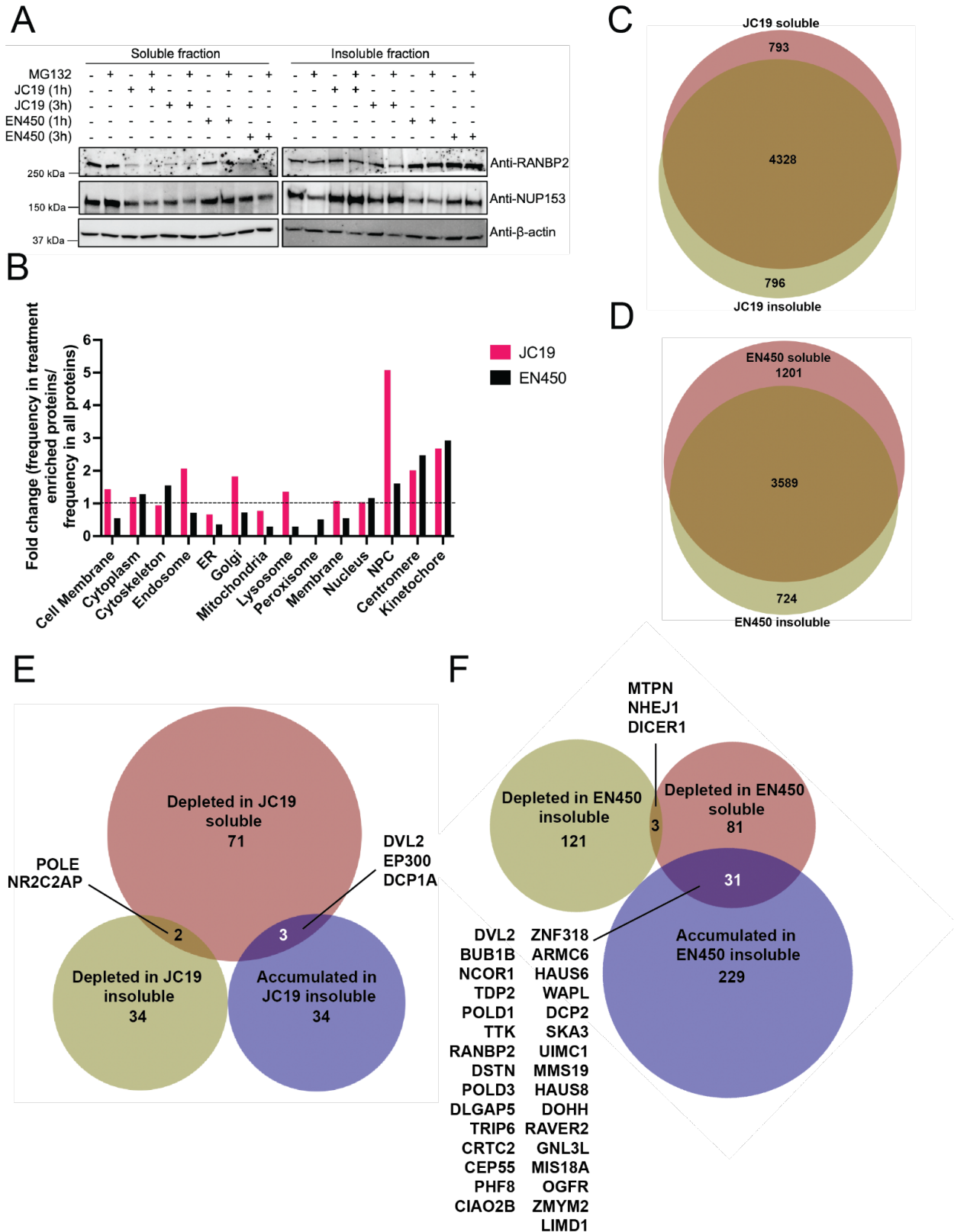
treated with either 100  $\mu$ M JC19 (n = 6) or 100  $\mu$ M JC19yne (n = 6) for 20 minutes. Cells were then lysed, proteins conjugated to biotin-azide via CuAAC, enriched on streptavidin, and on-resin trypsin digest performed to release peptides for LC-MS/MS. Label-free quantification was used to measure the intensity of each protein in all replicates. An unpaired Student's t-test was performed between control and treatment groups to generate a p-value, which was plotted against the average fold-change of each protein. Proteins with a  $\log_2(\text{fold-change}) > 1$  and p-value  $< 0.05$  are considered significantly enriched by JC19yne. All MS data can be found in **Supplementary Table 3** and **Supplementary Table 4**.



**Supplementary Figure 23. Effect of PDIA6 and PDIA3 siRNA knockdowns on nucleoporins RANBP2 and NUP153. Knockdown of PDIA6 does not have a widespread effect on the abundance of nucleoporins.** (A) Western blot showing RANBP2 and NUP153 abundance in response to SiRNA knockdown PDIA6 as compared to scramble siRNA in HEK293T nsp14-overexpressing cells (n = 3). (B) Western blot showing RANBP2 and NUP153 abundance in response to SiRNA knockdown PDIA3 as compared to scramble siRNA in HEK293T nsp14-overexpressing cells (n = 3 or 4).

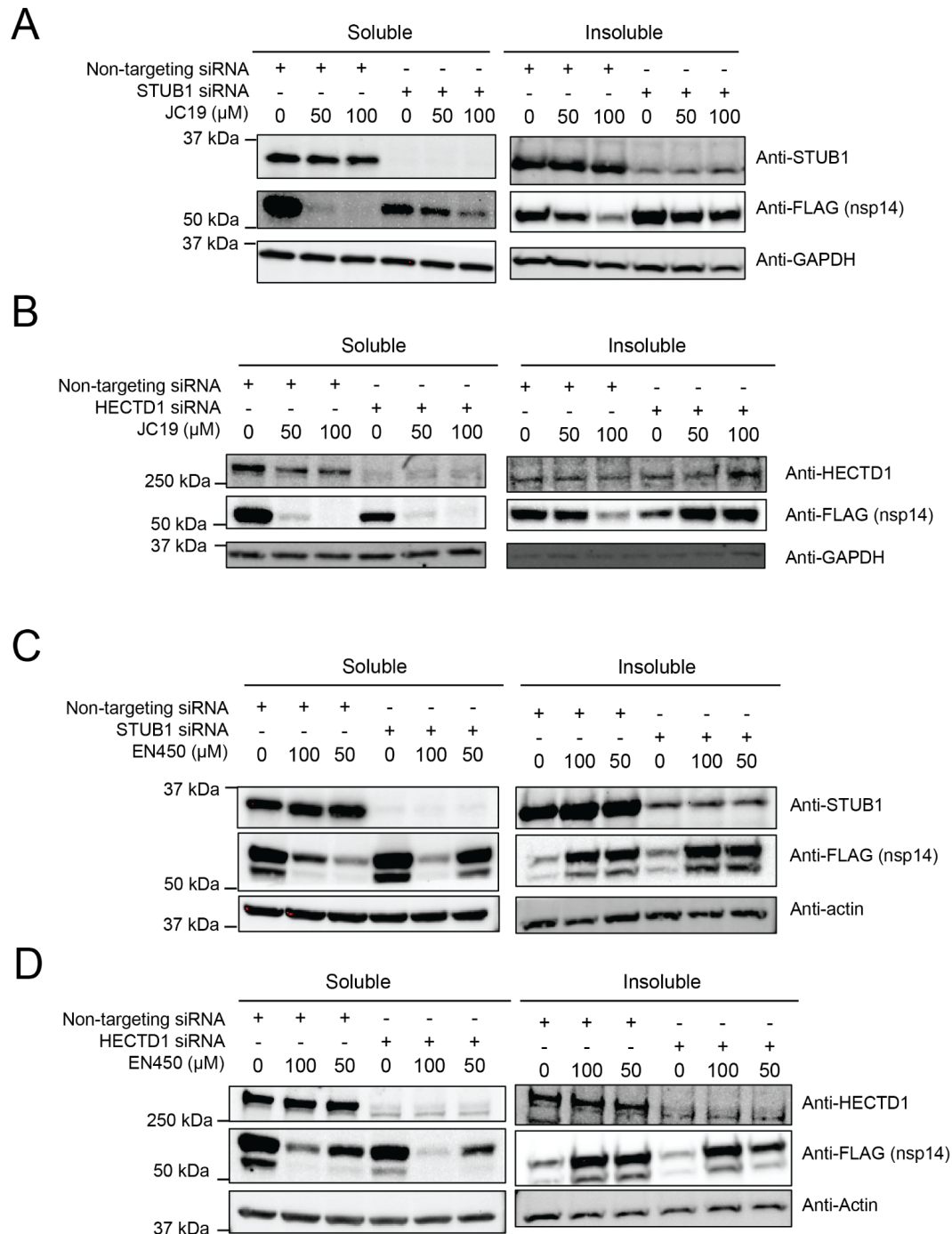


**Supplementary Figure 24. Effect of E3 ligase knockdowns on nuclear pore proteins.** (A) Immunoblot of siRNA knockdown of HUWE1 and its impact on RANBP2 and NUP153 abundance, both with and without JC19 treatment (25  $\mu$ M, 0.5 h). (B) Immunoblot of siRNA knockdown of UBR4 and its impact on RANBP2 and NUP153 abundance, both with and without JC19 treatment (25  $\mu$ M, 0.5 h). (C) Immunoblot of siRNA knockdown of STUB1 and its impact on RANBP2 and NUP153 abundance, both with and without JC19 treatment (25  $\mu$ M, 0.5 h). (D) Immunoblot of siRNA knockdown of TRIM21 and its impact on RANBP2 and NUP153 abundance, both with and without JC19 treatment (25  $\mu$ M, 0.5 h). (E) Immunoblot of siRNA knockdown of HECTD1 and its impact on RANBP2 and NUP153 abundance, both with and without JC19 treatment (100  $\mu$ M, 1 h). (F) HA-HECTD1 overexpression and its impact on NUP153 abundance upon JC19 treatment (100  $\mu$ M, 1 h) as compared to control. Knockdown blots for panels A-E are the same blots from **Supplementary Figure 6** re-blotted for host proteins. All western blot data is representative of two independent measurements.



**Supplementary Figure 25. The nuclear pore complex is an enriched subcellular compartment within proteins that accumulate in the insoluble fraction upon JC19 and EN450 treatment. (A) Immunoblot analysis of RANBP2 and NUP153 abundance within the**

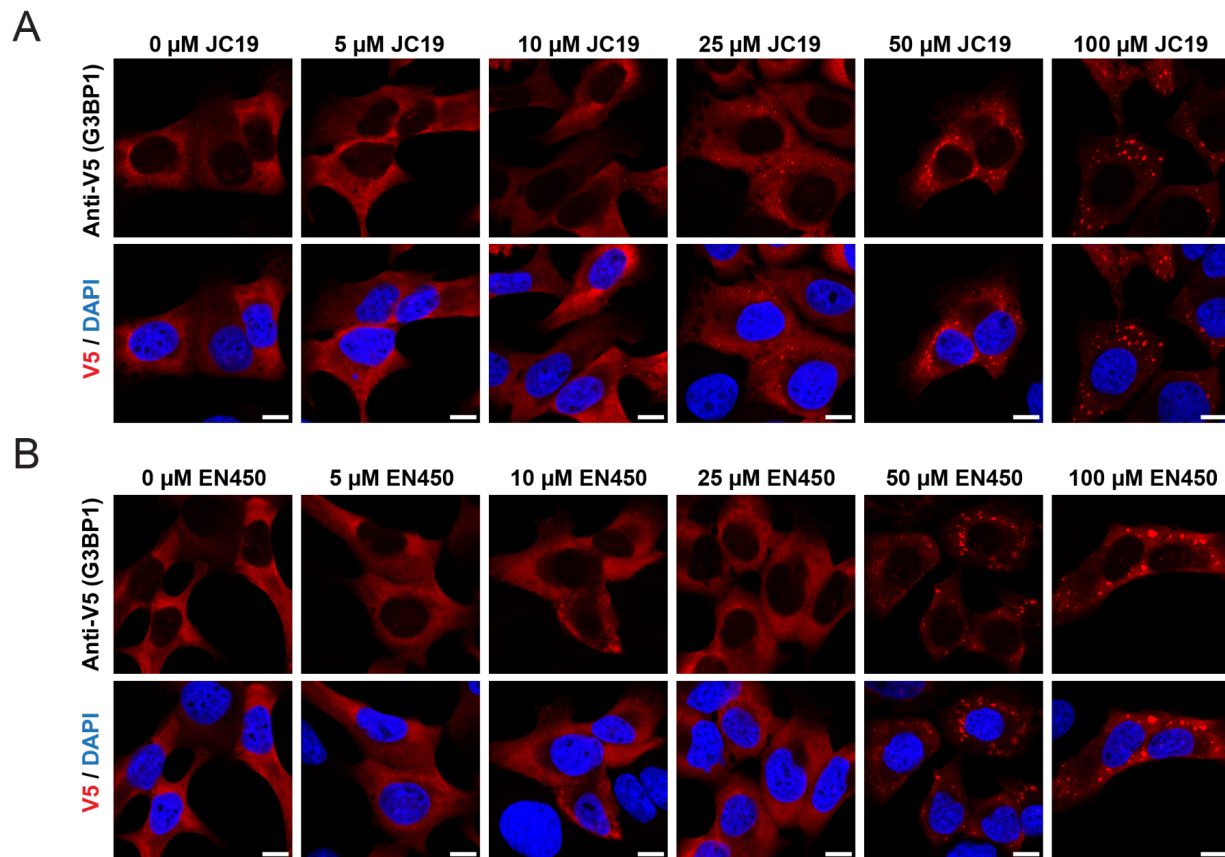
soluble and insoluble fractions in response to 100  $\mu$ M JC19 or EN450 treatment at the indicated time points, both with and without a MG132 pretreatment (25  $\mu$ M for 5 h). (B) All proteins identified from JC19 and EN450 insoluble label-free bulk proteome MS analysis annotated by subcellular location using Uniprot annotations. The total representation of subcellular location for all proteins that were identified was compared to representation of subcellular locations for the enriched proteins. Fold-change of those two values were plotted above. (C-D) The overlap of proteins identified in the bulk soluble and insoluble proteomics datasets for JC19 (C) and EN450 (D). (E-F) Overlap of proteins that were significantly depleted from the soluble dataset and significantly depleted or accumulated in the insoluble dataset for JC19 (E) and EN450 (F). All proteins in E-F were identified in both soluble and insoluble datasets for the corresponding compound. All MS data can be found in **Supplementary Table 4**. Panel a is representative of two independent measurements.



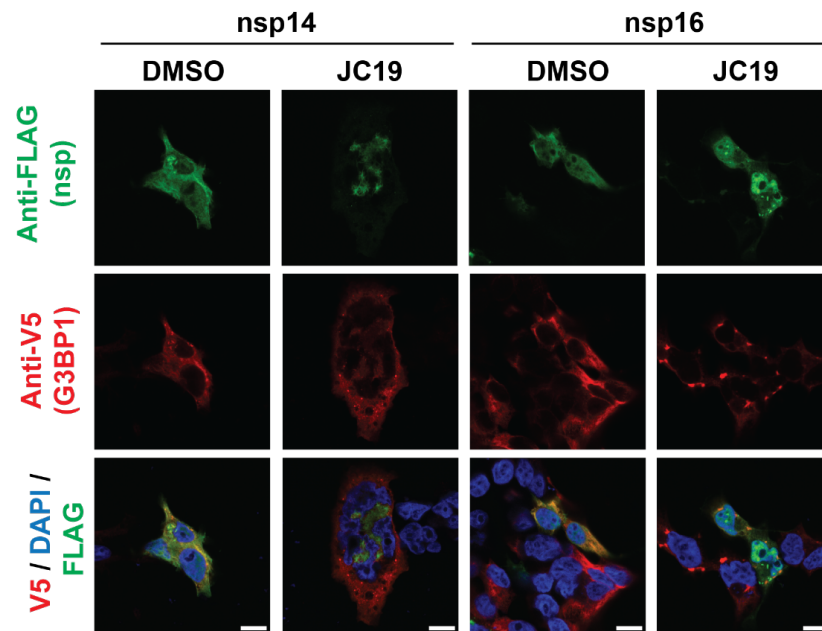
**Supplementary Figure 26. Genetic knockdown of E3 ligases STUB1 and HECTD1 protect nsp14 from JC19-mediated depletion from the insoluble fraction.** (A-B) HEK293T cells transiently transfected with nsp14-FLAG were subject to siRNA-mediated knockdown of STUB1 (A) or HECTD1 (B), treated with the indicated concentrations of JC19 for 1 h, and the soluble and insoluble lysate fractions assayed by immunoblot. (C-D) HEK293T cells transiently transfected with nsp14-FLAG were subject to siRNA-mediated knockdown of STUB1 (C) or HECTD1 (D), treated with the indicated concentrations of EN450 for 3 h, and the soluble and insoluble lysate



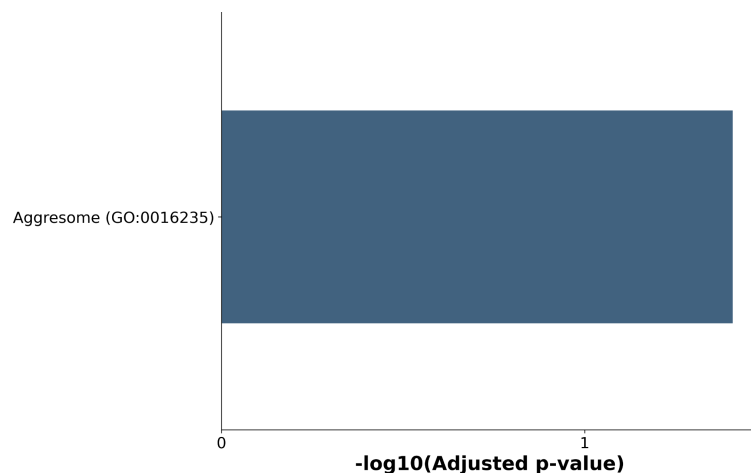
fractions assayed by immunoblot. All western blot data is representative of two independent measurements.



**Supplementary Figure 27. Dose-dependence of cysteine-reactive electrophile induction of stress granule formation.** (A) U2OS cells stably expressing V5-tagged G3BP1 were treated with the indicated concentrations of JC19 for 0.5 h, fixed, permeabilized, and subject to immunofluorescence microscopy. Images were acquired on an LSM880 confocal microscope at 63X objective and 2X manual zoom. All scale bars = 10 $\mu$ m. (B) U2OS cells stably expressing V5-tagged G3BP1 were treated with the indicated concentrations of EN450 for 0.5 h, fixed, permeabilized, and subject to immunofluorescence microscopy. Images were acquired on an LSM880 confocal microscope at 63X objective and 2X manual zoom. All scale bars = 10 $\mu$ m. All images are representative of two independent measurements.

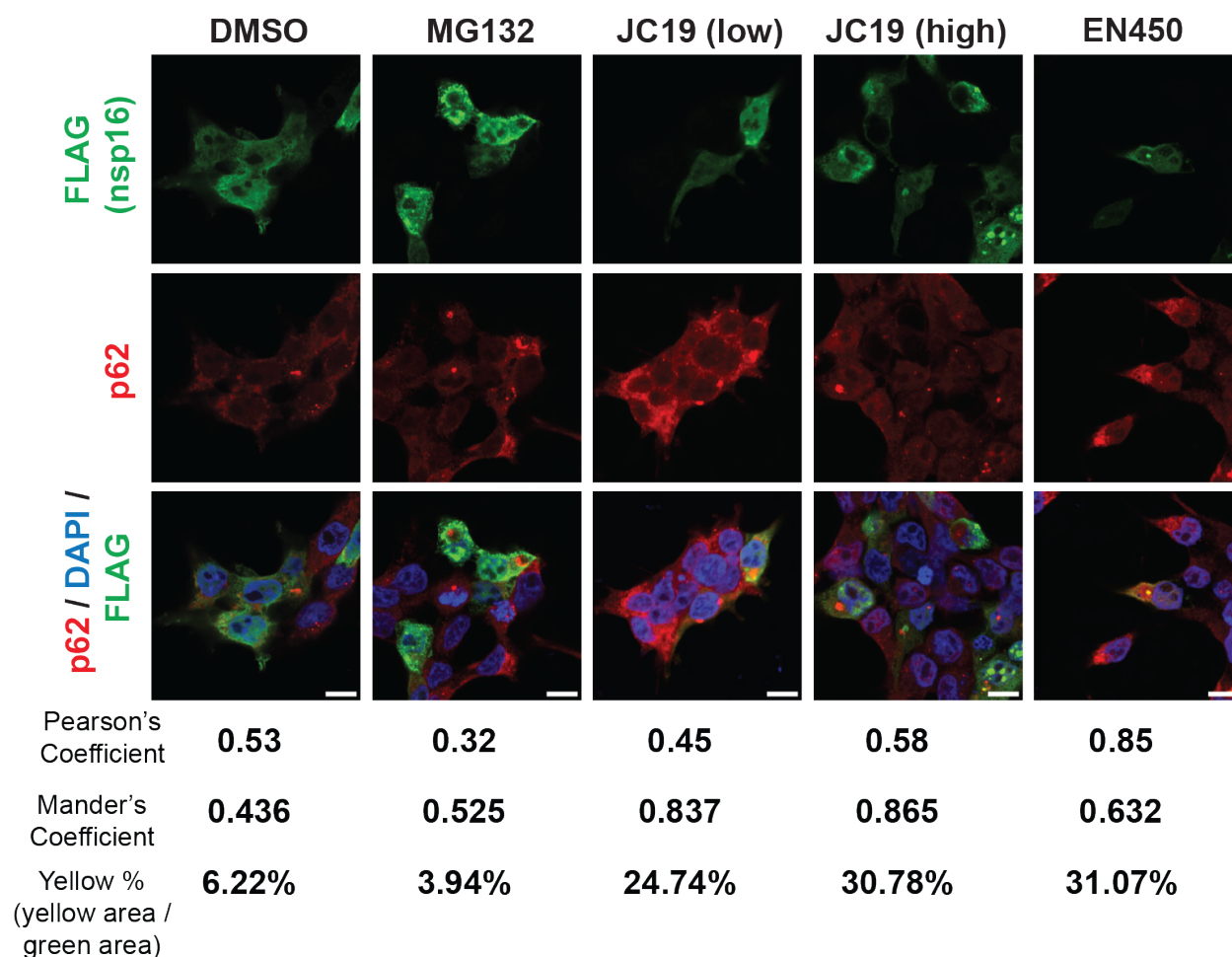


**Supplementary Figure 28. Nsp14 and nsp16 are not sequestered into stress granules (SGs) in response to JC19 treatment.** HEK293T cells stably expressing V5-tagged G3BP1 were transiently transfected with nsp14 or nsp16. Cells were then treated with either DMSO or 100  $\mu$ M JC19 for 1 h, fixed, permeabilized, and subject to immunofluorescence microscopy to visualize nsp14, nsp16, and G3BP1. Images were acquired on an LSM880 confocal microscope at 63X objective and 2X manual zoom. All scale bars = 10  $\mu$ m. All images are representative of two independent measurements.



**Supplementary Figure 29. The aggresome is enriched among proteins that accumulate in the insoluble fraction upon JC19 treatment.** Gene Ontology Cellular Compartment analysis (from GO Cellular Compartment 2023) on the proteins that significantly accumulated in the insoluble fraction upon JC19 treatment ( $\log_2(\text{fold-change}) > 1$  and  $p\text{-value} < 0.05$  relative to DMSO control) relative to all proteins identified in the dataset. P-values were computed from Fisher's exact test to determine significance of each enriched term. MS Data can be found in **Supplementary Table 5**.





**Supplementary Figure 30. nsp16 is recruited to the aggresome upon compound treatment.** HEK293T cells were transiently transfected with nsp16-FLAG, treated with DMSO, MG132 (10  $\mu$ M, 5 h), a low dose of JC19 (25  $\mu$ M, 0.5 h), a high dose of JC19 (100  $\mu$ M, 1 h), or EN450 (100  $\mu$ M, 1 h), fixed, permeabilized, and subject to immunofluorescence microscopy to visualize nsp16 and p62. Images were acquired on an LSM880 confocal microscope at 63X objective and 2X manual zoom. All scale bars = 10  $\mu$ m. All images are representative of two independent measurements.

**Table S1-5.** Datasets corresponding to each figure, provided in the attached supplementary files.

Supplementary Table 1: Figure 1

Supplementary Table 2: Figure 3

Supplementary Table 3: Figure 4

Supplementary Table 4: Figure 5

Supplementary Table 5: Figure 6

Supplementary Table 11: Files in Proteomics Identification Database (PRIDE) datasets

**Supplementary Table 6.** Antibodies used in this study.

| Name | Catalog number | Lot number | Concentration |
|------|----------------|------------|---------------|
|------|----------------|------------|---------------|

|                                                    |                          |           |                          |
|----------------------------------------------------|--------------------------|-----------|--------------------------|
| IRDye® 800CW Goat anti-Rabbit Secondary Antibody   | LI-COR (926-32211)       | D20803-10 | 1:5000 (IB)              |
| IRDye® 800CW Goat anti-Mouse Secondary Antibody    | LI-COR (926-32210)       | D1116-25  | 1:5000 (IB)              |
| IRDye 680RD Donkey anti-Rabbit Secondary Antibody  | LI-COR (926-68073)       | D00421-09 | 1:5000 (IB)              |
| IRDye 680RD Donkey anti-Mouse Secondary Antibody   | LI-COR (926-68072)       | D20322-25 | 1:5000 (IB)              |
| Alexa Fluor™ 546 donkey anti-mouse IgG (H+L)       | Invitrogen (A10036)      | 2160040   | 1:1000 (IF)              |
| Alexa Fluor™ 633 goat anti-rabbit IgG (H+L)        | Invitrogen (A21070)      | 2563844   | 1:1000 (IF)              |
| Donkey anti-Rabbit IgG (H+L) Alexa Fluor™ Plus 488 | Invitrogen (A32790)      | VC296619  | 1:1000 (IF)              |
| DYKDDDDK Tag (D6W5B) Rabbit mAb                    | Cell signaling (14793S)  | 7         | 1:3000 (IB); 1:1000 (IF) |
| β-Actin (8H10D10) Mouse mAb                        | Cell signaling (3700S)   | 21        | 1:3000                   |
| GAPDH (14C10) Rabbit mAb                           | Cell signaling (2118S)   | 16        | 1:3000                   |
| HA-Tag (C29F4) Rabbit mAb                          | Cell signaling (3724S)   | 10        | 1:3000                   |
| PSMB5 (D1H6B) Rabbit mAb #12919                    | Cell signaling (12919S)  | 1         | 1:3000                   |
| 6x-His Tag Monoclonal Antibody (HIS.H8)            | ThermoFisher (MA1-21315) | TK269015  | 1:3000                   |
| PDI (C81H6) Rabbit mAb                             | Cell Signaling (3501S)   | 3         | 1:3000                   |
| PDIA6 Polyclonal antibody                          | Proteintech (18233-1-AP) | 00055732  | 1:3000                   |
| PDIA6 Monoclonal antibody                          | Proteintech (66669-1-Ig) |           | 1:3000                   |

|                                  |                                  |            |        |
|----------------------------------|----------------------------------|------------|--------|
| ERp57 Rabbit pAb                 | Abclonal (A1085)                 | 0056780201 | 1:3000 |
| ERp72 Antibody                   | Cell Signaling<br>(2798S)        | 2          | 1:3000 |
| LC3A/B (D3U4C) XP®<br>Rabbit mAb | Cell Signaling<br>(12741S)       | 5          | 1:2000 |
| PARP (46D11) Rabbit<br>mAb       | Cell Signaling<br>(9532S)        | 10         | 1:3000 |
| NRF2 Rabbit pAb                  | Abclonal<br>(NO:A1244)           | 5500014741 | 1:3000 |
| BiP (C50B12) Rabbit mAb<br>#3177 | Cell Signaling<br>(3177S)        | 10         | 1:1000 |
| CHOP (L63F7) Mouse<br>mAb        | Cell Signaling<br>(2895S)        | 13         | 1:1000 |
| ATF-4 (D4B8) Rabbit mAb          | Cell Signaling<br>(11815S)       | 5          | 1:1000 |
| STUB1 Rabbit pAb                 | ABclonal (A11751)                | 0071170101 | 1:3000 |
| TRIM21/SS-A Rabbit mAb           | ABclonal (A22724)                | 6100001135 | 1:3000 |
| HUWE1 Rabbit pAb                 | ABclonal (A9433)                 | 5500008129 | 1:3000 |
| UBR4 Rabbit pAb                  | ABclonal (A12193)                | 0092800201 | 1:3000 |
| HECTD1 Rabbit pAb                | ABclonal (A9433)                 | 0101520101 | 1:3000 |
| RANBP2 (D-4)                     | Santa Cruz Biotech<br>(sc-74518) | I0821      | 1:1000 |
| NUP153 (E3I6Z)                   | Cell Signaling<br>(98559S)       | 1          | 1:3000 |
| NUP54 Rabbit PolyAb              | Proteintech<br>(16232-1-AP)      | 00059775   | 1:3000 |
| NFKB1 Rabbit Ab                  | Cell Signaling<br>(3035S)        | 6          | 1:3000 |

|                                                      |                               |            |                          |
|------------------------------------------------------|-------------------------------|------------|--------------------------|
| V5-Tag (E9H80) Mouse mAb                             | Cell Signaling (80076S)       | 3          | 1:3000 (IB); 1:1000 (IF) |
| Alexa Fluor™ 488 phalloidin                          | Invitrogen (A12379)           | 2282104    | 1:800 (IF)               |
| PSMB2 Rabbit PolyAb                                  | Proteintech (I5I54-I-AP)      |            | 1:300 (IF)               |
| Rb pAn to RanBP2                                     | Abcam (AB64276)               | 1024132-3  | 1:300 (IF)               |
| SQSTM1 (D-3)                                         | Santa Cruz Biotech (sc-28359) | G2822      | 1:100 (IF)               |
| RAB21 Rabbit pAb                                     | ABclonal (A12095)             | 0099420101 | 1:3000 (IB)              |
| P62,SQSTM1 Polyclonal antibody                       | Proteintech (18420-1-AP)      | 00115919   | 1:100 (IF)               |
| K48-linkage Specific Polyubiquitin Rabbit mAb (D9D5) | Cell Signaling (#8081)        |            | 1:1000 (IB)              |

Supplementary **Table 7**. siRNAs used in this study.

| Target Gene            | Manufacturer | Catalog Number    | Lot #     | Duplex Sequence                                                        |
|------------------------|--------------|-------------------|-----------|------------------------------------------------------------------------|
| HECTD1                 | IDT          | hs.Ri.HECTD1.13.1 | 538907239 | 5'-AUGGUCAGCAAGCUACAAUUAUUGAA-3';<br>5'-UUCAUAUUGUAGCUUGCUGACCAUCU-3'  |
| HECTD1                 | IDT          | hs.Ri.HECTD1.13.2 | 538907240 | 5'-ACUGAUGCAAGCUAUCCAUCAGUCA-3';<br>5'-UGACUGAUGGAUAGCUUGCAUCAGUAG-3'  |
| HUWE1                  | ambion       | s19595            | ASO2JTF1  | 5'-CAUUGGAAAGUGCGAGUUAtt-3';<br>5'-UAACUCGCACUUUCCAAU-3'               |
| Negative Control siRNA | IDT          | 51-01-14-03       | 663345    | 5'-CGUUAUUCGCGUAUAAUACGCGUAT-3';<br>5'-AUACGCGUAUUUAUACGCGAUUAACGAC-3' |

|       |        |                      |           |                                                                        |     |
|-------|--------|----------------------|-----------|------------------------------------------------------------------------|-----|
| PDIA3 | ambion | s6229                | ASO2KV80  | 5'-GAACUUGAGGGGAUAACUAtt-3';<br>UAGUUAUCCCUCAAGUUGCt-3'                | 5'- |
| PDIA3 | ambion | s6227                | ASO2KV8R  | 5'-GGAAUAGUCCCAUUAGCAAtt-3';<br>UUGUAAUGGGACUAUUCCT-3'                 | 5'- |
| PDIA4 | IDT    | hs.Ri.PDIA4.1<br>3.1 | 445103958 | 5'-GGUGUUUCACUUCAUGCUCUGAATA-3';<br>5'-UAUUCAGAGCAUGAAGUGAAACACCAA-3'  |     |
| PDIA4 | IDT    | hs.Ri.PDIA4.1<br>3.2 | 445103959 | 5'-ACCUGAGAGAAGAUUACAAAUUUCA-3';<br>5'-UGAAAUUUGUAAUCUUCUCUCAGGUUG-3'  |     |
| PDIA6 | IDT    | hs.Ri.PDIA6.1<br>3.1 | 455645784 | 5'-GGAUGCUACAGUCAAUUCAGGUUCTG-3';<br>5'-CAGAACCUGAUUGACUGUAGCAUCCAC-3' |     |
| PDIA6 | IDT    | hs.Ri.PDIA6.1<br>3.2 | 455645785 | 5'-AGGCUUUCUAUGUUGUUAUGAACC-3';<br>5'-GGUUCAUUAACAACAUAGAAAGCCUUG-3'   |     |
| PH4B  | IDT    | hs.Ri.P4HB.13<br>.1  | 448986282 | 5'-AUUUCUAUUUCACAAUCGAAUUGAA-3';<br>5'-UUCAAUUCGAUUGUGAAAUAGAAUUGC-3'  |     |
| PH4B  | IDT    | hs.Ri.P4HB.13<br>.2  | 448986283 | 5'-AGAUGAACUGUAAUACGCAAAGCCA-3';<br>5'-UGGCUUUGCGUAUUACAGUUCAUUUU-3'   |     |
| STUB1 | IDT    | hs.Ri.STUB1.1<br>3.1 | 544882524 | 5'-AGGUUAUUGACGCAUUCUUCUGA-3';<br>5'-UCAGAGAUGAUGCGUCAUAACCUCC-3'      |     |
| STUB1 | IDT    | hs.Ri.STUB1.1<br>3.2 | 544882525 | 5'-GACGAGCUUUUUUCUCAGGUGGATG-3';<br>5'-CAUCCACCUGAGAAAAAAGCUCGUCCA-3'  |     |

|        |     |                   |           |                                                                        |
|--------|-----|-------------------|-----------|------------------------------------------------------------------------|
| TRIM21 | IDT | hs.Ri.TRIM21.13.2 | 544882526 | 5'-GUCCACUGAAUAUUGGAUCACAAGG-3';<br>5'-CCUUGUGAUCCAAUAUUCAGUGGACAG-3'  |
| UBR4   | IDT | hs.Ri.UBR4.13.1   | 530799933 | 5'-GUCAACUGGAUAAAGGAUCACCUTA-3';<br>5'-UAAGGUGAUCCUUUAUCCAGUUGACAU-3'  |
| UBR4   | IDT | hs.Ri.UBR4.13.2   | 530799934 | 5'-GAGAAGCAGUACGAGCCAUUCCUACT-3';<br>5'-AGUAGAAUGGCUCGUACUGCUUCUCAU-3' |

Supplementary **Table 8.** Primers used in this study.

| Gene                 | Forward primer                                                                | Reverse primer                                                                  |
|----------------------|-------------------------------------------------------------------------------|---------------------------------------------------------------------------------|
| nsp14-FLAG Cys11Ala  | 5'-<br>GAGAACGTGACCGGTCTG<br>TTCAAGGACGCCTCTAAG<br>GTCATCACTGGTCTGCAC<br>C-3' | 5'-<br>GAGCCTGAGTTGGGTGCA<br>GACCAGTGATGACCTTAGA<br>GGCGTCCTTGAACAGACC<br>GG-3' |
| nsp14-FLAG Cys39Ala  | 5'-<br>CCAAGTTCAAGACCGAAG<br>GACTGGCCGTGGACATCC<br>CCGGTATCCCTAAG-3'          | 5'-<br>GTAGGTCATGTCCTTAGGG<br>ATACCGGGGATGTCCACG<br>GCCAGTCCTTCGG-3'            |
| nsp14-FLAG Cys94Ala  | 5'-<br>GATCGGTTTCGACGTGGA<br>AGGTGCCCACGCTACCAG<br>AGAAGCTGTGGG-3'            | 5'-<br>GCAGGTTGGTACCCACAG<br>CTTCTCTGGTAGCGTGGG<br>CACCTTCCACGTC-3'             |
| nsp14-FLAG Cys208Ala | 5'-<br>GTGAAGATCGGACCCGAA<br>CGTACCTGCGCCCTGTGC<br>GACAGACGTGCC-3'            | 5'-<br>GAGAAGCAGGTGGCACGT<br>CTGTGCGACAGGGCGCAG<br>GTACGTTCTGGG-3'              |
| nsp14-FLAG Cys208Val | 5'-<br>GTGAAGATCGGACCCGAA<br>CGTACCTGCGCCCTGTGC<br>GACAGACGTGCC-3'            | 5'-<br>GAGAAGCAGGTGGCACGT<br>CTGTGCGACAGGGCGCAG<br>GTACGTTCTGGG-3'              |
| nsp14-FLAG Cys309Ala | 5'-<br>CAACGCCGCTGCCAGAAA                                                     | 5'-<br>GCTGCACCTTTCTGGCAGC                                                      |

|                                                    |                                                                  |                                                               |
|----------------------------------------------------|------------------------------------------------------------------|---------------------------------------------------------------|
|                                                    | GGTGCAGCACATGGTG-3'                                              | GGCGTTGATCTTCAGTT-3'                                          |
| nsp14-FLAG Cys356Ala                               | 5'-<br>GGAGTGGAAGTTCTACGA<br>CGCTCAGCCCGCCTCCGA<br>CAAGGCCTAC-3' | 5'-<br>GTTCTCGATCTTGTAGGC<br>CTTGTGCGAGGCGGGCTG<br>AGCGTAG-3' |
| nsp14-FLAG Cys414Ala                               | 5'-<br>CCTGCCCCGGAGCCGACGG<br>CGGATCCCTGTACGTGAA<br>C-3'         | 5'-<br>CGCCGTGCGCTCCGGGCA<br>GGTTCAGGTTGCTCAGCAC<br>ACG-3'    |
| nsp14-FLAG Cys216Ala                               | 5'-<br>GACGTGCCACCGCCTTCT<br>CCACCGCCTCTGACACC-3'                | 5'-<br>GGCGGTGGAGAAGGCGGT<br>GGCACGTCTGTGCGACAG<br>G-3'       |
| nsp14-FLAG Cys272Ala                               | 5'-<br>CCACGTGGCTTCTGCCGA<br>CGCTATCATGACCAGATG<br>CCTG-3'       | 5'-<br>GTCATGATAGCGTCGGCA<br>GAAGCCACGTGGGCGTTA<br>CCG-3'     |
| nsp14-FLAG Cys279Ala                               | 5'-<br>CATGACCAGAGCCCTGGC<br>CGTGCACGAGTGC-3'                    | 5'-<br>CGTGCACGGCCAGGGCTC<br>TGGTCATGATAGCGTCGC-<br>3'        |
| nsp14-FLAG Cys340Ala                               | 5'-GCT ATC AAG GCC GTG<br>CCC CAG GCC GAC GTG G-<br>3'           | 5'-GGG GCA CGG CCT TGA<br>TAG CCT TTG GGT TTC CG-<br>3'       |
| G3BP1 (for cloning into<br>plasmid #15)            | 5'-<br>TAAGGGAAGCTTGGGCCG<br>CCACCATGGTGATGGAGA<br>AGCCT-3'      | 5'-<br>TGCTTAGCGGCCCTGCCG<br>TGGCGCAAGCC-3'                   |
| G3BP1-TurboID-V5 (for<br>cloning into plasmid #18) | 5'-<br>GAAGGGGGATCCCGCCAC<br>CATGGTGATGGAGAAGCC<br>TAG-3'        | 5'-<br>GGCGTGATCGATTTAGGTG<br>CTGTCCAGGCCCAGCAG-3'            |
| G3BP1-TurboID-V5 (for<br>cloning into plasmid #20) | 5'-<br>GAAGGGTCTAGACGCCAC<br>CATGGTGATGGAGAAGCC<br>T-3'          | 5'-<br>GTGCGGTGTACATTAGGTG<br>CTGTCCAGGCCCAG-3'               |

Supplementary **Table 9**. Plasmids used in this study.

| Plasmid number | Figure | Plasmid                                               | Source                                                      | Resistance                     |
|----------------|--------|-------------------------------------------------------|-------------------------------------------------------------|--------------------------------|
| 1              | 1      | pDONR223 SARS-CoV-2 NSP14_nostop                      | Addgene (#149316)                                           | Spectinomycin                  |
| 2              | 1      | pDONR223 SARS-CoV-2 NSP9_nostop                       | Addgene (#149312)                                           | Spectinomycin                  |
| 3              | 1      | pDONR223 SARS-CoV-2 NSP10_nostop                      | Addgene (#149313)                                           | Spectinomycin                  |
| 4              | 1      | pDONR223 SARS-CoV-2 NSP16_nostop                      | Addgene (#149318)                                           | Spectinomycin                  |
| 5              | 1      | GateWay compatible C-terminal FLAG destination vector | Gift from T Wucherpfennig                                   | Chloramphenicol and ampicillin |
| 6              | 1-6    | nsp14-FLAG expression vector                          | GateWay cloned with plasmid #1 (donor) and #5 (destination) | Ampicillin                     |
| 7              | 1      | nsp9-FLAG expression vector                           | GateWay cloned with plasmid #2 (donor) and #5 (destination) | Ampicillin                     |
| 8              | 1      | nsp10-FLAG expression vector                          | GateWay cloned with plasmid #3 (donor) and #5 (destination) | Ampicillin                     |
| 9              | 1, 6   | nsp16-FLAG expression vector                          | GateWay cloned with plasmid #4 (donor) and #5 (destination) | Ampicillin                     |
| 10             | 3      | pRK5-HA-Ubiquitin-WT                                  | Addgene (#17608)                                            | Ampicillin                     |
| 11             | 3      | Nsp14_K_R ENTR (pMB1)                                 | Twist Biosciences (custom sequence)                         | Kanamycin                      |
| 12             | 3      | Nsp14_K_R expression vector                           | GateWay cloned with plasmid #11                             | Ampicillin                     |



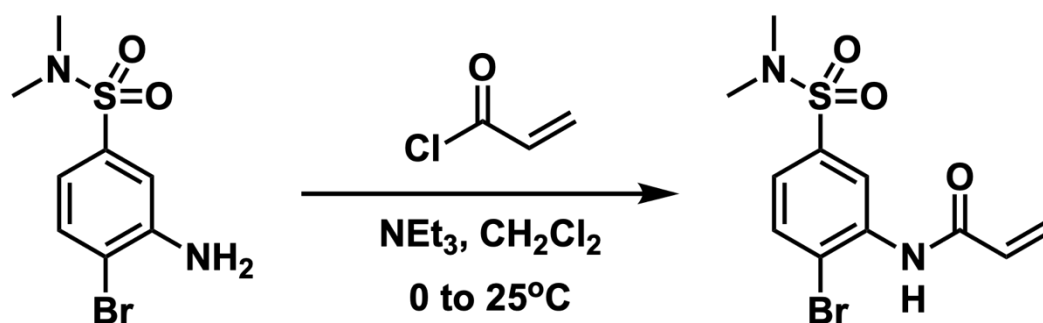
|    |     |                                    |                                                                     |            |
|----|-----|------------------------------------|---------------------------------------------------------------------|------------|
|    |     |                                    | (donor) and #5<br>(destination)                                     |            |
| 13 | 4   | PDIA3-bio-His                      | Addgene (#52070)                                                    | Ampicillin |
| 14 | 6   | pENTR4_G3BP1                       | Addgene (#127104)                                                   | Kanamycin  |
| 15 | 6   | C1(1-29)-TurboID-V5_pCDNA3         | Addgene (#107173)                                                   | Ampicillin |
| 16 | 6   | G3BP1-TurboID-V5                   | Restriction cloned<br>G3BP1 from<br>plasmid #14 into<br>plasmid #15 | Ampicillin |
| 17 | S27 | HA-HECTD1                          | Gift from Dr. Irene<br>Zohn at Children's<br>National               | Ampicillin |
| 18 | 6   | Empty Piggybac                     | Gift from Melody Li<br>at UCLA                                      | Ampicillin |
| 19 | 6   | Piggybac transposase               | Gift from Melody Li<br>at UCLA                                      | Ampicillin |
| 20 | S32 | pLenti CMV GFP Hygro (656-4)       | Addgene (#17446)                                                    | Ampicillin |
| 21 | 6   | G3BP1-TurboID-V5_Piggybac          | Restriction cloned<br>G3BP1 from<br>plasmid #16 into<br>plasmid #18 | Ampicillin |
| 22 | 6   | G3BP1-TurboID-<br>V5_pCMVHygro     | Restriction cloned<br>G3BP1 from<br>plasmid #16 into<br>plasmid #20 | Ampicillin |
| 23 | 4   | PDIA6 in pDONR221                  | DNASU<br>HsCD00451137                                               | Kanamycin  |
| 24 | 4   | pEGc1-DEST                         | DNASU<br>EvNO00534626                                               | Ampicillin |
| 25 | 4   | PDIA6-StrepTagII-His <sub>8x</sub> | Gateway cloned<br>plasmid #23 into<br>plasmid #24                   | Ampicillin |

Supplementary **Table 10**. Conditions of Liquid-chromatography (LC).

| Experiment Type                                                 | Parameter              | Condition                                                                                                                                                |
|-----------------------------------------------------------------|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|
| isoTOP-ABPP<br>-yne enrichment<br>Competition<br>SILAC<br>XRNAX | Column                 | 100 $\mu$ M ID fused silica capillary packed in-house with bulk C18 reversed phase resin (particle size, 1.9 $\mu$ m; pore size, 100 Å; Dr. Maisch GmbH) |
|                                                                 | Mobile phase           | Buffer A: water with 3% DMSO and 0.1% formic acid<br>Buffer B: 80% acetonitrile with 3% DMSO and 0.1% formic acid                                        |
|                                                                 | Gradient and flow rate | 0 – 5 min, 3 – 20% B, 220 nL/min                                                                                                                         |
|                                                                 |                        | 5 – 64 min, 20 – 47% B, 220 nL/min                                                                                                                       |
|                                                                 |                        | 64 – 70 min, 47 – 95% B, 250 nL/min                                                                                                                      |
|                                                                 | Run time               | 70 minutes                                                                                                                                               |
|                                                                 | Injection volume       | 5 $\mu$ L                                                                                                                                                |
| EN450 LFQ<br>Insoluble LFQ                                      | Column                 | 100 $\mu$ M ID fused silica capillary packed in-house with bulk C18 reversed phase resin (particle size, 1.9 $\mu$ m; pore size, 100 Å; Dr. Maisch GmbH) |
|                                                                 | Mobile phase           | Buffer A: water with 3% DMSO and 0.1% formic acid<br>Buffer B: 80% acetonitrile with 3% DMSO and 0.1% formic acid                                        |
|                                                                 | Gradient and flow rate | 0 – 3 min, 1 – 6% B, 220 nL/min                                                                                                                          |
|                                                                 |                        | 3 – 133 min, 6 - 35% B, 220 nL/min                                                                                                                       |
|                                                                 |                        | 133 – 153 min, 35 - 40% B, 220 nL/min                                                                                                                    |
|                                                                 |                        | 153 – 160 min, 40 – 95% B, 250 nL/min                                                                                                                    |
|                                                                 | Run time               | 160 minutes                                                                                                                                              |

## Supplemental Note: Synthesis of Compounds and Characterization

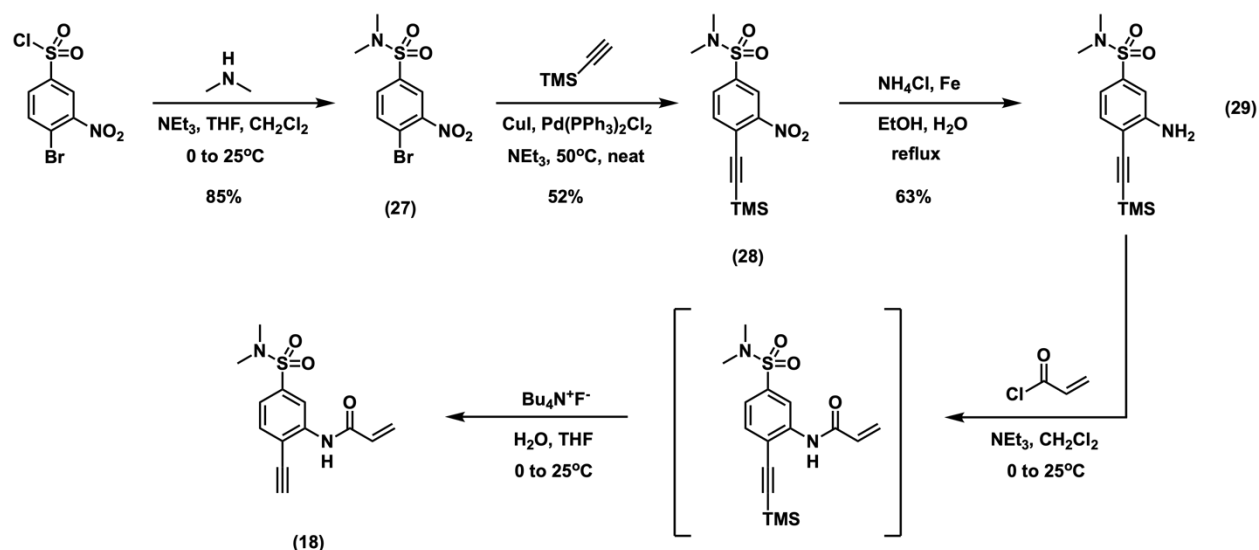
**General Methods.** All reactions were performed in dried glassware under an atmosphere of dry N<sub>2</sub> unless otherwise stated. Silica gel P60 (SiliCycle) was used for column chromatography. Plates were visualized by fluorescence quenching under UV light or by staining with iodine, KMnO<sub>4</sub>, or bromocresol green. Other reagents were purchased from Sigma-Aldrich (St. Louis, MO), Alfa Aesar (Ward Hill, MA), EMD Millipore (Billerica, MA), Fisher Scientific (Hampton, NH), Oakwood Chemical (West Columbia, SC), Combi-blocks (San Diego, CA) and Cayman Chemical (Ann Arbor, MI) and used without further purification. <sup>1</sup>H NMR and <sup>13</sup>C NMR spectra for characterization of new compounds and monitoring reactions were collected in CDCl<sub>3</sub>, CD<sub>3</sub>OD, D<sub>2</sub>O or DMSO-*d*<sub>6</sub> (Cambridge Isotope Laboratories, Cambridge, MA) on a Bruker AV 500 MHz spectrometer or Bruker AV 400 MHz in the Department of Chemistry & Biochemistry at The University of California, Los Angeles. All chemical shifts are reported in the standard notation of parts per million using the peak of residual proton signals of the deuterated solvent as an internal reference. Coupling constant units are in Hertz (Hz). Splitting patterns are indicated as follows: br, broad; s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; dd, doublet of doublets; dt, doublet of triplets. Low-resolution mass spectrometry was performed on an Agilent Technologies InfinityLab LC/MSD single quadrupole LC/MS (ESI source). High-resolution mass spectrometry was performed on a Waters LCT Premier with ACQUITY LC and autosampler (ESI source). All compounds named with prefix **KB** were purchased from MilliporeSigma or prepared according to our previous procedures<sup>3</sup>.



**(4)**

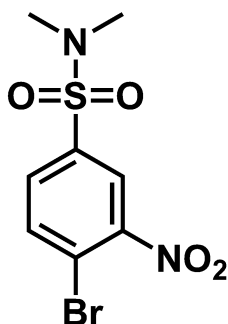
**N-(2-chloro-5-(N,N-dimethylsulfamoyl)phenyl)acrylamide (EN450 (4))**

To a solution of 3-amino-4-chloro-N,N-dimethylbenzenesulfonamide (100 mg, 1 Eq, 426  $\mu$ mol) in DCM (15 mL) was added triethylamine (86.2 mg, 2 Eq, 852  $\mu$ mol) dropwise. The reaction was then brought to 0°C and stirred for 5 minutes. Acryloyl chloride (57.8 mg, 1.5 Eq, 639  $\mu$ mol) was then added and the reaction was brought to 230 and left to stir for 16 hours. Mixture was quenched with H<sub>2</sub>O (5mL), extracted with DCM (3 x 5mL), organic layer was washed with brine (3mL), and dried with NaSO<sub>4</sub> and concentrated in vacuo. Column chromatography (eluting with 30% EtOAc in hexanes) afforded a white solid (62.0 mg, 50%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (d, J = 2.1 Hz, 1H), 7.56 (d, J = 8.3 Hz, 1H), 7.52 (d, J = 2.1 Hz, 1H), 6.50 (dd, J = 16.9, 1.0 Hz, 1H), 6.32 (dd, J = 16.9, 10.2 Hz, 1H), 5.89 (dd, J = 10.2, 1.0 Hz, 1H), 2.78 (s, 6H). All analyses were consistent with previously reported data <sup>4</sup>.



## Synthesis of EN450yne

### 4-bromo-N,N-dimethyl-3-nitrobenzenesulfonamide (27)



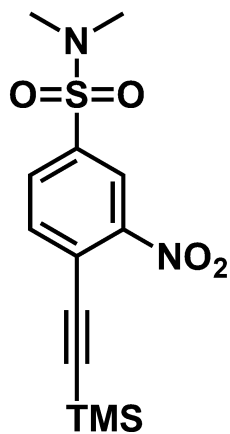
To an oven dried 25mL round-bottom flask was added dimethylamine (27.2 mg, 302  $\mu$ L, 2 molar, 1.21 Eq, 604  $\mu$ mol) in  $\text{CH}_2\text{Cl}_2$  (0.98 mL) and cooled to 0°C. Then with strong stirring, triethylamine (61.1 mg, 1.21 Eq, 604  $\mu$ mol) was added drop wise and stirred at 0°C for 50 min. 4-bromo-3-nitrobenzenesulfonyl chloride (150 mg, 1 Eq, 499  $\mu$ mol) in THF (0.43 mL) and  $\text{CH}_2\text{Cl}_2$  (0.85 mL) were then added, reaction was brought to room temperature and stirred for 5 minutes. Mixture was then diluted with  $\text{CH}_2\text{Cl}_2$  (3mL), quenched with  $\text{H}_2\text{O}$  (1mL), extracted with  $\text{CH}_2\text{Cl}_2$  (3 x 1mL), dried with  $\text{NaSO}_4$ , and concentrated in vacuo. Column chromatography (eluting with 30% EtOAc in hexanes) afforded a light-yellow solid (262 mg, 85%).

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.20 (d,  $J$  = 2.1 Hz, 1H), 7.95 (d,  $J$  = 8.3 Hz, 1H), 7.80 (dd,  $J$  = 8.4, 2.1 Hz, 1H), 2.79 (s, 6H).

$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  137.36, 136.31, 131.45, 124.63, 119.61, 37.94.

HRMS (ESI-MS)  $m/z$ : Calculated  $[\text{M}+\text{H}]^+ = 308.9536$ , Found  $[\text{M}+\text{H}]^+ = 308.9539$

### N,N-dimethyl-3-nitro-4-((trimethylsilyl)ethynyl)benzenesulfonamide (28)



To an oven dried round-bottom flask was added  $\text{Pd(PPh}_3)_2\text{Cl}_2$  (3.4 mg, 0.05 equiv, 4.9  $\mu\text{mol}$ ),  $\text{CuI}$  (0.92 mg, 0.05 equiv, 4.9  $\mu\text{mol}$ ), and 4-bromo-N,N-dimethyl-3-nitrobenzenesulfonamide (30 mg, 1 equiv, 97  $\mu\text{mol}$ ). The reaction flask was purged with argon followed by addition of dry triethylamine (0.15 mL). After a few seconds of stirring the reaction mixture turned yellow. Trimethylsilylacetylene (19 mg, 27  $\mu\text{L}$ , 2 equiv, 0.19 mmol) was then added dropwise turning the mixture black. Upon complete addition the reaction

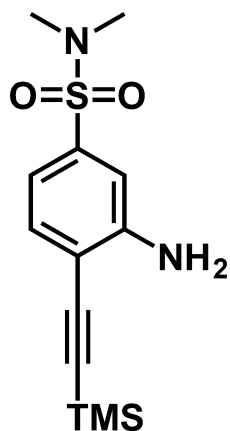
mixture was heated to 50 °C and let stir. After 2 hours the reaction mixture was cooled to ambient temperature and filtered over cotton and celite. The resulting flask and solids were rinsed with hexanes and the combined filtrate was concentrated down. The crude mixture was purified using flash column chromatography (0 to 70% ethyl acetate in hexanes) to yield the desired product as a beige solid (16mg, 52%).

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.37 (d,  $J$  = 1.8 Hz, 1H), 7.92 (dd,  $J$  = 8.2, 1.8 Hz, 1H), 7.81 (dd,  $J$  = 8.1, 0.4 Hz, 1H), 2.77 (s, 6H), 0.30 (s, 9H).

$^{13}\text{C}$  NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  150.16, 136.68, 135.94, 130.91, 123.64, 122.37, 108.88, 97.95, 37.84.

HRMS (ESI-MS)  $m/z$ : Calculated  $[\text{M}+\text{H}]^+ = 327.0829$ , Found  $[\text{M}+\text{H}]^+ = 327.0838$

### 3-amino-N,N-dimethyl-4-((trimethylsilyl)ethynyl)benzenesulfonamide (29)



To an oven dried round bottom flask was added ammonium chloride (31 mg, 6 equiv, 0.59 mmol), Iron Powder (33 mg, 6 equiv, 0.59 mmol), EtOH (0.75 mL), Water (0.75 mL), and heated to 60°C for 30 minutes.

After 30 minutes, N,N-dimethyl-3-nitro-4-((trimethylsilyl)ethynyl)benzenesulfonamide (32 mg, 1 equiv, 98 μmol) dissolved in EtOH (0.75 mL) and Water (0.75 mL) was added and mixture was refluxed at 80 °C. After 24 hours the mixture was quenched with H<sub>2</sub>O

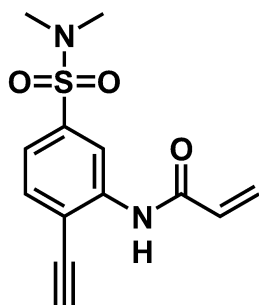
(5mL), extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x 5mL). The combined organic layers were then washed with brine (1x3mL), dried with NaSO<sub>4</sub> and concentrated in vacuo. Crude material was purified using flash column chromatography (20 to 50% ethyl acetate in hexanes) to afford the desired product as a white solid (18.0 mg, 62%).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.41 (d, J = 8.1 Hz, 1H), 7.07 (s, 1H), 7.01 (dd, J = 8.0, 1.5 Hz, 1H), 2.69 (s, 6H), 0.28 (s, 9H).

<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 136.09, 132.79, 116.35, 112.72, 38.00.

HRMS (ESI-MS) *m/z*: Calculated [M+H]<sup>+</sup> = 297.1087 , Found [M+H]<sup>+</sup> = 297.1146

#### N-(5-(N,N-dimethylsulfamoyl)-2-ethynylphenyl)acrylamide (EN450yne (18))



To an oven dried round-bottom flask containing 3-amino-N,N-dimethyl-4-((trimethylsilyl)ethynyl)benzenesulfonamide (19mg, 1 equiv, 63μmol) was added dry CH<sub>2</sub>Cl<sub>2</sub> (1mL). The flask was then capped and purged with argon followed by cooling to 0°C and addition of triethylamine (35μL, 4 equiv,

0.252 mmol). Next, acryloyl chloride (15μL, 3 equiv, 0.189mmol) was added and the reaction mixture allowed to warm to ambient temperature and stir for 16h. Upon completion the reaction mixture was quenched with water (1mL) and extracted with CH<sub>2</sub>Cl<sub>2</sub> (3x2mL). The combined

organic layers were dried over sodium sulfate, concentrated down, and filtered through a silica plug using 30% ethyl acetate in hexanes. The crude product was used directly for the next step.

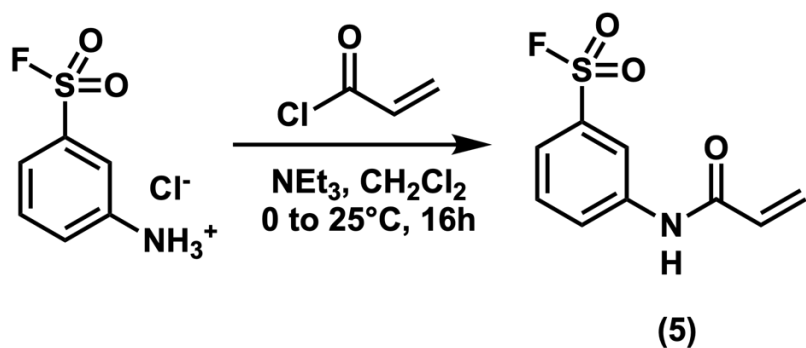
To the crude N-(5-(N,N-dimethylsulfamoyl)-2-((trimethylsilyl)ethynyl)phenyl) acrylamide in a dry round-bottom flask was added THF (0.7 mL) and water (2.3 mg, .0023 mL, 2.0 equiv, 0.13 mmol). The flask was capped, purged with argon, and cooled to 0 °C followed by dropwise addition of TBAF (16 mg, .063 mL, 1 molar, 1.0 equiv, 63  $\mu$ mol). Reaction mixture was then allowed to stir at 0°C for 10 minutes followed by warming to ambient temperature for 5 minutes. After 5 minutes, the mixture was quenched with H<sub>2</sub>O (2mL), extracted with EtOAc (3 x 5mL), organic layer, and dried with NaSO<sub>4</sub> and concentrated in vacuo. The crude material was purified by flash column chromatography (0 to 30% ethyl acetate in hexanes) to afford the desired compound as an off-white solid. (5mg, 28% over two steps)

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  8.92 (d, J = 1.7 Hz, 1H), 8.09 (s, 1H), 7.62 (d, J = 8.0 Hz, 1H), 7.49 (dd, J = 8.1, 1.8 Hz, 1H), 6.48 (dd, J = 16.9, 1.1 Hz, 1H), 6.30 (dd, J = 16.9, 10.3 Hz, 1H), 5.87 (dd, J = 10.2, 1.1 Hz, 1H), 3.70 (s, 1H), 2.78 (s, 6H).

<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)  $\delta$  163.50, 139.92, 137.30, 132.65, 130.73, 129.06, 122.53, 118.43, 114.89, 87.46, 77.94, 38.04.

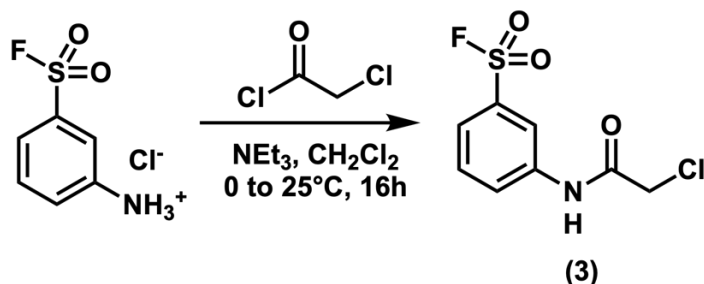
HRMS (ESI-MS) *m/z*: Calculated [M+H]<sup>+</sup> = 279.0798 , Found [M+H]<sup>+</sup> = 279.0830





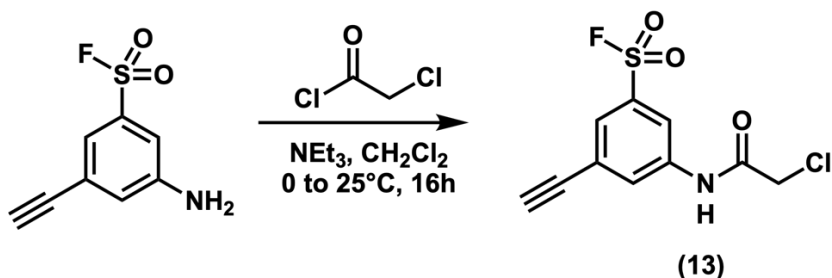
### 3-acrylamidobenzenesulfonyl fluoride (JC17 (5))

Following the previously described procedure<sup>5</sup>, to a solution of 3-aminobenzenesulfonyl fluoride hydrochloride (100 mg, 0.47 mmol, 1 equiv.) in 2 mL DCM was added triethylamine (198  $\mu$ L, 1.42 mmol, 3 equiv.), acryloyl chloride (77  $\mu$ L, 0.94 mmol, 2 equiv.) in an ice bath. Next the reaction was allowed to react for 3h at RT. The reaction was quenched with 30 mL saturated NaHCO<sub>3</sub> (aq), extracted with DCM (2  $\times$  20 mL), and was concentrated and purified by silica column chromatography (2:1 hexanes/ EtOAc) to afford product (88.9 mg, 82%) as a colorless solid. All analyses were consistent with previously reported data.



### 3-(2-chloroacetamido)benzenesulfonyl fluoride (JC19 (3))

Following the previously described procedure<sup>5</sup>, to a solution of 3-aminobenzenesulfonyl fluoride hydrochloride (100 mg, 0.47 mmol, 1 equiv.) in 2 mL DCM was added Triethylamine (198  $\mu$ L, 1.42 mmol, 3 equiv.), 2-chloroacetyl chloride (75  $\mu$ L, 0.94 mmol, 2 equiv.) in an ice bath. Next the reaction was allowed to react for 3 h at RT. The reaction was quenched with 30 mL saturated NaHCO<sub>3</sub> (aq), extracted with DCM (2  $\times$  20 mL), and was concentrated and purified by silica column chromatography (2:1 hexanes/ EtOAc) to afford product (108.1 mg, 91%) as a colorless solid. All analyses were consistent with previously reported data.



### 3-(2-chloroacetamido)-5-ethynylbenzenesulfonyl fluoride (JC19yne (13))

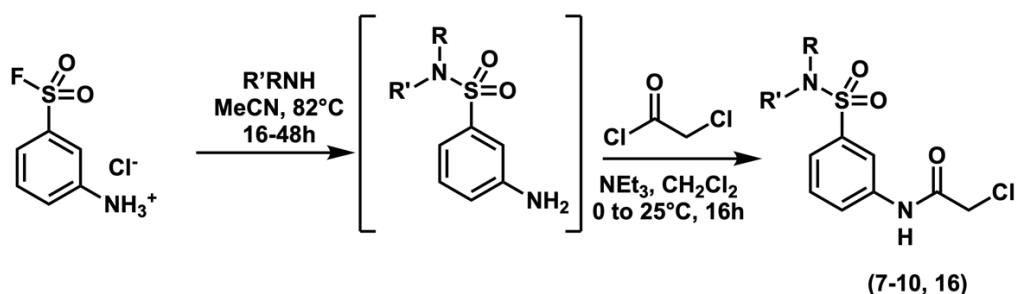
To a solution of 3-amino-5-ethynylbenzenesulfonyl fluoride hydrochloride (50 mg, 0.25 mmol, 1 equiv.) in 2 mL DCM was added triethylamine (104  $\mu$ L, 0.75 mmol, 3 equiv.), 2-chloroacetyl chloride (40  $\mu$ L, 0.50 mmol, 2 equiv.) in an ice bath. Next the reaction was allowed to react for 3 h at RT. The reaction was quenched with 30 mL saturated NaHCO<sub>3</sub> (aq), extracted with DCM (2  $\times$  20 mL), and was concentrated and purified by silica column chromatography (2:1 hexanes/ EtOAc) to afford product (63 mg, 92%) as an off-white solid.

$^1\text{H}$  NMR (400 MHz, Chloroform-*d*)  $\delta$  8.54 (s, 1H), 8.23 (t,  $J$  = 2.0 Hz, 1H), 8.10 (t,  $J$  = 1.8 Hz, 1H), 7.87 (t,  $J$  = 1.6 Hz, 1H), 4.24 (s, 2H), 3.27 (s, 1H).

$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  164.60, 138.39, 134.54, 134.29, 129.47, 127.89, 125.46, 119.31, 81.37, 80.39, 42.83.

$^{19}\text{F}$  NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  65.67.

HRMS (ESI-MS)  $m/z$ : Calculated  $[\text{M}-\text{H}]^-$  = 273.9741, Found  $[\text{M}-\text{H}]^-$  = 273.9761

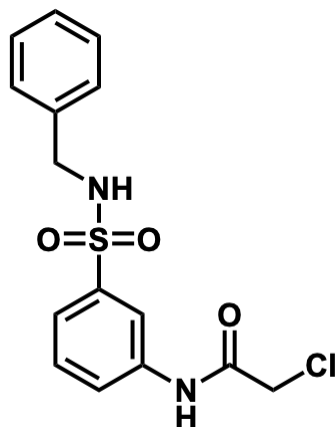


**General procedure for the synthesis of JC19 sulfonamide analogues (NB177, NB179, and NB001, NB92, NB82):**

Following a previously published procedure<sup>6</sup>, to a  $\mu$ wave vial was added 3-aminobenzenesulfonyl fluoride hydrochloride (1 Eq) and DMAP (0.05 Eq). The vial was capped and purged with argon followed by addition of dry acetonitrile (0.36M), triethylamine (2.1 Eq), and the specified primary or secondary amine (1.05 Eq). The reaction mixture was then left to stir at reflux (82°C) for 16-48h. Upon completion as determined by TLC (30% ethyl acetate in hexanes) the reaction mixture was concentrated down and ran through a silica plug (1:1 ethyl acetate:hexanes). These compounds were verified by LC-MS analysis and used directly for the following step.

The resulting anilines were then subjected to chloroacetylation using the described procedure above for **JC19** with the only deviation being the use of 2.0 Eq of triethylamine.

***N*-(3-(*N*-benzylsulfonyl)phenyl)-2-chloroacetamide (NB177 (8))**

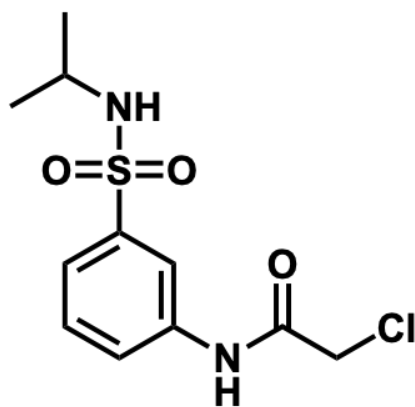


Prepared according to the general procedure with benzylamine and obtained as a yellow oil (7.5mg, 7% over two steps).

$^1\text{H}$  NMR (400 MHz, Chloroform- $d$ )  $\delta$  8.46 (br, 1H), 7.97 (t,  $J$  = 2.0 Hz, 1H), 7.89 (ddd,  $J$  = 8.2, 2.2, 1.0 Hz, 1H), 7.65 (ddd,  $J$  = 7.9, 1.8, 1.0 Hz, 1H), 7.49 (t,  $J$  = 8.0 Hz, 1H), 7.31 – 7.15 (m, 5H), 5.00 (t,  $J$  = 6.1 Hz, 1H), 4.19 (s, 2H), 4.17 (d,  $J$  = 6.1 Hz, 2H).

$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  164.36, 141.13, 137.77, 136.21, 130.20, 128.86, 128.11, 124.11, 123.52, 118.54, 47.53, 42.94.

HRMS (ESI-MS)  $m/z$ : Calculated  $[\text{M}-\text{H}]^-$  = 337.0414 , Found  $[\text{M}-\text{H}]^-$  = 337.0389



**2-chloro-*N*-(3-(*N*-isopropylsulfamoyl)phenyl)acetamide (NB179 (9))**

Prepared according to the general procedure with isopropylamine and obtained as a yellow oil (11mg, 7% over two steps)

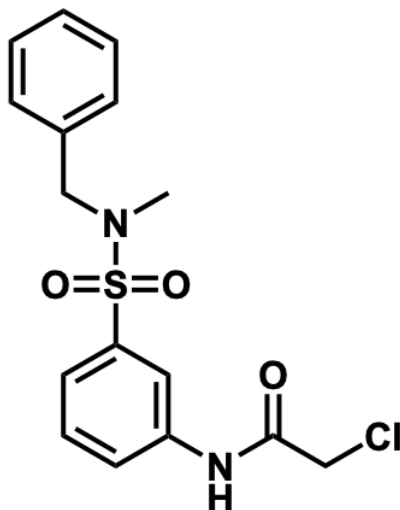
$^1\text{H}$  NMR (400 MHz, Chloroform- $d$ )  $\delta$  8.39 (br, 1H), 8.04 (t,  $J$  = 2.0 Hz, 1H), 7.86 (ddd,  $J$  = 8.1, 2.2, 1.0 Hz, 1H), 7.68 (ddd,  $J$  = 7.8, 1.8, 1.0 Hz, 1H), 7.51 (t,  $J$  = 8.0 Hz, 1H), 4.33 (d,  $J$  = 7.7 Hz, 1H), 4.22 (s, 2H), 3.51 (dp,  $J$  = 7.6, 6.5 Hz, 1H), 1.11 (d,  $J$

= 6.5 Hz, 6H).

$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  130.19, 123.86, 123.50, 118.43, 46.49, 42.93, 23.99.

HRMS (ESI-MS)  $m/z$ : Calculated  $[\text{M}-\text{H}]^-$  = 289.0414 , Found  $[\text{M}-\text{H}]^-$  = 289.0524

***N*-(3-(*N*-benzyl-*N*-methylsulfamoyl)phenyl)-2-chloroacetamide (NB001 (10))**



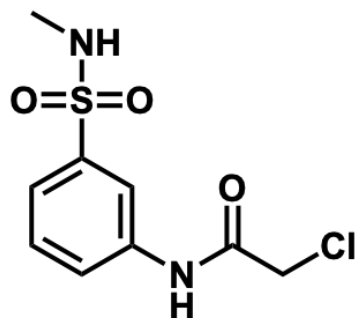
Prepared according to the general procedure with N-methyl-N-benzylamine and obtained as a white solid (55mg, 70% over two steps).

$^1\text{H}$  NMR (400 MHz, Chloroform-*d*)  $\delta$  8.58 (br, 1H), 7.97 (t,  $J$  = 1.9 Hz, 1H), 7.94 (ddd,  $J$  = 8.0, 2.2, 1.1 Hz, 1H), 7.61 (dt,  $J$  = 7.9, 1.3 Hz, 1H), 7.53 (t,  $J$  = 7.9 Hz, 1H), 7.34 – 7.27 (m, 5H), 4.21 (s, 2H), 4.16 (s, 2H), 2.62 (s, 3H).

$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  164.59, 138.35, 137.91, 135.48, 130.19, 128.80, 128.47, 128.11, 124.28, 123.74, 118.89, 54.30, 42.98, 34.56.

HRMS (ESI-MS)  $m/z$ : Calculated  $[\text{M}-\text{H}]^-$  = 351.0570 , Found  $[\text{M}-\text{H}]^-$  = 351.0521

#### 2-chloro-*N*-(3-(*N*-methylsulfamoyl)phenyl)acetamide (NB92 (7))



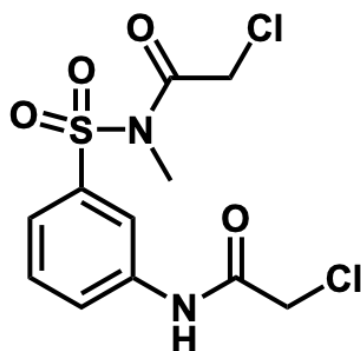
Prepared according to the general procedure with methylamine hydrochloride and 3.1 Eq of triethylamine and obtained as an orange solid (123mg, 40% over two steps).

$^1\text{H}$  NMR (400 MHz, Chloroform-*d*)  $\delta$  8.41 (br, 1H), 8.02 (t,  $J$  = 2.0 Hz, 1H), 7.87 (ddd,  $J$  = 8.1, 2.2, 1.0 Hz, 1H), 7.66 (ddd,  $J$  = 7.8, 1.8, 1.0 Hz, 1H), 7.53 (t,  $J$  = 8.0 Hz, 1H), 4.47 (d,  $J$  = 5.6 Hz, 1H), 4.22 (s, 2H), 2.70 (d,  $J$  = 5.4 Hz, 3H).

$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  164.35, 140.04, 137.73, 130.22, 124.21, 123.80, 118.68, 42.93, 29.57.

HRMS (ESI-MS)  $m/z$ : Calculated  $[\text{M}-\text{H}]^-$  = 261.0101 , Found  $[\text{M}-\text{H}]^-$  = 261.0344

#### 2-chloro-*N*-((3-(2-chloroacetamido)phenyl)sulfonyl)-*N*-methylacetamide (NB82 (16))



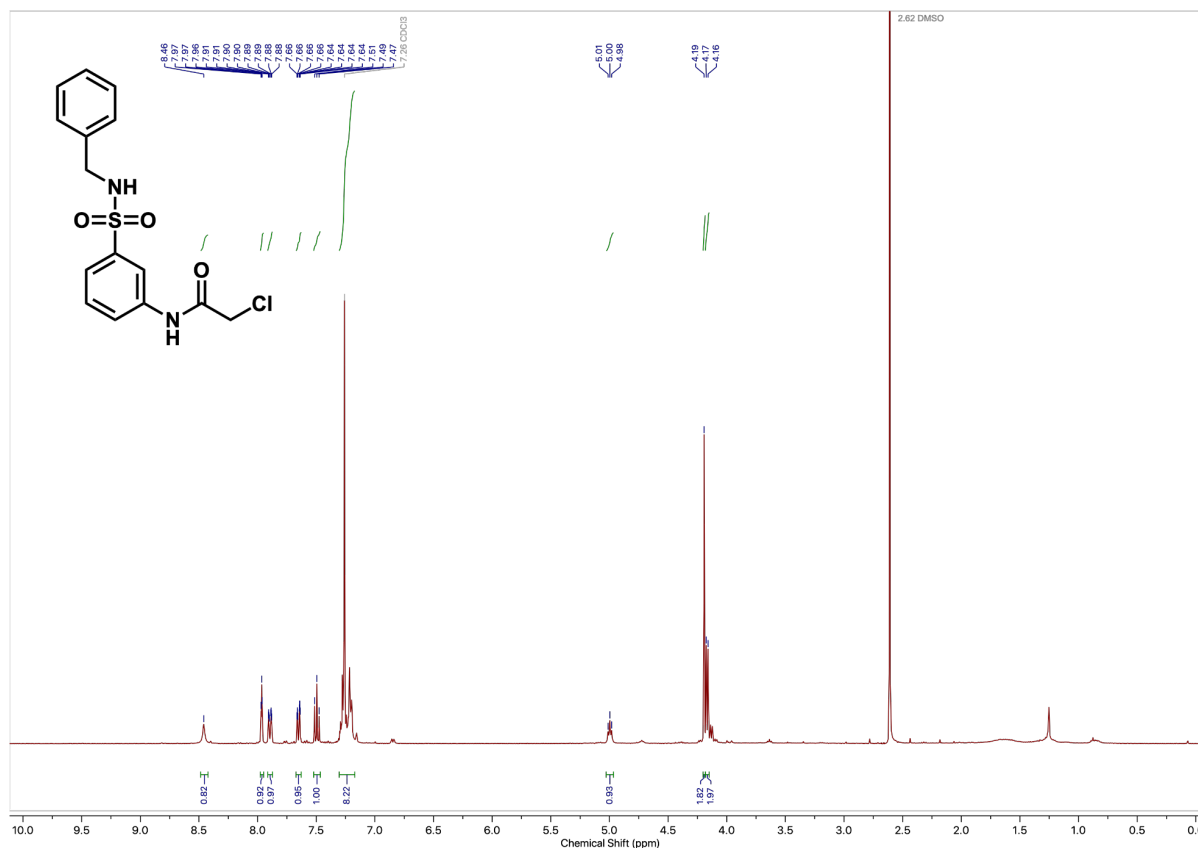
Prepared according to the general procedure with methylamine hydrochloride and 3.1 Eq of triethylamine and obtained as an orange solid (43mg, 11% over two steps).

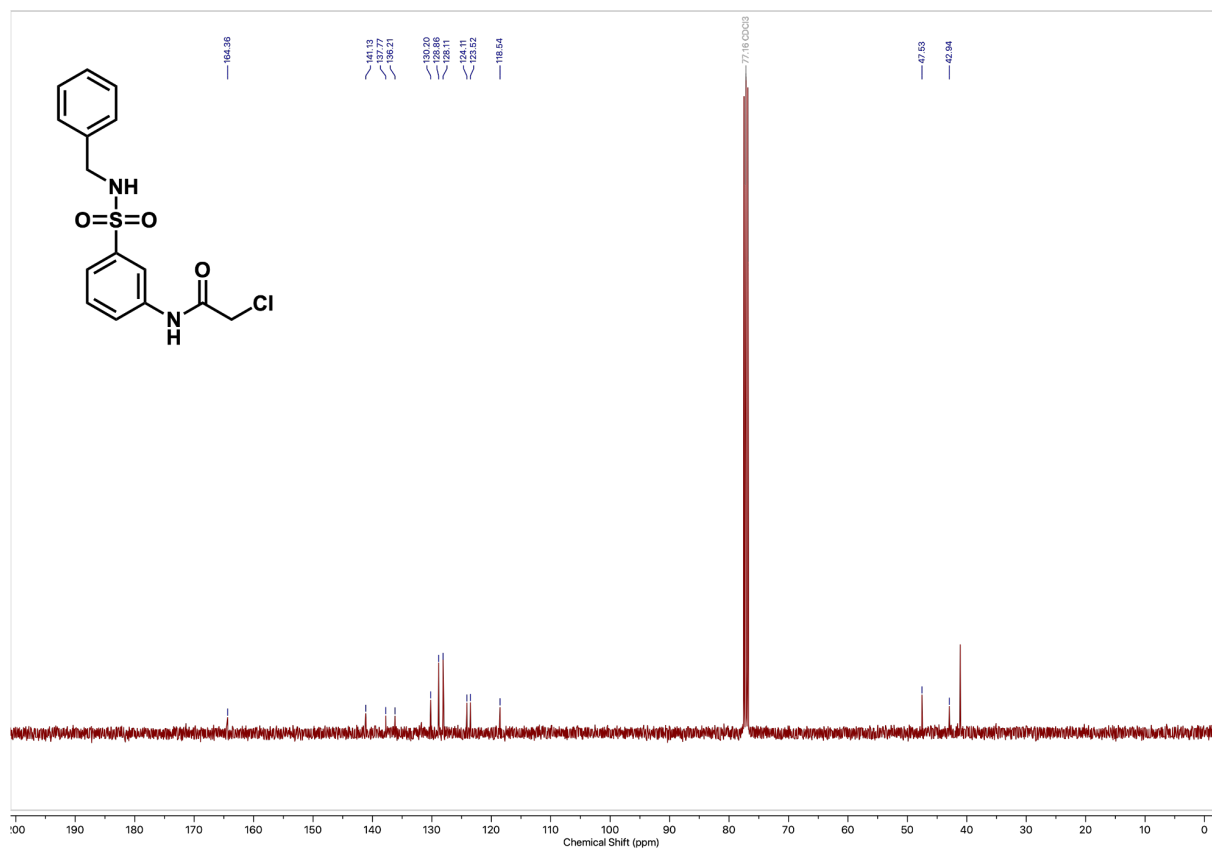
$^1\text{H}$  NMR (400 MHz, Chloroform-*d*)  $\delta$  8.53 (br, 1H), 8.22 (t,  $J$  = 2.0 Hz, 1H), 7.88 (ddd,  $J$  = 8.1, 2.2, 1.0 Hz, 1H), 7.69 (ddd,  $J$  = 7.9, 1.9, 1.0 Hz, 1H), 7.58 (t,  $J$  = 8.0 Hz, 1H), 4.60 (s, 2H), 4.22 (s, 2H), 3.30 (s, 3H).

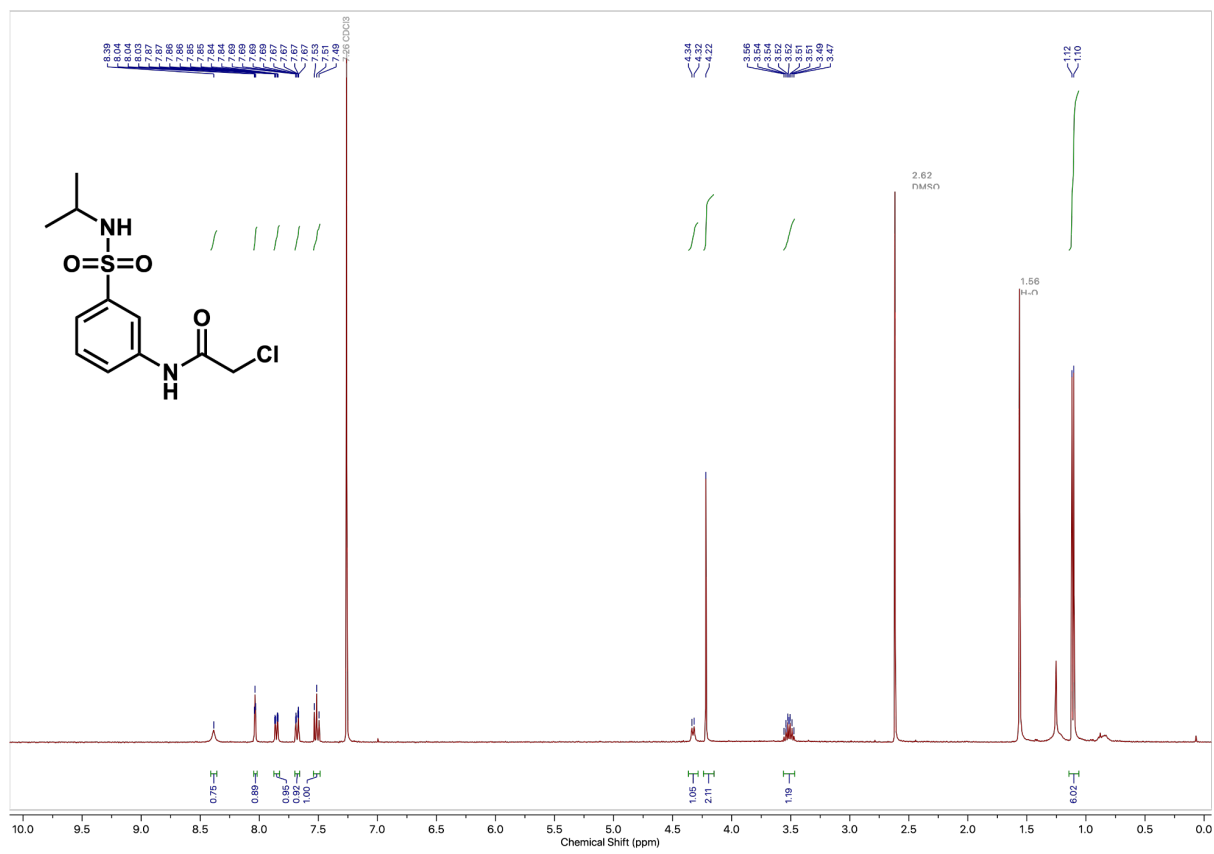
$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  166.43, 164.33, 138.88, 138.13, 130.46, 125.31, 123.62, 118.65, 44.33, 42.77, 33.67.

HRMS (ESI-MS)  $m/z$ : Calculated  $[\text{M}-\text{H}]^-$  = 336.9822, Found  $[\text{M}-\text{H}]^-$  = 336.9821

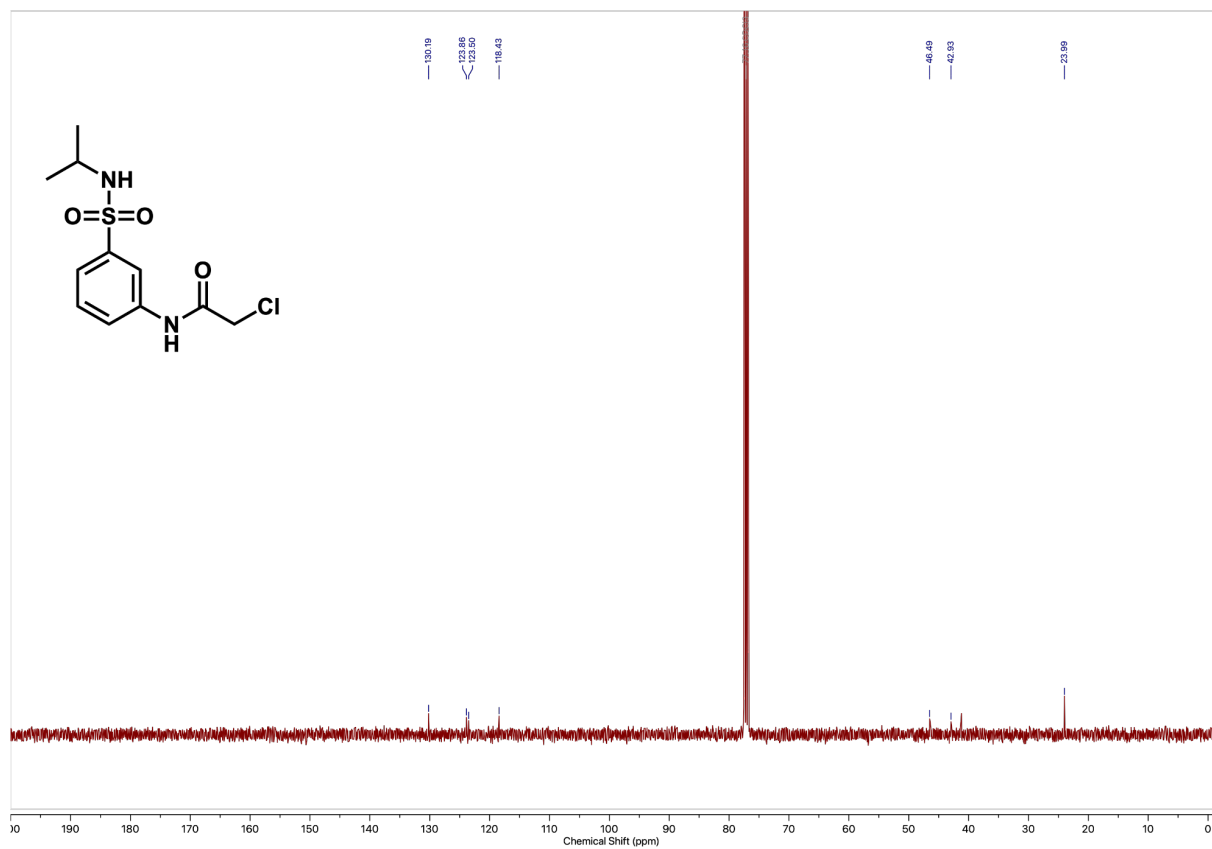
## NMR Spectra

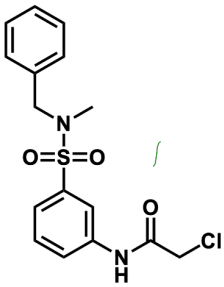


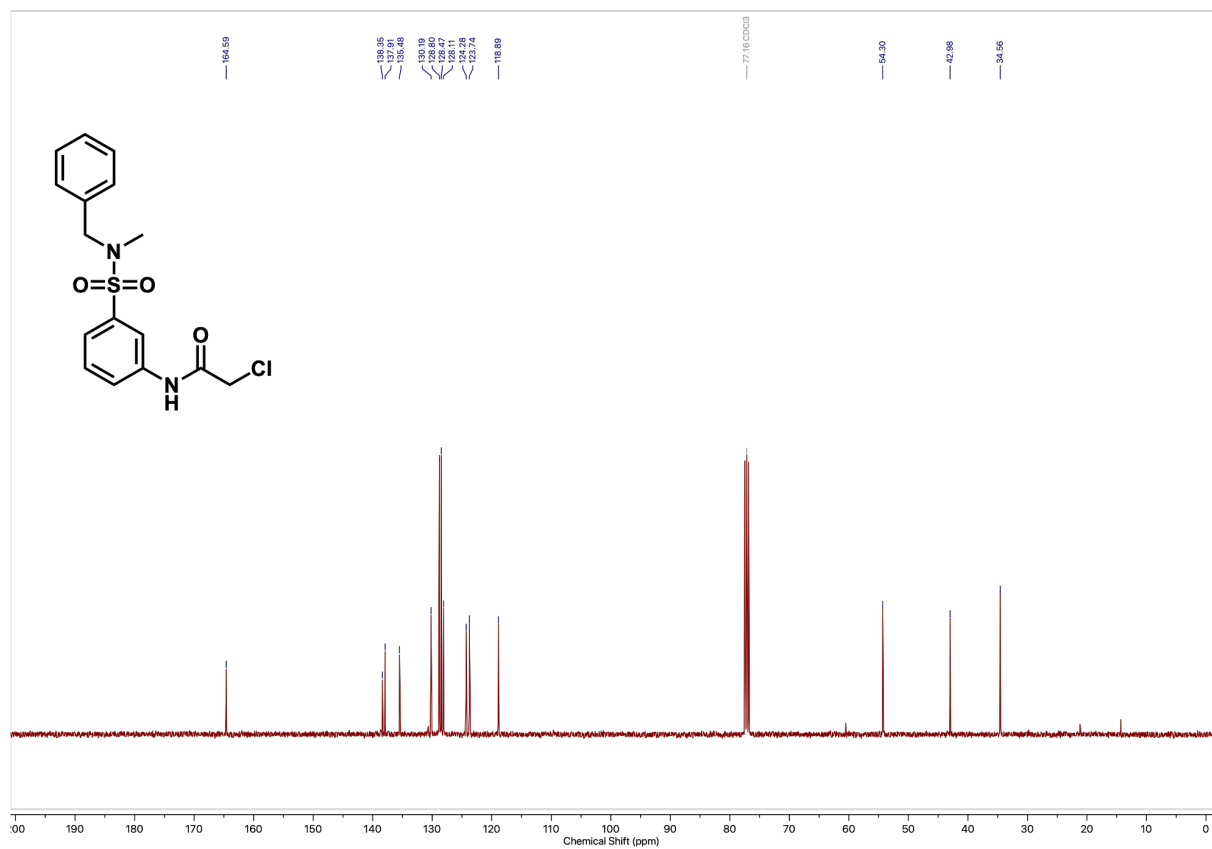


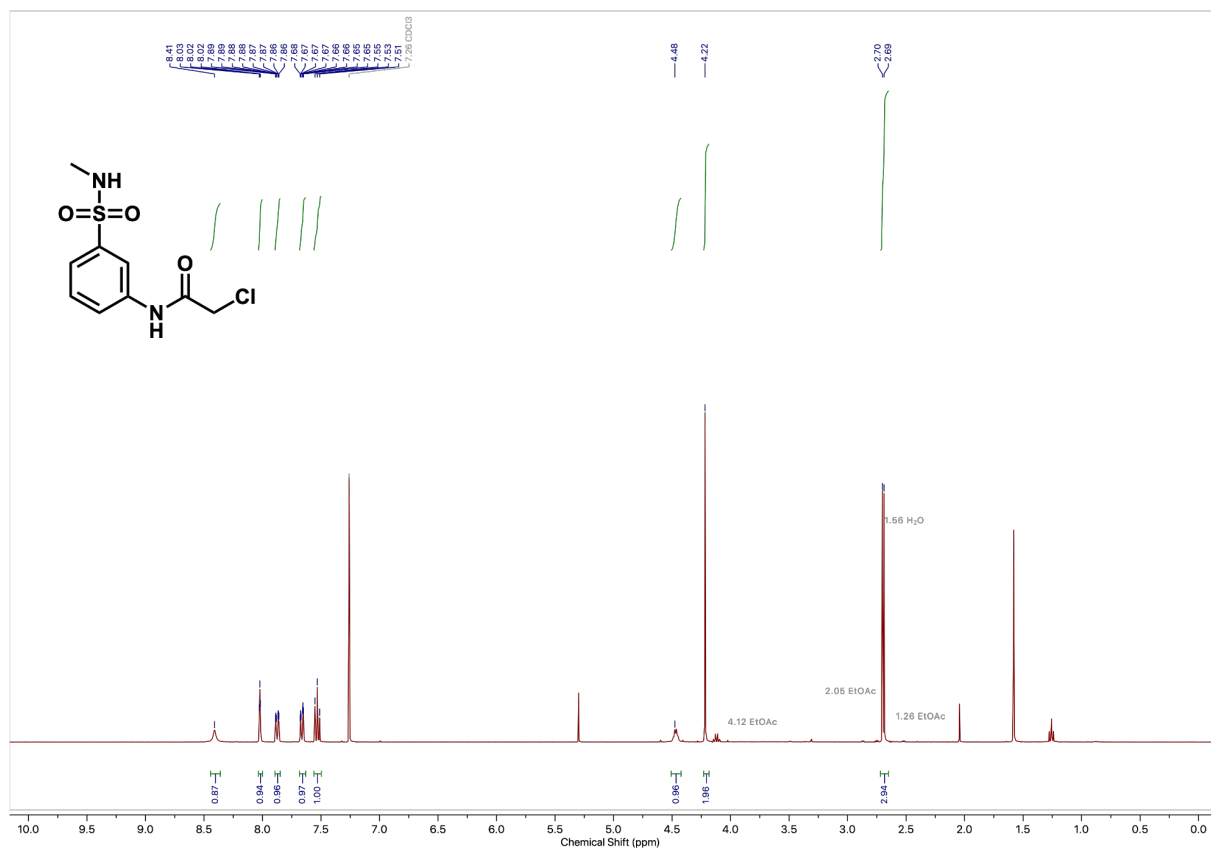


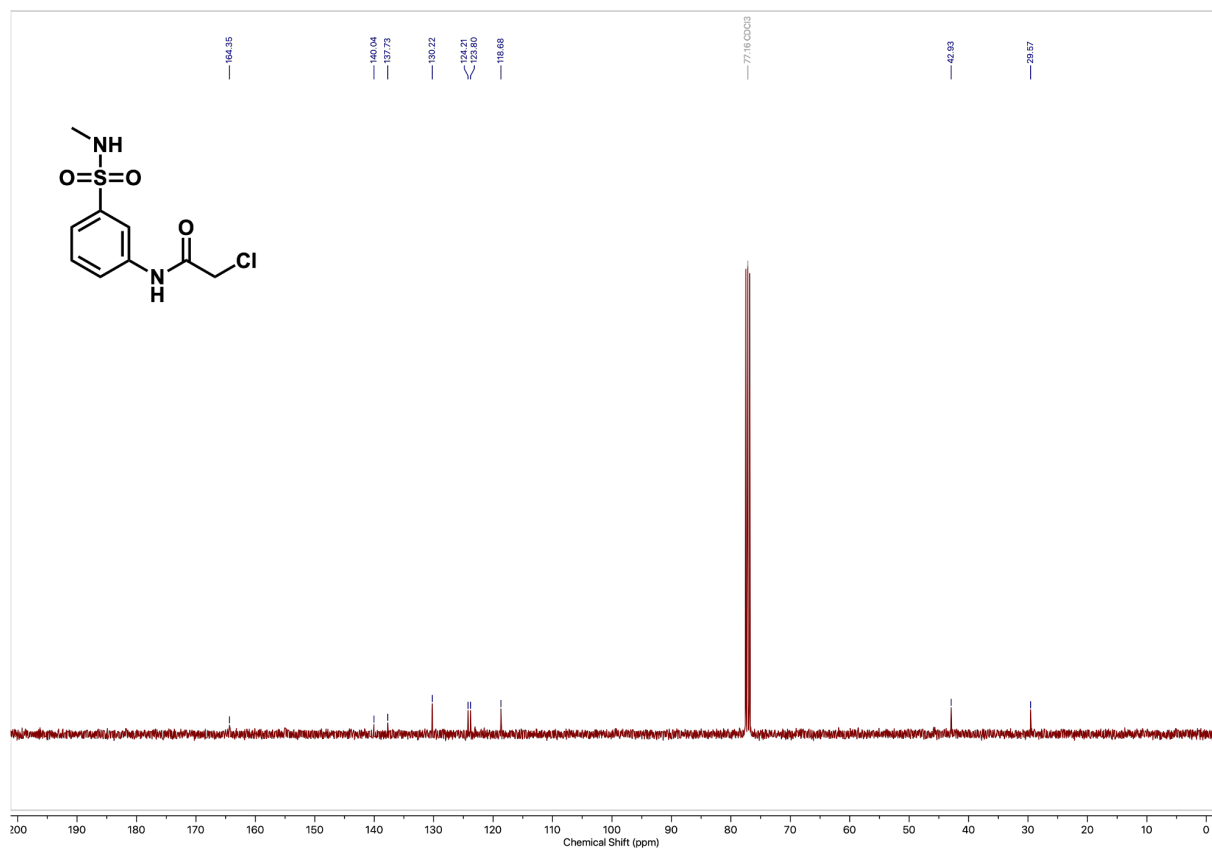


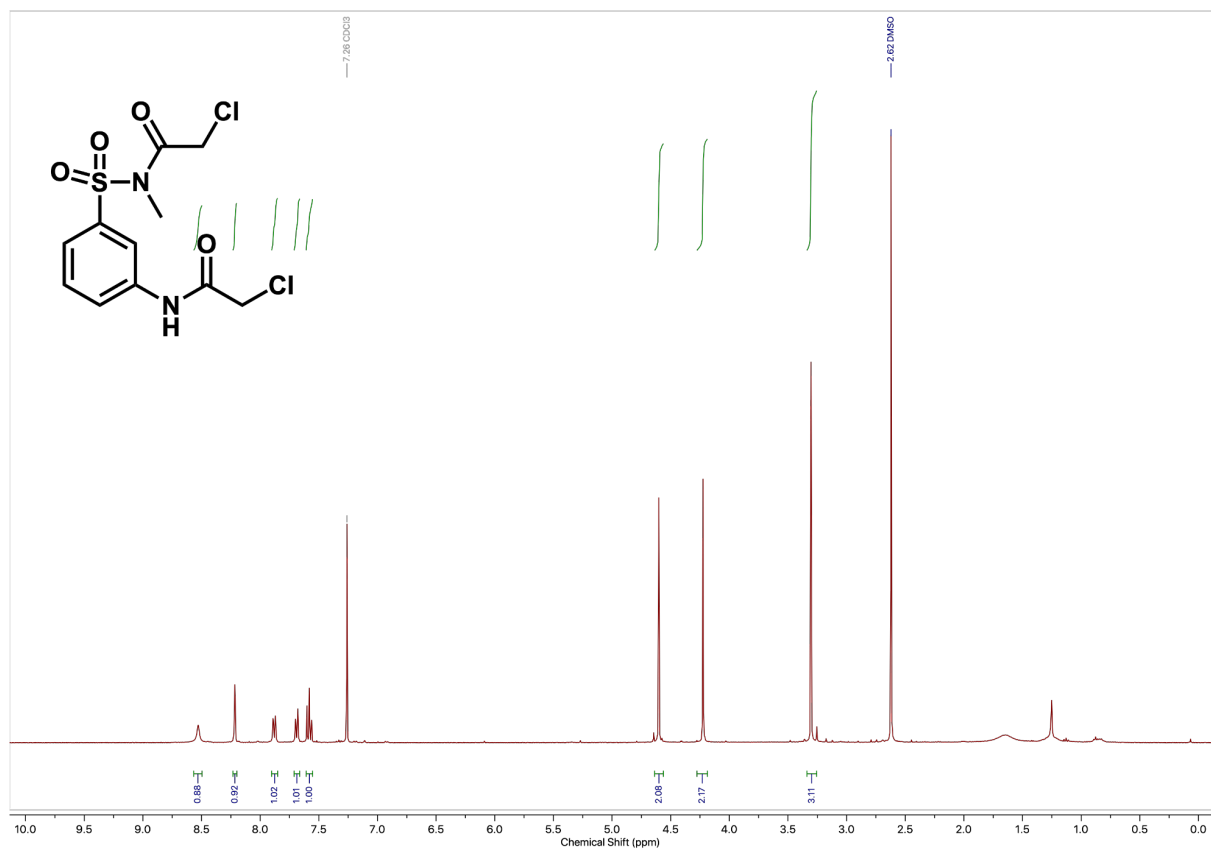


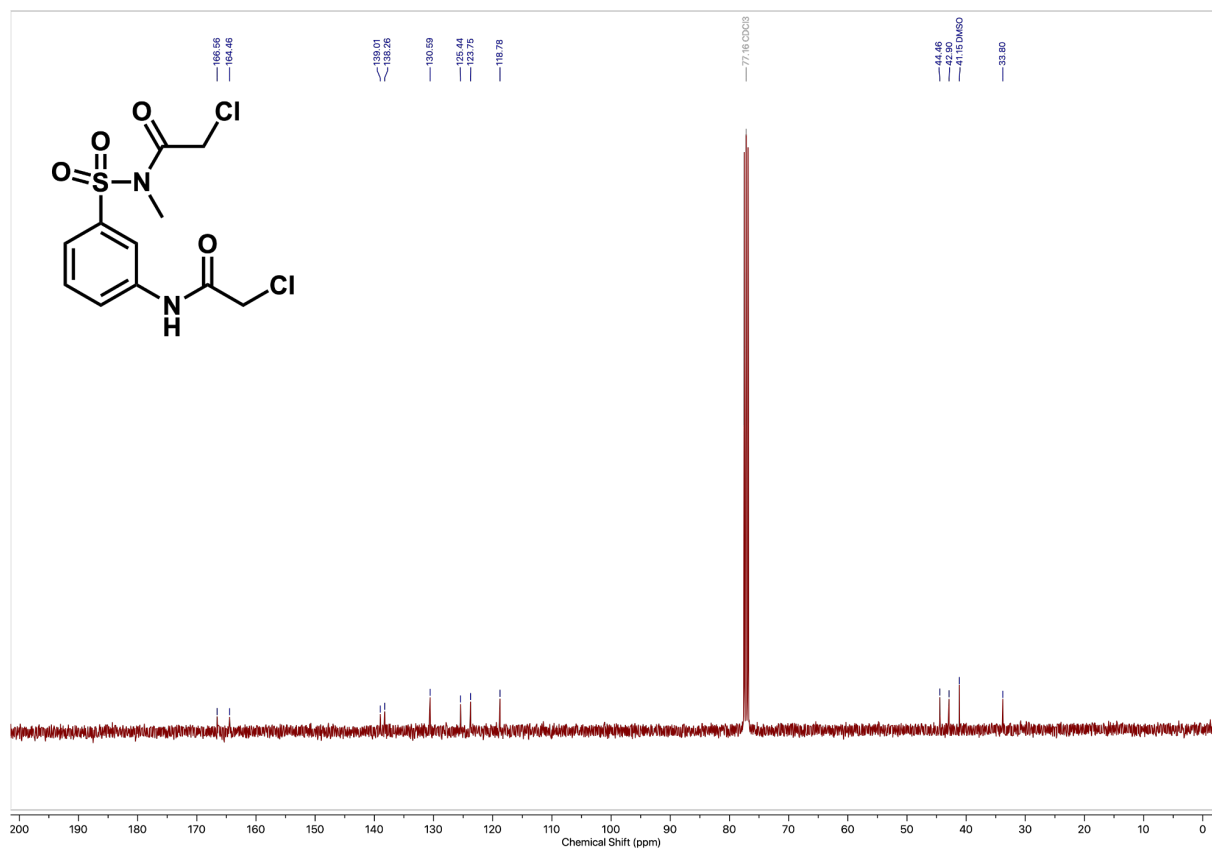


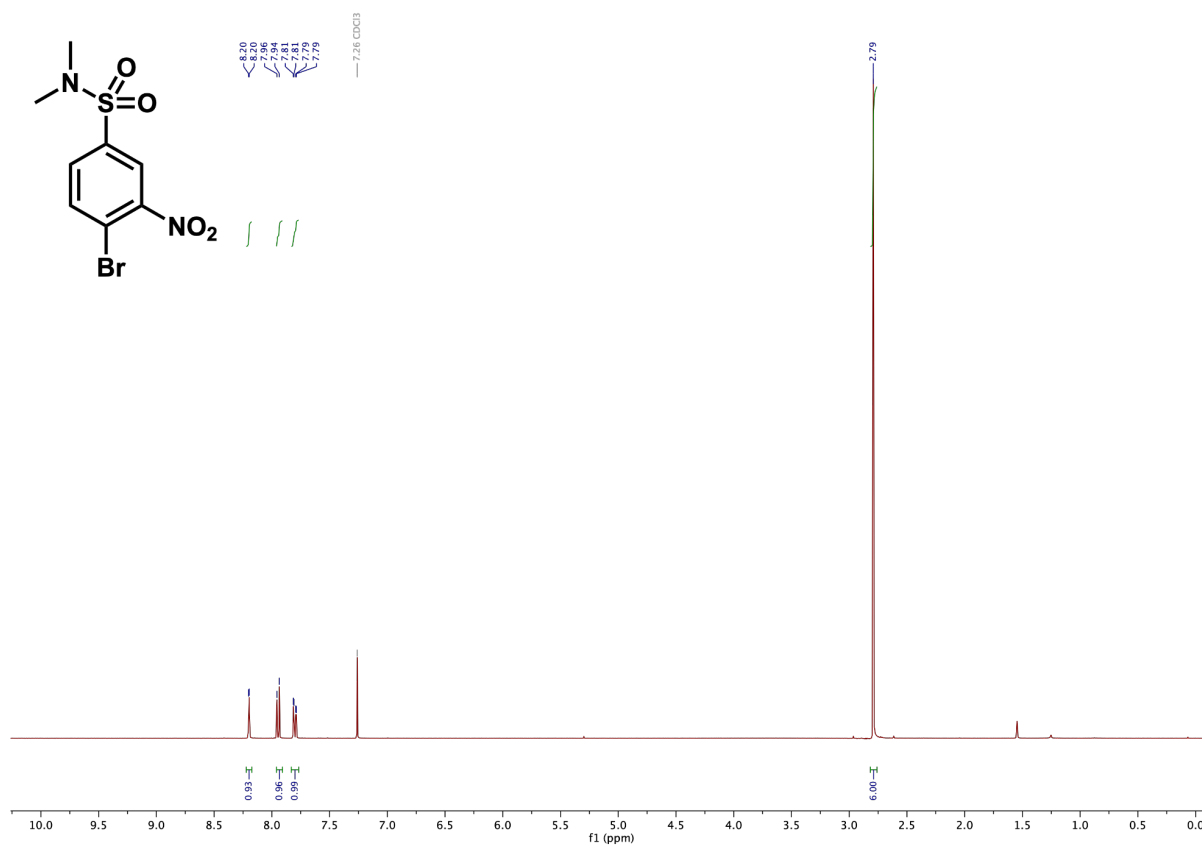




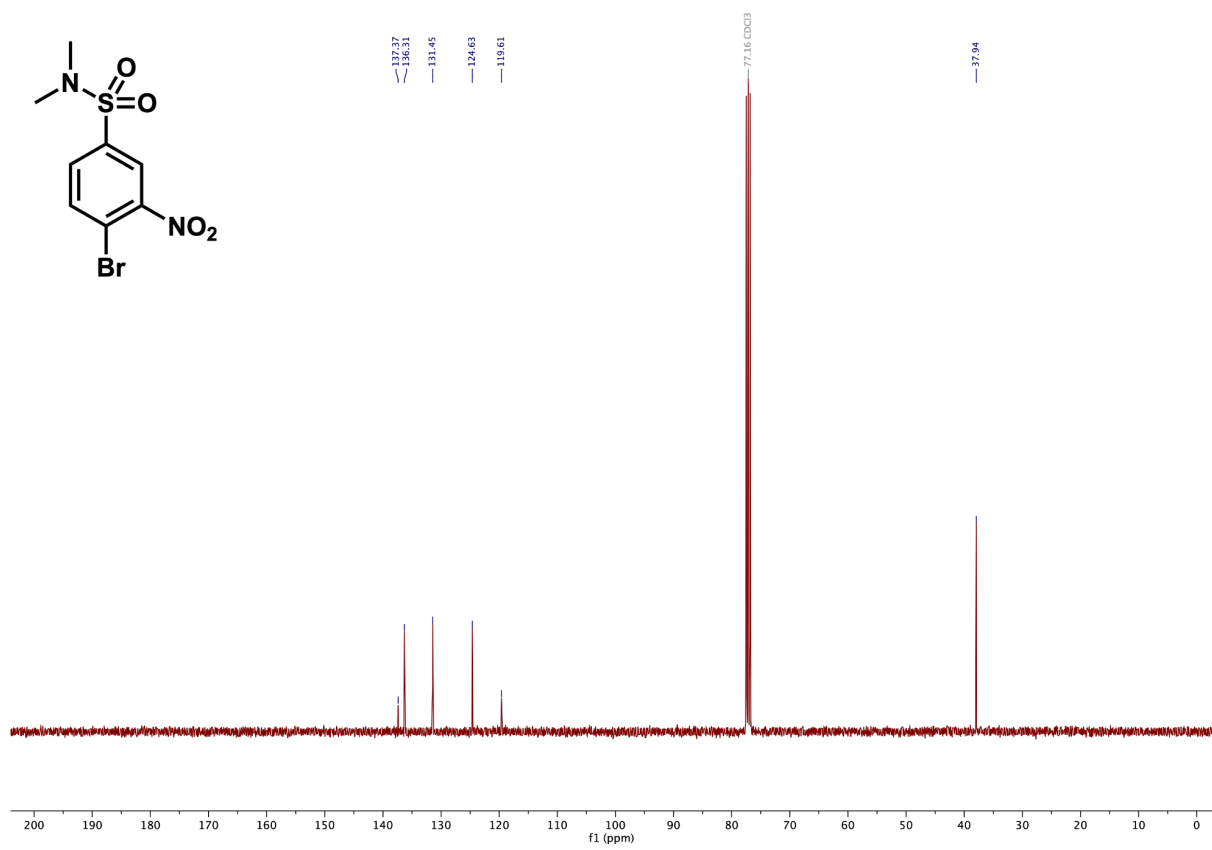


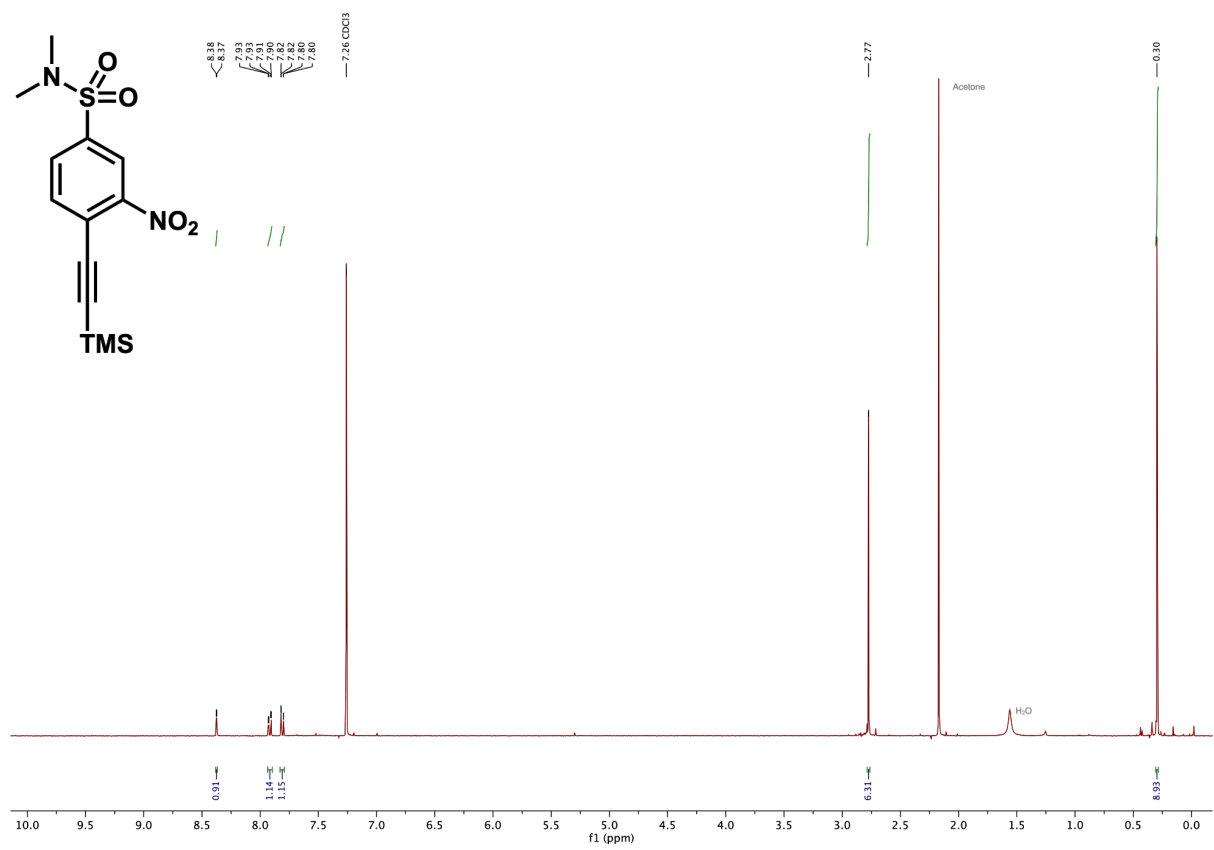


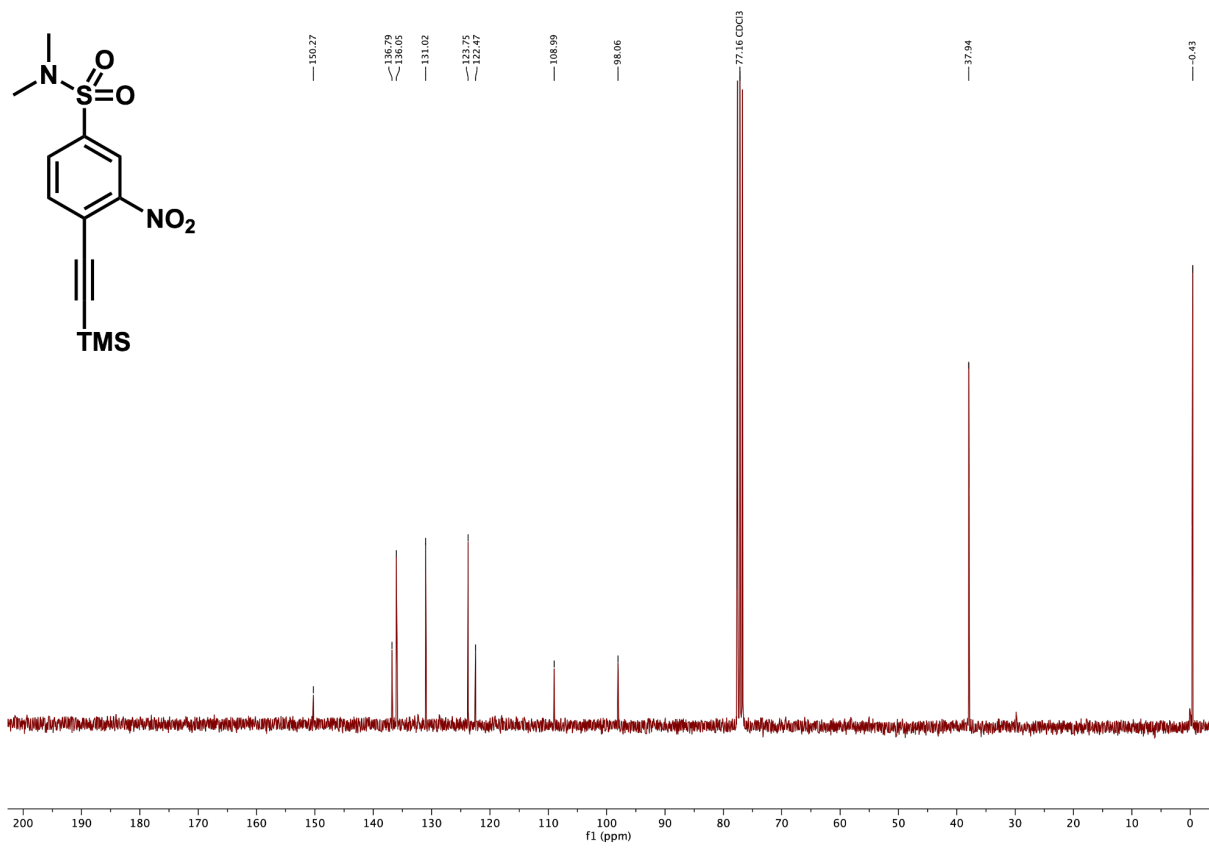


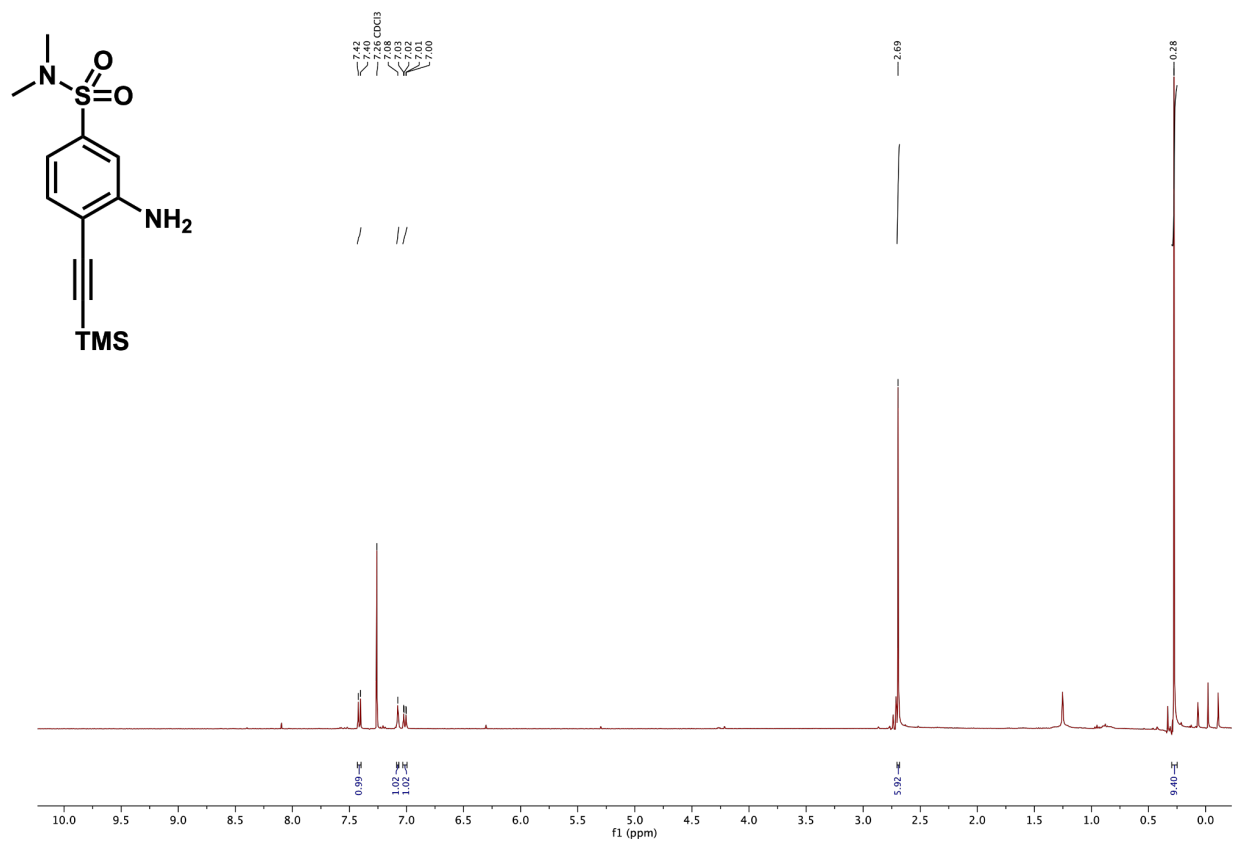


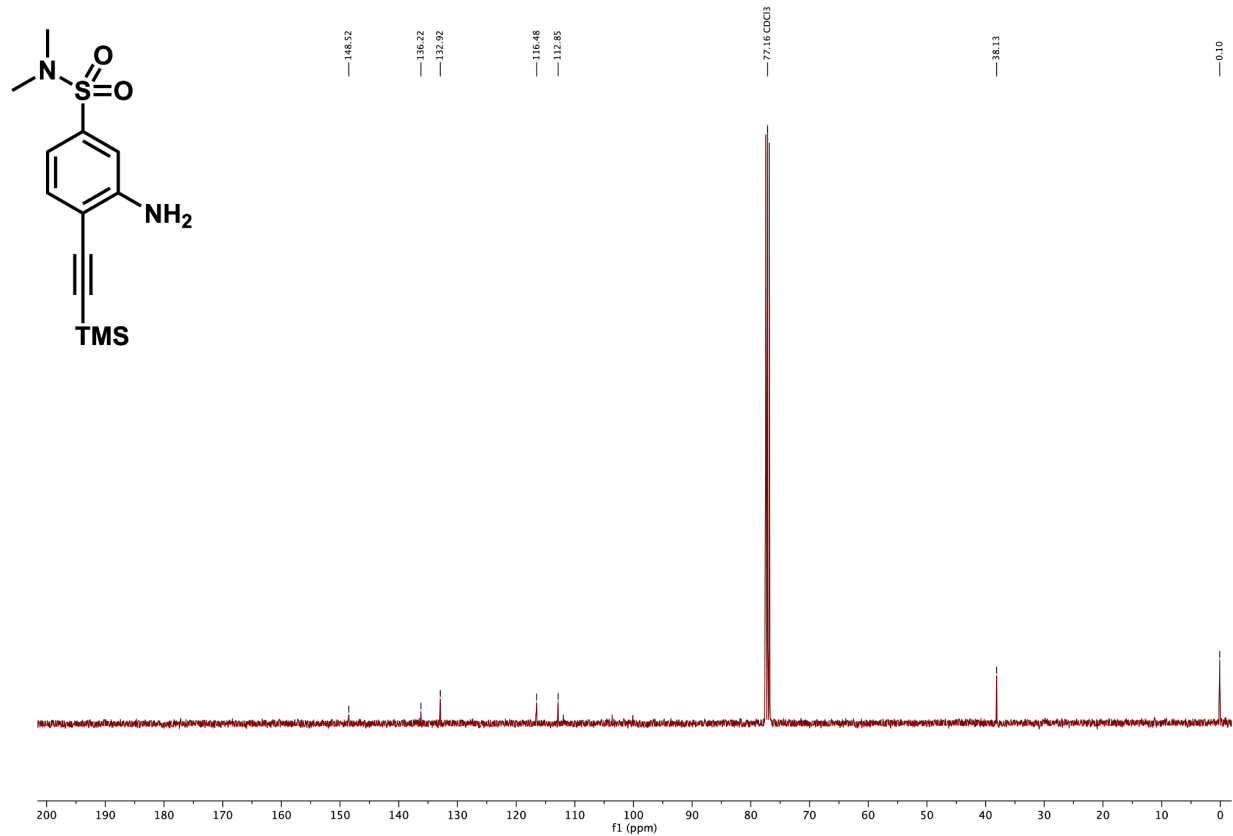


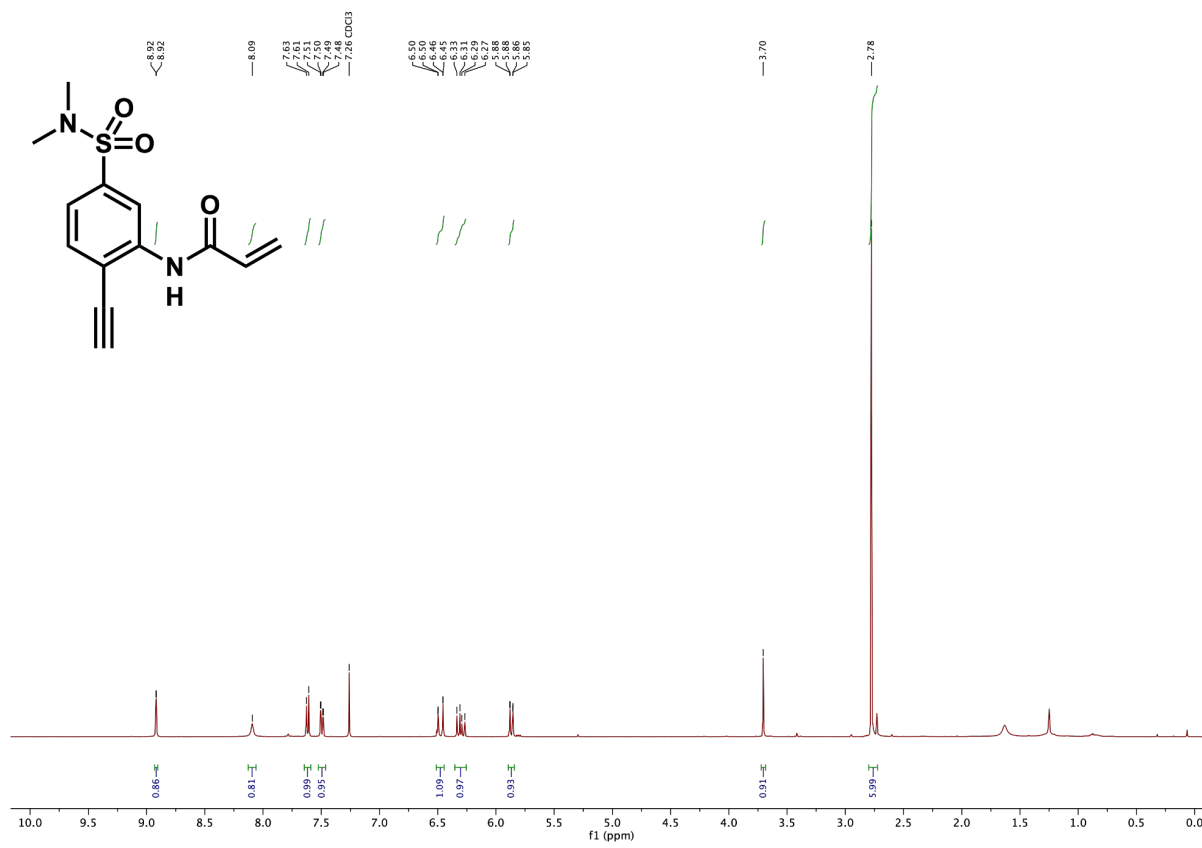


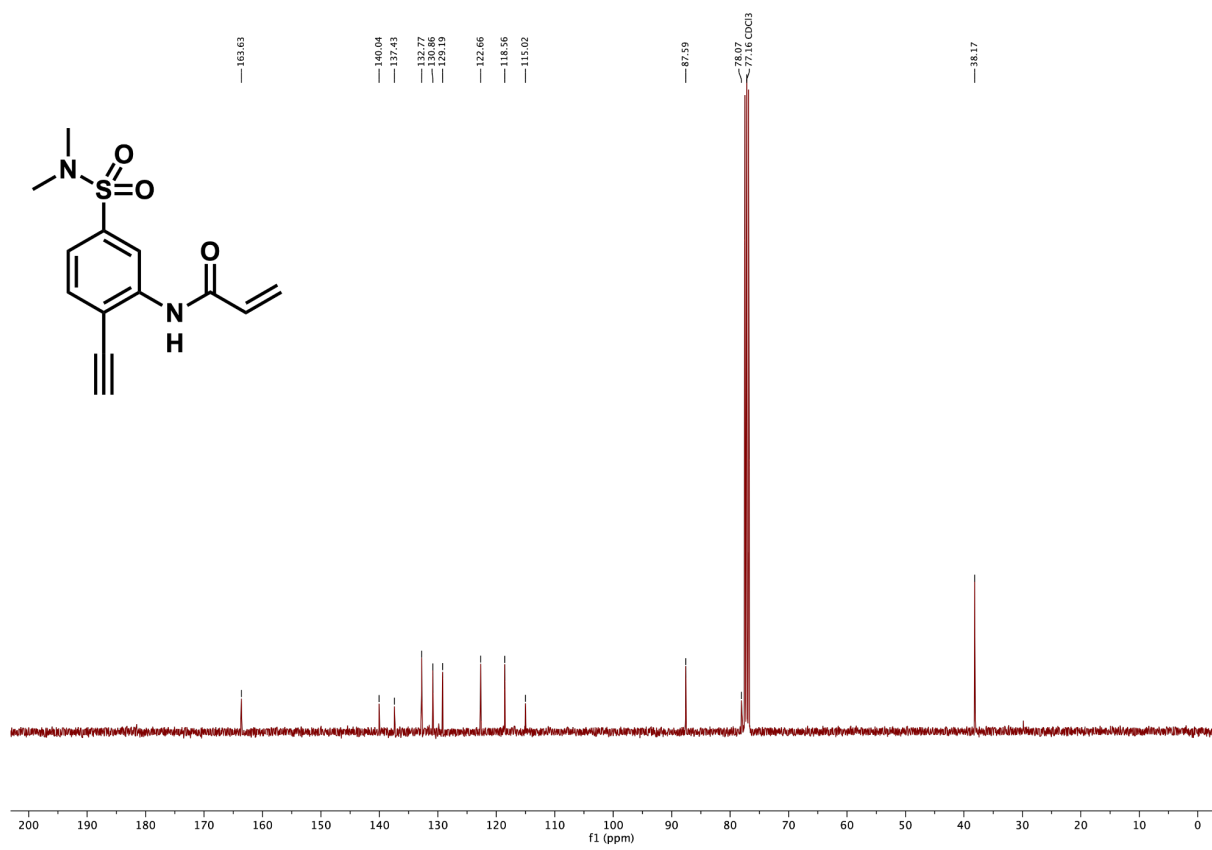


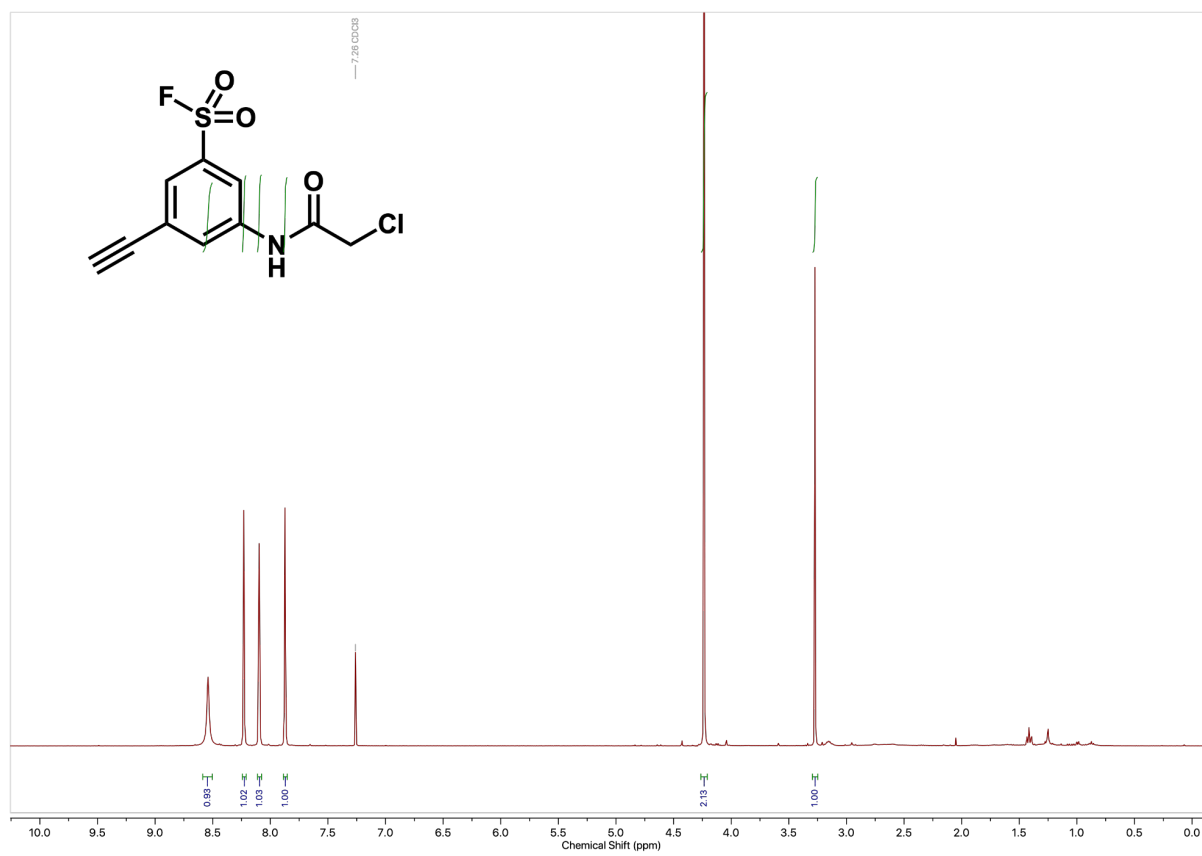




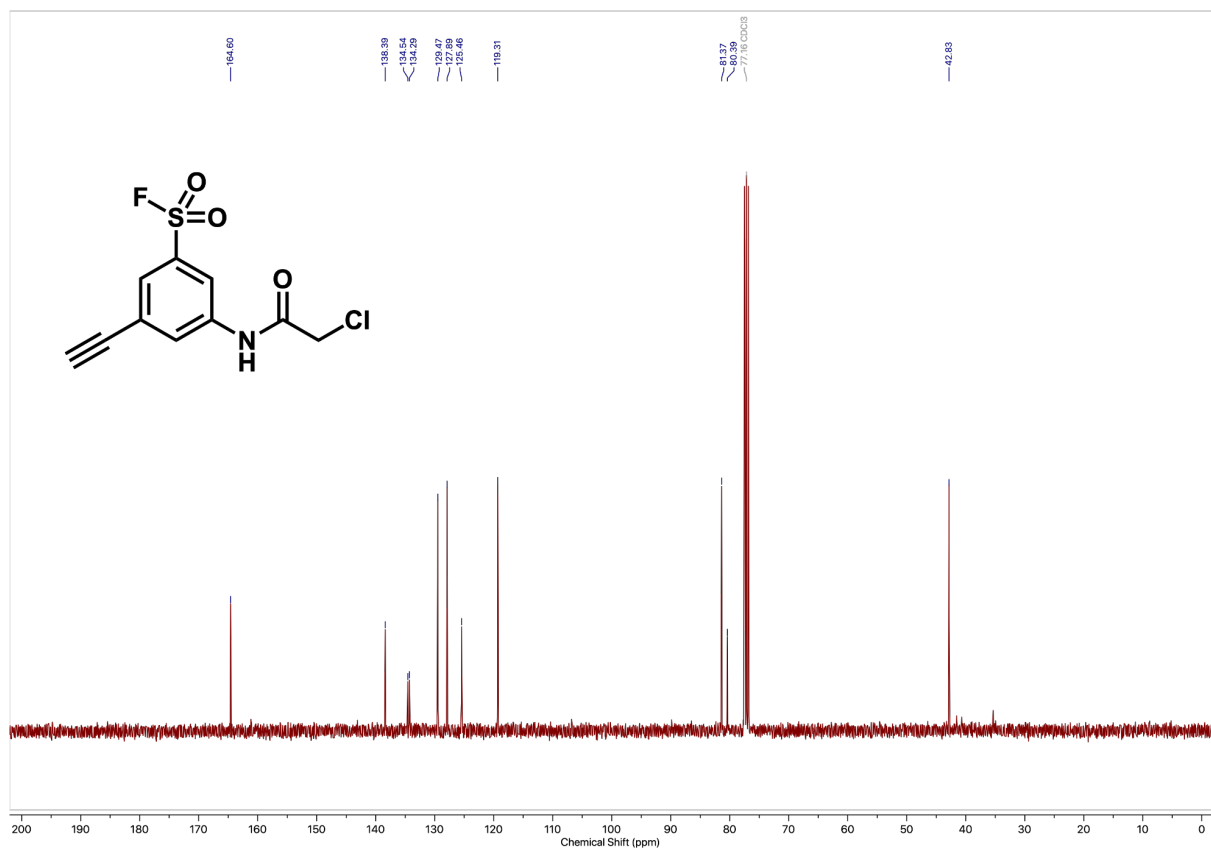


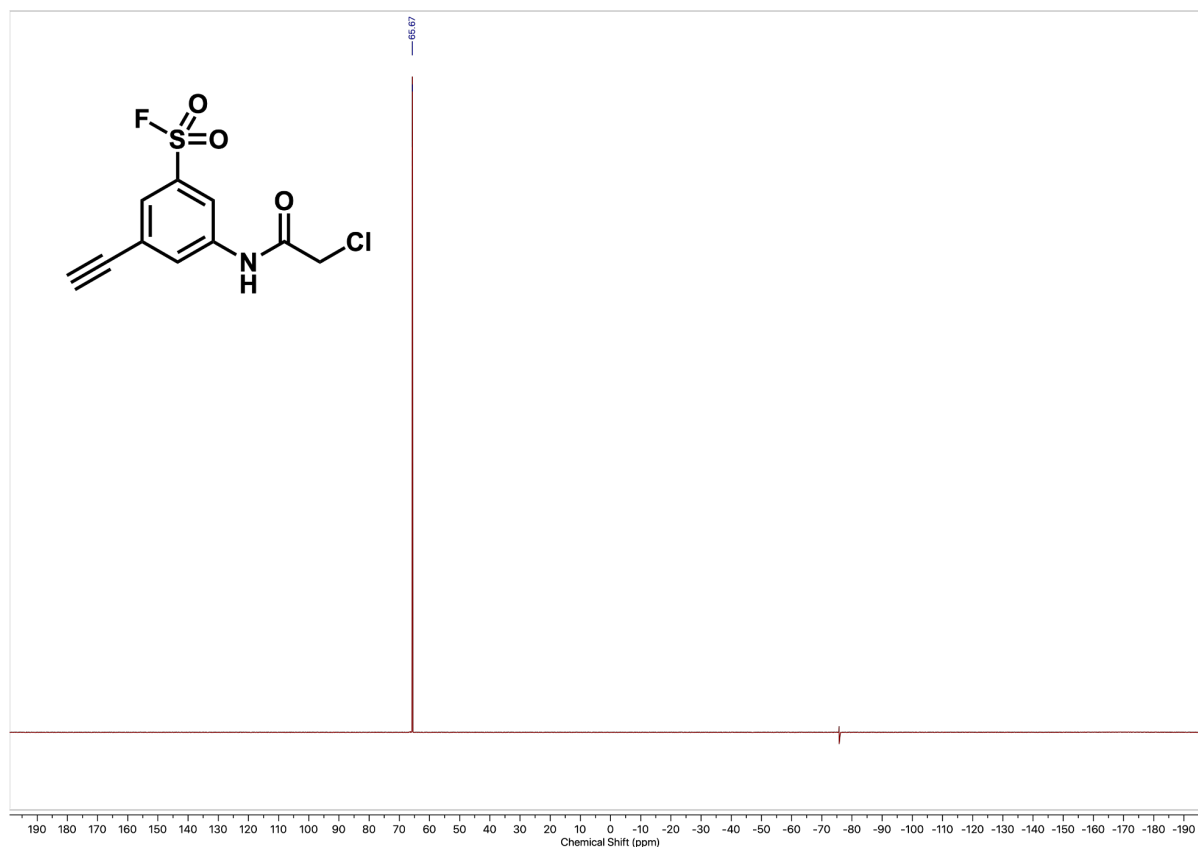












## References

1. Imprachim, N., Yosaatmadja, Y. & Newman, J. A. Crystal structures and fragment screening of SARS-CoV-2 NSP14 reveal details of exoribonuclease activation and mRNA capping and provide starting points for antiviral drug development. *Nucleic Acids Res.* **51**, 475–487 (2023).
2. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
3. Backus, K. M. *et al.* Proteome-wide covalent ligand discovery in native biological systems. *Nature* **534**, 570–574 (2016).

4. King, E. A. *et al.* Chemoproteomics-enabled discovery of a covalent molecular glue degrader targeting NF- $\kappa$ B. *Cell Chem. Biol.* **30**, 394-402.e9 (2023).
5. Cao, J. *et al.* Multiplexed CuAAC Suzuki-Miyaura Labeling for Tandem Activity-Based Chemoproteomic Profiling. *Anal. Chem.* **93**, 2610–2618 (2021).
6. Abibi, A. *et al.* The role of a novel auxiliary pocket in bacterial phenylalanyl-tRNA synthetase druggability. *J. Biol. Chem.* **289**, 21651–21662 (2014).

## Chapter 4: Assessing opportunities and challenges in proteogenomic thermal proteome profiling

Manuscript in preparation.

Flowreen Shikwana<sup>1,2</sup>, Keriann M.Backus<sup>1,2,3,4,5\*</sup>

<sup>1</sup>Department of Biological Chemistry, David Geffen School of Medicine, UCLA, Los Angeles, CA 90095 (USA); <sup>2</sup>Department of Chemistry and Biochemistry, UCLA, Los Angeles, CA 90095 (USA); <sup>3</sup>DOE Institute for Genomics and Proteomics, UCLA, Los Angeles, CA 90095 (USA); <sup>4</sup>Jonsson Comprehensive Cancer Center, UCLA, Los Angeles, CA 90095 (USA); <sup>5</sup>Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, UCLA, Los Angeles, CA 90095 (USA).

\*Corresponding Author: Keriann M. Backus, Biological Chemistry Department, David Geffen School of Medicine, UCLA, Los Angeles, CA, 90095, USA, E-mail: [kbackus@mednet.ucla.edu](mailto:kbackus@mednet.ucla.edu)

### Abstract

The impact of single amino acid variants (SAAVs) on protein activity is wide-ranging, spanning from no effect to highly damaging to cellular and organismal survival. Consequently, accurately assigning functional significance to SAAVs is of high value for both mechanistic and translational research. Here, we establish a framework for characterizing thermal stability changes in protein SAAVs. We first apply thermal proteome profiling (TPP) to a set of well characterized B-Raf variants, which confirms that protein-based cellular thermal shift assay (CETSA) data generally align with both protein- and peptide-level quantification via TPP using 16-plex isobaric tags. We then extend this methodology to a pooled B-Raf variant cell library, using affinity-purification and real time search (RTS) to boost coverage for B-Raf variants, as well as two-stage false discovery rate (FDR) error controlled proteomic search to support accurate assignment. In total we identify 47 B-Raf SAAVs and measure thermal stability for 19 of those variants. We observe clear

differences in mutation class-dependent predisposition to shifting B-Raf thermal stability, indicating that likely only a subset of mutations will induce thermal stability shifts for B-Raf. Extension of our TPP platform to capture 77 endogenous variants in Molt-4 cells and quantify 56 of those SAAVs illustrates the generalizability of the approach and nominates variants of uncertain significance (VUS) for future functional interrogation. Taken together, our study provides a roadmap for the implementation of TPP into pipelines aimed at functional assignment of SAAVs.

## **Introduction**

Missense variants, which result from single nucleotide variants (SNVs) leading to amino acid substitutions in proteins (single amino acid variants, SAAVs), can significantly impact protein function, stability, and interactions. Therefore, the characterization of missense variants in the human proteome is a crucial step in understanding the molecular underpinnings of various diseases and ultimately in the development of targeted therapeutic strategies. Advances in genomics and transcriptomics have enabled the identification of a multitude of missense variants in immortalized human cell lines and patient samples. Databases such as the Catalogue of Somatic Mutations in Cancer (COSMIC)<sup>1</sup>, ClinVar<sup>2</sup>, the database of single nucleotide polymorphisms (dbSNP)<sup>3</sup>, DepMap<sup>4</sup>, and the Genome Aggregation Database (GnomAD)<sup>5</sup> have allowed for the compilation of variants found in the general human population, both in disease and non-disease states. Aside from mutations with known effects or disease-relevance, a host of variants within the human population remain classified as variants of unknown significance (VUS). These VUS may not be associated with disease but may impact protein function(s) to a degree that may rationalize the variability in human responses to treatments or propensity to disease. For example, a person's genetic background (common variants) can interact with rare high effect variants to affect the variability of clinical symptoms<sup>6</sup>. Considering the hundreds of thousands of

variants found in human populations<sup>5</sup>, many are benign and do not contribute to negative health outcomes. However, characterizing each of these variants and distinguishing between benign and deleterious missense variants, especially those that are functionally relevant, remains a challenge.

Whole genome sequencing (WGS) and transcriptome sequencing have streamlined detection of mutations within the genome and transcriptome. With predictive tools like CADD<sup>7</sup>, PolyPhen<sup>8</sup>, and SIFT<sup>9</sup>, predictions of whether a variant is deleterious are also readily accessible. However, testing the functional consequences of encoded variants at the protein level remains less straightforward, particularly on a proteome-wide basis. For single genes, high throughput screening approaches using pooled libraries have begun to make inroads into this challenge<sup>10,11</sup>. However, scaling these approaches proteome-wide has not yet proven tractable. As the correlation between transcript and protein abundances are generally moderate at best<sup>12</sup>, assessing variant function at the protein level benefits from proteomic detection. Proteogenomics is a strategy that uses custom protein databases which are created from genomic and transcriptomic data to identify variant peptides within proteomic samples. This is often done through a two-stage search which uses a canonical protein database for the first search and the custom variant database for the second search. Identification of variants within the proteome is the first step towards studying the effects of SAAV. There is still a growing need for robust, high-throughput proteomic techniques that can provide a comprehensive assessment of multiple missense variant effects across the proteome.

In order to better understand these variants and delineate their effects on cellular function, techniques to profile and prioritize each of these proteoforms within the context of the whole proteome are needed. Mass spectrometry-based proteomics expands the scope of protein analysis, reaching a scale unattainable by traditional biochemical techniques. Thermal proteome profiling (TPP) is a widely utilized high-throughput proteomic methodology that provides an indication of altered protein functional states across the whole proteome by measuring changes

in protein thermal stability within the cell. TPP has proven useful for detecting drug binding<sup>13–15</sup>, changes in protein-protein<sup>16–20</sup>, protein-nucleic acid<sup>21</sup>, and protein-metabolite binding<sup>22</sup>, and monitoring post translational modification (PTM)-induced changes in protein states<sup>23,24</sup>. Measuring these changes in protein thermal stability serves as a proxy for functional state, which when applied to protein missense variants, can provide insight into important functional differences between wild-type (WT) and variant proteins. Highlighting this application of TPP, previous work has shown that TPP can identify both stabilizing and destabilizing mutational positions within yeast<sup>25</sup>.

TPP analyses typically rely on protein-level information which is first obtained by proteomic search software aggregating peptide spectral matches (PSMs) into peptide data, then to protein-level data for quantification of abundance. TPP uses tandem mass tag (TMT) systems for isobaric quantification across each temperature. The isobaric protein quantification, as done through FragPipe, first calculates reference normalized ratios for PSMs, then maps PSMs to peptide and protein groups, using unique and razor peptides to calculate protein group ratios. These ratios are then aggregated by median PSM ratio for each grouping and then transformed into abundances through MS1 intensities as calculated by a weight sum of the top 3 most intense peptide ions<sup>26</sup>. These abundances for each isobaric channel are then used to quantify protein thermal stability by calculating protein abundance over the temperature gradient. This approach uses more PSM data points due to this aggregation of PSM and peptide-level quantification in order to construct melt curve data per protein. The peptide-level information from this analysis is more limited in that some of these peptides may not be unique to one protein and may appear as lower abundance peptide ions, which would have more unreliable quantification. However, for SAAV-containing proteins, particularly heterozygous variants that are expressed endogenously, protein level quantification is likely not feasible. Instead, we expect that the thermal stability of such variants requires quantification at the peptide-level of the specific SAAV-containing peptide

sequences, as all other sequences will also be shared with the wild-type protein and, therefore, likely reflect an aggregate melting temperature of both the WT and mutant proteins, biased towards the more highly expressed protein species. Peptide-level TPP quantification is inherently noisier than protein-level quantification, where multiple peptides are “rolled up” into one protein intensity and batch effects can be particularly problematic<sup>27</sup>, introducing increased noise, which can result in identification of possibly spurious stabilization events. To combat this, the utilization of real time search (RTS) synchronous precursor selection (SPS) MS3 acquisition is important to ensure the minimization of ratio compression effects on individual peptides<sup>28,29</sup>. Additionally, utilization of RTS will aid in increased likelihood of variant peptide quantification across multiple variant-containing systems.

SAAVs are known to have important consequences within the cell, especially for certain classes of disease-associated proteins. One well-studied protein that has known disease-linked SAAVs is the serine/threonine-protein kinase B-Raf. Multiple oncogenic mutations have been observed in B-Raf, with over 80% of the oncogenic B-Raf proteoform being attributed to mutations at the V600 residue<sup>30</sup>. Other mutations throughout the B-Raf kinase domain have been reported across several cancer types and have been categorized under various classes of B-Raf mutations<sup>31</sup>. While the effects on B-Raf function as a result of many of these mutations have been studied and various inhibitors have been developed, especially for V600 mutants, not all cancer types with B-Raf variants have been observed to respond favorably<sup>32</sup>. Additionally, details of B-Raf SAAV activity are still being revised due to new studies providing more nuanced insight<sup>33</sup>. Due to the vast body of work towards studying the effects of B-raf variants, B-Raf SAAVs are not only an ideal model system for developing a method for monitoring thermal stability changes of SAAVs, but also for furthering the understanding of non-V600 variants. Recently, the LABEL-seq methodology was introduced as a way to profile a parallel pool of variants within a custom cell line containing ~1600 SAAV for B-Raf at positions known to have somatic mutations<sup>10</sup>. This



sequencing-based comprehensive analysis of B-Raf mutations raises an untapped opportunity to rigorously assess the capacity of TPP to discern functional variants. B-Raf variants represent a diverse pool of SAAVs covering both well-studied oncogenic driving mutations (class I), as well as mutations that are understudied in comparison, such as class II and III mutations and the >250 catalogued somatic mutations.

Here, we first establish thermal stability changes in protein SAAVs using B-Raf and a few of its more well characterized SAAVs (class I,II, and III variants) through blot-based cellular thermal shift assay<sup>34</sup> (CETSA) and proteomics-based TPP analyses. We establish that the peptide-level thermal profiling of each of these proteoforms are representative of the protein-level thermal profiling. After establishing the method, we moved into a pooled B-Raf variant system to build our variant TPP approach within a more complex, biologically relevant system and quantified 23 variants. Several cases within the group of 23 quantified variants highlight considerations for selecting WT comparators to variant peptides. Finally, to make variant TPP more broadly applicable to multiple cell lines, we applied our workflow to profile endogenously detected variants within acute lymphoblastic leukemia (ALL) Molt-4 cells, which contain a large number of variants due to being mismatch repair deficient. In combination with a FragPipe two-pass search approach for variant detection and quantification<sup>35</sup>, we quantified 56 unique variants, with 26 variant peptides containing matched WT pairs and 49 variant peptides that matched to protein-level thermal stability data. Our analyses also highlight important opportunities for future iterations of variant-directed TPP, including in addressing limited sequence coverage for variants, in both endogenous and pooled overexpression systems and remaining challenges with noisy peptide-level TPP quantification. We hope our study can serve as a roadmap for future efforts to deploy TPP in proteogenomics contexts and more broadly.

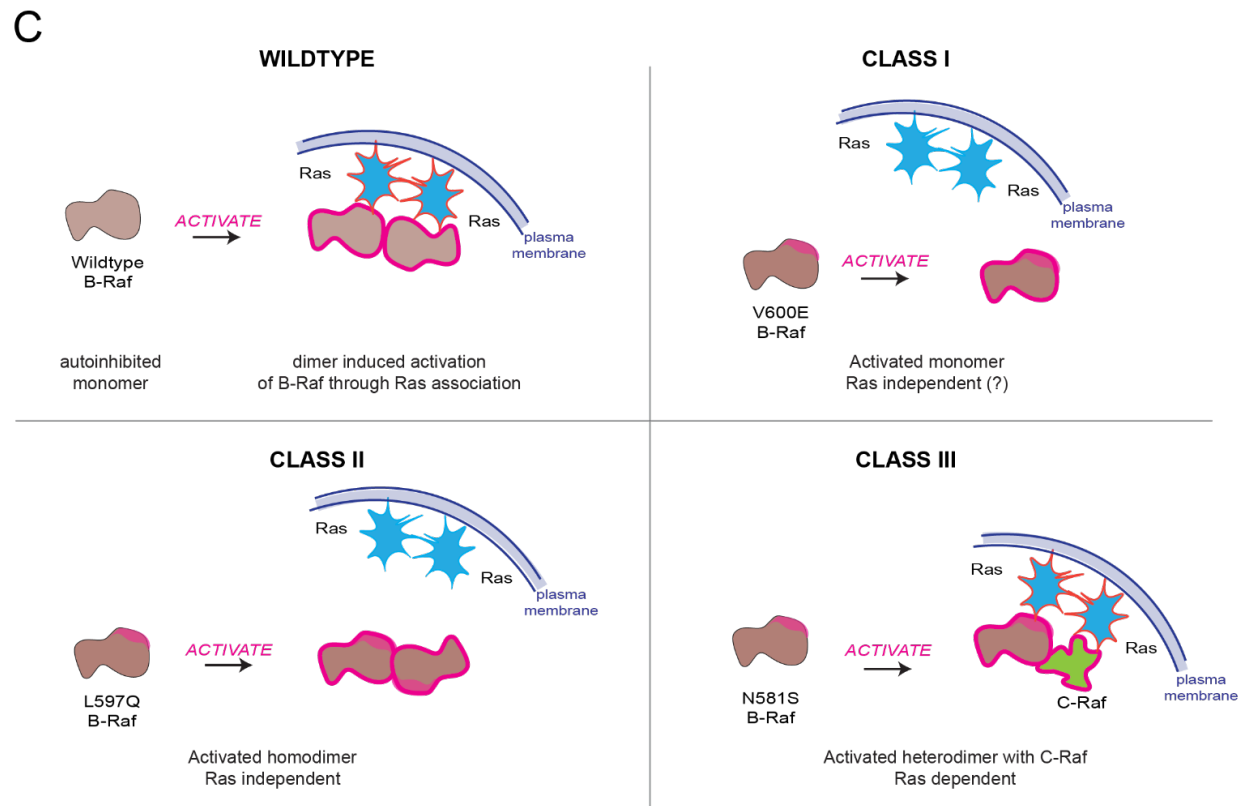
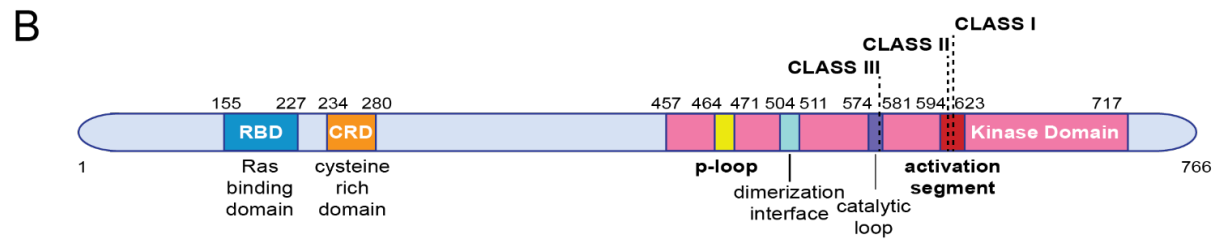
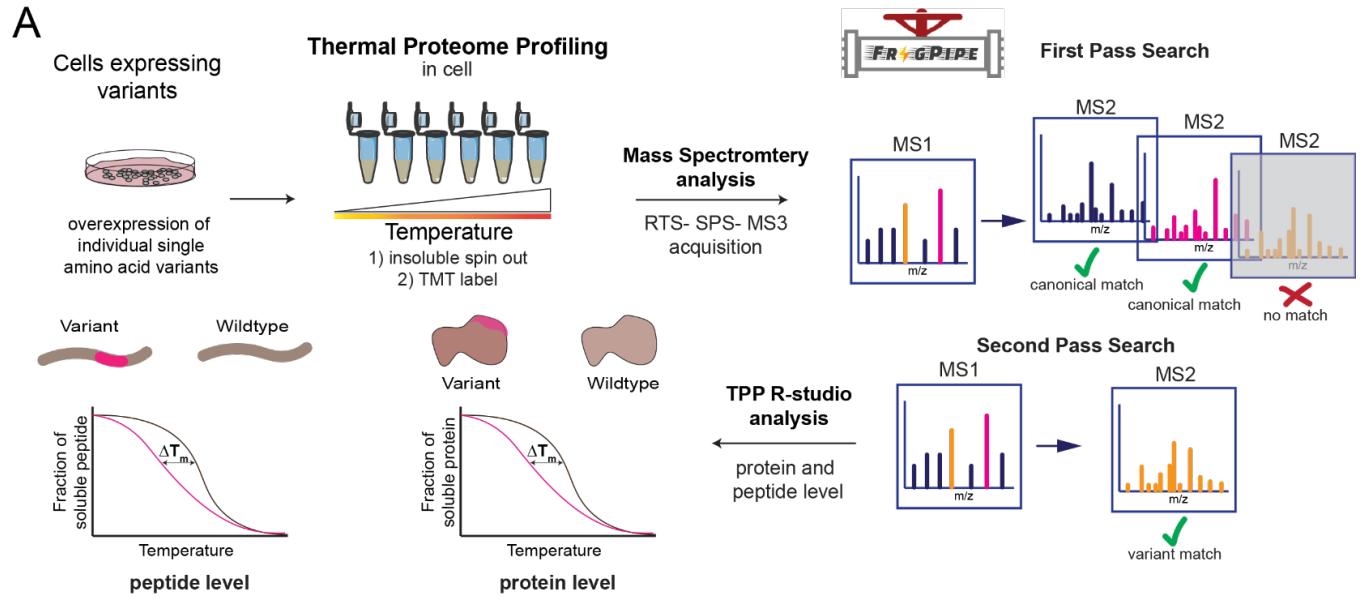
## Results

### **Establishing a model CETSA system to guide the development of variant TPP.**

Our first step towards achieving our overarching goal of proteome-wide variant TPP (**Figure 1A**) was to establish a model system on which we could rigorously test the impact of variant deleteriousness on protein thermal stability. For our initial proof-of-concept studies, we sought to confirm both that peptide level thermal stability was representative of protein level stability and that deleterious variants would impact protein stability in a manner detectable by a TPP readout. As we anticipated that peptide-level TPP could be hindered by poor coverage of sequence-matched WT and SAAV-containing peptides, we opted to establish a model system in which we would overexpress one SAAV-containing tagged protein and compare its protein-level melting temperature to the overexpressed WT construct. Such a model system would allow us to test whether thermal stability differences exist between WT and SAAV proteins as well as provide more easily detectable peptides to perform protein- and peptide-level comparisons. We additionally chose to incorporate “a two-pass” false discovery rate (FDR) controlled FragPipe search against a custom variant database to identify variants<sup>35,36</sup>, followed by TPP analysis through established TPP software packages<sup>37</sup> in order to quantify protein- and peptide- level thermal stability changes (**Figure 1A, Figure S1**).

For our model system, we selected the serine/threonine-protein kinase B-Raf, guided by the many well characterized variants, cellular phenotypes, and availability of the LABEL-seq cell lines for subsequent pooled screening<sup>10,30,31,38</sup>. We selected B-Raf SAAVs from each of the three classes of characterized variants (**Figure 1B,C**): (1) class I mutants at the V600 position that allow B-Raf to signal as a monomer<sup>30</sup>, (2) class II mutants that activate B-Raf as a Raf-independent dimer (L597Q)<sup>31</sup>, and (3) class III mutants which result in lower B-Raf activity as compared to WT and require C-Raf heterodimerization to become active (N581S)<sup>39</sup>. In addition to these well-studied

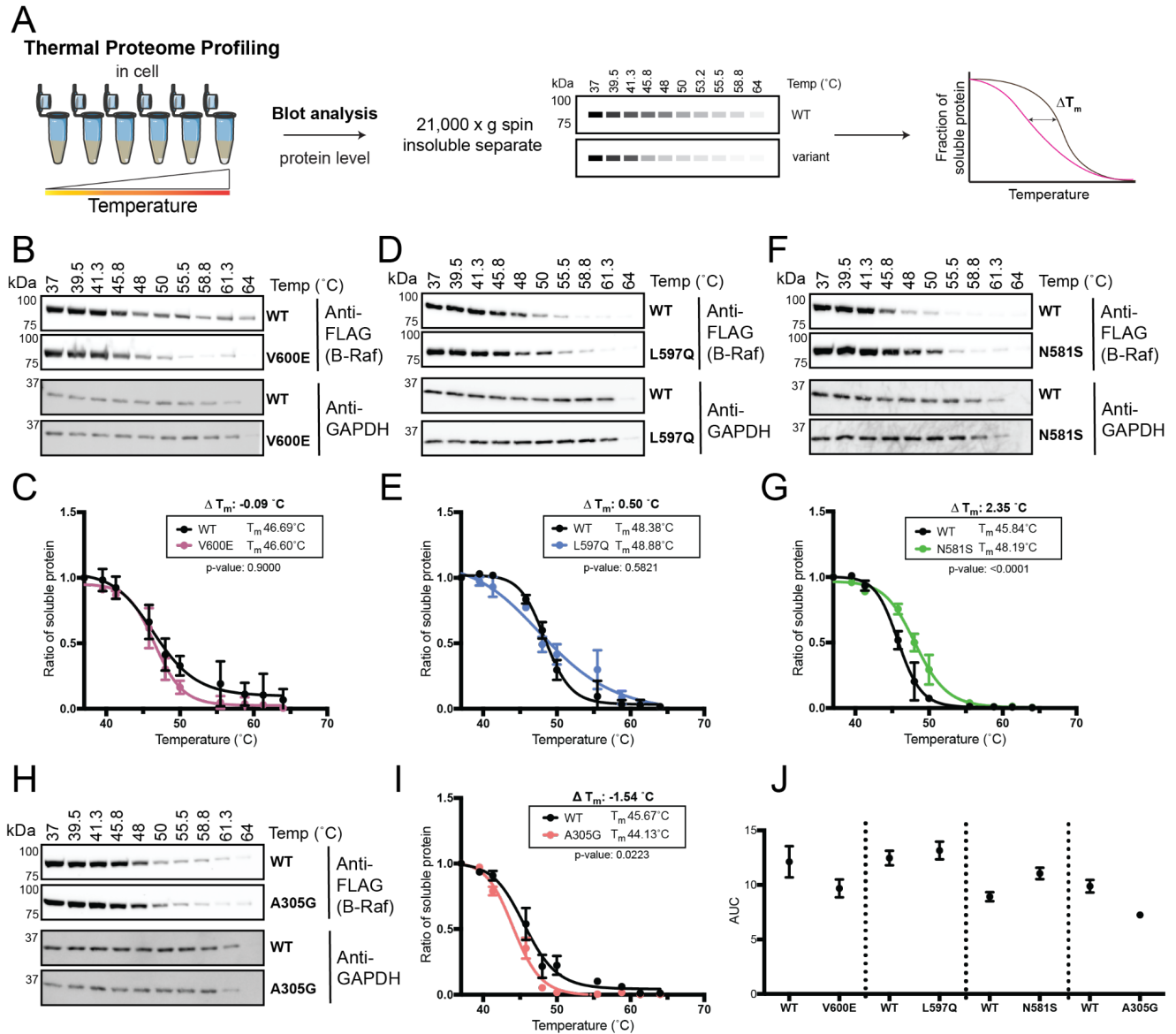
classes of B-Raf variants, we selected one benign variant, B-Raf A305G, as well as 2 other variants that are classified as likely benign variants based on computational predictions, B-Raf S323L and I326V, of which S323L is associated with Noonan syndrome and I326V is classified as a variant of unknown significance. Additionally, B-Raf variants A305G and I326V had LABEL-Seq data indicating no changes to interactions with binding partners C-Raf and MEK1, as well as no significant change in B-Raf activity<sup>10</sup>. To enable our initial benchmarking studies, we generated N-terminally epitope (FLAG)-tagged constructs for transient overexpression.



**Figure 1. Summary of TPP workflow for variant TPP.** A) Summary of the TPP workflow for assessing variant stability including a FragPipe two-stage FDR controlled search against a custom variant database<sup>35</sup> followed by TPP analysis through the TPP-TR package<sup>35,37</sup>. B) Sequence of B-Raf with relevant domains highlighted and Class I-III mutations mapped to sequence. C) B-Raf wildtype and class I-III variants in the context of known features about each class. B-Raf WT and its variants belonging to classes I-II will be used as a model system for insights into variant TPP as a methodology.

To test our hypothesis that thermal stability changes are detectable between WT and SAAVs, we selected a blot-based cellular thermal shift assay (CETSA) as our initial, orthogonal readout due to the relatively rapid throughput and decreased costs of the assay when compared to proteomics. We opted to use the widely utilized CETSA workflow with a centrifugation spin speed of 21,000 x g to separate the soluble from the insoluble fractions, consistent with established workflows<sup>34,40</sup> (**Figure 2A**). Using this blot-based workflow for our panel of mutants, we, unexpectedly, observed no significant shift ( $\Delta T_m = -0.09^\circ\text{C}$ , p-value = 0.0900) in the melting temperature of the class I V600E variant as compared to WT; however, visually by blot we observe the tailing of residual soluble protein observed in the WT B-Raf blot that cuts off earlier for V600E (**Figure 2B,C**). For the class II L597Q variant, again no significant shift was observed ( $\Delta T_m = +0.50^\circ\text{C}$ , p-value = 0.5821) —we did note a substantially distinct melt curve shape when compared to the wild-type protein, suggestive of a potential impact of the variant on a fraction of the protein (**Figure 2D,E**). The class III N581S variant showed a clear significant stabilization by blot (**Figure 2F**) and by quantification ( $\Delta T_m = +2.35^\circ\text{C}$ , p-value =  $<0.0001$ ; **Figure 2G**). For the seemingly benign A305G variant, a modest destabilization of  $-1.54^\circ\text{C}$  (p-value = 0.0223) was observed (**Figure 2I**). The other benign variants S323L and I326V showed insignificant thermal stability differences (**Supplementary Figure 1A-C**), as expected. These analyses provide initial evidence that a subset of functional variants likely alters protein thermal stability, with the extent of this effect likely not strictly correlated with variant deleteriousness.

We also note that the thermal stability differences we quantify between SAAVs (up to just under 2.5°C shift in our dataset) are much smaller than CETSA analyses for drug binding, for which we observe a > 5°C shift in thermal stability (**Supplementary Figure 2A-D**). This comparison suggests that single amino acid substitutions may produce less pronounced effects on protein thermal stability than drug binding effects. Previous reports have addressed this limitation by using area under the curve (AUC) analysis to quantify thermal stability rather than calculation of  $T_m$ <sup>25</sup>. Applying AUC analysis to our CETSA quantification, we observed similar trends as calculated by our melt curves of the blot quantification (**Figure 2I**). Given that we observed small thermal stability differences and distinct melt curve shapes, we considered that alternative melt curve analyses may be needed in moving forward with variant TPP. Alternative methods aside from  $T_m$  comparisons, such as comparison of melt curve inflection points<sup>41</sup>, nonparametric curve analyses<sup>42</sup>, and the use of Gaussian process models<sup>43</sup>, have been developed for such cases. Building upon the promising blot-based CETSA results showing us thermal stability differences for select SAAVs, we turned to using the more sensitive mass spectrometry-based TPP to investigate protein-level and peptide-level changes that may allow for better quantification of thermal stability for SAAVs.



**Figure 2. Blot-based CETSA analysis for individually overexpressed B-Raf SAAVs.** A) Workflow of CETSA analysis for each overexpressed FLAG-B-Raf construct. B) Comparison of blot analysis for B-Raf WT and V600E CETSA with 21k centrifugation to separate out insoluble proteins. C) Melt curve fit quantification of blot analysis for B-Raf V600E (n=3) CETSA as compared to WT B-Raf. D) Comparison of blot analysis for B-Raf WT and L597Q CETSA with 21k centrifugation to separate out insoluble proteins. E) Melt curve fit quantification of blot analysis for B-Raf L597Q (n=2) CETSA as compared to WT B-Raf. F) Comparison of blot analysis for B-Raf WT and N581S CETSA with 21k centrifugation to separate out insoluble proteins. G) Melt curve fit quantification of blot analysis for B-Raf N581S (n=2) CETSA as compared to WT B-Raf. H) Comparison of blot analysis for B-Raf WT and A305G CETSA with 21k centrifugation to separate out insoluble proteins. I) Melt curve fit quantification of blot analysis for B-Raf A305G (F, n=3) CETSA as compared to WT B-Raf. J) The mean area under the curve quantification of blot analysis for B-Raf V600E, L597Q, N581S, and A305G CETSA. Error bars represent the standard error as calculated by GraphPad Prism v 10.1 area under the curve analysis. FIJI (ImageJ quantification)<sup>44</sup>, plotted into GraphPad Prism v 10.1, and non-linear curve fitted with interpolation sigmoidal 4PL with no special handling for outliers. An extra-sum-of-squares F test was performed for the IC50 values ( $T_m$ ).

### Establishing Variant TPP at the Protein-Level with B-Raf Variant Overexpression

We next turned to TPP analysis of each variant, again using our transient overexpression system. We opted to implement a 100K rather than 20K centrifugation, guided by literature precedent<sup>37,45</sup>. We noted a slight decrease in the overall melting temperatures in pilot experiments, which resulted in poorly fit melting curves (**Supplementary Figure 3A-F**). Therefore, we chose to add additional temperature points at the lower end of the temperature range by transitioning from 10-plex to 16-plex TMT. With our newly modified TPP workflow, we analyzed our mass spectrometry protein-level data for each of the B-Raf variants. The V600E variant showed a +0.98°C stabilizing shift in the protein-level data (**Figure 3B**) as compared to the nonsignificant -0.09°C shift seen in the CETSA quantification (**Figure 2C**). When looking at the reported adjusted p-values calculated by the TPP-TR package, B-Raf V600E compared to WT B-Raf exhibited an adjusted p-value of 0.999 and 1 for replicates 1 and 2. To account for non-standard (sigmoidal) melt curve shapes and to account for melt curve comparisons outside of simple  $T_m$  estimations, we utilized the nonparametric analysis of response curves (NPARC)<sup>42</sup>. With this analysis, the adjusted p-value for V600E was calculated to be  $4.6 \times 10^{-3}$  (**Supplementary Figure 4A**). Plotting all of the calculated adjusted p-values for this dataset, we observe that the V600E melt curve differences are not significant (**Supplementary Figure 4B**). The L597Q B-Raf variant exhibited a stabilizing shift of +2.18°C (**Figure 3C**), which is in agreement with the slight stabilizing effect seen in the CETSA data for L597Q (**Figure 2E**); however, the shift appears to be greater in magnitude in the TPP analysis. NPARC analysis of the L597Q dataset determined significance with an NPARC adjusted p-value of  $4.0 \times 10^{-10}$  (**Supplementary Figure 4C-D**). Variant N581S exhibited a significant (NPARC adjusted p-value  $2.7 \times 10^{-11}$ ) shift of +2.05°C (**Figure 3D**, **Supplementary Figure 4E-F**), which is consistent with the CETSA data (**Figure 2G**). Protein-level TPP analysis for the benign A305G variant showed an insignificant +0.81°C stabilization (NPARC adjusted p-value =  $2.5 \times 10^{-3}$ ; **Supplementary Figure 5A-C**), differing from the slight destabilizing effect seen by CETSA quantification (**Figure 2I**).



We acknowledge that with the TPP data, the melt curves are fit using the TPP package on R<sup>37</sup>, while the melt curves for CESTA quantification were non-linear curve fitted with interpolation sigmoidal 4PL with GraphPad Prism. To address this, we fit the normalized output from the TPP package to compare to the  $T_m$  reported by the TPP package. With this analysis, the melting temperature differences between WT and variant proteins were generally consistent with the calculated  $T_m$  from the TPP package, with V600E showing a +1.96°C (p-value = <0.0001) stabilizing shift (**Supplementary Figure 6A**). The L597Q variant exhibited a shift of +0.58°C (p-value = 0.4049; **Supplementary Figure 6B**), N581S had a shift of +1.70°C (p-value = <0.0001) (**Supplementary Figure 6C**), and benign variant A305G had an insignificant shift of 0.49°C (p-value = 0.1367; **Supplementary Figure 6D**). The fitting of the curves and calculation of the  $T_m$  point is moderately different between the two methods, with the difference in  $T_m$  showing different magnitudes of stabilization. The main difference stems from the  $T_m$  calculation within GraphPad Prism arising from the point at which 50% of the soluble protein remains while the TPP R package calculates the  $T_m$  based on the point of inflection for the melt curve. Additionally, the TPP R package uses additional parameters such as the slope and plateau of the line, as well as additional filters to calculate an adjusted p-value for each comparison<sup>24</sup>. To stay consistent, from this point forward, we will only report the  $T_m$  as calculated from the TPP R package as it is a widely utilized analysis package for TPP experiments<sup>37</sup> and has been specifically designed to handle TPP data<sup>24</sup>. In the cases that the melt curves being compared pass the established filtering criteria for R-squared, plateau, and slope<sup>42</sup>, we will use the NPARC analysis to determine if the melt curves significantly differ from the null model.

### **Examining Peptide-Level Analysis for Variant TPP**

One potential confounder to our B-Raf heterologous overexpression data is the presence of endogenous B-raf, which could potentially interfere with and complicate the quantification of

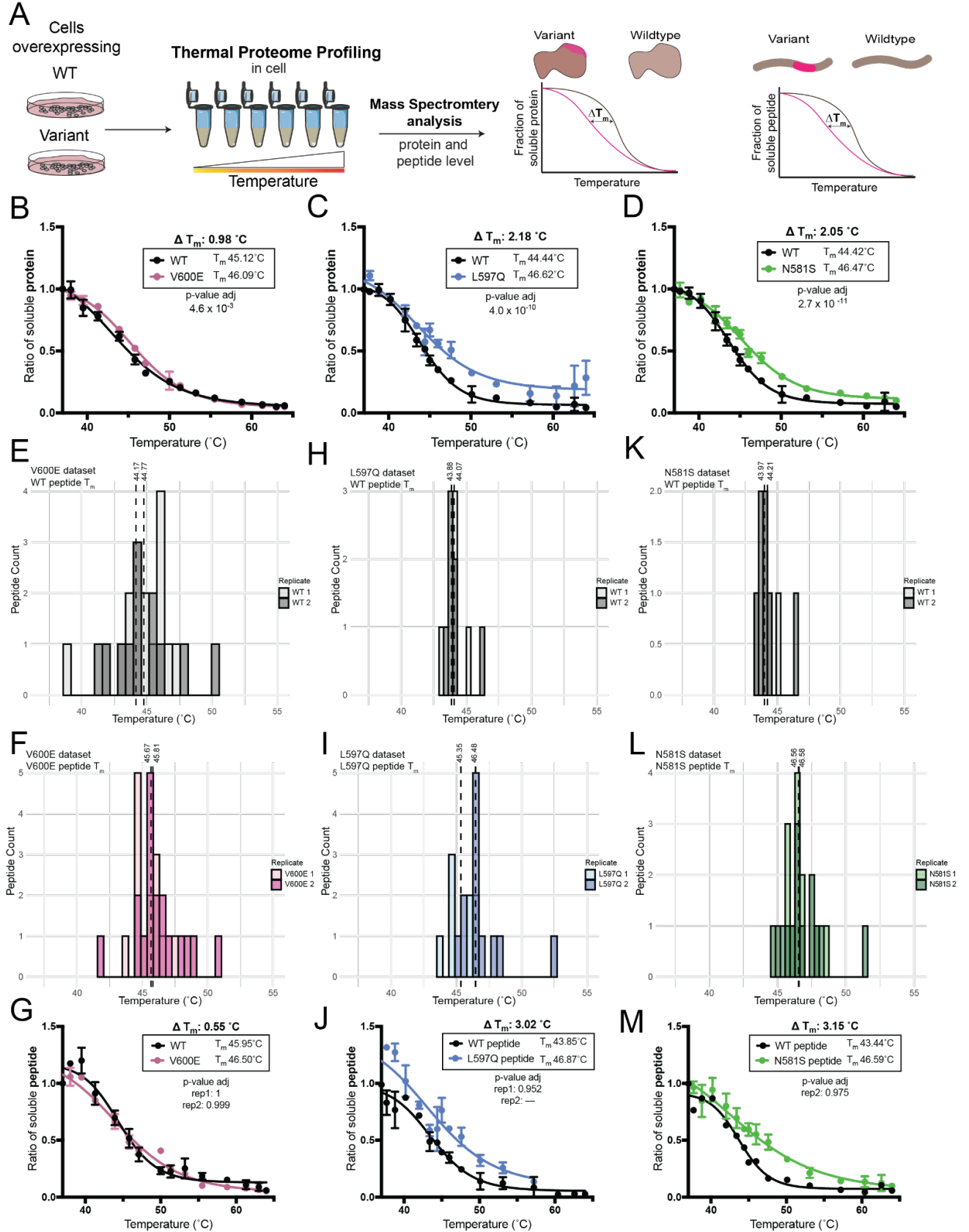
variant thermal stability at the protein-level. Since the goal application of this methodology would be to determine thermal stability differences between endogenous and variant proteins, the only way to navigate identifying variants by bottom-up mass spectrometry would be to exclusively consider the variant and WT peptide pairs. Therefore, peptide level quantification of both variant and WT peptides becomes necessary for a variant TPP approach.

We sought to validate peptide level quantification within the model B-Raf variants TPP data. Since each of these samples were expressing a different B-Raf SAAV, each dataset was searched against a database containing both the WT B-Raf sequence and the variant that was overexpressed. These data would include many overlapping B-Raf peptides in both WT and variant samples, while only the variant expressing sample would contain the variant peptide. In all cases, we were able to detect the WT counterpart to the variant peptide in the WT sample.

We fit all the peptide melt curves with the TPP R package<sup>37</sup>, as we had done with the protein-level data, to compare the WT and variant peptides, as well as examined each of the peptide  $T_m$  values to observe whether any outliers existed or if they all clustered around the protein  $T_m$ . For the WT comparison with the V600E variant, the median  $T_m$  across all the quantified peptides appears to closely align with the protein  $T_m$  (**Figure 3E**), as does the median  $T_m$  for the V600E quantified peptides (**Figure 3F**). When comparing the WT and variant peptide pairs for the V600E peptide-level analysis, a shift of only +0.55°C was observed (**Figure 3G**). The WT (**Figure 3H**) and L597Q (**Figure 3I**) median peptide  $T_m$  comparison for the L597Q dataset variant was consistent with the calculated protein-level  $T_m$ . Peptide-level analysis of the variant L597Q paired with the WT exhibited a shift of +3.02°C (**Figure 3J**). As seen with the other two datasets, the median  $T_m$  for the WT (**Figure 3K**) and N581S (**Figure 3L**) peptides in the N581S dataset are generally consistent with the protein level calculated  $T_m$ . The paired peptide analysis for N581S had a shift of +3.15°C (**Figure 3M**). With these data, we determined that the peptide  $T_m$  was generally representative of the overall protein  $T_m$ , with many of the peptides clustering around the measured  $T_m$  of the protein. More importantly, we can see similar effects between the protein-level and

peptide-level melt curve data within our model system, which gives us more confidence in utilizing peptide level thermal stability information from more complex systems containing multiple variants. Additionally, we foresee having to potentially rely on WT protein-level thermal stability in circumstances where a WT peptide counterpart to an endogenous variant is not identified or quantified. To illustrate what this might look like, we generated the melt curve comparisons of protein-level WT thermal stability and peptide-level variant thermal stability, which also align well with both the protein- and peptide-level analyses (**Supplementary Figure 7A-C**). In order to view these trends simultaneously across the multiple thermal stability comparisons for each of these variants, a summary table of all the  $T_m$  differences are aggregated in **Supplementary Table 1**.

We note that we were unable to analyze the A305G data at the peptide-level due to the lack of peptide identification/quantification for the variant containing region (FFEHHPIPQEEASLAETALTSGSSPSAPASDSIGPQILTSPSPSK). One of the limitations with relying on peptide-level quantification for variant TPP is that there will likely be proteins regions, such as with A305G, where the peptide of interest may be more difficult to identify due to the properties of the peptide or the stochastic nature of data-dependent acquisition. In the controlled model system where individual variants are overexpressed, we can rely on protein-level quantification of thermal stability in the absence of variant peptide identifications; however, this will not translate to variant TPP within systems containing multiple variants.



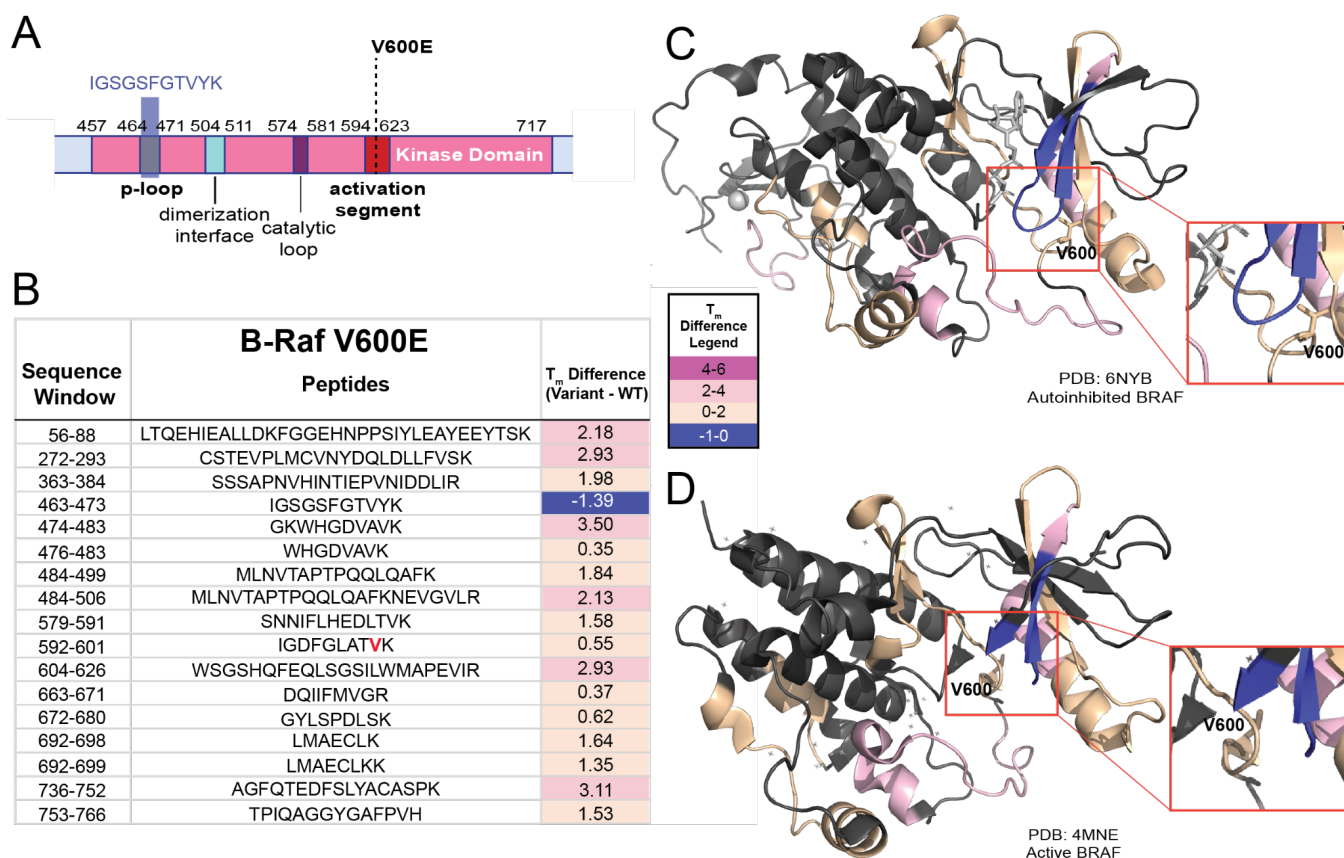
**Figure 3. Establishing variant TPP analysis with model B-Raf variants.** A) Overview of variant TPP workflow and analysis for a B-Raf SAAV overexpression model system to assess protein- and peptide-level thermal stability. B) Thermal stability analysis for B-Raf V600E mutant (pink) as compared to WT (black). C) Thermal stability analysis for B-Raf L597Q mutant (blue) as compared to WT (black). D) Thermal stability analysis for B-Raf N581S mutant (green) as compared to WT (black). E) Peptide-level analysis of the  $T_m$  of all WT peptides in the V600E comparison. Dotted lines represent the median  $T_m$  for each replicate. F) Peptide-level analysis of the  $T_m$  of all V600E peptides. Dotted lines represent the median  $T_m$  for each replicate. G) Peptide-level melt curves for matching WT peptide to the V600E variant peptide region. H) Peptide level analysis of the  $T_m$  of all WT peptides in the L597Q comparison. Dotted lines represent the median  $T_m$  for each replicate. I) Peptide level analysis of the  $T_m$  of all L597Q peptides. Dotted lines represent the median  $T_m$  for each replicate. J) Peptide level melt curves for matching WT peptide to the N581S variant peptide region. K) Peptide level analysis of the  $T_m$ s of all WT peptides in the N581S comparison. L) Peptide level analysis of the  $T_m$  of all L597Q peptides. Dotted lines represent the median  $T_m$  for each replicate. M) Peptide level melt curves for matching WT peptide to the N581S variant peptide region.  $n=1$  for WT peptide melt curve. All melt curves and  $T_m$  values are generated using the TPP R package<sup>37</sup> and the normalized values for ratio of soluble protein are exported into Prism for plotting and curve fitting. For all analyses within this figure  $n=2$  unless otherwise stated. Panels B,E,F,G correspond to the same LC-MS/MS dataset for V600E, panels C,H,I,J correspond to the same LC-MS/MS dataset for L597Q, and panels D,K,L,M correspond to the same LC-MS/MS dataset for N581S .

### Mapping Peptide $T_m$ Differences to B-Raf Active and Autoinhibited States

We next set out to determine if we could obtain more information about the effects of a SAAV on local protein intramolecular interactions using our peptide-level data. We first looked at the thermal stability differences across all the peptides detected for both B-Raf WT and each B-Raf variant. With this 2D view (**Figure 4A**) we observe V600E generally showing very small differences between melting temperatures, even for the peptide containing the SAAV, which only exhibited a difference in  $T_m$  of  $+0.55^\circ\text{C}$  (**Figure 4B**). Interestingly, the peptide IGSGSFSGTVYK (sequence region 463-473) appeared to be destabilized by  $-1.39^\circ\text{C}$  in the V600E variant as compared to WT (**Figure 4B**). The IGSGSFSGTVYK peptide exhibiting a destabilization effect was not in close 2D space to the V600 region containing the substitution. Since 2D space often greatly differs from 3D, we wanted to map these melting temperature differences to B-Raf onto the known structure of the protein. We took two PDB structures of B-Raf, a structure of the autoinhibited MAPK pathway complex (PDB: 6NYB) and an “active” BRAF:MEK1 complex (PDB: 4MNE), and mapped the peptide  $T_m$  differences to each. While the IGSGSFSGTVYK (sequence 463-473) region appears to be far from the V600 sequence in 2D space, it is in fact relatively close to the V600 residue in 3D space – the p-loop located within the 464-471 region about 6.8 Å away from the

V600 residue (**Figure 4C**). This p-loop is the loop which binds and stabilizes the ATP the phosphate groups<sup>46</sup>.

Interestingly, we are able to map this full sequence in the 6NYB autoinhibited structure (**Figure 4D**) but not in the 4MNE “active” structure as the 465-468 region of the p-loop is unresolved in this PDB structure (**Figure 4D**). The V600E mutation has been observed to increase the flexibility of the activation loop and that the V600 to E600 mutation mimics phosphorylation in the active loop<sup>47</sup>. For the L597Q variant, we also observe a similar destabilization of the IGSGSFQTVYK peptide (-1.02°C) (**Supplementary Figure 8A,C,E**) as well as the stabilization of the SNNIFLHEDLTVK peptide which is adjacent to the variant peptide we had already quantified as stabilized (**Figure 3J**). This is an interesting observation as the class I and class II variants both exhibit higher than normal (compared to WT) activity. We hypothesize that this destabilization we observe for the p-loop may be a difference in the activity state of B-Raf, while the stabilization we observe for L597Q may result from its requirement to signal as a homodimer as opposed to the V600E variant which signals as a monomer. In the N581S variant, we do not observe destabilization for the IGSGSFQTVYK peptide and instead observe the stabilization of the SNNIFLHEDLTVK variant region and the AGFQTEDFSQYACASPK peptide from the C-terminal region of the protein (**Supplementary Figure 8B,D,F**). We also hypothesize that this singular peptide destabilization within the protein thermal stability analysis may be due to sampling of a subset of the protein population which could be attributed to biased detection of the non-phosphorylated peptide IGSGSFQTVYK. While this region does not contain the major phosphorylation sites for B-Raf, phosphorylation of S465, S467, and T470 have been observed by others<sup>48-50</sup>. While this type of analysis is limited to overexpression systems, which tracks one variant at a time, it illustrates the utility of peptide level data for the assessment of thermal stability between SAAVs.



**Figure 4. Peptide T<sub>m</sub> mapped to active and inactive B-Raf states in 3D highlights variant-induced stability changes.** A) 2D sequence of B-Raf with highlighted regions of interest and V600E mutation. B) Table of B-Raf peptide sequence T<sub>m</sub> differences in B-Raf V600E compared to WT. V600 highlighted in hot pink. C) Difference in V600E vs WT peptide T<sub>m</sub> values mapped to a structure of autoinhibited B-Raf (PDB: 6NYB). D) Difference in V600E vs WT peptide T<sub>m</sub> values mapped to a structure of “active” B-Raf (PDB: 4MNE). Images created in PyMol<sup>51</sup>.

### Establishing Variant TPP in a cell line with pooled COSMIC B-Raf mutations

Our next objective was to apply our variant TPP methodology to a more complex system of B-Raf mutations, in which cells were engineered to express a mixed pool of B-Raf variants that were chosen from the COSMIC database<sup>10</sup>. We anticipated that this pooled model system could allow us to compare our variant TPP results for each of the variants to the previously reported characteristics of each of these variants through the LABEL-seq methodology. We constructed a custom database based on the 1437 B-Raf SAAVs that were engineered within this cell line to

include a total of 79 variant positions with every possible single amino acid replacement, including truncations (**Figure 5A**). Disappointingly, we first attempted to perform variant TPP within these cells using a bulk proteome (no immunoprecipitation) approach and searched them against our custom database containing all the expressed B-Raf variants. However, we were unable to detect any variants using our two-pass search to identify variants within our sample<sup>35</sup>. Taking a one-pass search approach yielded 1 B-Raf variant peptide per run and 2 B-Raf variants in aggregate (**Figure 5B**).

We postulated that coverage could be improved using an immunoprecipitation strategy. As prior work has indicated that introducing additional sample manipulation steps into TPP introduces additional variability<sup>27</sup>, we first assessed the performance of IP-TPP using a model protein, PTBP1 (**Supplementary Figure 9A**), which was readily available in our lab. We find generally good concordance between protein- and peptide-level melting temperatures (**Supplementary Figure 9B,C**), with peptides with R-squared values  $\geq 0.7$  giving a better representation of the protein-level melt curve data than the aggregated data prior to R-squared filtering (**Supplementary Figure 9D,E**). As we observed that the with R-squared filtering of  $\geq 0.7$  contained more peptides that were representative of protein-level data, we opted to adopt this cutoff as a filtering criteria for the peptide melt curves with the immunoprecipitated samples, which is already within the best practices for applying the TPP-TR package for protein-level TPP analysis<sup>37</sup>. The incorporation of an immunoprecipitation step proved necessary for the detection of multiple B-Raf variants within the cell line (**Figure 5B**).

To further improve our coverage of variants, we fractionated the HA-Braf IP samples using reverse phase high pH fractionation (**Figure 5A**). Our initial variant TPP analysis of these pooled B-Raf cells identified a limited number of variants, with only a fraction (~49%) of those being quantified with SPS-RTS-MS3. Across 2 technical replicates of 8 fractions we were able to identify



47 variants within our data through a FragPipe two-pass search with our custom variant database, with 23 of those containing quantified reporter ions from MS3 scans (**Figure 5B**). Even with fractionation, we noticed that the coverage for B-Raf SAAVs appeared to be on the lower end of what we expected since our pooled variant cell line contained 1437 SAAVs. Upon further evaluation, we noticed that only 25 of 79 mutated residue positions for B-Raf within this cell line are detectable based on our sequence coverage of B-Raf(**Supplementary Figure 10A**), with 25 of those variant positions distributed among 17 different peptides (**Supplementary Figure 10B**). We also note that many of our sequences are highly redundant and that high sequence coverage of each peptide needs to be achieved at MS2 for proper identification of these variants. We can see this within our spectra for the S467D variant (**Figure 5C**), where the WT peptide has greater b- and y-ion sequence coverage, likely due to its greater abundance within the bulk cell population in relation to the variant peptide. We hypothesize that due to the unique issue of multiple highly homologous sequences within our custom database for this engineered cell line, we are potentially identifying more variants at MS2 which are not sent to MS3 for quantification due to low peptide probability scores (XCorr) being calculated during the SPS-RTS-MS3 acquisition. An example of this can be seen with the variant G464D, where it is identified at MS2 in fraction 5 with a probability score of 0.9304 but within the RTS file, only the G464C variant is observed with an XCorr of 0.747 and a dCn score of 0.191, indicating more ambiguity in determining if it is a distinct peptide as compared to other similar peptides within the database. It also appears that the peptides that are being identified in the two-pass search at the MS2 level are being filtered out within FragPipe's TMT report for having a peptide probability of less than 0.90. As an example, we observe this with variant P192S, which has reported intensities within fraction 3 at the PSM-level but a probability score of 0.854, so it is excluded from the TMT report.

In addition to the peptide-level melt curve criteria, we noticed that the proteins being pulled down with B-Raf had increasing intensity values across increasing temperatures. This is likely due to

decreasing amounts of B-Raf available for pulldown post-thermal denaturation and ultracentrifugation, leading to more nonspecific interactions with the immunoprecipitation resin at higher temperatures. These increasing values among all the other proteins besides B-raf were affecting the normalization that the TPP-TR package applies to the melt curve fitting; therefore, we only processed the B-Raf peptides through the TPP-TR script.

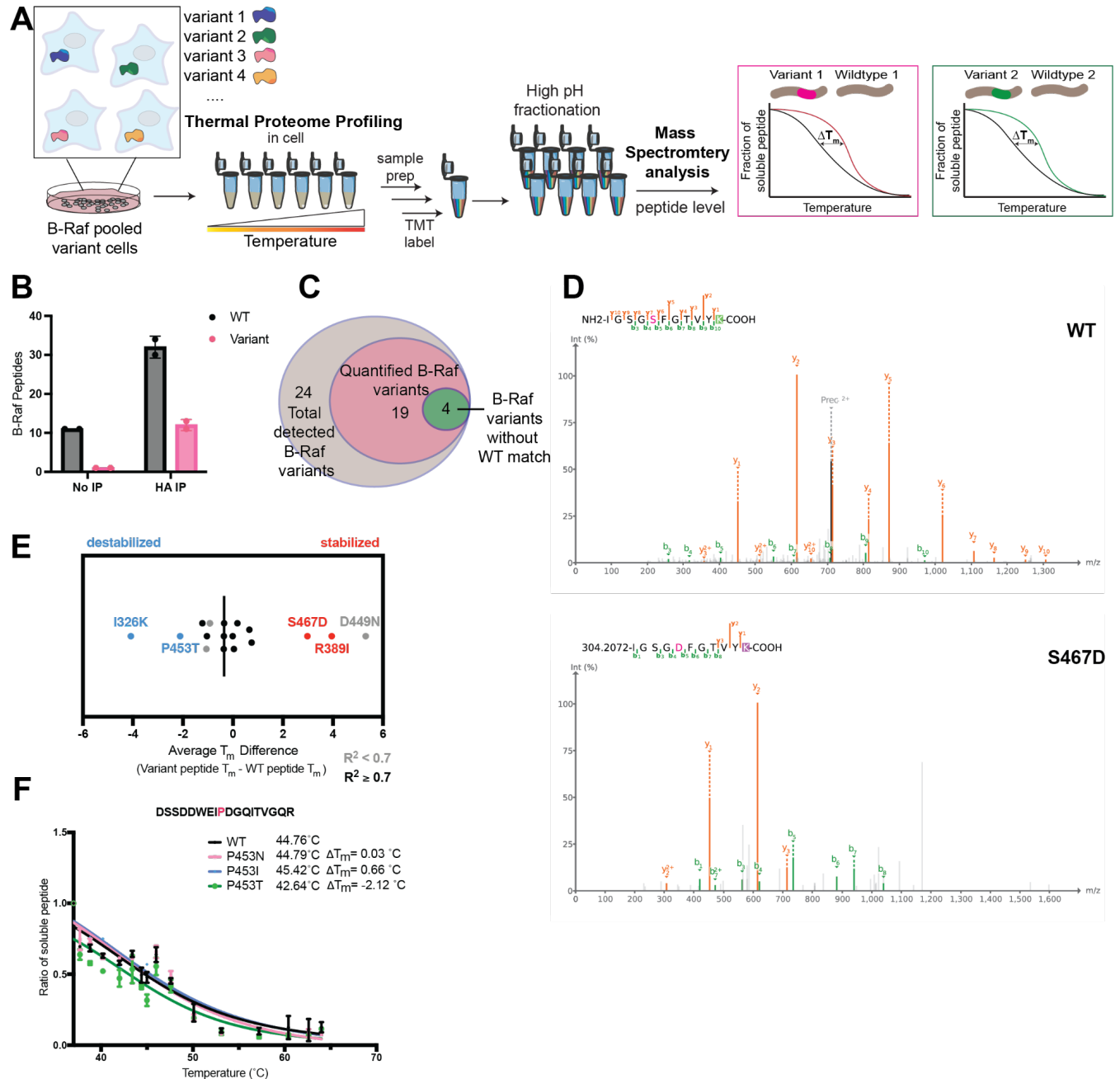
While our database contains 1437 SAAVs, 207 of those mutations are within tryptic peptides with a length of 6 or less amino acids and are therefore undetectable with our utilized search parameters, unless they are also associated with missed cleavage sites. Additionally, there are some variant peptides only observed as a result of a SAAV introducing a tryptic site. These variant peptides pose the unique challenge of not having a comparable WT counterpart, particularly if the tryptic sequence of the WT peptide is too long to detect (>50 aa). Within our analysis we identified 4 variants for which this was the case: S102K, S614K, E26R, and E80R (**Figure 5E**). For these variants, we were unable to calculate the thermal stability differences compared to the WT peptide so we have excluded these peptides from our analysis.

Looking at the results from the variant TPP analysis for 19 of the variants which were quantified alongside their WT peptide pair, we find that 5 of these variants appear to show thermal stability differences of 2 degrees or greater as compared to their counterpart peptide (**Figure 5E**). We detect two variants, P453T and I326K, that appear to be destabilizing (**Figure 5E**). At the P453 position we were able to quantify three different variants: P453N, P453I, P453T, with the P453T variant exhibiting the most destabilizing effect (**Figure 5F**). Comparing the LABEL-seq data for these variants, P453T is observed to be one of the more activating variants as compared to P453N and P453I<sup>10</sup>, consistent with our TPP analysis (**Figure 5D**).

Variant I326K, which appeared to be the most destabilized (**Figure 5E**), introduced a new tryptic site within a relatively long WT peptide of 45 amino acids which generated a 12-mer variant

peptide that does not contain the variant residue but can only exist if a new tryptic site was generated (**Supplementary Figure 11A**). The I326K variant can alternatively be attributed to the I326R variant as well. However, we were unable to make any conclusions about this variant due to the lack of quantification for these peptides in both of the technical replicates. This highlights that the reliable detection and quantification of variants across replicates is somewhat stochastic and that targeting variants within data acquisition would be beneficial.

We also observe three stabilizing variants, D449N, R389I, and S467D (**Supplementary Figure 11B**). The D449N variant appears to be most stabilized in the bulk analysis of the B-Raf variant thermal stability, and compared to other quantified variant peptides, D449W and D449M, D449N appears to have the greatest stabilization effect (**Figure 5E**). However, we observed that based on our melt curve filtering criteria for peptides, D449N does not pass the R-squared cutoff of  $>0.7$ , along with a couple other variants highlighted in grey within the bulk analysis of thermal stability differences (**Figure 5E**). In addition to the poor fit, D449N also suffers from the lack of quantification in both replicates. The other stabilizing variants R389I and S467D pass our peptide melt curve fit criteria, but also are not quantified in both replicates (**Supplementary Figure 11C,D**).



**Figure 5. Variant TPP analyses of a pooled B-Raf COSMIC variant cell line.** A) Workflow for pooled B-Raf variant cell line for variant TPP with fractionation. B) Improvement of B-Raf variant peptide coverage (pink) as compared to B-Raf WT peptides (grey) with addition of an immunoprecipitation step. Values reflect MS2 identified variants in a nonfractionated sample. C) Overview of total variant peptides identified at MS2 from a two-pass FragPipe search (24, tan), how many of those were quantified with SPS-RTS-MS3 (19, pink), and how many of the quantified variants lack a WT matched peptide (4, green). Quantified peptides only include peptides with a probability above 0.90. D) MS2 matched spectra for the IGSGSFGTVY WT peptide as compared to the variant peptide for S467D. Spectra generated from Fragpipe v22.0 PDV-viewer. E) Overview of average variant  $T_m$  subtracted from average WT  $T_m$  for the fully quantified variant-

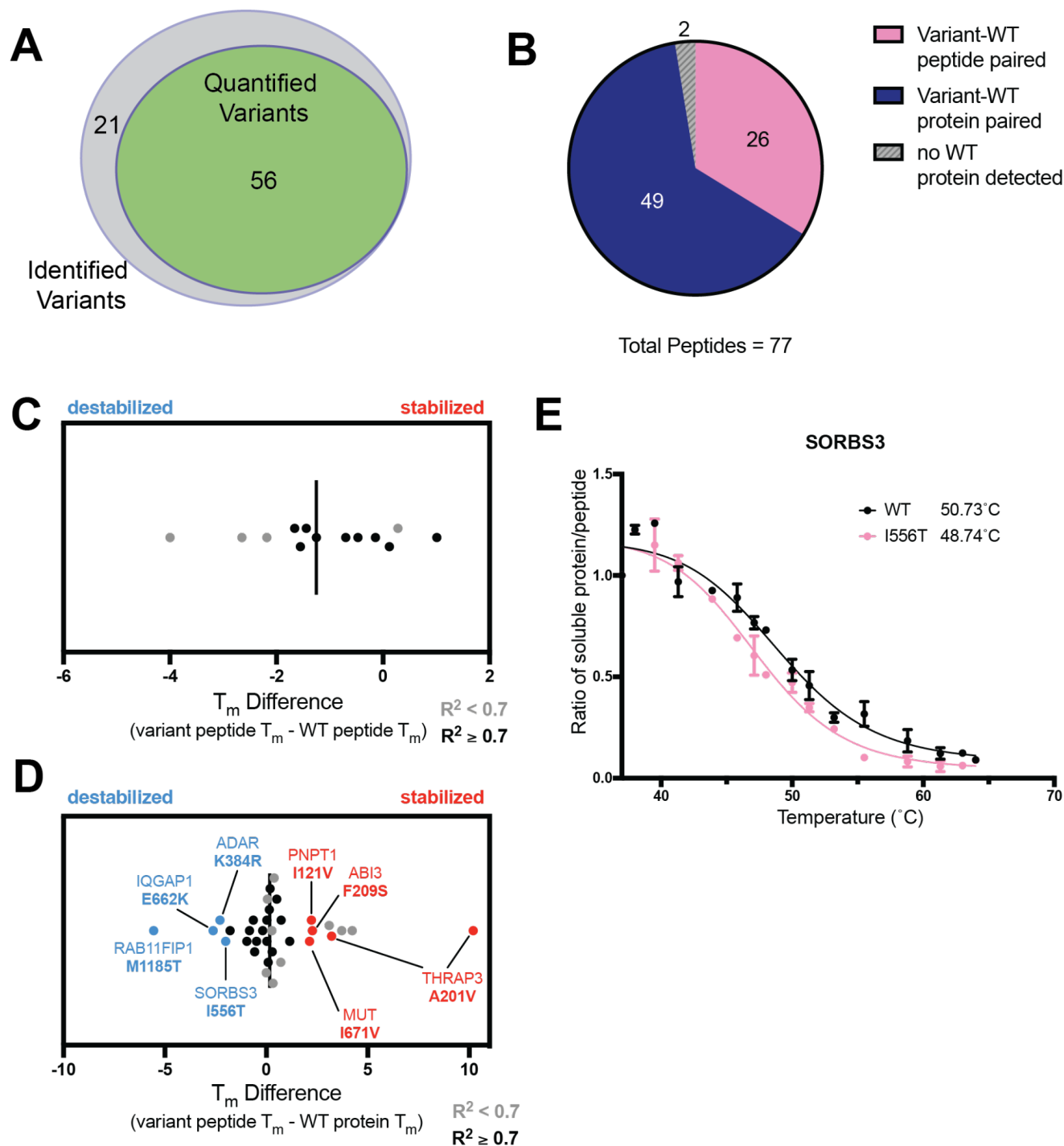
WT peptide pairs. Blue points represent variants exhibiting destabilization, red points represent variants showing stabilization, and grey points indicate peptides that had R-squared melt curve fits below 0.7. F) Comparisons of thermal stability for various mutants at the P453 position within the DSSDDWEIPDGQITVGQR peptide: P453N (pink), P453I (blue), and P453T (green). Variant position labeled in hot pink. Variant position labeled in hot pink. From n=2 technical replicates for all panels, unless noted that the peptide is not quantified in both replicates.

### **Applying Variant TPP to Molt-4 cells to detect thermal stability changes in endogenous variants**

We finally moved towards a cell line containing known endogenous variants, many of which are not sufficiently characterized. We chose to perform variant TPP on Molt-4 cells due to the large number of SAAVs found within this cell line and because of the readily available custom variant database generated from sequencing data<sup>35</sup>. Again, we chose to fractionate the TPP samples for the Molt-4 samples using reverse phase high pH fractionation and acquire the data SPS-RTS-MS3 to maximize both the number of detected variants and WT peptides we could get per protein to achieve as many WT-variant peptide pairs as possible. We achieved a total of 77 unique variants, with 21 of those only identified but not quantified and 56 of those quantified by MS3 (**Figure 6A**). There were 26 variant peptides which could be paired with a WT peptide counterpart and 49 variant peptides which needed to be compared to the WT protein-level thermal stability data (**Figure 6B**). Only 2 variants had no other peptides detected for that protein within the WT dataset and therefore could not be analyzed further.

When analyzing the peptide-level comparisons for variants with peptide comparators, we found the median  $T_m$  difference to be shifted towards destabilization. When adding in the filtering criteria for melt curve fits for each of these variants ( $R\text{-squared} \geq 0.7$ ), we observed that poor melt curve fits were likely contributing to this shift (**Figure 6C**). Within the variant peptide to protein-level comparisons, four protein variants showed a destabilizing effect: SORBS3 I556T, ADAR K384R, IQGAP1 E662K, RAB11FIP1 M1185T, and four showed a stabilizing effect: MUT I671V, PNPT1 I121V, ABI3F 209S (**Figure 6D**). SORBS3 I556T shows a destabilization of the SAAV with well

fitted curves across both replicates (**Figure 6E**). Vinexin (SORBS3) is known to be a vinculin-binding protein with three Src homology 3 (SH3) domains. The I556T mutation occurs between the second and third SH3 domains which are known to be the regions that bind to vinculin and the protein son of sevenless homolog 1 (SOS1), respectively<sup>52</sup>. This variant is also found within the COSMIC database, with the I556T substitution being the most detected variant for SORBS3 within COSMIC (18 entries logged)<sup>1</sup>. Thyroid hormone receptor-associated protein 3 (THRAP3) A201V is observed to be one of the more stabilizing SAAVs within this dataset (**Supplementary Figure 12A,B**). A variant peptide, DSRPSQAAGDNQGDEVKEQTFSGGTSQDTK, containing a missed cleavage site was detected in replicate 1 while a variant peptide DSRPSQAAGDNQGDEVK was detected in replicate 2. Both were compared to the protein-level melt curve for THRAP3 with the peptide DSRPSQAAGDNQGDEVKEQTFSGGTSQDTK showing a smaller stabilization effect than the DSRPSQAAGDNQGDEVK peptide (+3.22°C and +10.20°C, respectively), which can likely be attributed to the DSRPSQAAGDNQGDEVKEQTFSGGTSQDTK peptide having a better R-squared fit (0.82, **Supplementary Figure 12A**) as compared to the DSRPSQAAGDNQGDEVK peptide (0.72, **Supplementary Figure 12B**). However both peptides do show consistent stabilization for the A201V variant. Interestingly, THRAP3 A201V is also found within the COSMIC database and is also the most observed substitution within THRAP3 (13 entries logged)<sup>1</sup>. Studies have also implicated THRAP3 in multiple cancers, contributing to cell growth and proliferation<sup>53</sup>, as well as tumor progression<sup>54</sup>.



**Figure 6. Variant TPP analyses within the Molt-4 cell line.** A) Overview of variant peptides identified at MS2 from a two-pass FragPipe search (21, tan) and how many of those were quantified with SPS-RTS-MS3 (56, green). B) Overview of the number of variant peptides that were paired with a WT peptide comparator (26, pink) versus how many did not have a paired WT peptide and needed to be compared to the protein-level thermal stability (49, indigo). Two variant peptides did not have a WT protein detected to compare to (grey). C) Overview of average  $T_m$  differences for the fully quantified variant-WT peptide pairs. D) Overview of average  $T_m$  differences for variant peptide thermal stability paired with protein level thermal stability. From

n=2 replicates for all panels. E) Melt curve for SORBS3 I556T destabilization (pink) as compared to WT protein-level data for SORBS3 (black). n=2.

## Discussion

In establishing variant TPP using B-Raf as a model system, we observed thermal stability shifts between WT and SAAV through CETSA, protein-level and peptide-level TPP using mass spectrometry. Class II and class III variants, L597Q and N581S, exhibited stabilizing effects as compared to WT. Our analysis of these variants through both CETSA and TPP suggested that there does not appear to be a correlation between thermal stability shifts and the activity level of the B-Raf variant. The highly active V600E variant does not exhibit a change in thermal stability while the class III N581S, characterized by lower activity as compared to WT B-Raf, shows significant thermal stabilization. While we do observe detectable changes in SAAV thermal stability for select variants, we acknowledge that stabilization or destabilization does not correlate with altered activity, especially within the context of our small sample set of variants. Thermal stability changes are only a result of altered interactions that the protein makes with its surrounding environment. However, these altered interactions could provide useful clues into the effects of SAAVs, independent of protein activity.

We find that comparisons of both protein-level and peptide-level data yield reliable analysis for protein thermal stability comparisons in the context of comparing SAAVs to WT. This becomes useful within the context of analyzing endogenous variants, such as the variant TPP we applied to Molt-4 cells, as variant-peptide to WT-protein analyses yielded promising identifications of stabilizing and destabilizing variants. However, within a pooled system of multiple variants, such as the pooled B-Raf system that we explored, these peptide-to-protein melt curve analyses may not be as useful due to the high number of shared peptides between variants.

When analyzing our protein-level and peptide-level TPP data, we noticed that the selection of melt curve fitting and statistical analysis of the data made a difference in the significance calculations for these thermal stability shifts. The literature has shown that not all proteins have



the standard looking sigmoidal melt curves<sup>43</sup> and that simple comparison of estimated  $T_m$  values may result in failure to detect subtle changes in melting<sup>41,42</sup>.

Looking towards improving overall quantified variant coverage and improving paired variant and WT quantification, two areas can be addressed. First, our pooled B-Raf system highlighted challenges for identifying highly homologous peptide sequences as well as built the case for developing a variant TPP strategy that utilizes targeted acquisition softwares such as GoDig<sup>55,56</sup>. The challenge of quantifying variant and WT peptides across multiple replicates would be aided by spectral library generation for peptides of interest followed by targeted GoDig acquisition which would allow more robust PSM-level quantification of each variant. This could also be extended to endogenous variant systems with initial runs for spectral library generation followed by targeted proteomic acquisition. Further, the targeted acquisition approach could mitigate the poor melt-curve fits for variant peptides as PSM coverage per variant can be increased. Second, with many SAAVs either gaining or losing tryptic sites<sup>57</sup>, a multi-protease approach may also be useful but in a limited capacity to maintain compatibility with TMT reagents for TPP analyses.

## **Acknowledgements**

We thank Jessica J. Simon and Dustin J. Maly (Department of Chemistry, University of Washington, Seattle, WA, USA) for gifting the B-Raf COSMIC pooled cell line. We additionally thank all the members of the Backus lab for their helpful comments and suggestions.

## **Conflicts of interest**

K. M. B. is a member of the advisory board at Matchpoint Therapeutics. All other authors declare no competing interests.

## **Author contributions**

F. S. and K. M. B. writing and editing; F. S. investigation, analysis, methodology; F. S. and K. M. B. data curation; F. S. and K. M. B. conceptualization; K. M. B. supervision; K. M. B. project administration; K. M. B. funding acquisition.

## References

1. Tate, J. G. *et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res* 47, D941–D947 (2019).
2. Landrum, M. J. *et al.* ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res* 46, D1062–D1067 (2018).
3. Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 29, 308–311 (2001).
4. Arafeh, R., Shibue, T., Dempster, J. M., Hahn, W. C. & Vazquez, F. The present and future of the Cancer Dependency Map. *Nat Rev Cancer* 25, 59–73 (2025).
5. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581, 434–443 (2020).
6. Wei, A. *et al.* Investigating the sources of variable impact of pathogenic variants in monogenic metabolic conditions. *Nat Commun* 16, 5223 (2025).
7. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 46, 310–315 (2014).
8. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat Methods* 7, 248–249 (2010).
9. Ng, P. C. & Henikoff, S. Accounting for human polymorphisms predicted to affect protein function. *Genome Res* 12, 436–446 (2002).
10. Simon, J. J., Fowler, D. M. & Maly, D. J. Multiplexed profiling of intracellular protein abundance, activity, interactions and druggability with LABEL-seq. *Nature Methods* 21,

2094–2106 (2024).

11. Esposito, D. *et al.* MaveDB: an open-source platform to distribute and interpret data from multiplexed assays of variant effect. *Genome Biol* 20, 223 (2019).
12. Gygi, S. P., Rochon, Y., Franza, B. R. & Aebersold, R. Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol* 19, 1720–1730 (1999).
13. Savitski, M. M. *et al.* Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* 346, 1255784 (2014).
14. Kalxdorf, M. *et al.* Cell surface thermal proteome profiling tracks perturbations and drug targets on the plasma membrane. *Nat Methods* 18, 84–91 (2021).
15. Perrin, J. *et al.* Identifying drug targets in tissues and whole blood with thermal-shift profiling. *Nat Biotechnol* 38, 303–308 (2020).
16. Tan, C. S. H. *et al.* Thermal proximity coaggregation for system-wide profiling of protein complex dynamics in cells. *Science* 359, 1170–1177 (2018).
17. Kurzawa, N. *et al.* Deep thermal profiling for detection of functional proteoform groups. *Nat Chem Biol* 19, 962–971 (2023).
18. Dai, L. *et al.* Modulation of Protein-Interaction States through the Cell Cycle. *Cell* 173, 1481–1494.e13 (2018).
19. Rodina, A. *et al.* Systems-level analyses of protein-protein interaction network dysfunctions via epichaperomics identify cancer-specific mechanisms of stress adaptation. *Nat Commun* 14, 3742 (2023).
20. Mateus, A. *et al.* Thermal proteome profiling in bacteria: probing protein state. *Mol Syst Biol* 14, e8242 (2018).
21. Becher, I. *et al.* Pervasive Protein Thermal Stability Variation during the Cell Cycle. *Cell* 173, 1495–1507.e18 (2018).
22. Sridharan, S. *et al.* Proteome-wide solubility and thermal stability profiling reveals distinct regulatory roles for ATP. *Nat Commun* 10, 1155 (2019).

23. Potel, C. M. *et al.* Impact of phosphorylation on thermal stability of proteins. *Nat Methods* 18, 757–759 (2021).
24. Mateus, A. *et al.* Thermal proteome profiling for interrogating protein interactions. *Mol Syst Biol* 16, e9232 (2020).
25. Rodriguez-Mias, R. A. *et al.* Proteome-wide identification of amino acid substitutions deleterious for protein function. *bioRxiv* (2022) doi:10.1101/2022.04.06.487405.
26. Chang, H.-Y. *et al.* Analysis of isobaric quantitative proteomic data using TMT-Integrator and FragPipe computational platform. *bioRxiv* (2025) doi:10.1101/2025.05.27.656447.
27. Smith, I. R. *et al.* Identification of phosphosites that alter protein thermal stability. *Nat Methods* 18, 760–762 (2021).
28. Furtwängler, B. *et al.* Real-Time Search-Assisted Acquisition on a Tribrid Mass Spectrometer Improves Coverage in Multiplexed Single-Cell Proteomics. *Mol Cell Proteomics* 21, 100219 (2022).
29. Ting, L., Rad, R., Gygi, S. P. & Haas, W. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nat Methods* 8, 937–940 (2011).
30. Davies, H. *et al.* Mutations of the BRAF gene in human cancer. *Nature* 417, 949–954 (2002).
31. Dankner, M., Rose, A. A. N., Rajkumar, S., Siegel, P. M. & Watson, I. R. Classifying BRAF alterations in cancer: new rational therapeutic strategies for actionable mutations. *Oncogene* 37, 3183–3199 (2018).
32. Mechahougui, H., Gutmans, J., Gouasmi, R., Smekens, L. & Friedlaender, A. BRAF Targeting Across Solid Tumors: Molecular Aspects and Clinical Applications. *Int J Mol Sci* 26, (2025).
33. Yuan, J. *et al.* The dimer-dependent catalytic activity of RAF family kinases is revealed through characterizing their oncogenic mutants. *Oncogene* 37, 5719–5734 (2018).
34. Martinez Molina, D. *et al.* Monitoring drug target engagement in cells and tissues using the

- cellular thermal shift assay. *Science* 341, 84–87 (2013).
35. Desai, H. *et al.* Chemoproteogenomic stratification of the missense variant cysteinome. *Nat Commun* 15, 9284 (2024).
  36. Jagtap, P. *et al.* A two-step database search method improves sensitivity in peptide sequence matches for metaproteomics and proteogenomics studies. *Proteomics* 13, 1352–1357 (2013).
  37. Franken, H. *et al.* Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nat Protoc* 10, 1567–1593 (2015).
  38. Wan, P. T. C. *et al.* Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell* 116, 855–867 (2004).
  39. Yao, Z. *et al.* Tumours with class 3 BRAF mutants are sensitive to the inhibition of activated RAS. *Nature* 548, 234–238 (2017).
  40. Chernobrovkin, A. L. *et al.* A Tale of Two Tails: Efficient Profiling of Protein Degradation by Specific Functional and Target Engagement Readouts. *SLAS Discov* 26, 534–546 (2021).
  41. McCracken, N. A., Peck Justice, S. A., Wijeratne, A. B. & Mosley, A. L. Inflect: Optimizing computational workflows for Thermal Proteome Profiling data analysis. *J. Proteome Res.* 20, 1874–1888 (2021).
  42. Childs, D. *et al.* Nonparametric Analysis of Thermal Proteome Profiles Reveals Novel Drug-binding Proteins. *Mol Cell Proteomics* 18, 2506–2515 (2019).
  43. Hevler, J. F., Verma, S., Soijtra, M. & Bertozzi, C. R. Thermal Tracks: A Gaussian process-based framework for universal melting curve analysis enabling unconstrained hit identification in thermal proteome profiling experiments. *ArXiv* (2025).
  44. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9, 676–682 (2012).
  45. Jafari, R. *et al.* The cellular thermal shift assay for evaluating drug target interactions in

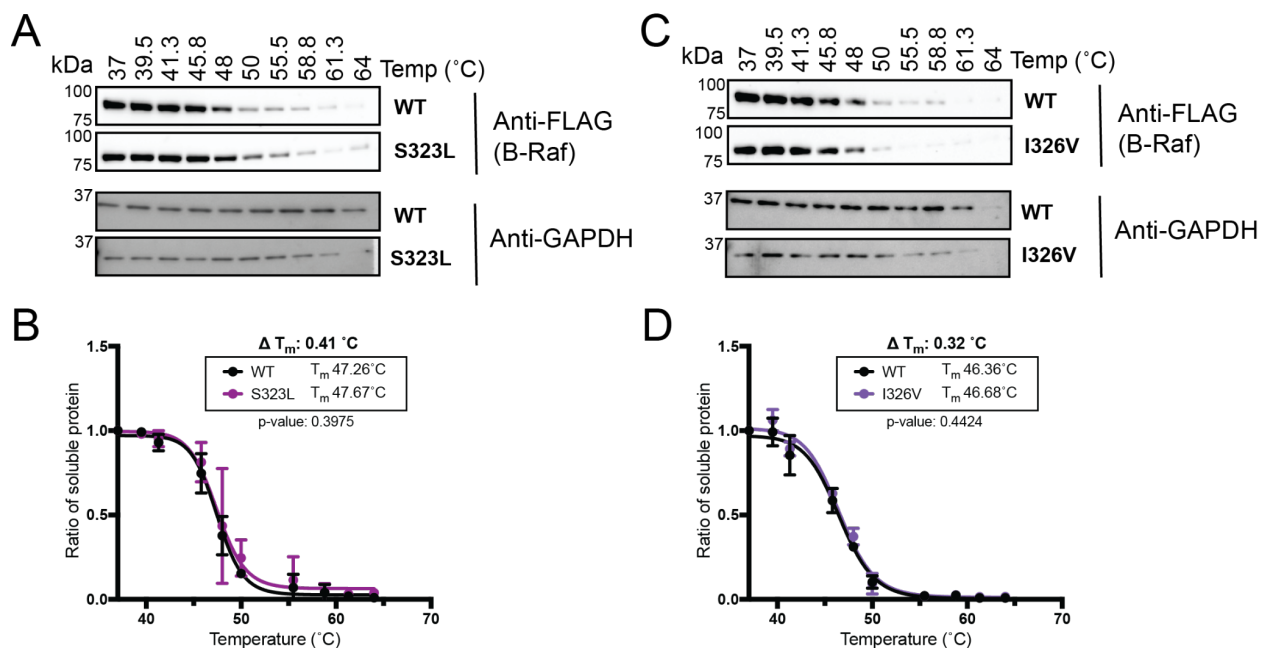
- cells. *Nat Protoc* 9, 2100–2122 (2014).
46. Holderfield, M., Deuker, M. M., McCormick, F. & McMahon, M. Targeting RAF kinases for cancer therapy: BRAF-mutated melanoma and beyond. *Nat Rev Cancer* 14, 455–467 (2014).
  47. Maloney, R. C., Zhang, M., Jang, H. & Nussinov, R. The mechanism of activation of monomeric B-Raf V600E. *Comput Struct Biotechnol J* 19, 3349–3363 (2021).
  48. Mertins, P. *et al.* Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature* 534, 55–62 (2016).
  49. Cook, J. G., Bardwell, L. & Thorner, J. Inhibitory and activating functions for MAPK Kss1 in the *S. cerevisiae* filamentous-growth signalling pathway. *Nature* 390, 85–88 (1997).
  50. Eisenhardt, A. E. *et al.* Phospho-proteomic analyses of B-Raf protein complexes reveal new regulatory principles. *Oncotarget* 7, 26628–26652 (2016).
  51. Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. Preprint at (2015).
  52. Kioka, N. *et al.* Vinexin: a novel vinculin-binding protein with multiple SH3 domains enhances actin cytoskeletal organization. *J Cell Biol* 144, 59–69 (1999).
  53. Kang, H. J. *et al.* Thrap3 promotes R-loop resolution via interaction with methylated DDX5. *Exp Mol Med* 53, 1602–1611 (2021).
  54. Ino, Y. *et al.* Phosphoproteome analysis demonstrates the potential role of THRAP3 phosphorylation in androgen-independent prostate cancer cell growth. *Proteomics* 16, 1069–1078 (2016).
  55. Yu, Q. *et al.* Sample multiplexing-based targeted pathway proteomics with real-time analytics reveals the impact of genetic variation on protein expression. *Nat Commun* 14, 555 (2023).
  56. Shuken, S. R., Yu, Q. & Gygi, S. P. Inserting Pre-analytical Chromatographic Priming Runs Significantly Improves Targeted Pathway Proteomics with Sample Multiplexing. *J Proteome*

Res 23, 1834–1843 (2024).

57. Tsuber, V., Kadamov, Y., Brautigam, L., Berglund, U. W. & Helleday, T. Mutations in Cancer Cause Gain of Cysteine, Histidine, and Tryptophan at the Expense of a Net Loss of Arginine on the Proteome Level. *Biomolecules* 7, (2017).

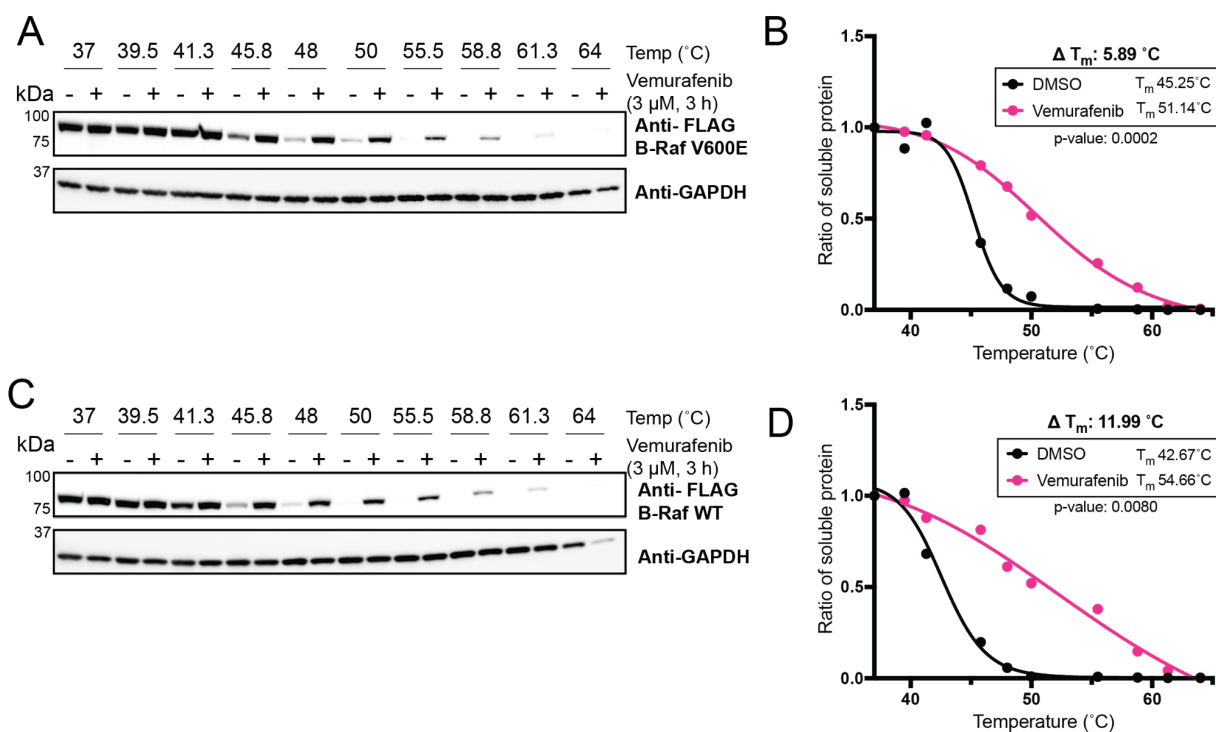
## Chapter 4 Supporting Information

### Supplementary Figures

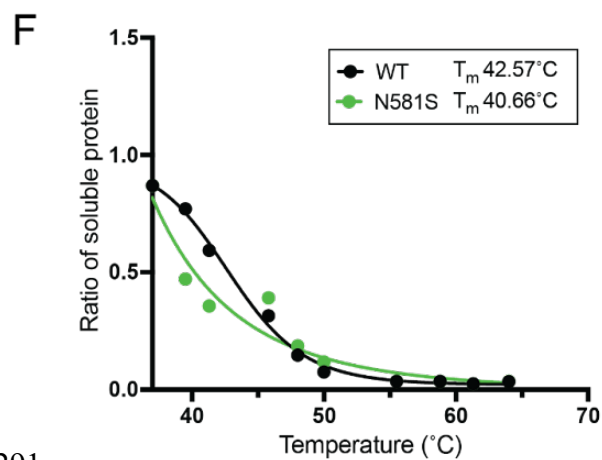
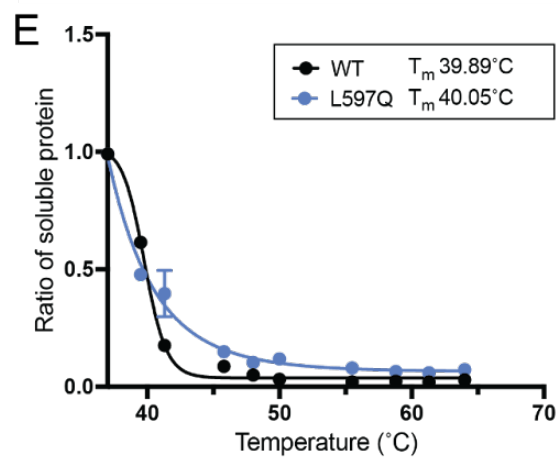
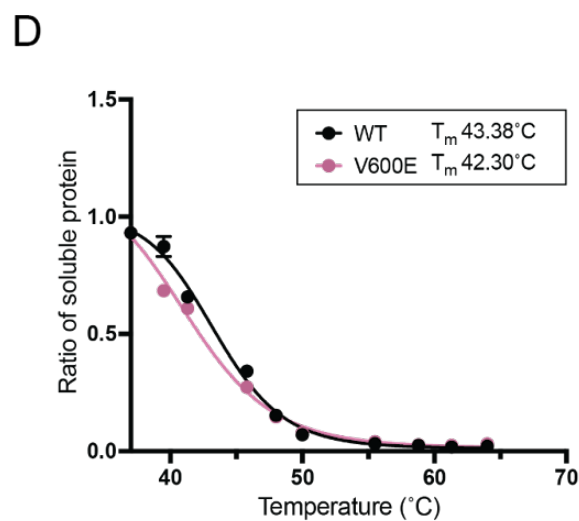
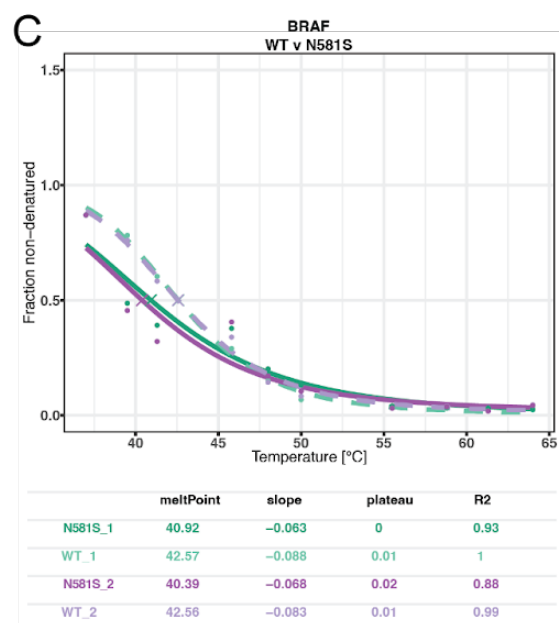
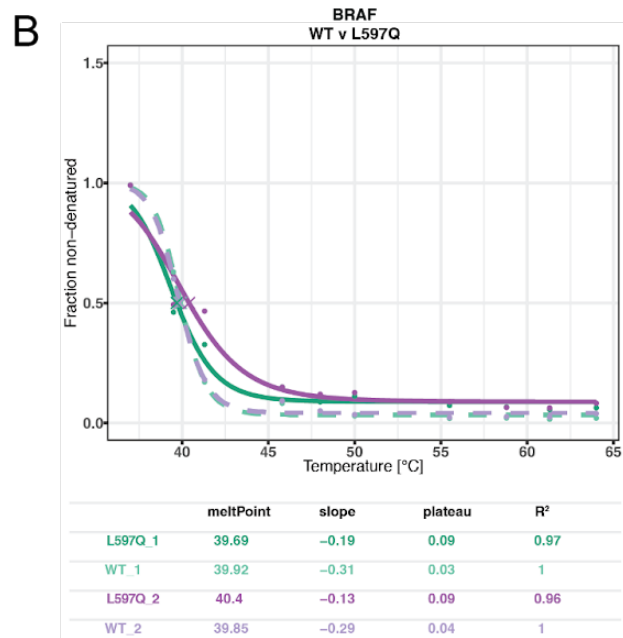
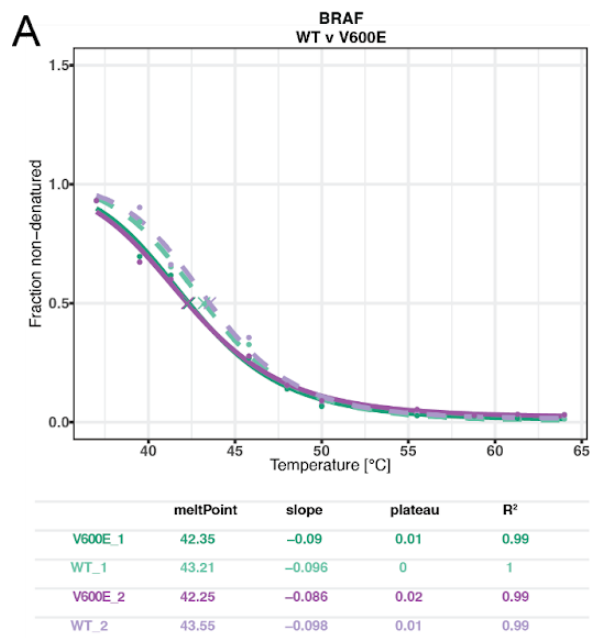


**Supplementary Figure 1. Blot-based CETSA analysis for transiently expressed FLAG-B-Raf variants classified as benign.** A) Comparison of blot analysis for B-Raf WT and S323L CETSA with 21k centrifugation to separate out insoluble proteins. B) Melt curve fit quantification of blot analysis for B-Raf S323L (n=2) CETSA as compared to WT B-Raf. C) Comparison of blot analysis for B-Raf WT and I326V CETSA with 21k centrifugation to separate out insoluble proteins. D) Melt curve fit quantification of blot analysis for B-Raf I326V (n=2) CETSA as compared to WT B-Raf. FIJI (ImageJ quantification)<sup>1</sup>, plotted into GraphPad Prism v 10.1, and non-linear curve fitted with interpolation sigmoidal 4PL with no special handling for outliers. An extra-sum-of-squares F test was performed for the IC<sub>50</sub> values ( $T_m$ ).

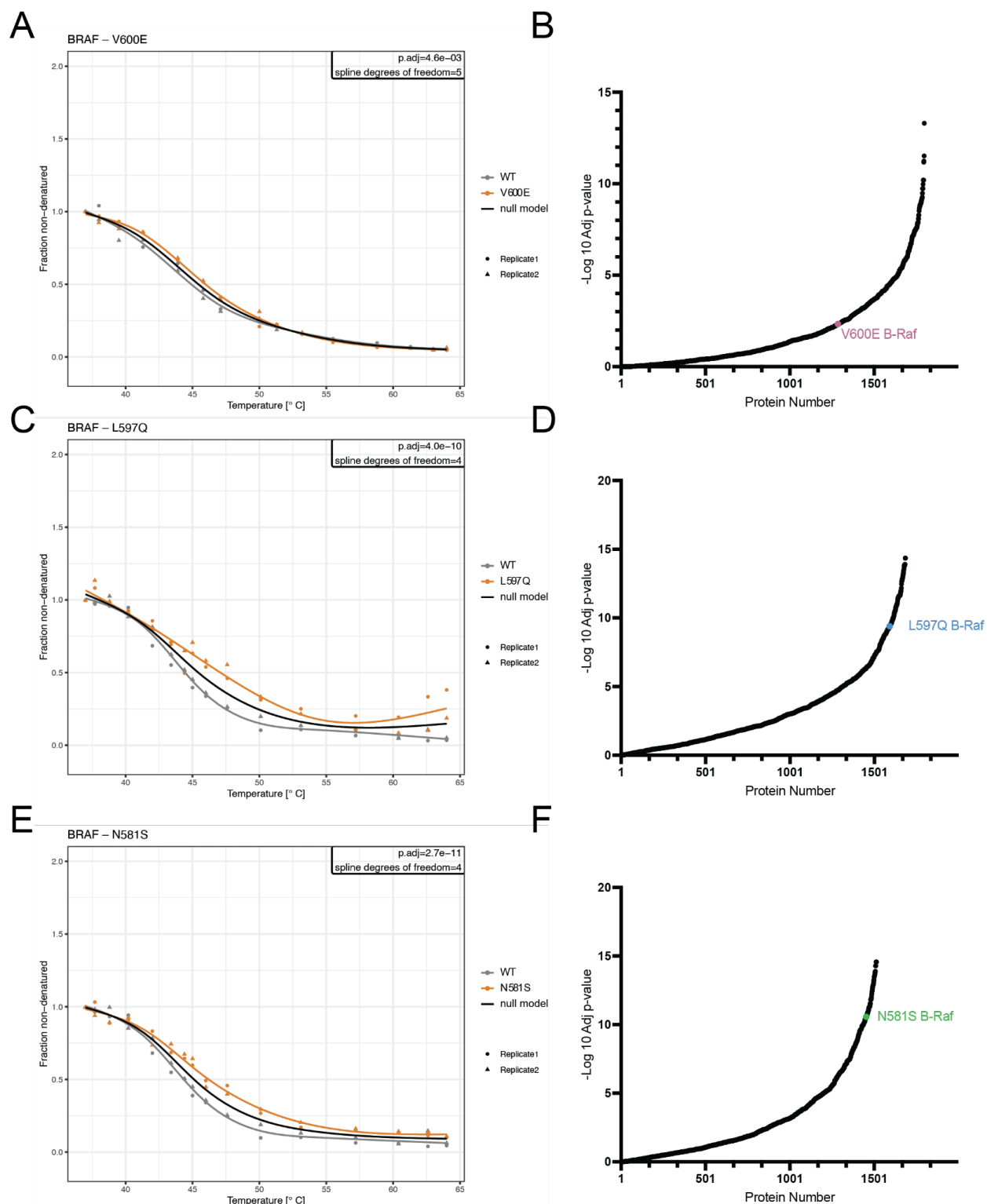




**Supplementary Figure 2. Drug binding TPP induces a much larger  $T_m$  shift by CETSA.** A) HEK cells overexpressing V600E FLAG-BRAF which are treated with either DMSO or 3  $\mu$ M Vemurafenib<sup>2</sup> (PLX4032) for 3 h then subjected to CETSA analysis. B) Western blot quantification of B-Raf V600E Vemurafenib treatment (n=1). C) HEK cells overexpressing WT FLAG-BRAF which are treated with either DMSO or 3  $\mu$ M Vemurafenib (PLX4032) for 3 h then subjected to CETSA analysis. D) Western blot quantification of B-Raf WT Vemurafenib treatment (n=1). FIJI (ImageJ quantification), plotted into GraphPad Prism v 10.1, and non-linear curve fitted with interpolation sigmoidal 4PL with no special handling for outliers. An extra-sum-of-squares F test was performed for the IC50 values ( $T_m$ ).

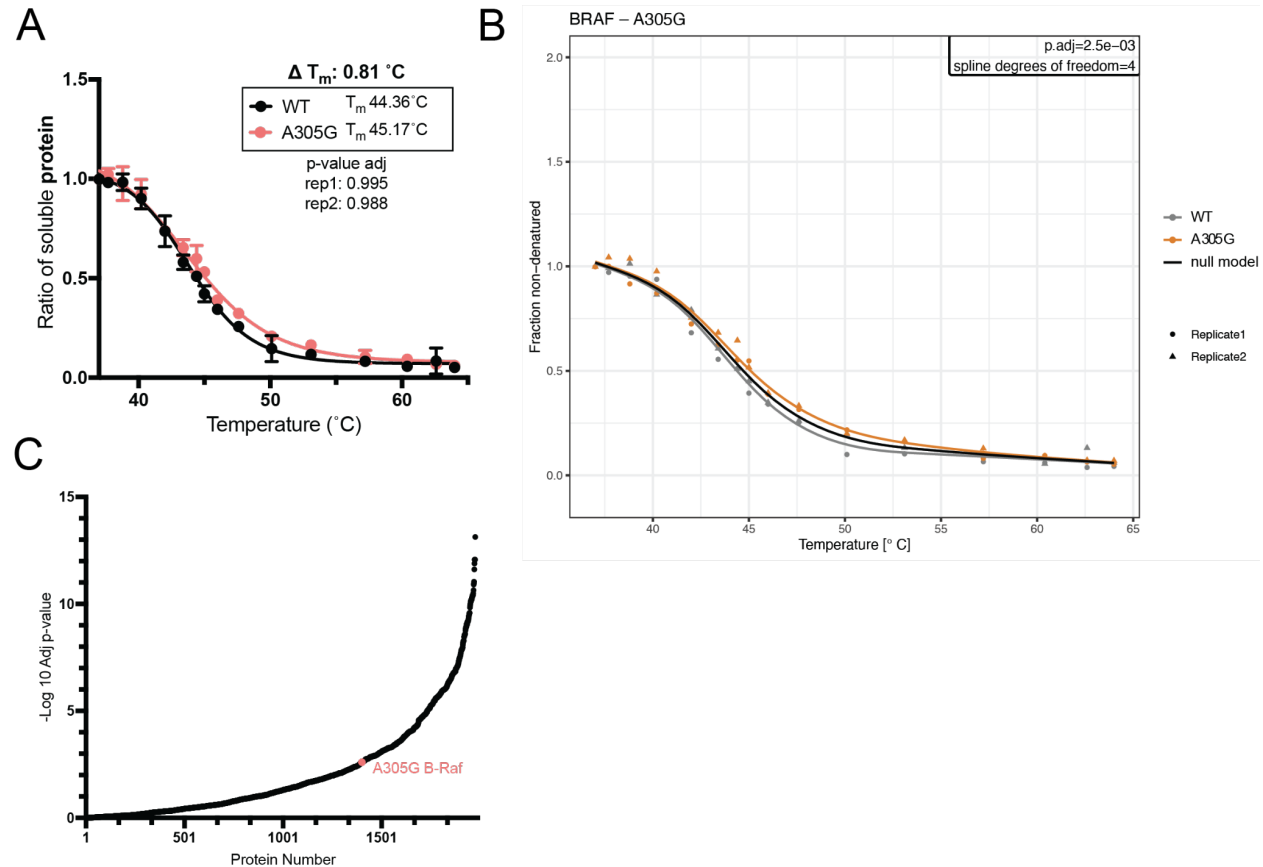


**Supplementary Figure 3. Quantified TPP with FLAG IP of B-Raf melt curves from mass spectrometry analysis.** A) TPP analysis for B-Raf WT as compared to B-Raf V600E with FLAG immunoprecipitation (n=2) as generated by the R package for TPP-TR. B) TPP analysis for B-Raf WT as compared to B-Raf L597Q with FLAG immunoprecipitation (n=2) as generated by the R package for TPP-TR. C) TPP analysis for B-Raf WT as compared to B-Raf N581S with FLAG immunoprecipitation (n=2) as generated by the R package for TPP-TR. D-F) TPP analysis for B-Raf WT as compared to B-Raf V600E (D), L597Q (E), and N581S (F) with FLAG immunoprecipitation (n=2) as generated by GraphPad Prism v 10.1 with data from the TPP-TR R package<sup>3</sup> and fitted with interpolation sigmoidal 4PL with no special handling for outliers.

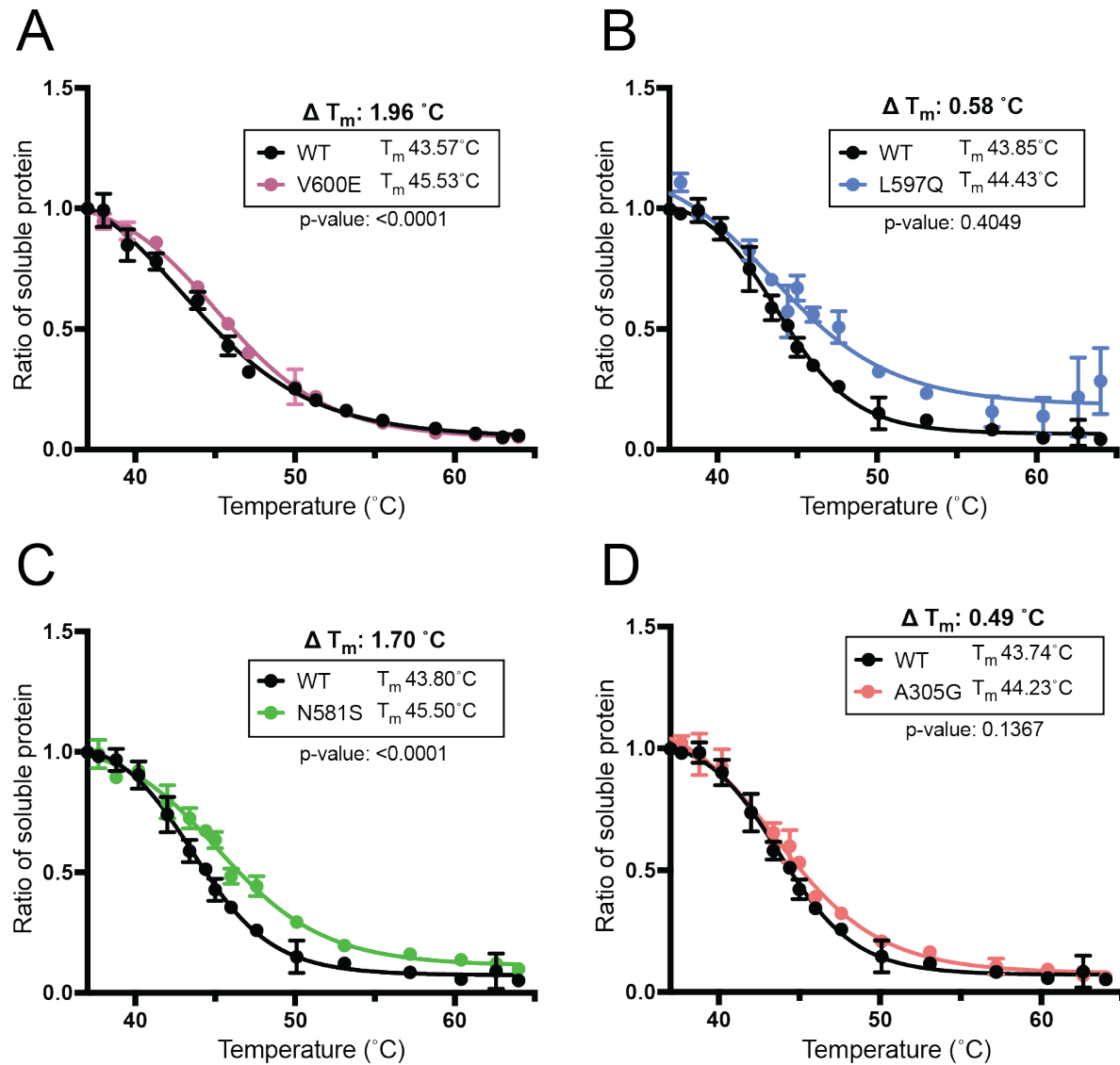


**Supplementary Figure 4. TPP-TR NPARC analysis of protein-level melt curves for transiently expressed FLAG-B-Raf variants.** A) B-Raf V600E variant protein-level analysis. B) B-Raf V600E adjusted p-value calculation from NPARC analysis as compared across global calculated adjusted p-values within the protein dataset. C) B-Raf L597Q variant protein-level analysis. D) B-Raf L597Q adjusted p-value calculation from NPARC analysis as compared

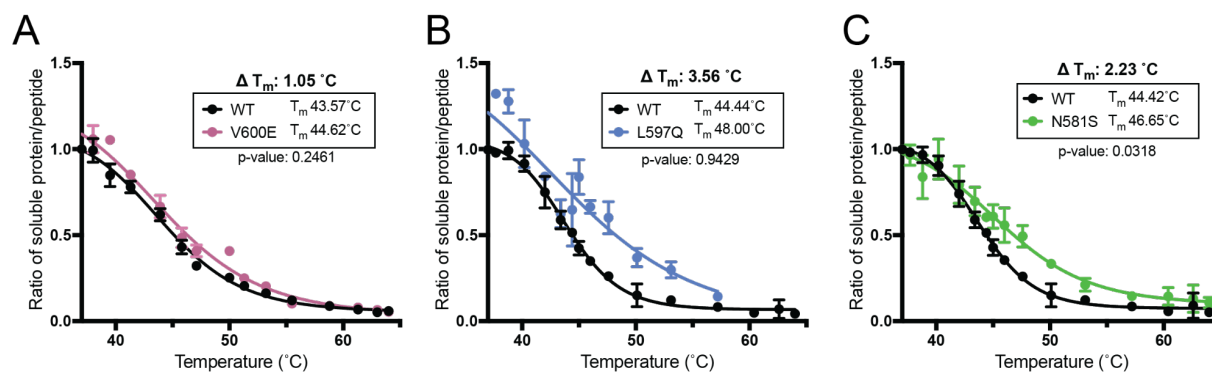
across global calculated adjusted p-values within the protein dataset. E) B-Raf N581S variant protein-level analysis. F) B-Raf N581S adjusted p-value calculation from NPARC analysis as compared across global calculated adjusted p-values within the protein dataset. Melt curves were fitted with “splinefit” method within the TPP-TR package for NPARC analysis<sup>4</sup>.



**Supplementary Figure 5. TPP-TR analysis of protein-level melt curves for transiently expressed FLAG-B-Raf A305G.** A) B-Raf A305G variant protein-level analysis fitted with “meltcurvefit” method within the TPP-TR package<sup>3</sup>. B) B-Raf A305G variant protein-level analysis fitted with “splinefit” method within the TPP-TR package for NPARC analysis<sup>4</sup>. C) B-Raf A305G adjusted p-value calculation from NPARC analysis as compared across global calculated adjusted p-values within the protein dataset.



**Supplementary Figure 6. Quantified TPP protein-level melt curves for transiently expressed FLAG-B-Raf variants** TPP-TR fitted melt curves to melt curves fitted with GraphPad Prism non-linear curve sigmoidal 4PL interpolation. A) B-Raf WT as compared to class I V600E mutant. B) B-Raf WT compared to class II L597Q mutant. C) B-Raf WT compared to class III N581S mutant. D) B-Raf WT compared to A305G mutant. Melt curves generated by GraphPad Prism v 10.1 with data from the TPP-TR R package<sup>3</sup> and fitted with interpolation sigmoidal 4PL with no special handling for outliers.

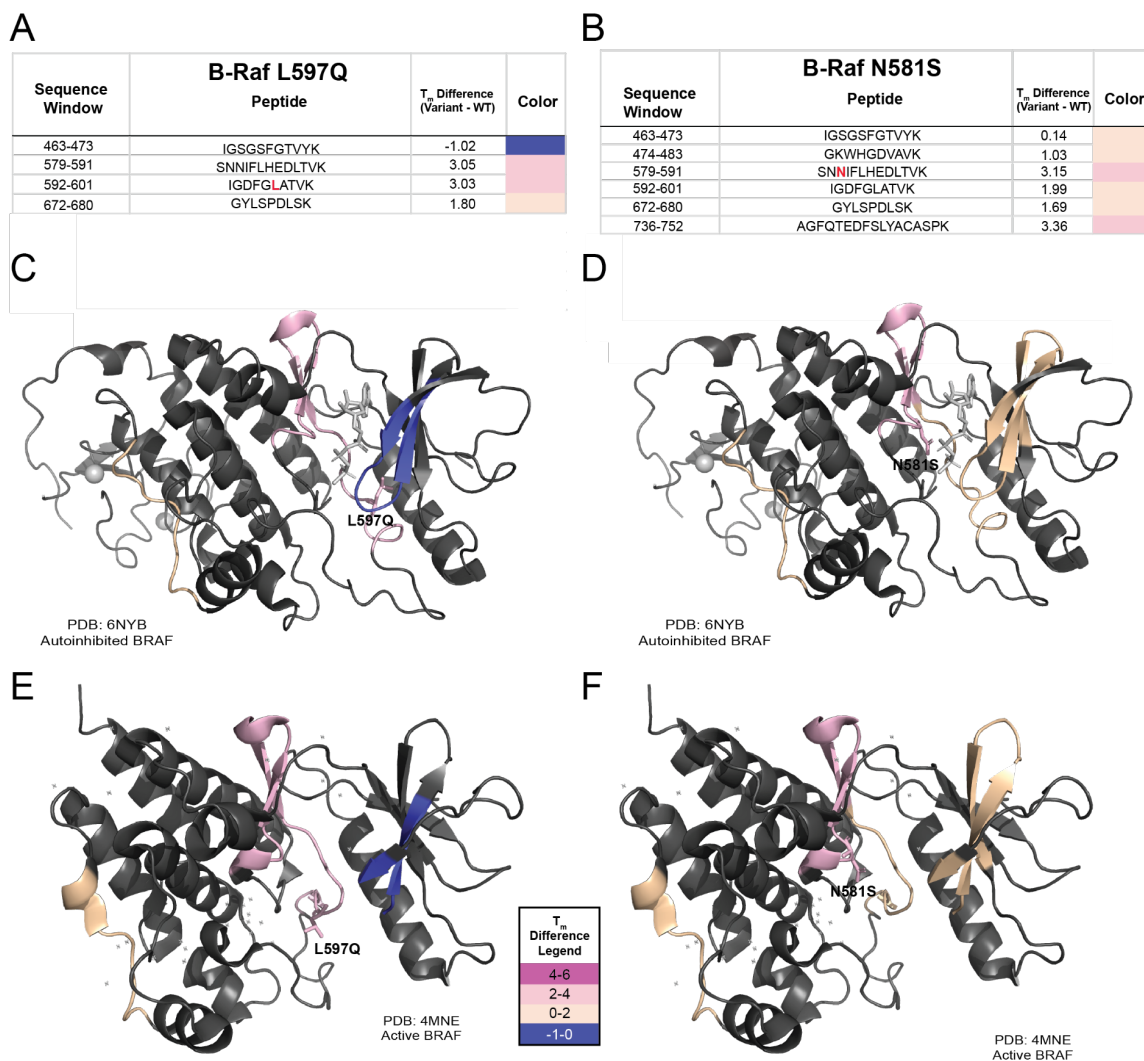


**Supplementary Figure 7. Quantified variant TPP WT protein-level melt curves compared to variant peptide-level melt curves.** A) B-Raf WT as compared to class I V600E mutant. B) B-Raf WT compared to class II L597Q mutant. C) B-Raf WT compared to class III N581S mutant. Melt curves generated by GraphPad Prism v 10.1 with data from the TPP-TR R package<sup>3</sup> and non-linear curve fitted with interpolation sigmoidal 4PL with no special handling for outliers.

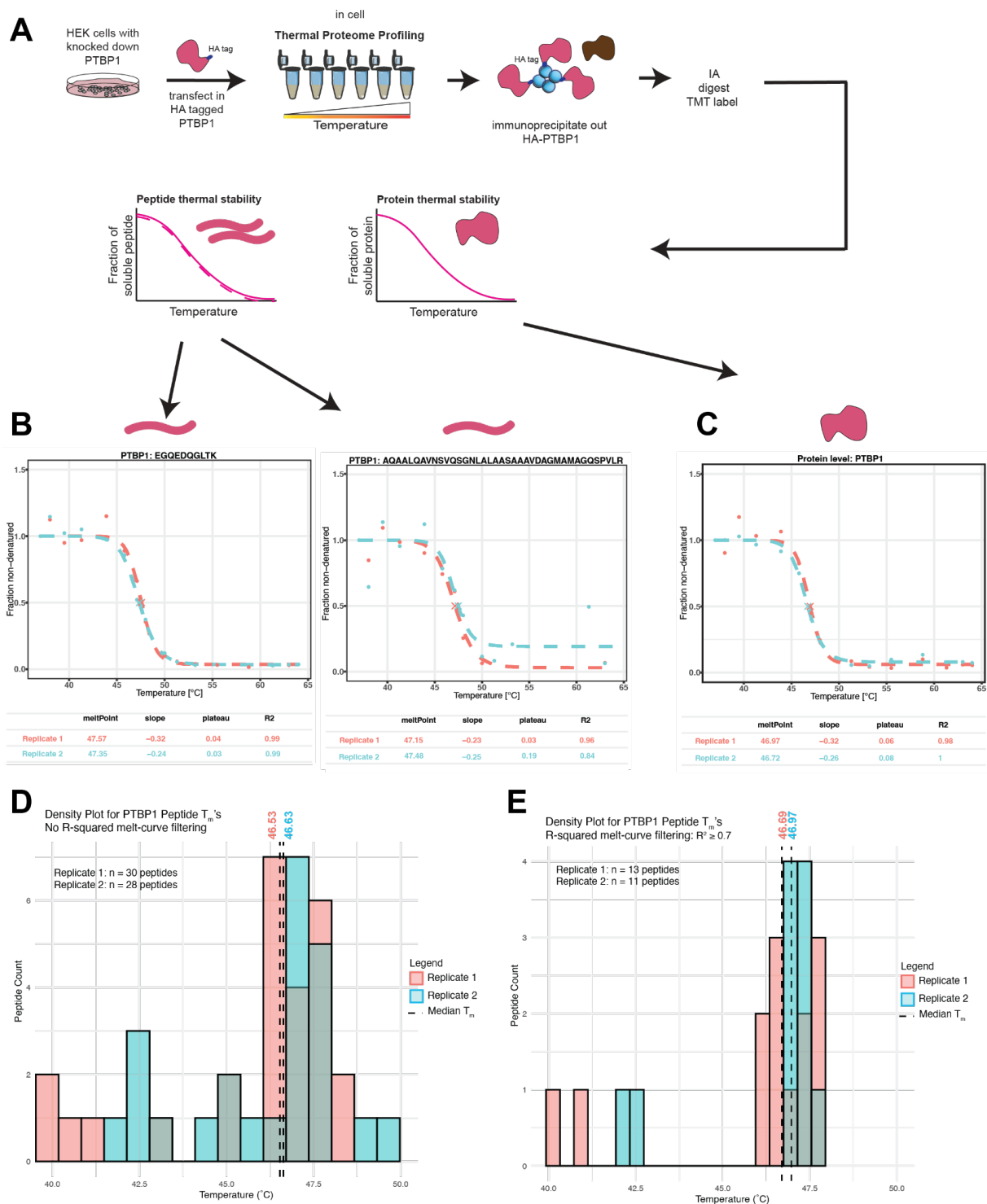
| B-Raf Proteoform | Experiment Type     | T <sub>m</sub> Difference (°C) | p-value | Adjusted p-value             | NPARC adjusted p-value |
|------------------|---------------------|--------------------------------|---------|------------------------------|------------------------|
| V600E            | CETSA               | -0.09                          | 0.9000  | —                            | —                      |
|                  | TPP - protein level | 0.98                           | <0.0001 | Rep 1: 0.999<br>Rep 2: 1     | 4.6x10 <sup>-3</sup>   |
|                  | TPP - peptide level | 0.55                           | 0.8941  | Rep 1: 1<br>Rep 2: 0.99      |                        |
| L597Q            | CETSA               | 0.50                           | 0.5821  | —                            | —                      |
|                  | TPP - protein level | 2.18                           | 0.4049  | p-value adj:<br>0.963, 0.546 | 4.0x10 <sup>-10</sup>  |
|                  | TPP - peptide level | 3.02                           | 0.6566  | p-value adj:<br>0.952, —     |                        |
| N581S            | CETSA               | 2.35                           | <0.0001 | —                            | —                      |
|                  | TPP - protein level | 2.05                           | <0.0001 | p-value adj:<br>0.522, 0.893 | 2.7x10 <sup>-11</sup>  |
|                  | TPP - peptide level | 3.15                           | 0.0685  | p-value adj: NA,<br>0.975    |                        |
| A305G            | CETSA               | -1.54                          | 0.0223  |                              | —                      |
|                  | TPP - protein level | 0.81                           | 0.1367  | p-value adj:<br>0.995, 0.998 | 2.5x10 <sup>-3</sup>   |
| S323L            | CETSA               | 0.41                           | 0.3975  | —                            | —                      |
| I326V            | CETSA               | 0.32                           | 0.4424  | —                            | —                      |

**Supplementary Table 1. Melting temperature differences for each B-Raf variant across all thermal profiling approaches utilized.** Aggregation of all the calculated T<sub>m</sub> differences (variant T<sub>m</sub> - WT T<sub>m</sub>) for CETSA analysis and both protein- and peptide- level TPP analyses. P-values calculated with the extra sum-of-squares F test in GraphPad Prism v10.1 with normalized data from the TPP-TR R package<sup>3</sup> non-linear curve fitted with interpolation sigmoidal 4PL with no special handling for outliers. The adjusted p-values are calculated from the TPP-TR results output from the R package analysis.





**Supplementary Figure 8. Quantified TPP peptide level melt temperature differences from mass spectrometry analysis mapped to B-Raf 3D protein structure.** A) 2D representation of peptide sequence T<sub>m</sub> differences in B-Raf L597Q compared to WT. B) 2D representation of peptide sequence T<sub>m</sub> differences in B-Raf N581S compared to WT. C) B-Raf L597Q mutant peptide melting temperature differences mapped onto an autoinhibited B-Raf structure (PDB 6NYB). D) B-Raf N581S mutant peptide melting temperature differences mapped onto an autoinhibited B-Raf structure (PDB 6NYB). E) B-Raf L597Q mutant peptide melting temperature differences mapped onto an active B-Raf structure (PDB 4MNE). F) B-Raf N581S mutant peptide melting temperature differences mapped onto an active B-Raf structure (PDB 4MNE).



**Supplementary Figure 9. Establishing the incorporation of immunoprecipitation for TPP of the HA-B-Raf COSMIC cell line using HA-PTBP1 as a model case study.** A) Workflow for proof of concept pulldown TPP experiment to validate peptide and protein thermal stability correlation using overexpressed HA-PTBP1 as the target protein in cells with knocked down PTBP1. B) Two representative thermal stability curves for two different peptides identified for PTBP1 with good melt curve fits. Across 2 replicates, a total of 41 peptides were detected (33 quantified) and 614 total spectral counts observed for PTBP1 with pulldown. C) Protein thermal

stability curves for PTBP1. D) Histogram plot for melt temperatures for all the PTBP1 peptides without filtering for R-squared values. E) Histogram plot for melt temperatures for PTBP1 peptides passing R-squared criteria for melt curves. n=2 replicates for HA-PTBP1 TPP with pulldown. Melt curves generated using the TPP R package<sup>3</sup>.

**A**

```

MAALSGGGGG GAEPGQALFN GDMPEEAGAG AGAAASSAAD PAIPEEVWNI KQMIKLTQEH IEALLDKFGG
EHNPPSIYLE AYEEYTSKLD ALQREQQLL ESLNGTDFS VSSASMDTV TSSSSSSLSV LPSSLSVFQN
PTDVARSNPK SPQKPIVRVF LPNKQRTVVP ARCGVTVRDS LKKALMMRGL IPECCAVYRI QDGEKKPIGW
DTDISWLTGE ELHVEVLENV PLTTHNFVRK TFFTLAFCDF CRKLLFQGFR CQTCGYKFHQ RCSTEVPLMC
VNYDQLDLLF VSKFFEHHPI PQEEASLAET ALTSGSSPSA PASDSIGPQI LTSPSPSKSI PIPQPFRRPAD
EDHRNQFGQR DRSSAPNVH INTIEPVNID DLIRDQGFRG DGGSTTGLSA TPPASLPGSL TNVKALQKSP
GPQERKSSS SSEDRNRMT LGRRDSSDDW EIPDGQITVG QRIGSGSFGT VYKGKWHGDV AVKMLNVTAP
TPQQLQAFKN EVGVLRKTRH VNILLFMGYS TKPQLAIVTQ WCEGSSLYHH LHIETKFEM IKLIDIARQT
AQGMDYLHAK SIIHRDLKSN NIFLHEDLTV KIGDFGLATV KSRWSGSHQF EQLSGSILWM APEVIRMQDK
NPYSFQSDVY AFGIVLYELM TGQLPYSNIN NRDQIIFMVG RGYLSPDLSK VRSNCPKAMK RLMAECLKKK
RDERPLFPQI LASIELLARS LPKIHRSASE PSLNRAGFQT EDFSLYACAS PKTPIQAGGY GAFPVH

```

**Legend**

Undetected sequence

Detected peptide

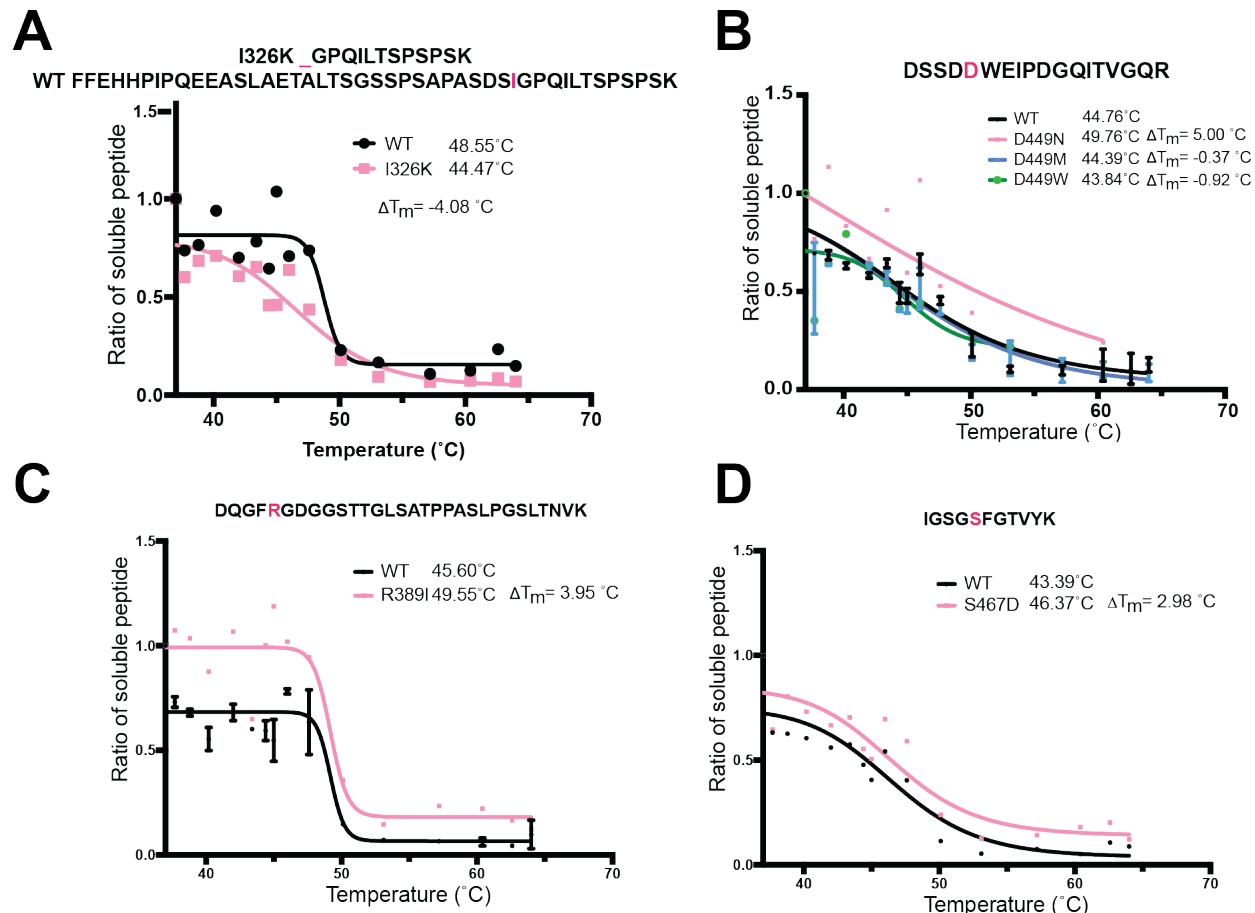
Variant Position: within detected peptide

Variant Position: not within detected sequence

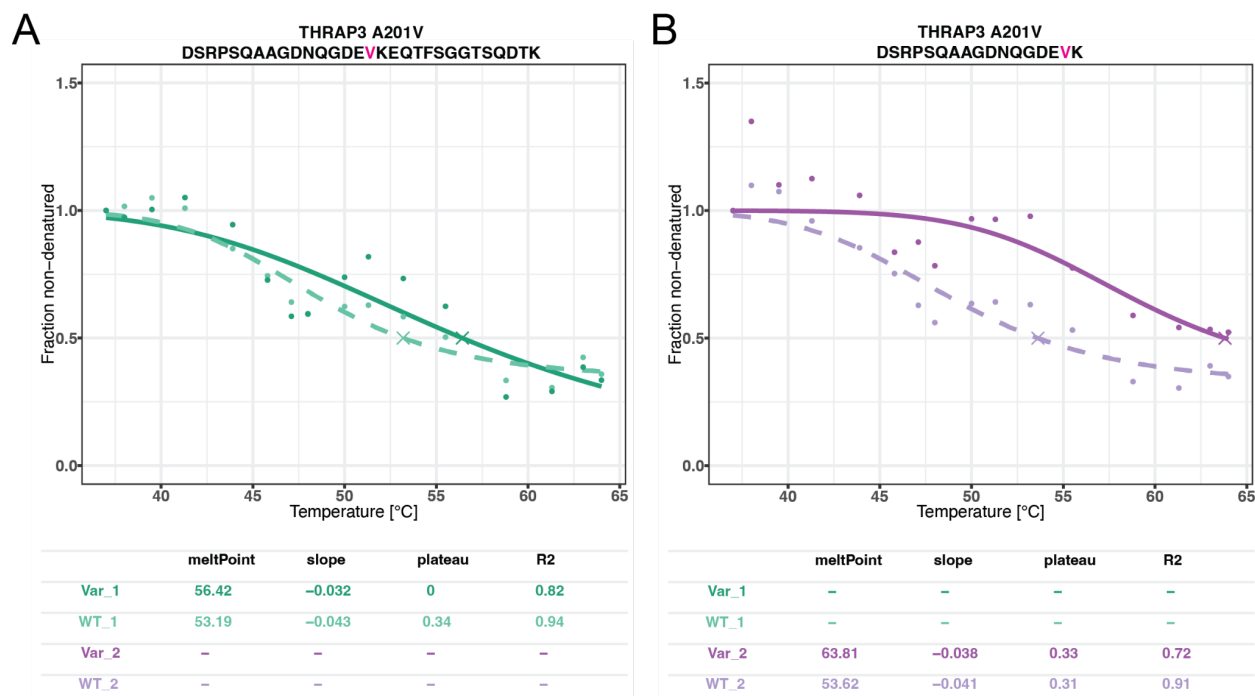
**B**

| residue #<br>start | Peptide                                       | residue #<br>end |
|--------------------|-----------------------------------------------|------------------|
| 189                | GLIPECCAVYR                                   | 199              |
| 254                | LLFQGFR                                       | 260              |
| 261                | CQTCGYK                                       | 267              |
| 294                | FFEHHPIPQEEASLAETALTSGSSPSAPASDSIGPQILTSPSPSK | 338              |
| 339                | SIPIQPFR                                      | 347              |
| 363                | SSSAPNVHINTIEPVNIDDLIR                        | 384              |
| 385                | DQGFR                                         | 389              |
| 444                | RDSSDDWEIPDGQITVGQR                           | 462              |
| 463                | IGSGSFGTVYK                                   | 473              |
| 484                | MLNVTAPTPQQLQAFK                              | 499              |
| 500                | NEGVLR                                        | 506              |
| 579                | SNNIFLHEDLTVK                                 | 591              |
| 592                | IGDFGLATVK                                    | 601              |
| 663                | DQIIFMVGR                                     | 671              |
| 691                | RLMAECLKK                                     | 699              |
| 727                | SASEPSLNR                                     | 735              |
| 753                | TPIQAGGYGAFPVH                                | 766              |

**Supplementary Figure 10. B-Raf variant positions in relation to quantified peptides by TPP analysis.** A) Visualization of B-Raf sequence and the variants that could possibly exist within the sequences that were detected and quantified by the variant TPP workflow and as quantified by SPS-RTS-MS3. Pink regions represent detected and quantified, green bolded amino acids represent potential variant sites within the detected regions, and blue bolded positions represent variant positions within peptides not quantified by the variant TPP workflow. B) List of all the detected and quantified B-Raf WT peptides with the bolded green amino acids being variant positions that can be detected.



**Supplementary Figure 11. B-Raf variant peptides compared to matched WT quantified peptides by TPP analysis.** A) Variant I326K compared to WT peptide with much longer sequence due to introduction of tryptic site with SAAV (variant peptide GPQILTSPSPSK) within the FFEHHPIPQEEASLAETALTS GSSPSAPASDSIGPQILTSPSPSK peptide. Variant position labeled in hot pink. Variant and WT peptide each identified in one replicate of 2, n=1. B) Comparisons of thermal stability for various mutants at the D449 position within the DSSDDWEIPDGQITVGQR peptide: D449W (green), D449M (blue), and D449N (pink). Variant position labeled in hot pink. C) Comparison of thermal stability for R389I within the DQGFRGDGGSTTGLSATPPASLPGLTNVK peptide. Variant position labeled in hot pink. Variant identified in one replicate of 2, n=1, WT peptide identified in both replicates, n=2. D) Comparison of thermal stability for S467D within the IGSGSFGTVYK peptide. Variant position labeled in hot pink. Variant and WT peptide each identified in one replicate of 2, n=1.



**Supplementary Figure 12. THRAP3 melt curves for A201V variant detected within the Molt-4 variant TPP data.** A) Variant A201V for THRAP3 with peptide DSRPSQAAGDNQGDEVKEQTFSGGTSQDTK (variant peptide shown) containing a missed cleavage site detected in replicate 1 and compared to the protein-level melt curve for THRAP3. B) Variant A201V for THRAP3 with peptide DSRPSQAAGDNQGDEVK (variant peptide shown) detected in replicate 2 and compared to the protein-level melt curve for THRAP3. Variant position is shown in hot pink on peptide sequence.

## Supplementary Tables

**Table S1-11.** Datasets corresponding to each figure, provided in the attached supplementary files.

Table S8: Figure 3 Proteomics Data

Table S9: Figure 5 Proteomics Data

Table S10: Figure 6 Proteomics Data

Table S11: Files in Proteomics Identification Database (PRIDE) datasets

**Table S4.** Antibodies used in this study

| Name | Catalog number | Lot number | Concentration |
|------|----------------|------------|---------------|
|------|----------------|------------|---------------|

|                                                  |                          |           |             |
|--------------------------------------------------|--------------------------|-----------|-------------|
| IRDye® 800CW Goat anti-Rabbit Secondary Antibody | LI-COR (926-32211)       | D20803-10 | 1:5000 (IB) |
| IRDye 680RD Donkey anti-Mouse Secondary Antibody | LI-COR (926-68072)       | D20322-25 | 1:5000 (IB) |
| DYKDDDDK Tag (D6W5B) Rabbit mAb                  | Cell signaling (14793S)  | 7         | 1:3000 (IB) |
| GAPDH Mouse mAb                                  | Proteintech (60004-1-Ig) | 10080731  | 1:3000 (IB) |

**Table S5.** Primers used in this study

| Gene            | Forward primer                                    | Reverse primer                                  |
|-----------------|---------------------------------------------------|-------------------------------------------------|
| FLAG-BRAF WT    | 5'-<br>CCACAGAGACCTCAAGAG<br>TAATAATATATTTCTTC-3' | 5'-<br>GAGGTCTTCATGAAGAAAT<br>ATATTATTACTCTT-3' |
| FLAG-BRAF A305G | 5'-<br>GGCAGTCTCTGCTAAGGA<br>GCCCTCTTCCTGTGG-3'   | 5'-<br>CCAATACCACAGGAAGAG<br>GGCTCCTTAGCAG-3'   |
| FLAG-BRAF S323L | 5'-<br>CCAATAGAGTCCAGGGCG<br>GGTGCGG-3'           | 5'-<br>CCCTTCCGCACCCGCCCT<br>GGA CTCTATTG-3'    |
| FLAG-BRAF I326V | 5'-<br>GAGAATTTGGGGCCCCAC<br>AGAGTCCGAG-3'        | 5'-<br>CGCCTCGGACTCTGTGGG<br>GCCCCAAATTC-3'     |

**Table S6.** Plasmids used in this study

| Plasmid number | Figure | Plasmid             | Source                    | Resistance    |
|----------------|--------|---------------------|---------------------------|---------------|
| 1              | 2-3    | pDONR223 BRAF V600E | Addgene kit (#1000000103) | Spectinomycin |
| 2              | 2-3    | pDONR223 BRAF L597Q | Addgene kit (#1000000103) | Spectinomycin |
| 3              | 2-3    | pDONR223 BRAF N581S | Addgene kit               | Spectinomycin |

|    |     |                                                       |                                                             |                                |
|----|-----|-------------------------------------------------------|-------------------------------------------------------------|--------------------------------|
|    |     |                                                       | (#1000000103)                                               |                                |
| 4  | 2-3 | GateWay compatible N-terminal FLAG destination vector | Gift from T Wucherpfennig                                   | Chloramphenicol and ampicillin |
| 5  | 2-3 | FLAG-BRAF V600E expression vector                     | GateWay cloned with plasmid #1 (donor) and #4 (destination) | Ampicillin                     |
| 6  | 2-3 | FLAG-BRAF L597Q expression vector                     | GateWay cloned with plasmid #2 (donor) and #4 (destination) | Ampicillin                     |
| 7  | 2-3 | FLAG-BRAF N581S expression vector                     | GateWay cloned with plasmid #3 (donor) and #4 (destination) | Ampicillin                     |
| 8  | 2-3 | FLAG-BRAF WT expression vector                        | Site-directed mutagenesis of plasmid #7                     | Ampicillin                     |
| 9  | 2-3 | FLAG-BRAF A305G expression vector                     | Site-directed mutagenesis of plasmid #8                     | Ampicillin                     |
| 10 | 2-3 | FLAG-BRAF S323L expression vector                     | Site-directed mutagenesis of plasmid #8                     | Ampicillin                     |
| 11 | 2-3 | FLAG-BRAF I326V expression vector                     | Site-directed mutagenesis of plasmid #8                     | Ampicillin                     |

**Table S7.** Conditions of Liquid-chromatography (LC)

| Experiment Type                                    | Parameter | Condition                                                                                                                                                |
|----------------------------------------------------|-----------|----------------------------------------------------------------------------------------------------------------------------------------------------------|
| SPS-RTS-MS3 of B-Raf overexpressed WT and Variants | Column    | 100 $\mu$ M ID fused silica capillary packed in-house with bulk C18 reversed phase resin (particle size, 1.9 $\mu$ m; pore size, 100 Å; Dr. Maisch GmbH) |

|                                   |                        |                                                                                              |
|-----------------------------------|------------------------|----------------------------------------------------------------------------------------------|
|                                   |                        |                                                                                              |
|                                   | Mobile phase           | Buffer A: water and 0.1% formic acid<br>Buffer B: 80% acetonitrile and 0.1% formic acid      |
|                                   | Gradient and flow rate | 0 – 110 min, 12 – 33% B, 250 nL/min                                                          |
|                                   |                        | 110 – 120 min, 33% – 50% B, 250 nL/min                                                       |
|                                   |                        | 120 – 130 min, 50% – 95% B, 250 nL/min                                                       |
|                                   |                        | 130 – 140 min, 95% B, 250 nL/min                                                             |
|                                   | Run time               | 140 minutes                                                                                  |
|                                   | Injection volume       | 1 uL                                                                                         |
| GoDig Spectral Library Generation | Column                 | EASY-Spray PepMap Neo UHPLC 75 um x 500 mm column (ThermoScientific Cat. No. ES75500PN)      |
|                                   | Mobile phase           | Buffer A: water and 0.1% formic acid<br>Buffer B: 80% acetonitrile with and 0.1% formic acid |
|                                   | Gradient and flow rate | 0 – 110 min, 12% – 33% B, 250 nL/min                                                         |
|                                   |                        | 110 – 120 min, 33% - 50% B, 250 nL/min                                                       |
|                                   |                        | 120 – 130 min, 50% - 95% B, 250 nL/min                                                       |
|                                   |                        | 130 – 140 min, 95% B, 250 nL/min                                                             |
|                                   | Run time               | 140 minutes                                                                                  |

## Experimental Procedures

**Cell lines, culture conditions.** Cell culture reagents including Dulbecco's phosphate-buffered saline (DPBS), Dulbecco's Modified Eagle Medium (DMEM) media, and penicillin/streptomycin



(Pen/Strep) were purchased from Fisher Scientific. Fetal Bovine Serum (FBS) was purchased from Avantor Seradigm (lot #214B17). All cell lines (except for the gifted BRAF COSMIC library cell line) were obtained from ATCC and all were maintained at a low passage number (<20 passages). HA-BRAF COSMIC library cells were provided by Dr. Dustin Maly and used with author's permission<sup>5</sup>. HEK293T (ATCC: CRL-3216) cells were cultured in DMEM supplemented with 10 % FBS and 1 % antibiotics (Penn/Strep, 100 U/mL). HA-BRAF COSMIC library cells were cultured in DMEM supplemented with 10% FBS, 1% antibiotics (Penn/Strep, 100 U/mL), 2µg/mL doxycycline (Sigma Aldrich Fine Chemicals Biosciences; D98911G), and 100 nM SCH772984 (ERK inhibitor; MedChemExpress HY-50846). All media was filtered (0.22 µm) prior to use. Cells were maintained in a humidified incubator at 37 °C with 5 % CO<sub>2</sub>. Cell lines were tested monthly for mycoplasma using the Mycoplasma Detection Kit (InvivoGen).

**Cloning of plasmids.** List of plasmids with detailed information used in this study can be found in **Supplementary Table S6**. pDONR223 BRAF vectors containing sequences of variants V600E, L597Q, and N581S were a part of the Target Accelerator Pan-Cancer Mutant Collection from the labs of Jesse Boehm, William Hahn, Matthew Meyerson, and David Root (Addgene kit #1000000103) and subcloned using GateWay cloning into N-terminal FLAG destination vector, which was a kind gift from T Wucherpennig. For GateWay Cloning, 300 ng of both the donor and destination vector were combined in a microcentrifuge tube, following addition of TE buffer up to 8 µL. Then, 2 µL of Gateway LR Clonase II enzyme mix (Invitrogen; 11791020) was added to the mixture, vortexed briefly, and allowed to incubate for 2 hours at room temperature before being transferred to 4°C for 12-16 hours. The enzymes were then inactivated at 65°C for 15 minutes, and the reaction mixture transformed into competent TOP10 cells, colonies grown up in LB media containing 100 µg/mL ampicillin, and successful cloning confirmed by sequencing. The BRAF WT plasmid was generated by PCR-based site-directed mutagenesis, as previously described{ref}, with the N-terminal FLAG BRAF N581S as the template plasmid. The BRAF A305G, S323L, and

I326V plasmids were generated by PCR-based site-directed mutagenesis, with the BRAF WT plasmid as the template. All primers used to generate mutant constructs can be found in **Supplementary Table S5**. Mutant plasmids were then transformed into competent TOP10 cells, colonies selected and grown in SOC media at 37°C for >2 hours, and successful mutagenesis confirmed by sequencing.

**Transfections.** Cells were transfected at 70-80% confluency. For a 10-cm plate, plasmid (3 µg), serum-free DMEM (350 µL) and PEI MAX- Transfection Grade Linear Polyethylenimine Hydrochloride (MW 40,000) (Polysciences, Inc., 24765-1) (21 µL of 1 mg/mL, pH 7.4) were mixed and incubated for 15 minutes at room temperature. The transfection cocktail was then added dropwise to the cells and incubated for 24 hours. Transfection reagents were scaled according to volume of media in plates of various sizes.

**Western Blots.** A portion of lysate was aliquoted out of each temperature point and 4X SDS loading dye (plus 10% 2-mercaptoethanol) (BioRad) was added so that the final concentration of loading dye was 1X. The samples were run on either 4-20% or 8-16% SDS-PAGE gels (BioRad) at 120V. Gels were transferred to nitrocellulose or activated PVDF membranes (BioRad) using a semi-dry transfer system (BioRad) and membranes blocked in 5% (w/v) milk in Tris-buffered saline (TBS) for 30 minutes by rocking at room temperature. Membranes were rocked with primary antibodies (**Supplementary Table 4**) in 5% (w/v) milk in TBS overnight (14-16 hours) at 4°C, then washed 3 times with TBS for 5 minutes. Membrane was imaged on a BioRad ChemiDoc Imaging System. FIJI ImageJ<sup>1</sup> was used to normalize and quantify band intensities.

**Thermal denaturation in cells for TPP or CETSA.** Cells were washed with cold PBS, lifted into cold PBS and harvested by centrifugation at 1800 x g for 3 minutes. Cell pellets were then washed twice with cold PBS. Cells were counted with a Thermo Fisher Countess II and cell counts

normalized within each comparative set. The cells were then resuspended in a total volume of PBS for 50  $\mu$ L/temperature. The cell suspension was then aliquoted into PCR tubes and placed into a pre-warmed thermocycler programmed to run desired thermal gradients. Cells were heated for 3 minutes then taken off the thermocycler to cool at ambient temperature for 3 minutes. Halt protease and phosphatase inhibitor cocktail (Thermo Fisher Scientific, Cat. No. 78440) was added to 1X final concentration to each temperature point. Cells were then lysed via 3 rounds of freeze thaw lysis with liquid nitrogen. After lysis, cellular debris and insoluble portions were clarified by either centrifugation at 21,100 x *g* for 45 minutes or ultracentrifugation at 100,000 x *g* for 20 minutes and the soluble fraction transferred to a fresh low-bind microcentrifuge tube.

**Immunoprecipitation of soluble TPP fractions.** Thermo Scientific, Pierce anti-HA magnetic beads (Cat. No. 88837) were washed 3x with ultrapure water and once with 0.3% CHAPS in PBS then resuspended with 0.3% CHAPS in PBS. Each of the soluble fractions post-ultracentrifugation in fresh low-bind microcentrifuge tubes were adjusted to a final concentration of 0.3% CHAPS. 10  $\mu$ L of anti-HA magnetic beads were added to each of the samples and allowed to rotate at 4°C overnight. The next day, the flowthrough was aspirated and the beads were washed with PBS + 0.3% CHAPS 2x and PBS 2x. Then, 40  $\mu$ L of 8M urea in PBS was added to the washed resin and incubated at 37°C for 20 minutes. Eluent was collected and taken through sample cleanup and preparation for mass spectrometry.

**Sample cleanup and preparation for Mass Spectrometry.** For samples not immunoprecipitated with anti-HA resin, about 10  $\mu$ g of material from each of the temperatures from the soluble TPP fractions was placed into fresh low-bind microfuge tubes and 8M urea in PBS was added to each of the tubes for a final concentration of 2M urea. For samples immunoprecipitated with anti-HA resin, the eluent was diluted with PBS to a final concentration of 2M urea. Each of the samples were then reduced with 10 mM dithiothreitol (DTT) for 20 minutes

at ambient temperature. Following reduction, the samples were alkylated with 20 mM iodoacetamide (IA) for 20 minutes at 37°C. The samples were then subjected to cleanup with either Sera-Mag SpeedBeads Carboxyl Magnetic Beads (GE Healthcare, 65152105050250, 50 µg/µl, total 1 mg and GE Healthcare, 45152105050250, 50 µg/µl, total 1 mg), or DNA cleanup magnetic beads obtained from SergiLabSupplies (Cat. No. 1040, Lot S22122801). For samples prepared with the Ser-Mag beads, 7.5 µL of each bead type was combined per sample for a total bead volume of 15 µL. For samples prepared with SergiLabSupplies DNA cleanup magnetic beads, 40 µL of beads were used per sample. Before use, each of these beads were washed 3x with ultrapure water then resuspended with ultrapure water. Magnetic beads were added to each sample and shaken at 1200 rpm on an Eppendorf thermoshaker for 5-10 minutes at ambient temperature. Then, 100% ethanol was added to a final v/v of >60% ethanol and the samples were shaken at 1200 rpm for 10 minutes at ambient temperature. Magnetic beads were separated and the liquid was aspirated. The beads were then washed with 80% ethanol three times. The beads were then resuspended in 50 µL of a digest solution (2M urea in PBS containing 4 mM CaCl<sub>2</sub> and 0.5 µL of 1 mg/mL trypsin for non-IP samples and 0.2 µL of 1 mg/mL trypsin (Worthington Biochem LS003740) for IP samples). Samples were incubated at 37°C with shaking (300 rpm) overnight. The next morning, 1200 µL of 100% acetonitrile (ACN) was added to each of the samples and left to shake at 1200 rpm for 10-15 minutes at ambient temperature. The samples were then placed on a magnetic rack to separate the beads and all the liquid was aspirated. The beads were then washed three times with 100% ACN. The beads were then resuspended with 40 µL of 2% DMSO in ultrapure water and incubated at 37°C with shaking at 1200 rpm for 30 minutes. The peptides were then collected in a fresh low-bind eppendorf tube, frozen, and dried down using a speedvac.

**TMT labeling of TPP peptides.** The dried down peptides were resuspended in 20 µL of 100 mM triethylammonium bicarbonate (TEAB; Thermo Scientific Cat. No. 90114) and 5 µL of 100% ACN was added to each tube. 1µL (IP samples) or 2 µL (non-IP) samples of TMTpro 16-plex

(ThermoFisher Scientific Cat. No. A44520) reagents at 20 µg/µL in ACN were added to each of the samples. Samples were lightly vortexed and incubated at ambient temperature for 1 hour. After incubation, samples were quenched with ~1.73 µL of 5% hydroxylamine (for a final concentration of ~0.3%) (ThermoFisher Scientific Cat. No. 90115), vortexed lightly, and incubated at ambient temperature for 15 minutes. Then, 1.5 µL of 100% formic acid (final concentration of 5%) was added to each sample, lightly vortexed, and incubated for 5 minutes at ambient temperature. Each of the samples from the TPP temperature range were combined into 1 sample which was then dried down (SpeedVac) and reconstituted in 100 µL of 1% trifluoroacetic acid (TFA) in ultrapure water. The sample was then cleaned up with Pierce C18 spin tips (Thermo Fisher Scientific, 87784) according to the manufacturer's protocol. The sample was then dried (SpeedVac), and either reconstituted in MS solvent (5% acetonitrile and 1% formic acid in ultrapure water) or fractionated by high pH fractionation (see section **High pH Reversed-Phase Peptide Fractionation**) then analyzed by liquid chromatography–tandem mass spectrometry (LC–MS/MS).

**High pH Reversed-Phase Peptide Fractionation.** Samples were resuspended in 300 µL of 0.1% TFA in ultrapure water. Peptides were fractionated using the Pierce High pH Reversed-Phase Peptide Fractionation Kit (ThermoFisher Scientific Cat. No. 84868) according to the manufacturer's protocol.

**Proteomics acquisition.** Samples were analyzed by LC-MS/MS(/MS) using a Thermo Fisher Scientific Orbitrap Eclipse Tribrid mass spectrometer. Peptides were fractionated online using an 18 cm long, 100 µM inner diameter (ID) fused silica capillary packed in-house with bulk C18 reversed phase resin (particle size, 1.9 µm; pore size, 100 Å; Dr. Maisch GmbH) or an EASY-Spray PepMap Neo UHPLC 75 µm x 500 mm column (ThermoScientific Cat. No. ES75500PN). The 140-min water/acetonitrile gradient was delivered using a Thermo Fisher Scientific EASY-

nLC 1200 system at different flow rates (buffer A: water and 0.1% formic acid and buffer B: 80% acetonitrile and 0.1% formic acid). Detailed gradient information can be found in **Supplementary Table S7**. Data were acquired using a data-dependent acquisition method consisting of a full MS1 scan (resolution = 120,000) with a 2 sec cycle time. For data acquired using SPS-RTS-MS3, MS1 scan range was set to 400-1600 m/z with a normalized AGC target of 25% and a maximum injection time (IT) of 50 ms. Precursor fit window was set to 1.2 m/z with a 50% fit threshold and dynamic exclusion for 60s. MS2 scans were performed in the orbitrap at a resolution of 15,000 and NCE of 28% HCD. Normalized AGC target was 200% with a max IT of 22 ms. Real-Time Search through the Thermo Method Editor software was enabled using a custom database consisting of the NCBI Taxonomy ID 9606 Human Sapiens protein database and our B-Raf variant sequences. Static modifications TMTpro16plex at sites Kn and carbamidomethyl at sites C were enabled. Variable modification M oxidation was enabled with maximum variable modifications per peptide set to 1 and maximum missed cleavages set to 1. Precursor Ion Exclusion was set to  $\pm 25$  ppm and Isobaric Tag Loss Exclusion was enabled for the TMTpro reagent. For MS3 scans, 10 SPS precursors were selected with an MS isolation window of 0.7 m/z, NCE was set to 50% HCD and scans were performed in the orbitrap at 50,000 resolution. Scan range was set to 100-500 m/z with a normalized AGC target of 200% and a max IT of 105 ms.

**Proteomics data analysis.** For samples run with untargeted proteomics (non-GoDig samples), all were searched using FragPipe v22.0<sup>1,6</sup>, MSFragger v4.1<sup>1,6,7</sup>, IonQuant v1.10.38<sup>8</sup>, and Philosopher v5.1.1<sup>9</sup> after converting all files to .mzml format using ProteomeWizard (MSConvert, peak picking 1-3)<sup>10</sup>. For the B-Raf overexpressed samples, the standard TMT-16plex-MS3 FragPipe workflow was used with a reverse concatenated UniProtKB Human database corresponding to the 18432 CCDS release with a release date of 01-01-2020 with 18,556 entries and 18,556 decoys. The variant tryptic sequence was included for the variant B-Raf protein of

interest. For the B-Raf cosmic cells, the data was first searched against our canonical UniProtKB Human database (corresponding to the 18432 CCDS release with a release date of 01-01-2020 with 18,556 entries and 18,556 decoys), decoys were printed and submzml files were generated through FragPipe. The submzml files were then searched in a second pass search against the custom B-Raf database containing 1437 variant sequences. For the MOLT-4 samples, the data was again searched in a two pass search, however the second pass search was done against a custom variant database which was generated as described in Desai et al.<sup>11</sup>.

**Melt Curve Fitting and Statistics.** All TPP FragPipe outputs were formatted into acceptable input for the R package “TPP”<sup>3</sup> and the melt curves fits were fitted and normalized using the “meltcurvefit” method under the analyzeTPP-TR function. NPARC was applied using the TPP-TR (“splinefit” method) and NPARC packages<sup>4</sup>. GraphPad Prism v10.1 was also used to fit melt curves from CETSA quantification or TPP-TR normalized values and p values were calculated from an extra-sum-of-squares F test was performed for the IC50 values ( $T_m$ ).

## References

1. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676–682 (2012).
2. Tsai, J. *et al.* Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proc Natl Acad Sci U S A* **105**, 3041–3046 (2008).
3. Franken, H. *et al.* Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nat Protoc* **10**, 1567–1593 (2015).
4. Childs, D. *et al.* Nonparametric Analysis of Thermal Proteome Profiles Reveals Novel Drug-binding Proteins. *Mol Cell Proteomics* **18**, 2506–2515 (2019).

5. Simon, J. J., Fowler, D. M. & Maly, D. J. Multiplexed profiling of intracellular protein abundance, activity, interactions and druggability with LABEL-seq. *Nat Methods* **21**, 2094–2106 (2024).
6. Chang, H.-Y. *et al.* Analysis of isobaric quantitative proteomic data using TMT-Integrator and FragPipe computational platform. *Nat Commun* (2026) doi:10.1038/s41467-026-70118-7.
7. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D. & Nesvizhskii, A. I. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat Methods* **14**, 513–520 (2017).
8. Yu, F., Haynes, S. E. & Nesvizhskii, A. I. IonQuant Enables Accurate and Sensitive Label-Free Quantification With FDR-Controlled Match-Between-Runs. *Mol Cell Proteomics* **20**, 100077 (2021).
9. da Veiga Leprevost, F. *et al.* Philosopher: a versatile toolkit for shotgun proteomics data analysis. *Nat Methods* **17**, 869–870 (2020).
10. Adusumilli, R. & Mallick, P. Data Conversion with ProteoWizard msConvert. *Methods Mol Biol* **1550**, 339–368 (2017).
11. Desai, H. *et al.* Chemoproteogenomic stratification of the missense variant cysteinome. *Nat Commun* **15**, 9284 (2024).



## Chapter 5: Conclusion

This work explored the use of mass spectrometry-based proteomic approaches for identifying and assessing chemical and genetic amino acid modifications. One main underlying theme between these chapters was the improvement of proteomic workflows to increase the throughput of detecting modified residues while also providing analyses for characterizing the effects of the modifications across the proteome.

In Chapter 2, we developed a proteomics sample preparation workflow to transition covalent cysteine-reactive small molecule screening from a smaller scale tube-based workflow to a 96-deep-well plate-based protocol. Applying our CySP3-96 plate-based workflow, we screened 16 cysteine-reactive compounds in duplicate while also including comparisons of select compounds across a 5-point concentration range. In aggregate, we identified 2633 ligandable cysteines and 21 ligandable cysteines not previously captured within CysDB. Within the set of 16 cysteine-reactive compounds, 2 were atropisomer compounds each screened at 5 different concentrations which showed stereoselective cysteine ligandability across 67 targets. We developed CySP3-96 to be a widely compatible format for chemoproteomic studies using different custom capture reagents. To make this workflow more accessible and lower cost, we found that generic DNA cleanup magnetic beads could be substituted for Cytiva magnetic SP3 beads with no effect on overall coverage or biases for protein capture on the beads. These beads also made bead handling within a 96-deep-well plate easier due to smaller bead volumes. CySP3-96 serves as a user-friendly, low cost, and reproducible screening platform for a wide range of chemoproteomic applications.

In Chapter 3, we explored the consequences of cysteine-reactive covalent compounds on overall cellular proteostasis processes. We discovered that treatment with cysteine-reactive compound JC19 results in the depletion of the SARS-CoV-2 nsp14 protein in a cysteine-dependent manner. Through biochemical and proteomic analyses, we find that direct labeling of nsp14 by JC19 and ubiquitylation of nsp14 is required for its depletion. While we were unable to determine which E3

ubiquitin ligase was implicated in the proteasomal degradation of nsp14, we observed that JC19 and other cysteine-reactive compounds appeared to activate the proteasome. We also observed the modification of cysteines on PDI proteins which we found to also be implicated in nsp14 depletion. The knockdown or inactivation through labeling of PDI3 and PDI6 also leads to nsp14 depletion, with PDI overexpression having a protective effect on nsp14 depletion. We also observed, more broadly, that JC19, along with multiple other cysteine-reactive compounds, induced stress granules and aggresome-like structures throughout the cell. Within this study we found that nsp14 depletion, driven by cysteine labeling, ubiquitylation, and PDI involvement, proceeded through aggregation of nsp14 within p62-containing aggresome-like structures followed by degradation of the nsp14 protein. The results of this work laid the framework for establishing guidelines for the delineation of widespread impacts of cysteine-reactive compounds on cellular proteostasis from more specific impacts of cysteine-reactive compounds of specific protein depletion/degradation. We anticipate that these findings will better guide future studies with covalent degrader molecules.

In Chapter 4 of this work, we set out to explore the use of thermal proteome profiling (TPP) towards characterizing SAAVs throughout the entire proteome. We leveraged a small model system of B-Raf variants to validate thermal stability changes resulting from SAAV proteoforms. Within this model system, we validated the use of protein-level and peptide-level quantification for TPP in the context of proteogenomic workflows where variants are identified only through variant-containing peptides. When applying TPP to a system of pooled B-Raf variants, we identified challenges within a variant TPP study and explored data analysis approaches to improve the robustness of our TPP quantification for variants. Using a two-stage FDR controlled proteomic search against a custom variant data for the pooled B-Raf variant system, we identified a total of 47 B-Raf SAAVs and quantified 19 of those variants by TPP. We find 2 stabilizing mutations and 2 destabilizing mutations. We also extended our variant TPP approach to profiling endogenous variants within Molt-4 cells. We also used a two-stage FDR controlled proteomic

search for this system, however, we searched against a custom database containing variants identified within the cell line through WGS and transcriptomic analyses. Variant TPP of the Molt-4 cells captured 77 unique variants and quantified 56 of those variants. From our peptide-level and protein-level comparisons, we identified 4 stabilizing and 4 destabilizing variants. We highlighted two of these SAAVs, both found within the COSMIC database, as potentially interesting follow up case studies. This work provides a starting point for further improvement and implementation of a variant TPP pipeline to characterize functional consequences of SAAVs.