

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY

Chemoproteomic target-class drug discovery against the deubiquitinating enzymes

Citation

Chan, Wai Cheung. 2021. Chemoproteomic target-class drug discovery against the deubiquitinating enzymes. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.

Link

https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37371107

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFlOWI2M2EwYTg

Accessibility

https://accessibility.huit.harvard.edu/digital-accessibility-policy

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story

HARVARD UNIVERSITY
Graduate School of Arts and Sciences



DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of Chemical Biology

have examined a dissertation entitled

Chemoproteomic target-class drug discovery against the deubiquitinating enzymes

presented by Wai Cheung Chan

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature _____

Typed name: Prof. Philip Cole

Signature _____

Typed name: Prof. Hanno Steen

Signature _____

Typed name: Prof. Eric Fischer

Signature _____

Typed name: Prof.

Signature _____

Typed name: Prof.

Date: Nov. 16th, 2021.

Chemoproteomic target-class drug discovery against the deubiquitinating enzymes

A dissertation presented

by

Wai Cheung Chan

for the degree of

Doctor of Philosophy

in the subject of

Chemical Biology

Harvard University

Cambridge, Massachusetts

November 2021

@ 2021 Wai Cheung Chan
All rights reserved

Dissertation Advisor: Jarrod Marto, Sara Burhlage Wai Cheung Chan

Chemoproteomic target-class drug discovery against the deubiquitinating enzymes

Abstract

Modern small molecule drug discovery is dominated, and subsequently limited by a paradigm of reversible non-covalent inhibitors. The deubiquitinating enzymes (DUBs) is a class of ~100 proteases that regulate normal and pathophysiological processes through cleaving ubiquitin marks from protein substrates and is underserved by this existing paradigm. Despite intense interest, a lack of selective inhibitors and design principles impede biological investigation. Here, we develop a DUB target-class inhibitor discovery platform consisting of a tailored covalent small molecule library, mass spectrometry chemoproteomics, and a suite of orthogonal validation assays. We first demonstrate small-scale proof-of-concept and subsequent improvements (chapter 2), then apply the platform for a screening campaign to identify and validate novel DUB covalent ligands (chapter 3). This has led to a revision of statistical approaches for handling chemoproteomic screening data (chapter 4). Our systematic interrogation provides a methodological roadmap for future target-class campaigns, guiding principles for DUB-targeting medicinal chemistry, attractive chemical starting points for diverse DUBs and a first-in-class probe for studying the understudied DUB VCPIP1.

Table of Contents

Title page i

Copyright ii

Abstract iii

Table of Contents iv

Acknowledgements viii

Glossary x

Body of text ..... 1 - 131

1. Background 1

1.1. Challenging paradigms in small molecule drug discovery 1

1.1.1 Target-centric drug discovery 2

1.1.2 Target-class drug discovery 3

1.1.3 Reversible Inhibitor modality 4

1.1.4 Alternative modalities 5

1.2. The deubiquitinating enzymes 8

1.2.1 Introduction to ubiquitin signaling 8

1.2.2 Introduction to the deubiquitinating enzymes (DUBs) 9

1.2.3 Therapeutic potential of deubiquitinating enzymes (DUBs) 12

1.2.4 The current landscape of DUB inhibitor development 16

1.2.5 Bottleneck in DUB inhibitor development 17

1.2.6 Towards DUB target-class inhibitor discovery 18

1.3. Chemoproteomic methods for covalent inhibitors 20

1.3.1 Activity-based probes and Bachovchin et al. 21

1.3.2 Residue-based probes 23

1.3.3 Inhibitor-based probes 25

1.3.4 Statistics for chemoproteomic studies 26

1.4. Towards chemoproteomic DUB target-class covalent drug discovery 30

1.4.1 Assays used in this thesis 30

1.4.2 Towards chemoproteomic DUB target-class covalent drug discovery 33

2. Foundations for a library versus library screen 35

2.1. Design and synthesis of a DUB-focused covalent small molecule library 35

2.2. Development of an ABPP-based primary screening assay 40

2.2.1 Previous work 40

2.2.2 Foundations of DUB ABPP 41

2.2.3 Validating competition 44

2.3. Pilot Screen 46

2.4. Further method development 48

2.4.1 Optimizing pulldown conditions 49

2.5. Sequential pulldown for interrogating quantitative DUB enrichment 53

2.5.1 Streptavidin Capacity 55

2.5.2 Scaling down the assay 56

2.5.3 Incubation times for streptavidin pulldown 57

2.5.4 Titrating down TMT 58

2.6. The final screening assay 58

2.7. Conclusions 59

3. Library versus library screen 60

3.1. Primary Screening 60

3.2. Hit Validation/Characterization 64

3.2.1 VCPIP1 Hits 71

3.2.2 USP48 Hits 73

3.2.3 UCHL1 Hits 73

3.2.4 UCHL3 Hits 74

3.2.5 USP10 Hits 75

3.2.6 USP3, USP9X, USP22 76

3.3. Target/target-class optimization of VCPIP1 hit 76

3.3.1 A potent and selective first-in-class VCPIP1 inhibitor 77

3.3.2 Additional targets of the scaffold 80

3.4. Discussion 81

4. Statistical analysis of chemoproteomic screening data 85

4.1. Observed challenges in chemoproteomic data analysis 85

4.2. Statistical modelling of baseline variance 89

4.2.1 INPROBE procedure 90

4.2.2 Preliminary analysis of baseline variance 92

4.3. Impact on hit identification 93

4.4. Discussion 96

5. Future directions and conclusions 100

5.1. Future directions 100

5.1.1 Second-generation active-site focused library 100

5.1.2 Beyond the active site: noncovalent library versus library screening 102

5.1.3 Enhancing the DUB ABPP chemoproteomic primary screening assay 105

5.2. Overall conclusions 106

5.2.1 Contributions to DUB inhibitor discovery 107

5.2.2 MS chemoproteomic primary screening 109

5.3. Towards new paradigms of therapeutic discovery 111

6. Methods and materials 112

6.1. Methods and regents 112

6.1.1 Synthetic procedures for library compounds 112

6.1.2 Assay methods 217

7. References 225

Acknowledgements

First and foremost, I would like to thank my research advisers Sara Buhrlage and Jarrod Marto for their constant guidance in academic, profession, and personal matters in the past five years. Your advice on all matters big and small has motivated, encouraged, comforted, and guided me through the world of academic and industrial research science. Thank you for listening to what I have to say, it is your faith in my own self-direction that enabled chapter 4 in this dissertation.

I would like to thank staff scientists Scott Ficarro and Guillaume Adelmante for their teaching and unwavering support across all matters since I have joined the Marto Lab. Thank you for teaching me everything I know about mass spectrometry proteomics, affinity purification, and sample preparation, in addition to the interpersonal dynamics of academic science.

I would like to thank all members across the Buhrlage and Marto labs for their intellectual, scientific contributions to the work described in this dissertation. I have received tremendous help, ideas, and encouragements particularly from Robert Magin, Xiaoxi Liu, Nathan Schauer, Christopher Browne, Brad Palanski, Shabnam Sharifzadeh, He (Eric) Zhu, Isidoro Tavares, Joseph Card, and William M. Alexander through the years.

I would like to thank Phil Cole, Hanno Steen, and Wade Harper of my Dissertation Advisory Committee for their support in scientific and professional matters in the past few years. I would also like to thank Catherine Dubreuil of the Therapeutics Graduate Program for the many chats through the years, which helped me navigate through the course of my PhD work.

I would like to thank my roommates and friends for moral support. We have celebrated milestones large and small across the course of the past five years, giving me the energy to keep going.

Finally, I would like to thank sources of funding which supported this work: research described in this dissertation was mostly supported by the Mark Foundation for Cancer Research and the NIH. I would also like to acknowledge my PhD funding, in chronological order: The Forris Jewett Moore Fellowship in Chemistry, the Fujifilm Fellowship in Therapeutic Discovery, and the Chleck Family Fellowship. The research described in this document would not have happened without their general support.

Glossary

ABP: Activity-based probe

ABPP: Activity-based protein profiling

CE-MS: Capillary electrophoresis-mass spectrometry

CR: Competition ratio

DUBs: Deubiquitinating enzymes

FDA: United States Food and Drug Administration

IBP: Inhibitor-based probe

IC_{50}: Half-maximal inhibitory concentration

LC-MS: Liquid chromatography-mass spectrometry

INPROBE: Individualized protein baseline extraction

MW: Molecular weight

PSM: Peptide spectrum match

PTM: Post-translational modification

RBP: Residue-based probe

SWATH-MS: Sequential Window Acquisition of All Theoretical Mass Spectra

TMT: Tandem mass tags

Ub: Ubiquitin

Ubls: Ubiquitin-like proteins

To my friends who could not start, or did not complete their PhD studies.

1. Background

Author's note: Portions of this section have been adapted from a review manuscript published in Chemical Society Reviews.

Small molecule drugs bind to cellular proteins, genetic material, and other cellular machinery to modulate their function, thereby achieving therapeutic benefit. Among those listed, proteins are the most frequently targeted by small molecule drugs, with the set of proteins amenable to small molecule-mediated modulation termed the "druggable proteome". The work described in this thesis are fundamentally motivated by a need to expand the druggable proteome, and thereby expand the scope of diseases which can be treated with small molecule therapeutics.

In this manuscript, I discuss my contributions to expanding the druggable protein space within the deubiquitinating enzymes, as well as the development of inhibitor discovery approaches for doing so. At the highest level, my work challenges established paradigms in small molecule therapeutics development by both approach and modality.

1.1. Challenging paradigms in small molecule drug discovery

Two paradigms are dominant in modern small molecule drug discovery: (1) a target-centric discovery approach, and (2) the reversible inhibitor modality. While these paradigms have led to considerable success and advances in modern medicine, they are also inherently limiting. Targets which are compatible with such drug discovery campaigns (e.g. kinases, proteases) are termed "druggable", and incompatible targets (e.g. phosphatases, transcription factors) are

termed “undruggable”. This is detrimental for drug discovery and human health, as a limitation in addressable drug targets translates into a limitation in which diseases can be treated.

In order to develop novel therapeutics against increasing challenging diseases, there is a need to expand the realm of actionable disease targets. To do so would require breaking existing paradigms of drug discovery and developing new approaches for drug discovery as well as new drug modalities. In the following paragraphs, I will discuss each current paradigm and their associated limitations, then alternatives for each approach.

1.1.1 Target-centric drug discovery

Modern drug discovery across academia and industry is primarily carried out in a target-centric approach.[1] A molecular target (most often a protein), whose modulation is likely to induce a therapeutic benefit, is selected based on disease-specific knowledge. Target-specific assays are developed to enable high-throughput screening, in which chemical matter possessing the desired activity is identified. The active chemical matter, “hits”, is then further developed in “hit-to-lead/probe” efforts to improve potency and selectivity against the target. The resulting probe is then used to pharmacologically validate that modulating the target protein does indeed lead to therapeutic benefit in disease models. With successful validation, the lead is further developed into a drug candidate.

Despite its power, a target-centric approach holds numerous shortcomings due to its discrete, target-specific nature. First, each target-centric drug discovery campaign requires separate investment into assay development, reagent generation, and biological investigation, making them very labor- and resource- intensive. A target-centric approach also suffers from its

inherent specificity: in considering only one target from the outset, a target-centric approach ignores the possibility of modulating related enzymes. This leads to three problems: (1) selectivity of lead molecules is usually only evaluated far down the discovery process, after significant investment; (2) opportunities that a lead scaffold can be leveraged for other targets in the same class are often neglected; (3) there is no risk management for when modulation of the supposed target ultimately proves to be therapeutically untenable.

Thus, while a target-centric approach has been immensely successful with well-known enzyme classes, alternative discovery strategies are needed for less well-studied enzyme classes requiring more attention on selectivity.

1.1.2 Target-class drug discovery

In recent years, there has been increasing advocacy for target-class drug discovery as a complementary for target-centric approaches.[1] In contrast to the one-at-a-time philosophy of target-centric approaches, a target-class approach attempts to leverage a common base of scientific knowledge and physical resources against all members of a selected target class. The target class is most often defined as a structurally or mechanistically related set of enzymes with therapeutic relevance. Focused chemical libraries with bias for chemical motifs against the target class enable more cost-efficient, information-rich screening. Class-wide assay platforms inform in-family selectivity and reduces cost for iterative optimization. Since multiple targets are pursued at the same time, risk is greatly diminished since back-ups exist in the forms of other target class members should one fail.

Despite the many potential benefits, target-class approaches can be practically challenging. Choice of the target class stands at the forefront: for a target-class drug discovery approach to be worthwhile, the target class has to be of sufficient therapeutic interest but underexplored so insights gained can lead to maximum payoff. The availability of knowledge surrounding chemical design principles for inhibiting members of the target class can also be challenging. If too little is known, focused compound libraries cannot be designed. However, for sophisticated target classes with well-known inhibitor design principles, there is little motivation to start such a campaign. In addition, while the notion of a class-wide assay is alluring, the practicalities of devising and setting up such assays can be prohibitive. Hence, target-class approaches are best taken for emerging drug target classes with proof-of-concepts inhibitors and some supporting technology. In section 1.2, I will nominate the deubiquitinating enzymes (DUBs) family as a promising target class warranting such an approach.

1.1.3 Reversible Inhibitor modality

While target-centricism is one paradigm in drug discovery approach, another paradigm exists: small molecule leads/drugs are conventionally of the reversible inhibitor modality. Reversible inhibitors make non-covalent interactions with the target protein, blocking its function through competition for active site against native substrate, or noncompetitively by inducing catalytically incompetent conformations. The dominance of reversible inhibitors is one of intention: therapeutic programs mostly set off to develop reversible inhibitors, leading to most discovered small molecule drugs being noncovalent.

Physiologically, reversible inhibitors face a number of challenges.[2] Since binding is reversible, the local drug concentration at the targeted organ must remain above a certain level. This

places stringent constraints on the pharmacokinetic properties of each candidate noncovalent drug. There is also considerable difficulty in mapping reversible inhibitor off-target binding at both the site- and proteome-levels. Short of a tell-tale phenotype, we often have little insight on what off-targets each reversible inhibitor binds to, again generating risk for downstream failure. For structural information of the inhibitor binding site, X-ray crystallography and cryo-EM can be used; both techniques require significant optimization for each individual target, and neither guarantees information.

Thus, while reversible inhibitors do hold multiple advantages and have historically been immensely successful, it is worth exploring other small molecule therapeutic modalities to further expand the realm of druggable targets.

1.1.4 Alternative modalities

Outside of the conventional reversible inhibitor modality, there has been a shift in attention in recent years to small molecule drugs which achieve efficacy by other mechanisms. Below I present two alternative strategies: (1) covalent inhibition and (2) targeted protein degradation.

Covalent inhibitors achieve exquisite potency and durable target engagement through a combination of noncovalent interactions and covalent bond formation with the target protein.[2] As a function of this sustained target engagement, covalent drugs are less prone to effects of drug washout due to metabolism/excretion, and might require less frequent/smaller quantity dosing. Despite the portrayal of covalent inhibitors as an "alternative modality" in this section, they are cornerstones of modern medicine. Aspirin and penicillin, among the earliest drugs known, both act through a covalent mechanism.[2]

Despite the numerous advantages, the pharmaceutical industry has been reluctant to engage in the intentional discovery of covalent drugs, due to worries about idiosyncratic toxicity and hyperreactivity.[3–5] Thus, most currently approved covalent drugs had their covalent mechanism of action discovered serendipitously post-approval.

Since the early 2010s, there has been a shift in this view: FDA approval of multiple successful covalent drugs (ibrutinib, afatinib, telaprevir among others) has led to a resurgence of interest in this class of therapeutics.[2] For the first time in decades, the field is engaging in rational design and intentional discovery of covalent mechanism-of-action drugs. This is also supported by the emergence of chemoproteomic technologies to identify off-target binding partners and mitigate hyperreactivity liabilities, which will be discussed in a later section.

Targeted protein degradation is another emerging therapeutic strategy. Instead of simply inhibiting enzymatic activity, the goal is to induce the destruction of the target protein through repurposing parts of the cellular machinery, such as the proteasome or the lysosome. This is particularly important when noncatalytic functions of an enzyme contribute to disease.

For example, bifunctional degrader molecules engage an E3 ligase on one end and the protein target on the other, leading to hyperubiquitination and subsequent proteasomal degradation of the protein target.[6] Therapeutic application was first demonstrated in Winter et al. where the BRD4-targeting small molecule JQ1 is linked to a phthalimide moiety to recruit the E3 ligase cereblon for polyubiquitination and degradation of BRD4, leading delayed progression in a leukemia mouse model.[7] In the five years since this proof-of-concept, targeted protein degradation has developed into a vibrant subfield with active participation from academic labs and biotech/pharamaceutical companies large and small: more than 15 molecules of this

modality has entered the clinic as of Fall 2021.[8] Indeed, targeted protein degradation is well-poised to improve access to hitherto intractable targets including STAT3, IKZF1/3, and the androgen receptor.[8]

The work described in this thesis is centered around developing covalent inhibitors targeting the deubiquitinating enzymes to achieve targeted degradation of pathogenic proteins. This will be explained in the following paragraphs.

1.2. The deubiquitinating enzymes

In this subsection, I will introduce the family of deubiquitinating enzymes (DUBs) as central regulators of both normal and pathophysiological biological processes and an attractive frontier for targeted protein degradation. I start by providing background on ubiquitin signaling, which DUBs are a part of. Next, I discuss the members and function of the DUB enzyme family and establish their therapeutic relevance. Afterwards, I present a brief history of DUB inhibitor discovery and discuss the current state of the field. I conclude by making a case for why target-class covalent inhibitor discovery would be fruitful for the DUBs.

1.2.1 Introduction to ubiquitin signaling

Ubiquitination is a reversible posttranslational modification based on the covalent attachment of ubiquitin, a 76-amino acid polypeptide.[9,10] The carboxyl-terminus of ubiquitin is conjugated to lysine residues by a triad of activating (E1), conjugating (E2), and ligase (E3) enzymes, and removal is catalyzed by the deubiquitinating enzymes (DUBs).[11] Ubiquitin is attached onto the ε-amino group of lysine residues on the substrates: the addition of a single ubiquitin modification is known as monoubiquitination. Ubiquitin groups can be further appended onto any of the seven lysine residues (K6, K11, K27, K29, K33, K48, K63) or the N-terminus on ubiquitin to generate polyubiquitin chains of various topology.[9]

In addition to ubiquitin, there are also an array of ubiquitin-like proteins (Ubls) which are appended onto lysine residues in a similar fashion as ubiquitin.[9] Examples of Ubls include small ubiquitin-like modifier (SUMO), neuronal precursor cell-expressed developmentally

downregulated protein 8 (NEDD8), and interferon stimulated gene 15 (ISG15). Ubls can form chains of their own or hybrid chains containing multiple kinds of Ubls and ubiquitin.

As a fundamental ubiquitous PTM, ubiquitination lies at the center of both normal and pathophysiological biological processes.[9] Most prominently, ubiquitination is involved in protein homeostasis: polyubiquitinated substrates are degraded by either the ubiquitin-proteasome system or the autophagy-lysosome systems.[9] K48 polyubiquitin chains are associated with proteasomal degradation, and K63 polyubiquitination is linked to lysosomal degradation.[9] Ubiquitination also regulates protein activity, degradation, localization, and protein-protein interactions, influencing transcription, DNA repair, cell cycle progression, and stress signaling.[9] Disruption of ubiquitination has been linked to the onset and progression of various oncology, metabolic, autoimmune, and neurodegenerative indications.[11]

1.2.2 Introduction to the deubiquitinating enzymes (DUBs)

The deubiquitinating enzymes (DUBs) are a family of proteases which catalytically cleave ubiquitin marks from substrate proteins, thereby regulating a large fraction of the proteome.[11,12] There are approximately 100 DUBs, divided into six families. Of those, five consist of cysteine proteases: the USP, UCH, OTU, MJD, and MINDY family.[11] [Figure 1a] The sixth JAMMs family consists of zinc-dependent metalloproteases. Recently, novel DUBs possessing little homology to the known families have also been identified, raising the possibility that there might be as-yet unidentified DUBs.[13–15]

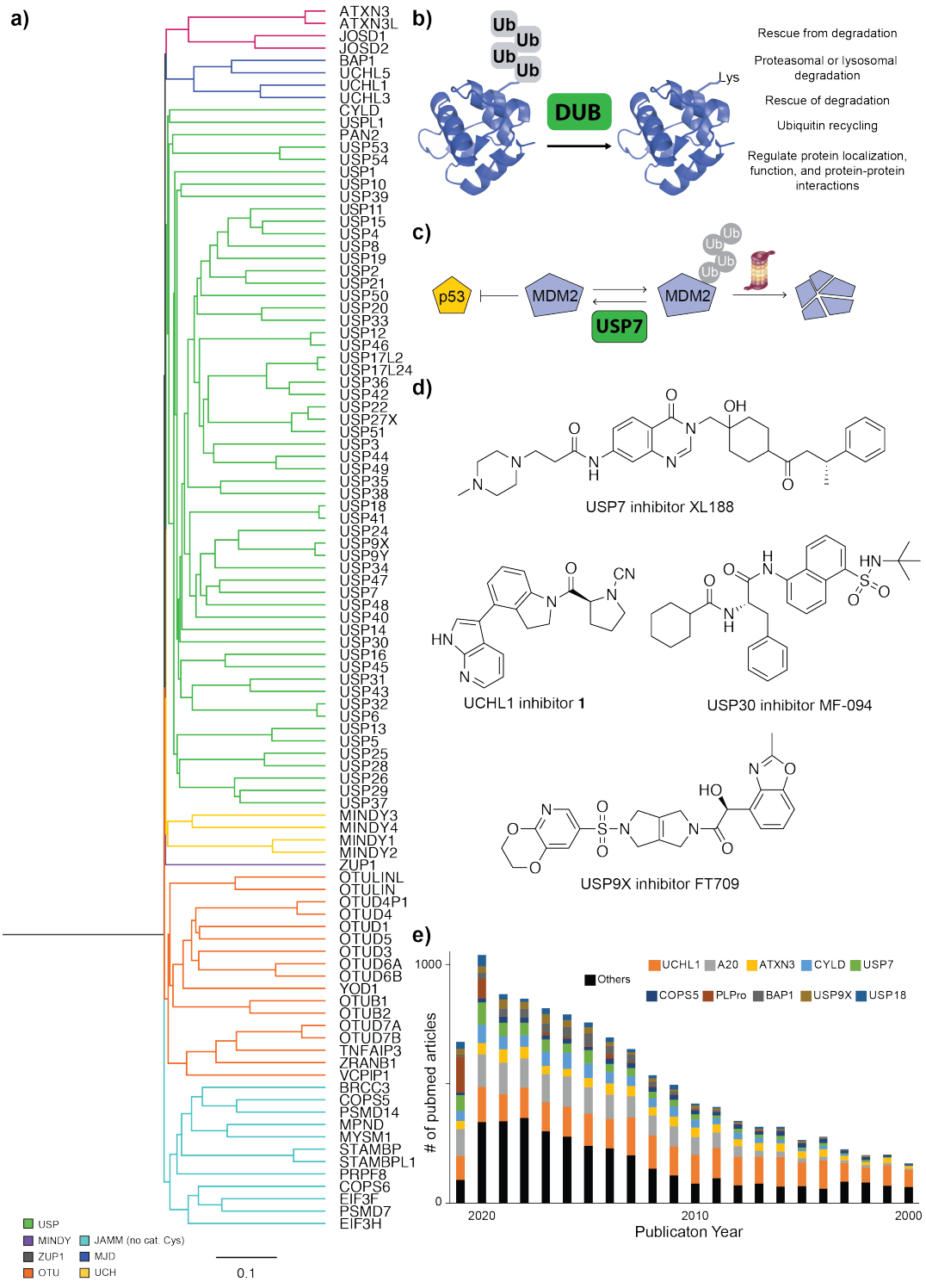


Figure 1. Introduction to the DUBs. a) Phylogenetic tree of the DUB family. b) DUBs regulate most fundamental cellular processes by cleaving ubiquitin groups from substrate. c) USP7 modulates p53 signaling by regulating MDM2 ubiquitination and degradation. d) Chemical

(continued from previous page) structures of validated DUB inhibitors. e) DUBs have been receiving increasing research interest in the past twenty years, but attention is focused on a cluster of family members.

The work described in this thesis focuses on the cysteine-protease DUBs. Most cysteine protease DUBs possess a catalytic triad consisting of a cysteine, a histidine, and an aspartate residue.[11] The aspartate residue polarizes the histidine, which lowers the pK_a of the hydrogen atom on the cysteine thiol. During catalysis, the deprotonated cysteine thiolate ion attacks the carbonyl of the scissile isopeptide bond, forming an oxyanion intermediate of the DUB and the ubiquitinated substrate. The intermediate collapses to release the substrate protein, leaving the ubiquitin C-terminus linked to the DUB catalytic cysteine via a thioester bond. Next, a water molecule attacks the carbonyl carbon, releasing ubiquitin and free DUB enzyme.

As the negative regulators of ubiquitin signaling, DUBs regulated all processes that ubiquitination has been linked to. [Figure 1b] Therapeutically, what's most interesting is DUBs' roles in protein homeostasis: through cleaving ubiquitin chains from substrates targeted to the proteasome, DUBs rescue substrates from degradation. This gives an exciting potential to inhibit DUBs to achieve targeted protein degradation of pathogenic DUB substrates. By inhibiting the cognate DUB for a protein-of-interest, we can induce hyper-ubiquitination of the POI, leading to proteasomal degradation.

Other specific functions of DUBs include the maintenance of the cellular free ubiquitin pool, by cleaving ubiquitin off proteasome substrates or processing of ubiquitin precursor chains. Outside of humans, many pathogens have evolved deubiquitinating enzymes to hijack the ubiquitin-

proteasome system for their benefit.[16] For example, the SARS-CoV-2 viral protease PLPro has been shown to cleave ubiquitin and ISG15 chains to regulate the host innate immune response.[17]

1.2.3 Therapeutic potential of deubiquitinating enzymes (DUBs)

In recent years, there has been significant interest in DUBs as disease targets.[18,19] Interest is two-pronged: (1) for targeted protein degradation of known substrates, and (2) for DUBs as therapeutic targets of their own right.

First, inhibition of DUBs brings an attractive potential of targeted protein degradation of pathogenic protein substrates. While particular targets might be deemed conventionally undruggable due to incompatibility with noncovalent small molecule inhibition (e.g. lack of targetable pockets, difficulty in drugging protein-protein interactions), inhibiting its DUB can provide an alternative way for pharmacological access.[18] This is further supported by preliminary evidence that some DUBs might have specificity for pathogenic mutants over wild-type protein.[20,21] Thus, inhibitor programs have been started from a protein of interest, identifying its cognate DUB, then attempting to discovery inhibitors targeting this DUB of interest.

Secondly, many DUBs have been identified as therapeutic targets of their own right; there is less mechanistic detail in these cases, but genetic knockouts and other animal models have led to the identification of these DUBs are disease targets. DUB inhibitors for these pathogenic DUBs might have direct therapeutic potential.[22–24] Below, we will discuss one examples of each to illustrate potential applications of DUB inhibitors.

USP7 and the p53/MDM2 signaling axis

p53 is a fundamental tumor suppressor protein known as the "guardian of the genome".[25] In a quest to understand the precise mechanisms through which the abundance of p53 is regulated, Li et al. identified USP7 (also known as HAUSP) as an interaction partner through affinity chromatography coupled to mass spectrometry.[26][Figure 1c] The authors went on to show that p53 abundance and ubiquitination levels are modulated by USP7 catalytic activity. In the following years, it became clear that USP7 directly deubiquitinates and regulates the abundance of MDM2, E3 ligase which ubiquitinates p53 and directs it for degradation.[27]

Thus, USP7 inhibitor programs emerged around the following hypothesis: USP7 inhibition leads to MDM2 degradation, which in turns leads to rescue of p53 from degradation, and finally therapeutic benefit.[18] Early inhibitors struggled with poor potency and selectivity, a course mirrored by the rest of the enzyme class. P22077 and HBX41108 were two examples of early USP7 inhibitors; both were identified using high-throughput screening, and have mid to low micromolar biochemical IC_{50}s.[28,29] These early first-generation USP7 inhibitors were widely applied in USP7 signaling studies, but ultimately were not potent nor selective enough to provide useful information. In 2017-2018, three highly homologous USP7 inhibitors were reported by three independent groups: XL188, ALM5, and FT671.[30,31] [Figure 1d] They were identified by different ways but converged upon a common hydroxypiperidine scaffold. An unrelated compound, GNE-6776 was also reported by Genentech to achieve USP7 inhibition by interfering with ubiquitin binding.[32]

In 2019, our group reported XL177A, a third-generation covalent USP7 inhibitor based on XL188 which targets the catalytic cysteine residue with a chlorotetrahydroacridine warhead.[33] Using XL177A, our group was able to show p53-mediated cell killing across a panel of Ewing sarcoma

cancer cell lines. Susceptability to XL177A was dependent on p53 mutational status, supporting that functional p53 is required for therapeutic benefit and supporting our initial hypothesis. XL177A treatment followed by proteomics revealed upregulation of p53 and downregulation of MDM2, confirming the therapeutic hypothesis.[33]

The USP7 example serves as an exemplar for the potential opportunities in targeted protein degradation afforded when deep DUB mechanistic research is combined with availability of potent and selective inhibitors. This is counterbalanced by unanswered questions in leveraging DUBs for targeted protein degradation, such as extent of functional redundancy across DUBs and layers of DUB activity regulation. In the case of USP7, the availability of small molecule inhibitors has led to pharmacological validation of therapeutic benefits from inhibition and confirmation of downstream mechanism of action. In addition to USP7, other DUB/substrate pairs identified for potential therapeutic intervention include USP28/Myc, USP8/EGFR, and USP1/ID1.[34–36]

USP19 in muscle wasting

USP19 was first identified as a potential protein of interest due to its upregulation in multiple condition of muscle wasting.[37] To investigate USP19's role in muscle wasting, Bedard et al. generated USP19 mice knockout and found them to be resistant against glucocorticoid-induced and denervation-induced muscle atrophy.[37] To obtain mechanistic understanding, the authors attempted to correlate USP19 expression to muscle wasting factors, but found the effect sizes to be generally quite small (20-30%) and limited to only a few proteins.[37] The authors concluded by nominating USP19 as a potential therapeutic target for muscle wasting.

In a follow-up study, Coyne et al. painted a mechanistic picture of how USP19 exerts its muscle-protective effects.[22] They identified that the effect was due to an increase in muscle synthesis and not a reduction in atrophy rates. They then linked that observation to an increase in insulin signalling and decrease in glucocorticoid signaling in USP19KO mice. Finally, they identified that glucocorticoid receptor (GR) levels are down in USP19KO mice, raising a possibility that USP19 might be directly controlling the ubiquitination levels and thus cellular degradation of GR.

Through first observing a phenotype of USP19 knockout and then identifying a potential mechanism, these studies have motivated the development of USP19 inhibitors in the biotechnology industry. Almac Discovery Limited filed a series of patents across 2018-2020 detailing the design/synthesis of USP19 inhibitors and applications in therapy of muscle wasting, obesity, and metabolic syndrome.[38,39] Similar to the USP7 inhibitors, the Almac USP19 inhibitors are based upon a 4-hydroxypiperidine core.[39]

This USP19 example serves as an example for DUB-focused phenotypic drug discovery. In this case, the authors investigated the role of USP19 with a mice knockout model. Alternatively, a USP19 small molecule inhibitor could have been an invaluable tool if one were available. One can imagine the use of a well-annotated library of small molecule DUB inhibitors to achieve reverse chemical genetic screen for identifying a DUB of interest in a particular phenotype. In addition to USP19, active inhibitor programs are also under way for USP30 in mitochondrial diseases and PLPro for CoVID-19.[23,40]

1.2.4 The current landscape of DUB inhibitor development

As with any enzyme class, inhibitors for DUBs would be invaluable as tools to investigate DUB basic biology and to pharmacologically validate DUBs are therapeutic targets. The history of DUB inhibitor development can be roughly split into before and after 2016.

Before 2016, whether DUBs were druggable remained an open question.[18] While more than 30 DUB small molecule inhibitors had been published at the time, very few compounds were comprehensively characterized across biochemical inhibition, cellular target engagement, biophysical binding, and selectivity profiling.[18] As such, most DUB inhibitors reported before 2017 were later shown to not inhibit DUBs, fail to engage target in cells, have poor selectivity, or act via unspecific mechanisms.[41] This is symptomatic of overreliance on one single assay for HTS screening or cellular signaling readout in the early days of DUB inhibitor discovery.

As the field matured, highly optimized DUB inhibitors characterized across multiple orthogonal assays started being reported. This is exemplified by the publication of potent and selective USP7 inhibitors in the 2017, where each report leveraged several assays to characterize inhibitors with the support of extensive structure-activity studies.[19] This provided much-needed proof of concept that DUBs are indeed druggable. In subsequent years, highly optimized, comprehensively characterized inhibitors for UCHL1, USP30, and USP9X were reported in the peer reviewed literature.[23,42–44] In aggregate, these reports bolster confidence that DUBs might indeed be broadly druggable.

1.2.5 Bottleneck in DUB inhibitor development

Despite recent progress, significant unexplored opportunity exists in the development of novel inhibitors targeting the important but understudied DUB enzyme family. Even with tremendous investment and interest from the past five years, DUB inhibitor discovery remains a slow one-at-a-time undertaking. The current bottleneck surrounding DUB inhibitor development is two-fold: (1) the lack of design principles for targeting DUBs; (2) poor biological understanding of DUBs.

As of September 2021, fewer than five DUBs possessed selective inhibitors that were validated for both biochemical inhibition and in-cell target engagement reported in the peer-reviewed literature.[23,30,42–44] With such few precedents, it is difficult to derive design principles for the target class. Aside from a common hydroxypiperidine scaffolds being used in USP7 and USP19 inhibitors, there is little evidence whether any given scaffold has affinity across multiple DUBs. At the same time, while studies between 2015-2020 have shown that it is indeed possible to develop potent and selective inhibitors for individual DUBs, the chemical tractability for inhibiting the target class at large is still an open question.[19] As a result, DUB inhibitor design principles remain elusive and DUB inhibitor discovery remains a slow one-at-a-time undertaking.

At the same time, biological substrates and function remain completely uncharacterized for the vast majority of the 100-member DUB family. Outside of the aforementioned well-studied few, very little is known about the remaining 90+ members of the DUB target class.[19] Indeed, Pubmed analysis revealed that 66% of DUB-related publications from 2000-2021 concern the top-10 most-studied DUBs (UCHL1, TNFAIP3, ATXN3, CYLD, USP7, COPS5, PLPro, BAP1, USP9X,

USP18). UCHL1, the DUB with the largest number of publications, alone accounts for 22% of all DUB-related Pubmed articles. [Figure 1e] As long as the biological function and substrate scope remains poorly understood for most DUBs, there is little motivation for the pharmaceutical industry to develop DUB inhibitors against the remainder of the target class. As such, inhibitor programs continue to narrowly focus on the aforementioned handful of DUBs that emerged as drug targets over a decade ago, neglecting potentially pharmacologically accessible family members of emerging interest.

With these two problems, a vicious cycle results. On one hand, the lack of inhibitor design principles slows down inhibitor discovery and in turn pharmacological elucidation of DUB biological function. Without biological knowledge, there is little therapeutic interest for most DUBs, demotivating inhibitor discovery targeting less-studied DUBs. Without studying the targetability of the target class at large, inhibitor design principles remain elusive.

In order to break the vicious cycle, a systematic attempt to discover and thoroughly validate DUB inhibitors to identify design principles synthesis is needed.

1.2.6 Towards DUB target-class inhibitor discovery

Against this backdrop, a target-class approach for DUB covalent inhibitor discovery holds immense appeal for accelerating DUB inhibitor discovery. As mentioned, target class approaches leverage targeted chemical libraries, tailored screening/validation assays, and structural insight to discover inhibitors for multiple members of structurally and mechanistically related enzyme classes in one campaign.

Highly conserved active sites and divergent folds architectures outside of the active site signal potential for selective covalent targeting of individual DUBs.[11] General homology and shared ubiquitin-binding features allow the use of common scaffolds in a library, while sequence divergence provides opportunities to form unique interactions for targeting specific DUBs selectively. At the same time, existence of an invariant catalytic cysteine residue on cysteine protease DUBs signals opportunity for covalent targeting. Covalent inhibitors also hold particular appeal as they can be modularly designed across noncovalent building blocks, electrophilic warheads, and linkers.[2] Our recent success at developing selective covalent USP7 inhibitor XL177A highlights that selective covalent targeting of individual DUBs at the catalytic cysteine residue is indeed possible.[33]

Despite the appeal, the path towards a target class approach for DUB inhibitor discovery is laden with challenges. At the start of this study, there was no DUB-targeted small molecule library, and very little robust historical data to inform design. Unlike in the case of kinases and ATP, there is no opportunity for ligand-based design: peptidomimetic inhibitors based on native ubiquitin substrates were attempted in the early 2000s, but poor potency terminated the inquiry.[45] Scarcity of selectivity profiling means whether divergent DUBs could be targeted by diversifying on some basic scaffold remains largely unknown. An invariant catalytic cysteine also raises questions about feasibility of selective covalent targeting. As such, design and synthesis of a first-generation DUB-focused library involves much risk. On the assaying side, existing DUB assays were only applied in low-throughput profiling of discrete compounds. Sophisticated screening and assaying platforms similar to Kinative or KINOMEscan for the kinases do not exist

for DUBs. The lack of an integrated, quantitative assaying platform made assessment, triaging, and development of chemical matter against DUBs very difficult.

1.3. Chemoproteomic methods for covalent inhibitors

To enable a systematic attempt to discover and validate DUB inhibitors, we turn to the power of chemoproteomics. Chemoproteomic techniques seek to understand the interactions between a given inhibitor and proteins in a proteome and are often paired with mass spectrometry for an unambiguous, quantitative readout. The irreversible nature of covalent inhibitor binding, as well as the resultant physical mass shift make this class of inhibitors particularly suited to mass spectrometry-based chemoproteomic analysis. As such, chemoproteomic techniques hold immense appeal in our quest to systematically discover and validate DUB inhibitors across the target class.

Most proteome-wide chemoproteomic techniques involve the use of an enrichment reagent. Each enrichment reagent biochemically captures, most often covalently, a defined set of cellular proteins, thereby focusing mass spectrometry analytical bandwidth on that defined set. This circumvents dynamic range limitations posed by the 12 orders of magnitude of cellular protein expression, enabling the detection of even low-abundance species. The set of proteins or residues captured by each enrichment probe constitutes a distinct “addressable chemical space” for each reagent. Importantly, it is the addressable chemical space (ACS) of the reporter probe, and not the inhibitor of interest, which dictates the parameters and the overall scope of the chemoproteomic experiment.

For inhibitor target profiling, inhibitors are competed against the enrichment probe for binding to cellular targets. This functions across target identification, hit identification, and lead characterization. Vehicle-only control and inhibitor-treated samples are incubated with the ABP, followed by subsequent enrichment and quantification. If signal of a particular family member diminishes from the control to the inhibitor-treated sample, it can be inferred that the inhibitor is competing against the probe for binding to that particular enzyme. This technique is commonly applied in target profiling for covalent inhibitors as well as examining the mechanism and target space of natural products for all three of probes discussed below.[46–49]

1.3.1 Activity-based probes and Bachovchin et al.

Since its introduction in the late 1990s, activity-based protein profiling (ABPP) has been used to characterize the function of enzyme families in complex proteomes.[50,51] ABPP leverages covalent probes (ABPs) which react in a mechanism-based manner with active-site residues of related enzymes. Thus, a classical ABP binds to a family of mechanistically related enzymes. ABPs consist of (1) a reactive group that selectively binds active site residues of the targeted enzyme class, (2) a reporter group, such as a fluorophore or an affinity tag for identification and enrichment purposes, and (3) a target class recognition moiety (i.e. ATP, Ubiquitin), which is connected to the warhead or reporter group. The use of ABPs enables visualization of enzyme activity and localization, selectivity profiling of inhibitors, and as a physical enrichment reagent to generate pools of an enriched enzyme class for further processing.[52]

By competing a given inhibitor against the ABP for target protein binding, ABPs offer a comprehensive profile of inhibitor activity against a given enzyme family of interest. Shortcomings of ABPP technology includes limited scope: ABPP does not capture compound

binding outside of the specified protein target class. In addition, it is often difficult to identify the ABP-modified amino acid. As a result, site-level binding data are often not available with ABP-based chemoproteomic methods.

Despite most applications being late-stage lead characterization for selectivity profiling, there is one example in which ABPP was applied in a hit discovery context. In 2010, Bachovchin et al. from the Cravatt Group reported the use of a fluorophosphonate-based ABP for inhibitor discovery against the serine hydrolases.[53] They syntheized a focused library consisting of 140+ carbamate-containing small molecules and screened it against a library of 70+ mammalian serine hydrolases overexpressed in HEK293T cell lysates. Competition against a fluorescent ABP for SH binding enabled readout of inhibitory activity by 1D SDS-PAGE. The authors coined this approach as “library versus library screening”, since a focused small molecule library is screened against a library of protein targets. The authors were able to identify hits for 40+ seriene hydrolases, achieve selective inhibition of homologous enzymes, and were able to quickly optimize one hit into a potent and selective probe for ABHD11.

This type of library versus library screen holds multiple advantages as a target-class-wide primary screen over traditional biochemical screening. First, since each molecule is screened against a library of enzyme class members, selectivity information is available upfront to facilitate hit triage and prioritization. Secondly, library versus library screening is uniquely poised to provide information about the target class at large. These include tractability of chemical targeting for the target class at large, or identification of “privileged scaffolds” that can be broadly leveraged to target members of the target class.

Despite its demonstrated power, Bachovchin et al.'s ABP-based inhibitor screening has not been widely adapted in the field, nor has there been attempts to develop MS-ABPP primary screening assays. There are multiple potential reasons for this. First and foremost, an ABPP-based primary screening assay coupled to a MS readout is very resource intensive. To enable such a screen, there needs to be significant investment in method development, instrumentation, and a dedicated small molecule library. As such, it is difficult to justify such a hit discovery campaign in the absence of a poorly understood enzyme class with high therapeutic potential. In addition, well-validated ABPs suitable for such screening attempts simply are not available for many enzyme classes, further discouraging such attempts. It is also of note that this approach goes against the current paradigm of target-based drug discovery: without the ability to specific a single target, library vs library primary screens are inherently target-class focused, demanding a shift from established pipelines of drug discovery.

1.3.2 Residue-based probes

In 2010, Weerapana et. al introduced isoTOP-ABPP (isotopic Tandem Orthogonal Proteolysis – Activity-Based Protein Profiling) as a method to enrich and identify accessible (reactive) cysteine residues in the proteome.[54] isoTOP-ABPP leverages an iodoacetamide probe as a broad cysteine alkylating agent that labels cysteine residues with an appropriate reactivity profile. The probe also includes an alkyne functional group for click chemistry incorporation of a biotin enrichment handle, and a cleavable, isotopically tagged peptide sequence to enable relative quantification of target peptides from multiple samples by mass spectrometry. Proteins with one or more cysteines annotated as 'reactive' represent a potential opportunity for covalent inhibitor development. The strategy also allows for identification of hyperreactive cysteine

residues, which are enriched at a level that is independent of the probe concentration. This hyperreactivity was interpreted by Weerapana et al. as indicative of specific functional role, such as catalysis.[54] In subsequent years, variations of the method for more focused characterization of functional cysteines within subcellular compartments or in particular oxidation states have also been reported.[55–58]

RBPs are typically used to identify off-targets for covalent compounds in a cellular proteome. Competing covalent compounds against the probe allows relative quantification of covalent compound binding against each of the amino acid sites captured by the probe. Compared to ABPs which is limited to a particular enzyme class, RBPs are potentially proteome-wide in scope: any residue which is captured by the probe is within scope for off-target identification. First introduced in 2014 by Wang et al. to quantify inhibitor binding to cellular cysteine residues, competitive RBPs and similar reagents are now commonly applied in target profiling for covalent inhibitors,[59] as well as examining the mechanism and target space of natural products.[46–49]

While RBP's wider ACS might initially appear more comprehensive and hence preferable over more focused techniques, it leads a more complex sample and complicates analysis. Consider the case of cysteine-targeting probes. There are predicted to be more than 200,000 cysteines in the human proteome.[60] The sheer size and enormous complexity inherent to the cellular cysteinome presents a significant challenge for robust detection and quantification of inhibitor targets by these residue-based methods. Hence, common cysteine-reactive probes routinely captures only 3000-4000 cysteine sites out of 200,000 (2%) in the proteome.[49,54,61] The challenge of identifying DUB cysteines in such studies has been explicitly discussed in many

such studies.[62] These problems are further exacerbated for residues that occur more frequently in the proteome, such as lysine. Thus, residue-based chemoproteomic data likely comprise only a subset of peptides targeted by the compound under study, and additional targets may exist.

1.3.3 Inhibitor-based probes

Inhibitors can be fashioned into enrichment probes through elaboration with an affinity enrichment handle or bioorthogonal chemical handle for subsequent coupling. To the extent that these modifications only modestly alter selectivity, the binding behavior of the IBP will closely mimic that of the native inhibitor. In this way, competing parent inhibitor against the IBP for enrichment provides a high-fidelity readout of the pharmacologic binding activity of the native inhibitor.

In principle, repurposing a selective inhibitor as a reporter probe offers several advantages. The ACS of the IBP is dictated by the parent inhibitor and is expected to be modest in size. Therefore, from an analytical perspective the LC-MS/MS analysis can be more straightforward. These promising aspects of IBPs are accompanied by specific challenges. Unlike the two previous categories where a generic probe can be used across different inhibitors, each native inhibitor is chemically modified to create a paired IBP. This elaboration process is not trivial: it demands structural data to determine a suitable exit vector and additional validation to confirm that the IBP retains suitable binding activity against the intended target. Finally, the size and complexity of IBPs often complicate MS/MS identification of probe-modified peptides.

IBPs can come as biotinylated inhibitor probes or alkynylated inhibitor probes, whose main difference lies in bioavailability. Due to the large size of the biotin tag and linker, probes

fashioned from direct installation of biotin usually cannot cross mammalian cell membranes. In contrast, the minimal size of the alkyne handle often allows live-cell treatment. Owing to its small, bioinert chemical footprint, the alkyne handle is less likely to perturb on- and off-target binding compared to a biotin modification. Alkynylated inhibitor probes were first reported by Wright et al. in 2007, where an alkyne handle was appended onto cytochrome P450 inhibitor 2-ethylnaphthalene.[63] Since then, the aforementioned advantages have led to the widespread application of alkyne functionalization followed by click chemistry and pulldown in off-target profiling for covalent small molecules.

Out of the three types of chemoproteomic probes introduced, the DUB target class-focused work described in this thesis primarily centers around classical ABPs. I build upon the ABPP-based approach by Bachovchin et al. to achieve library versus library primary screening coupled to a mass spectrometry readout. This type of methodology is ideal for our campaign to identify covalent inhibitors for the DUB target class with rich unexplored opportunities. In addition, this can also address unanswered questions surrounding druggability for the target class at large and identification of privileged scaffolds. We leverage RBPs and IBPs downstream to identify proteome-wide off-targets for selected inhibitors.

1.3.4 Statistics for chemoproteomic studies

As chemoproteomics is still in active development, the field has not settled on a common methodology for determining thresholds for probe competition and inhibitor selectivity. As such, reported characterization of probe binding varies widely across studies. The metric for quantifying competition itself is straightforward: fold change in signal across DMSO and

inhibitor-competed conditions. However, significance tend to be defined by basic integer fold-change thresholds, and use of statistical tests is limited.[61,64–68]

As the field progressed, the field started borrowing statistical methods from the field of genetics to evaluate changes in protein abundance levels. The predominating statistical method used for controlling chemoproteomic datasets concatenates competition ratios for each detected protein species into a single list to generate a distribution with which each selected data point (CR for a single protein species) can be tested against.[59,61] This approach is widely applied because it requires a small number of experimental replicates for high statistical power, since each detected protein serves as a single data point, essentially generating a 10,000-member null distribution with every replicate that captures 10,000 proteins.

In essence, this workflow aims to compare competition ratios across all detected proteins, which holds various troubling implications. The primary shortcoming of this model is that in concatenating all proteins into a single list, it assumes that baseline distribution of each and every protein is similar, and that the same competition ratios are statistically comparable across protein species. This is often untrue: relative baseline signal variation for low abundance species can be much larger than that for high abundance species. As such, the same "2-fold change" for two protein species can be vastly different in terms of statistical significance. This can result in high false positive rates and wasted time in hit follow-up.

Another, more technical problem surrounds the statistical question being asked. In an inhibitor assay development context, it is of paramount importance to distinguish signal due to compound-induced inhibition from background noise due to baseline random fluctuation in signal. Indeed, that is the underlying impetus for undertaking statistical testing. However, in

concatenating competition ratios across protein species, the statistical question being asked is different. In testing a given competition ratio against a distribution of all other competition ratios, one is asking how much a particular protein-inhibitor interaction is standing out from the rest. While this might sound like a subtle technical differentiation, it holds significant implications. For one, concatenation methodologies will fail in the context of a promiscuous inhibitor, where the distribution of competition ratios flatten out.

An alternative approach would be to establish a baseline for each detected protein species, then determine the statistical significance of the competition ratio observed in a treatment condition relative to the baseline from null conditions. By carrying out statistical analysis in a protein-by-protein fashion, competitive ratios in treatment conditions can be compared against the background noise specific to the exact same protein in null conditions for a more case-specific control. There is no assumption of equivalency across different proteins.

As of June 2021, there are no reported datasets of multiple null chemoproteomic experiments in the literature, for a number of reasons. First, a large number of (50+) null replicates are required to obtain such a baseline with any statistical power. In the context of one-off chemoproteomic experiments as is routine in inhibitor target characterization, such numbers of null replicates simply are not carried out. In addition, proteins need to be consistently detected for such a baseline to be drawn. Due to the aforementioned problems with replicability in RBP-based methods from the stochastic nature of LC-MS/MS data acquisition, many proteins of interest simply are not detected with sufficient reproducibility for such a baseline to be drawn. The final reason is one of practice. In the few reports of large-scale chemoproteomic screens, only competition ratios are reported; baseline raw signal for the null conditions are simply

omitted. Thus, even when large numbers of replicates are carried out, such baseline raw signals are not reported for public analysis.

As an extension of our goal to explore applications of mass spectrometry chemoproteomics as a novel approach for covalent fragment screening, we will also explore measures of statistical control. Theoretically, the large number of planned replicates will enable calculation of a baseline distribution for each detected protein species, which can then be compared with the predominant methodology of concatenating competition ratios across protein species. Our hypothesis is that this will reveal large protein-by-protein variation in the magnitude of fold change required to be statistically significant and inform on best practices in hit identification for future chemoproteomic primary screening.

1.4. Towards chemoproteomic DUB target-class covalent drug discovery

In addition to the aforementioned proteome-level proteomic techniques, assays for DUB biochemical activity, target engagement, and other tests for covalency are needed to achieve DUB target-class drug discovery. Below I introduce each assay leveraged in this thesis, then conclude the introduction by summarizing each chapter briefly.

1.4.1 Assays used in this thesis

Below, I will discuss DUB assays frequently used in this thesis. A basic understanding of the operating principles, advantages, and disadvantages are needed for comprehending the work presented in this thesis.

Western blot target engagement with DUB ABPs

I utilize a ubiquitin-based DUB ABP to confirm cellular target engagement by each inhibitor.[28,33] In the absence of compound, ABP binding to DUBs lead to a +10kDa shift in mass, which can be easily visualized by Western blotting. Competition by compound for DUB binding leads to disappearance of the +10kDa band and appearance of a -10kDa band, signifying target engagement.

The limit of detection for this assay stems from limited resolution for the 10kDa shift in mass for ABP-binding. Using conventional 4-12% Bis-Tris gels and SDS-PAGE conditions, it becomes difficult to resolve the +10kDa band for DUBs larger than 150kDa. As such, inhibitor cellular target engagement of large DUBs (e.g. USP9X) cannot be probed using this method.

DUB biochemical activity assays

Ubiquitin-rhodamine is a commonly used fluorescent substrate for assaying DUB catalytic activity.[69] It comprises of a rhodamine fluorophore conjugated to the C-terminus of ubiquitin protein, where conjugation to substrate lysines occurs. Deconjugation catalyzed by DUBs releases free rhodamine, which then fluoresces at 535nm when excited at 485nm.[69] In our kinetic biochemical assay, purified recombinant DUBs are first pre-incubated with inhibitor over a 10-point dose curve. After substrate addition, fluorescence is continuously monitored. The rate of fluorescence increases and thus product generation can then be plotted as function of the inhibitor to calculate a biochemical IC_{50} of the inhibitor.

The main limitations of these DUB in vitro biochemistry assays involve availability of the purified DUB enzyme and activity towards the synthetic substrate. The DUB of interest must be able to be recombinantly expressed and maintain proper folding and activity in the conditions of the assay. While this sounds trivial, this is not the case for many DUBs. In addition, the DUBs tested must possess protease activity against the synthetic monoubiquitin-rhodamine substrate. As some DUBs recognize specific polyubiquitin chain types and are not active monoubiquitin, the commercially available Ub-Rho substrate is not suitable against those family members.

In addition to these in-house assays, I also utilize the DUBprofiler™ service provided by Ubiquigent Ltd. Using the same ubiquitin-rhodamine substrate, this services tests a given inhibitor at one single dose against a panel of 41 DUBs. This is particularly helpful for biochemical selectivity profiling against DUBs we do not have in stock.

Intact protein mass spectrometry

In general, most strategies for analyzing purfied proteins using mass spectrometry can be roughly divided into “intact protein mass spectrometry” where proteins are analyzed without digestion, and methods where main analytes are peptides derived from proteolytic digestion. In intact protein MS, purified protein(s) are reacted with electrophile-containing compounds of interest and are then typically analyzed by liquid chromatography (LC) or capillary electrophoresis (CE) coupled directly to a mass spectrometer with electrospray ionization (LC-MS or CE-MS).[70,71] The protein of interest is detected as a series of multiply-charged ions, which can be computationally deconvoluted to yield a nominal mass.[72]

A shift in mass corresponding to that of the inhibitor minus any leaving group between compound-treated and control samples indicates covalent labelling of the protein. This direct observation of the inhibitor-protein adduct provides information about covalent mechanism of action. Additionally compounds which indiscriminately label multiple residues on a protein to more than a one-to-one ratio can be triaged for hyperreactivity. Despite its usefulness, intact protein MS has an upper limit for protein analyze size and purity.

Peptide-level MS analysis

In many cases the same inhibitor-protein reaction mixture used for intact protein mass spectrometry can be processed for peptide-level analysis.[73] Generally, intact protein mass spectrometry and peptide level analysis go hand-in-hand and provide complementary information. While intact protein mass spectrometry offers confirmation of covalent bond formation and protein-level stoichiometry, peptide level analysis offers information about the

precise site(s) of modification as well as inhibitor site occupancy. Considering inhibitor-specific fragmentation behavior can improve confidence sequence identification of the labelled peptide.[74]

The strength of LC-MS/MS analysis at the peptide level lies in an unbiased view for identifying the targeted residue with site occupancy information. Unlike mutagenesis-based approaches, no *a priori* information about the site of binding is needed. This unbiased nature also enables compound triage: compounds that label target proteins with apparent one-to-one stoichiometry may in fact be nonspecifically reacting with different residues on each target protein molecule. If multiple residues were observed to be labelled by the inhibitor, this could indicate a hyperreactivity liability.

When digesting protein-inhibitor complexes for peptide-level analysis, it is important to consider the nature of the covalent linkage. Some reversible covalent inhibitor-protein complexes are known to dissociate in the reducing environment of protein digestion. This has been observed with the N-cyanopyrrolidine warhead commonly used in this thesis.[42]

1.4.2 Towards chemoproteomic DUB target-class covalent drug discovery

With the vision of a library versus library screen and a toolbox of both DUB-focused and unbiased chemoproteomic, biochemical, and cellular assays, we set off to revolutionize DUB inhibitor discovery.

Our goal was to validate targetability of the entire class outside of the limitations of the one-target-at-a-time paradigm. A covalent compound library tailored for the DUB active site maximizes positive hit rate, while diversification expands DUB scope. Consistent, reliable

quantification of compound potency across the target class provides insight on structure-activity relationships, privileged chemical scaffolds, and selectivity for individual hits.

In chapter 2, I will discuss how I built the foundations to enable a library versus library screening campaign, including: (1) the design and synthesis of a DUB-focused covalent small molecule library, and (2) building, evaluating, and validating a mass spectrometry ABPP primary screening assay. I tested out both aspects of our vision in a small 50-compound pilot screen, which informed on weaknesses in the workflow for further improvement. I act on those observations and make changes to protocol to address shortcomings.

In chapter 3, I embark on a full screen of the 150-member covalent small molecule library. In this chapter, I summarize the results of the screen, then identify, validate, and characterize hits against 15 DUBs from 4 subfamilies. Then, I describe medicinal chemistry optimization on hits against VCPIP1, which resulted in a potent, selective, first-in-class inhibitor.

In chapter 4, I describe parallel advances in statistical analysis of chemoproteomic primary screening data sets. I introduce individualized protein baseline extraction (INPROBE) as a novel method for protein-by-protein statistical hit identification.

In chapter 5, I conclude by discussing the future outlook on two fronts: development of DUB inhibitors as novel targeted protein degradation therapy and application of chemoproteomic covalent inhibitor screening in hit identification.

2. Foundations for a library versus library screen

At the start of this project, neither the DUB-focused library nor the DUB-wide screening assay in our vision to revolutionize DUB inhibitor discovery existed. This chapter is about building, test-running, and refining the library and methodology for a full library versus library screen.

The first two parts of this chapter focus on (1) design and synthesis of a DUB-focused covalent small molecule library, (2) development of an ABPP-based primary screening assay. With both components of the screen in hand, I tested the methodology in a pilot screen. Opportunities for improvements were identified, and adjustments to the protocol were tested accordingly.

2.1. Design and synthesis of a DUB-focused covalent small molecule library

A DUB target-class approach necessitates a DUB-focused small molecule library. Since none existed when we initiated our study, we designed and synthesized a first-in-kind DUB-focused covalent small molecule library.

We began by identifying foundational chemotypes on which library design can be based. Criteria for foundational chemotypes compounds included the following: (1) Compound must be potent with IC_{50} <1μM against a target DUB; (2) Compound must be covalent; (3) Compound must be modular in makeup and synthetic route to facilitate downstream diversification. By these criteria, we identified and validated three foundational chemotypes as putative DUB inhibitors. [Figure 2b] SB1-F-22 is reported in a patent published by Mission Therapeutics as a UCHL1 inhibitor. XL177A is developed in-house in a structure-guided campaign based on an X-ray co-structure of USP7 bound by our noncovalent USP7 inhibitor XL188.[33] AV12 was reported by Ward et al. to bind to USP4, USP16, among other USP subfamily DUBs.[75] Each compound

was validated for biochemical inhibition, covalent bond formation, and in-cell target engagement; where unreported, we performed the validation ourselves.

Next, we assembled a first-in-kind DUB-focused covalent small molecule library by diversifying upon the foundational chemotypes. The expectation was that diversification upon these basic scaffolds would expand the scope of inhibition to other DUB family members, hopefully with good selectivity. The library was mostly made up of *de novo* compounds designed and synthesized in house, supplemented with molecules from commercial libraries and patent/peer-reviewed literature. Keeping in mind that this was a first-generation library, we strived for maximal diversity to facilitate exploration of chemical space.

Members in this first-generation library mostly followed a three-piece modular design, containing: cysteine-reactive electrophilic warheads, chemical linkers, and reversible binding moieties. [Figure 2c]

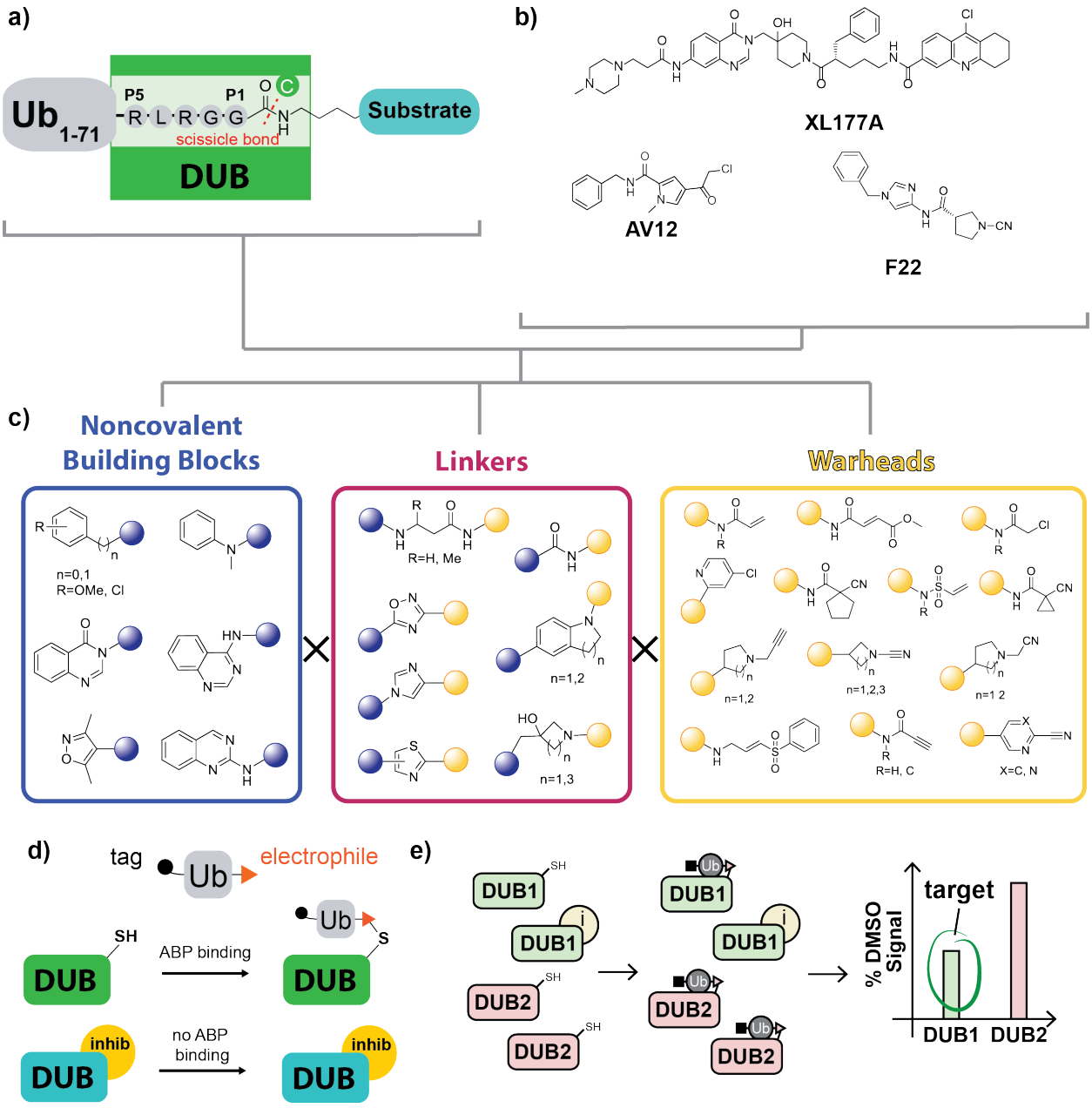


Figure 2.1. The DUB target-class screening platform. a) DUB catalytic domains feature a long and narrow channel leading up to the catalytic cysteine residue. b) Structures of foundational chemotypes AV12, F22, and XL177A. The C-terminus of the substrate ubiquitin occupies this channel. c) Library compounds follow a three-piece modular design, where noncovalent building blocks, linkers, and warhead moieties are diversified as shown to generate the library.

(continued from previous page) d-e) DUB ABPs consist of a ubiquitin moiety, a C-terminus electrophile and a N-terminal biotin tag. They react with the catalytic cysteine residue on active DUBs only. Inhibitor-bound DUBs do not react with the probe. This differential labelling enables indirect readout of compound binding to each DUB.

In designing the electrophilic warheads, we considered electrophile reactivity, reaction mechanism, and geometry of the glycine-lysine isopeptide bond that is natively cleaved by DUBs. Common cysteine reactive moieties with varying reactivity such as acrylamides, vinyl sulfonamides, and α-chloroacetamides are included. With reference to validated activity against the wider cysteine proteases superfamily outside of DUBs, cyano, alkynyl, and halogenated pyridyl electrophilic groups were included as well.[76] In terms of mechanistic diversity, cysteine thiols can attack these warheads by S_N2, nucleophilic aromatic substitution, and Michael addition. As triazoles have been reported as isosteres for the ubiquitin-substrate isopeptide bond, we posited that the similar warhead-bearing pyrrolidine ring on F22 might function in a similar manner.[77,78] We elaborated on this electrophile-bearing ring motif by altering ring size, geometry, degree of saturation, and electrophile position. Some compounds possess chains extending beyond the warhead to mimic the lysine side chain from the ubiquitinated substrate. To enable effective comparison across different electrophiles, we designed series of compounds in which everything except the electrophile was kept constant. In designing the linker, we were inspired by the C-terminal peptide sequence of ubiquitin (LRGG). [Figure 2.1a] This C terminal tail has been demonstrated to traverse a narrow channel leading up to the reactive catalytic cysteine residue between the palm and the thumb of ubiquitin specific

proteases, and provides a model for our compounds to do the same (Figure 2b).[79] Hence, carbon chain length and amide bonds on the linkers were designed to mimic the sterics and hydrogen bond donor/acceptor placement of the GG C-terminus motif. We further diversified on these aspects by length and structural flexibility in the hopes of capitalizing on differences in this channel across different DUBs. We spaced the linker and the warhead with a methylene group in some compounds to enable linker-independent warhead rotation in the hope of expanding access to catalytic cysteines with different placement in the active site across different DUB families.

Building blocks were diversified from reported DUB inhibitors and optimized with insight from DUB-inhibitor co-crystal structures. Current inhibitor-DUB cocrystal structures have blocking loop 2 in diverse conformations in USP family members.[30,80] To explore and leverage this plasticity, we incorporated a diverse set of building blocks on leftmost part of the molecule. We attempted to occupy the leucine-binding S4 site by placing mostly rigid hydrophobic moieties at that position. Rings of different sizes, hydrogen bonding capacities, and geometries were incorporated for maximal exploration of chemical space. Various heterocycles and fused rings systems featured in other reported DUBs inhibitors are also incorporated. In total, the Buhrlage Lab, with efforts led by synthetic chemist Xiaoxi Liu, synthesized approximately 150 compounds; structures of in-house compounds are summarized in supplemental information.

To further enhance chemical diversity, selected representatives from a commercial covalent small molecule library from Ubiquigent Ltd. were included to further enhance structural and electrophile diversity. Reported covalent DUB inhibitors were also included to a) serve as internal

validation for the primary screening assay, and b) explore additional DUB targets outside limited profiling information from the original report.

2.2. Development of an ABPP-based primary screening assay

As discussed in the introduction, this work is partially inspired by Bachovchin et al.'s work in library versus library screening for serine hydrolase inhibitors.

We aimed to leverage activity-based protein profiling (ABPP) technology for a target-class primary screening assay to read out library compound binding to endogenous DUBs. DUB activity-based probes (ABPs) contain a ubiquitin moiety with a N-terminal biotin and a C-terminal electrophile. [Figure 2.1d-e] The electrophile reacts with the catalytic cysteine residues on active DUBs, thereby irreversibly labelling the DUB with the probe. The biotin handle then allows DUBs to be pulled down with streptavidin affinity purification. Since only DUBs bound by the ABP are ligated to biotin, inhibitor-bound DUBs would not be enriched. Hence, differential enrichment of DUBs across DMSO and inhibitor-bound conditions allow indirect quantification of compound binding to each target DUB.

2.2.1 Previous work

There were multiple opportunities for further improvement of Bachovchin's strategy. First, the use of a mass spectrometry readout instead of SDS-PAGE enabled unambiguous target protein identification. Due to limitations in protein MW resolution on a 1D SDS-PAGE gel, Bachovchin et al. screened compounds in pools of several purified SHs at a time. We elected to read out compound competition against endogenous proteins derived from mammalian cell lysates,

which allowed single-pot competition against all target class members in an endogenous context.

Immediately before the start of this study, Lawson et al. reported using a cocktail of DUB ABPs to profile the DUB targets of isothiocyanate natural products in HeLa cell lysates.[46] They used stable isotope labelling with amino acids in cell culture (SILAC) for quantification: populations of HeLa cells grown in light and heavy nutrition media were treated with DMSO and test compound respectively. After lysis, they treated with a 1:1 mixture of DUB ABP biotin-ubiquitin-vinylmethylester and biotin-ubiquitin-propargylamine, combined, enriched with streptavidin, digested, and analyzed by MS. They identified 35 DUBs in the mixture and saw labelling blockage of 10 DUBs. Methodologically, this study is significant for its use of DUB ABPs to profile compound off-targets in a one-off fashion and provide a good starting point on which we can base our DUB ABPP primary screening assay methodology.

2.2.2 Foundations of DUB ABPP

We identified two primary limitations in Lawson et al.'s workflow: (1) coverage of DUBs, and (2) limits in number of channels per experiment. The critical downfall is that the experiment only captured 35 DUBs, approximately one-third of DUBs in the human genome. With such a limited scope, it is difficult to make any conclusions about inhibitor selectivity and target class-wide chemical principles. Another problem concerned the use of SILAC-based quantitation, which only permits a maximum of three channels per experiment. With the DMSO negative control taking up one slot and positive control taking up another, this leaves only one for a test compound. This low test-to-control ratio diminishes the practical efficiency of the screen. Thus,

efforts were made to: (1) expand the range of DUBs quantified with the assay, (2) shift to TMT-based quantitation to accommodate more channels per experiment.

To expand the range of DUBs quantified with the assay, we began by identifying optimal ABPs for the screening system. Based on RNA-seq data, we identified the HEK293 cell lines to express 72 out of 77 cysteine protease DUBs in the human genome.[81] The quick doubling time of these cells also supports their use as a source for input cellular material. To minimize live cell-associated artifacts such as variations in compound concentration due to differing cell permeability and downstream modulation of protein abundance, we elected to carry out compound treatments in cell lysates.

We carried out a first-pass competitive pulldown with conditions similar to Lawson et al using a panel of three commercial DUB ABPs with varying warheads. [Figure 2.2a-c] HEK293T cells were lysed, then incubated in DMSO or compound for 5 hours. This was followed by addition of each ABP with 90 minutes of incubation. Proteins were denatured by boiling in 1.2% SDS, followed by streptavidin pulldown, reduction, alkylation, and overnight digest. Resulting peptides were desalted and TMT-labelled. 5% of each sample was pooled, desalted, and analyzed by LC-MS/MS to ensure TMT labelling exceeds 95% before final analysis. After passing the TMT percentage labelling screen, 50% of each TMT-labelled sample was pooled, desalted, and analyzed by LC-MS/MS. TMT reporter ion intensity is compared across DMSO and compound-treated channels to obtain a percentage value of ABP labelling blocked, a proxy measure of compound binding to each DUB. A first-pass pulldown without compound competition was able to identify 40 DUBs from a HEK293T cell line from the samples that used the biotin-ubiquitin-vinylmethylester and biotin-ubiquitin-propargylamine provide the widest range of DUB capture.

The complementary DUB profile of the two probes led us to proceed with a 1:1 cocktail of the two.

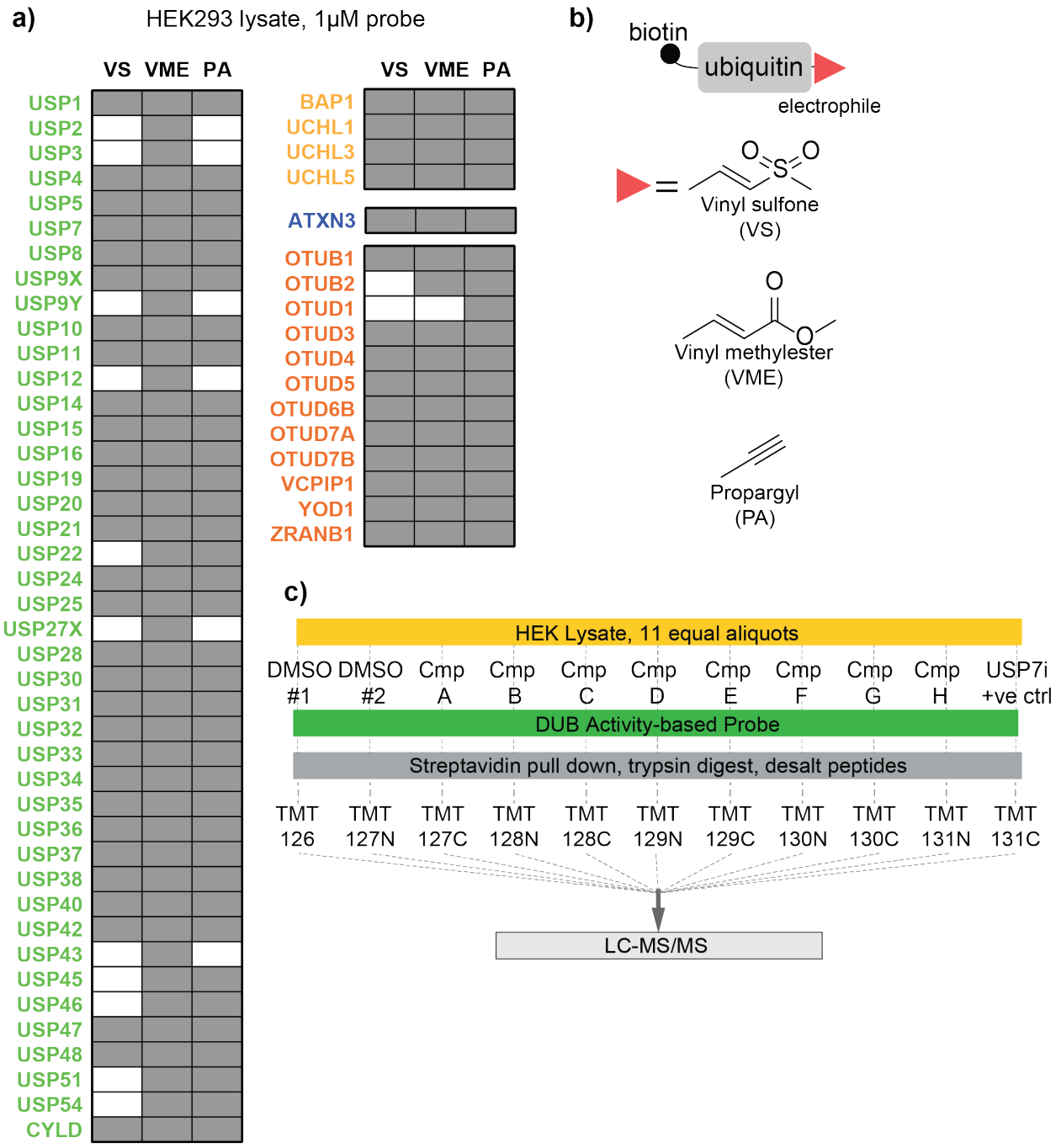


Figure 2.2. Exploring the DUB ABPP Assay. a) DUBs identified after treatment and enrichment using three ubiquitin-based ABPs. b) Chemical structures of the three electrophiles on the ABPs

(continued from previous page) used. c) Schematic of the DUB-ABPP assay. HEK 293T lysates are treated with DMSO/inhibitor for 6 hours, then incubated with ABP for 90 minutes. This is followed by streptavidin pulldown, reduction, alkylation, and overnight tryptic digestion. Resultant peptides are desalted, labelled with TMT, and undergo one final desalt step before LC-MS/MS analysis.

2.2.3 Validating competition

With 40 DUBs identified in the pulldown, our next goal was to validate specificity of the probe-mediated DUB enrichment. Towards this purpose, we competed the probes against warhead-less biotin-ubiquitin in a self-competition experiment. If enrichment was indeed specific, we should observe a dose-dependent decrease of DUB enrichment and signal as concentration of biotin-ubiquitin increases. Indeed, that was the case, showing probe-mediated pulldown was indeed specific for a broad swathe of the DUBs detected in the assay. [Figure 2.3a]

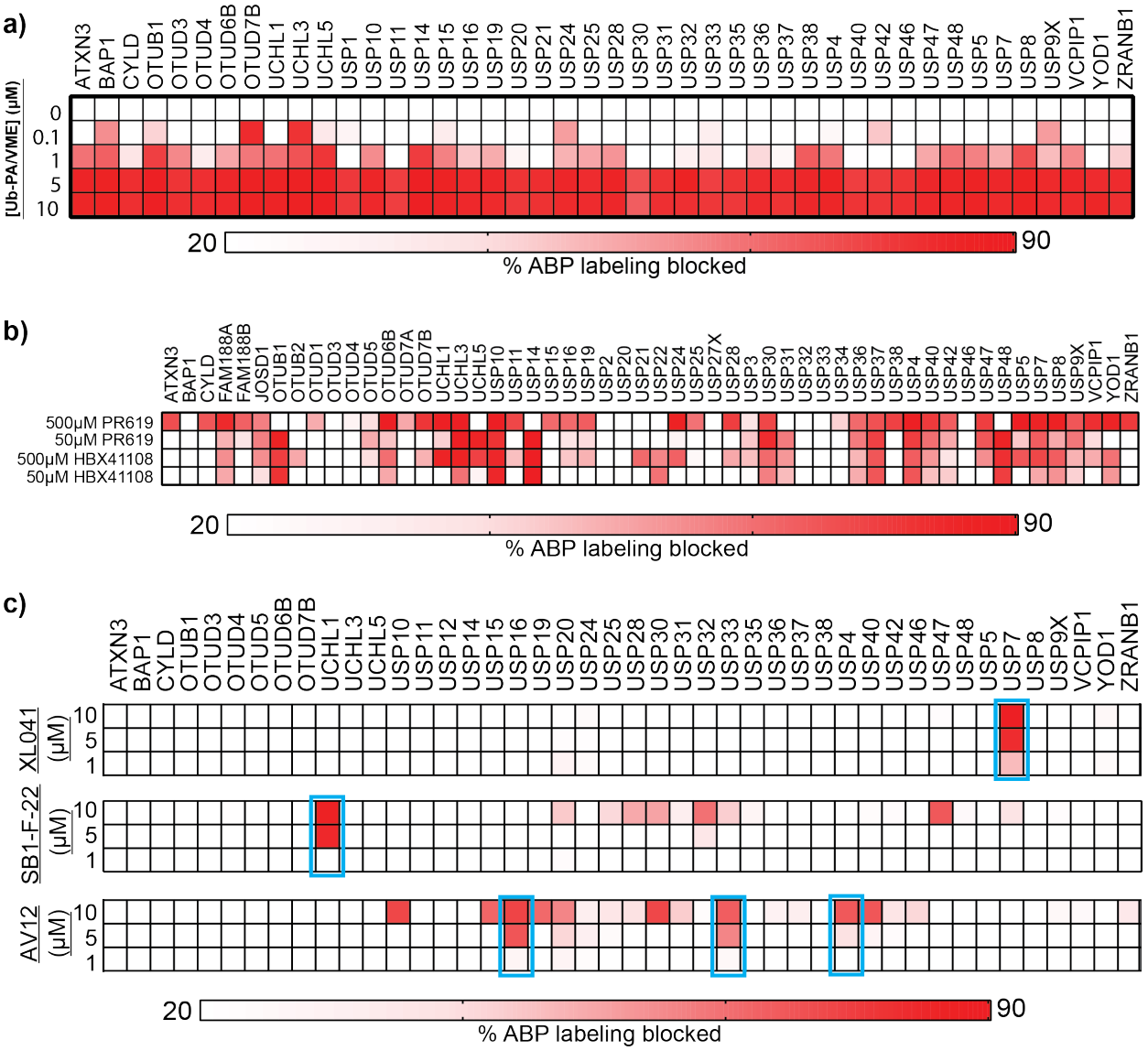


Figure 2.3. Validation of the ABPP experiment. a) A self-competition experiment with biotin-free ubiquitin-propargylamine/vinylmethylester compounds supported specific enrichment of DUBs by the DUB ABPs. b) Competition against promiscuous DUB inhibitors showed the assay could read out small molecule binding against most members of the target class. c) Competition against foundational chemotypes demonstrated that DUB ABPP could read out selective inhibition of known DUB target for each compound.

Next, we validated that small molecule blockage of the DUB catalytic cysteine residue can indeed be read out via the assay in a reproducible fashion. We leveraged two promiscuous DUB inhibitors, HBX41108 and PR619 to induce blockage of the catalytic cysteine residue generally across the DUB target class in HEK293T cell lysates. Results from the MS-ABPP assay showed widespread blockage of most DUBs detected, consistent with literature characterization of these compounds.[41] [Figure 2.3b]

To examine whether the assay was indeed able to read out selective binding, we carried out a competition assay with three dose points of each well-characterized foundational chemotype compound. Satisfyingly, we observed consistent probe binding blockage patterns that were consistent with the reported targets of each foundation chemotype compound: AV12 blocked USP4, 10, 16, 19, and 40; F22 blocked UCHL1, and XL177A blocked USP7.[33,43,75] With validated ability to read out both selective and promiscuous compound activity, we concluded that this assay is ready for a pilot screen. [Figure 2.3c]

2.3. Pilot Screen

As an initial evaluation of our DUB-focused library and activity-based protein profiling platform, we performed a pilot screen of 49 compounds at a single high concentration against cellular DUBs. [Figure 2.4a]

We were pleased to observe that 71% (35 of 49) of the compounds were 'active', defined as blocking ≥ 40% of ABP binding for at least a single DUB. [Figure 2.4b] Encouragingly, target coverage was broad; 30 DUBs were bound by at least one compound. As anticipated for a first-

generation library, we observed a range of selectivity for our compounds, from selective for one DUB to targeting 20% of all detected DUBs. These results suggested discrete selectivity categories to prioritize compounds for validation studies: (i) selective for a single DUB, (ii) binds to a small number (2-5) of DUBs, (iii) multi-targeted within one DUB family and (iv) multi-targeted spanning multiple DUB families. As summarized, each of the five selectivity classes were represented by more than one compound in our library. These data lend considerable confidence to the notion that our DUB-targeted library is populated with a sufficient number and diversity of DUB ligands to generate hits suitable for further development of lead compounds.

Even at this early stage, data from our pilot screen reveal valuable clues for rational design of DUB-targeting compounds. For example, USP25 and USP28 along with UCHL1 and UCHL3 are the two pairs of DUBs with the highest sequence homology within their respective catalytic domains; interestingly our pilot screen identified hit compounds which preferentially target USP28 or UCHL1 within each high-homology DUB pair. These results are particularly encouraging given that similar discovery efforts using reversible ligands have failed to identify compounds that are selective for USP28 over USP25. As a second example, compounds XL027A, XL027B and XL027C contain the same noncovalent building block with different warheads. [Figure 2.4c] We found that these compounds shared a subset of DUB targets, with each compound also binding DUBs uniquely from the other two. These data validate our approach in both library design and screening, supporting the fruitfulness in screening a full library.

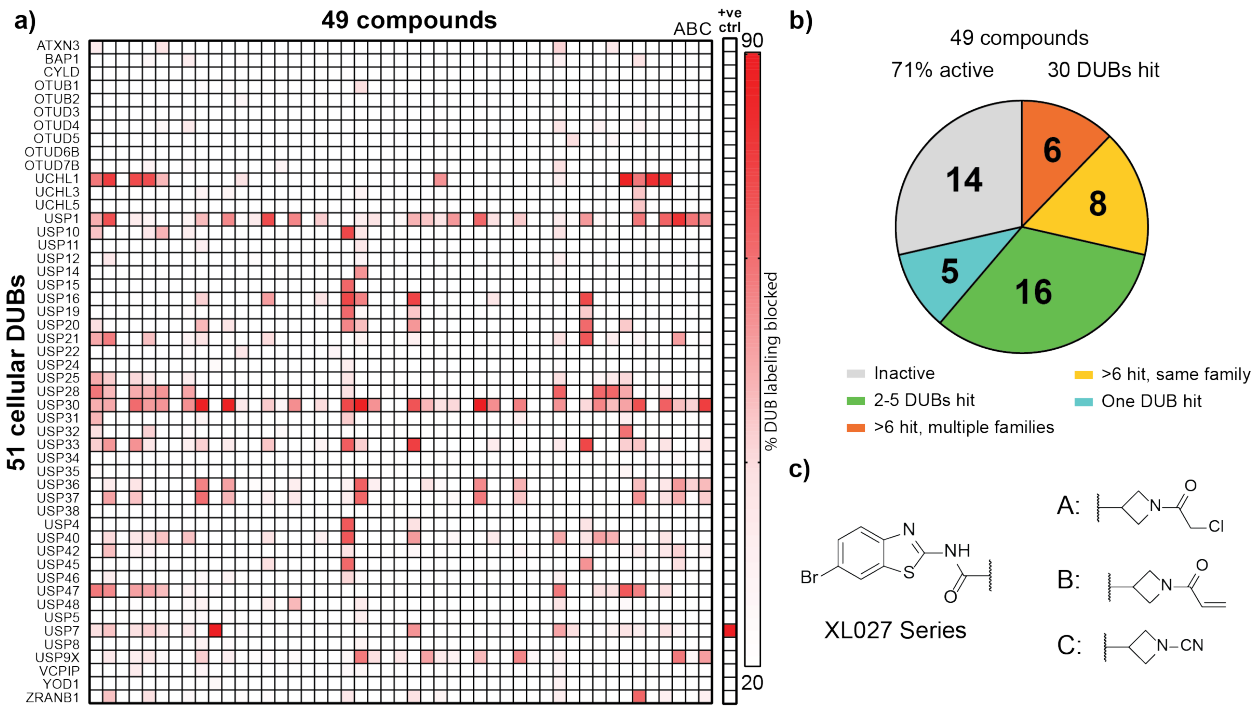


Figure 2.4. Pilot screen. a) Heatmap showing binding data for 49 compounds against 50 total endogenous DUBs. b) Library compounds displayed a wide range of activity against DUBs. c) The XL027 series highlighted how diversification at one part of the molecule can attenuate the DUB inhibition profile of resultant compounds.

2.4. Further method development

In the process of the pilot screen, I noticed several points for improvement surrounding the cost, sensitivity, and reproducibility of the assay. In this subsection, I will discuss the efforts I made towards improving the screening assay by experimenting with different pulldown reagents, examining the efficiency of DUB capture, and titrating down reagents for cost-reduction.

2.4.1 Optimizing pulldown conditions

Over the course of the pilot screen, it came to my attention that the streptavidin pulldown step was problematic for two reasons: (1) variable bead loss across samples during washing introduced considerable noise; (2) there was significant background signal in my pulldown samples, with some identifiable components being streptavidin and trypsin peptides. To combat these problems, I tested various pulldown and digestion conditions in TMT-free pulldowns, then evaluated the result of each using by the number of DUB detected, number of DUB PSMs, and the magnitude of contaminant PSMs as a percentage of the total. For simplicity, representative data from a single run is shown for each condition.

I first made the switch from non-magnetic to magnetic streptavidin beads to improve handling and reduce the minimum working amount of beads for each pulldown. Up till this point, I had been using the Thermo Pierce non-magnetic high-capacity streptavidin beads. Due to the non-magnetic nature of the beads, I had to use a large 40-fold excess of streptavidin beads to obtain a reasonable volume (25μL) for handling. This additional unnecessary streptavidin then generates contaminant peptides during on-bead digest. We reasoned that the use of magnetic beads with a lower concentration of streptavidin on the beads would lead to improved handling, reduced streptavidin contamination, and hopefully wider, more consistent DUB detection.

As shown in Table 2.1, use of magnetic streptavidin beads did indeed lead to improved performance. The number of DUBs detected increased by 17% to 46, while the number of DUB PSMs surged 70%, from 717 to 1219. Meanwhile, the number of streptavidin PSMs decreased moderately by 20%, from 722 to 565. This translates to a decrease from 16% to 8.4% in terms

of the relative percentage of PSMs that streptavidin takes up. Taken together, this suggested that the shift to magnetic streptavidin beads led to improved performance of the assay.

Table 2.1. Results comparison for DUB enrichment using original non-magnetic streptavidin beads versus new magnetic streptavidin beads.

	Non-magnetic beads	Magnetic streptavidin beads
Number of DUBs detected	39	46
Number of DUB PSMs	717	1219
Total number of proteins identified	722	1253
Total number of PSMs	4473	6732
Number of streptavidin PSMs	722	565

Separately, with updated instructions from the vendor, I started reconstituting the ABP cocktail with a new procedure. Instead of dissolving the dry powder in DMSO, then adding 50mM Tris pH 8 and 150mM NaCl on top for a 200μM solution, the new protocol called for adding the DMSO solution to 50mM pH 4.5 sodium acetate for proper re-folding of the ubiquitin moiety. With this change, we observed greatly enhanced performance in DUB enrichment, with 61 DUBs detected (33% increase), 1936 DUB PSMs (58% increase), while the relative percentage of streptavidin PSMs decreased from 8.4% to 4.8%. [Table 2.2]

Table 2.2. Results comparison for DUB enrichment using old versus new protocols for ABP preparation.

	Old ABP preparation	New ABP preparation
Number of DUBs detected	46	61
Number of DUB PSMs	1219	1936
Total number of proteins identified	1253	1491
Total number of PSMs	6732	8378
Number of streptavidin PSMs	565	401

In the same experiment, I also compared pulldown systems other than biotin-streptavidin which I had been using thus far. Additional candidates included an anti-biotin antibody, monomeric avidin, and a FLAG epitope-based pulldown system. These three provided the advantage of specific elution from the enrichment reagent, with biotin for the anti-biotin antibody and monomeric avidin, and with TFA in the case of the FLAG pulldown system.

Table 2.3. Results comparison for DUB enrichment using various affinity enrichment systems.

	Magnetic streptavidin	Monoavidin	Anti-biotin antibody	FLAG pulldown
Number of DUBs detected	61	43	47	67
Number of DUB PSMs	1936	656	826	2360
Total number of proteins identified	1491	2042	1878	1688
Total number of PSMs	8378	9461	9214	9236
Number of streptavidin PSMs	401	28	N/A	N/A

We quickly eliminated the monoavidin and anti-biotin antibody options due to vastly inferior DUB enrichment, both by number of DUBs detected and relative specificity of the enrichment.

[Table 2.3] While the FLAG pulldown appeared to capture a few more DUBs than the magnetic streptavidin condition, upon closer inspection, a number of the additional DUBs uniquely detected in the FLAG condition were zinc metalloprotease JAMM DUBs, reducing the difference between the two. In the end, we decided to proceed with the magnetic streptavidin enrichment system due to its commercial accessibility.

Another hypothesis surrounded the denaturation step before streptavidin pulldown. I suspected denaturation might lead to an increase in nonspecific interactions between non-biotinylated proteins and the solid bead support, leading to enhanced background. At the suggestion of a colleague, I also investigated whether pre-washing of the beads for three times in 0.2% SDS before addition to the biotinylated sample might improve enrichment specificity.

Table 2.4. Results comparison for DUB enrichment with and without a denaturation before ABP addition, and with/without pre-washing the magnetic streptavidin beads before addition.

	Denatured Pre-washed	No denature Pre-wash	Denatured No pre-wash	No denature No pre-wash
Number of DUBs detected	64	56	64	59
Number of DUB PSMs	1506	1998	1839	2482
Total number of proteins identified	785	750	568	777
Total number of PSMs	4711	6250	4218	7554
Number of streptavidin PSMs	289	251	320	277

Removing the denaturation step before addition of ABP improved the specificity of the enrichment for cysteine protease DUBs. [Table 2.4] While 4 zinc metalloprotease JAMM DUBs were identified in the denatured, no pre-wash condition, none were identified in either of the conditions without the denaturation step. The pre-wash step seemed reduce DUB signal somewhat. We decided to proceed with the no denature, no pre-wash condition.

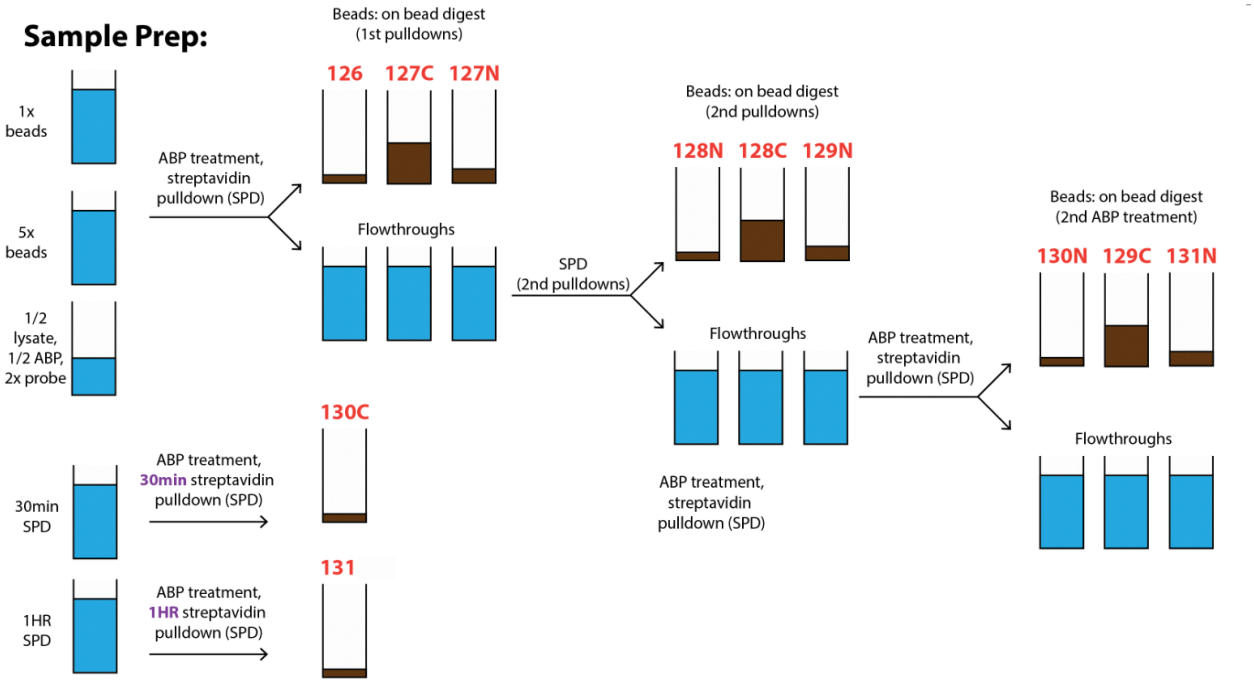
Over the course of the optimization, we were able to increase the number of DUBs detected in a single run from 39 to 60 DUBs (50% increase), reduce the relative percentage of streptavidin PSMs from 16% to 3.6% of the total, while increasing the relative percentage of DUB PSMs from 16% to 33% of the total. Thus, we were able to greatly improve the DUB coverage of the assay while improving on the specificity of the enrichment as well.

2.5. Sequential pulldown for interrogating quantitative DUB enrichment

While the pilot screen worked well, a central cluster of questions surrounding ABP labelling and pulldown efficiency were left unanswered. Are we capturing all DUBs which could be captured by the probes at current treatment concentrations? Is the capacity of streptavidin beads sufficient for capturing all ABP-labelled DUBs? How do the dynamics of DUB probe binding and inhibitor competition work out in the case of partial labelling? The last question, about the equilibrium and kinetics of DUB probe binding in a partial labelling scenario, would be particularly difficult to answer. As such, we decided to circumvent this issue by ascertaining that we were indeed capturing all DUBs quantitatively in this assay.

To answer these questions, I devised a sequential pulldown experiment, whereby I treated HEK293T lysates with ABP, then performed streptavidin pulldown with 25μL or 125μL of

magnetic streptavidin beads. Next I took the flowthroughs and carried out another round of pulldown with 25µL or 125µL of beads. Finally, I treated the flowthroughs with ABP for a second time and performed one final streptavidin pulldown. Another set three with two sequential pulldowns and leftover beads was prepared with half the HEK293T input. Each set of pulldowns were digested on-bead and labelled with TMT to enable relative quantitation across across 9 channels. To make-up an 11-plex experiment, two additional channels were prepared with different incubation times for streptavidin pulldowns: instead of two hours, one was incubated for 30 minutes while the other for 1 hour only. Since this was a TMT-assisted multiplexed experiment, proteins were quantified by total TMT reporter ion signal.



(continued from previous page) Figure 2.4. Sequential pulldown experiment. An 11-plex experiment was planned to explore streptavidin capacity, quantitative ABP labelling, feasibility of scaling down total lysate input, and reducing incubation times for the streptavidin pulldown.

2.5.1 Streptavidin Capacity

The two sequential pulldowns after the first ABP treatment was for probing streptavidin bead capacity. If the volume of beads was sufficient to capture all ABP-labelled species, we would expect the first pulldown after the first ABP treatment (126, 127C) to contain much higher DUB signal than the second pulldown (128N, 128C). The second ABP treatment was designed to explore whether 2μM of ABP was indeed sufficient for quantitative DUB labelling. If one round of ABP labelling was sufficient to capture all DUBs, little DUB signal would be expected after the second ABP treatment (130N, 129C).

As shown in tables 2.5 and 2.6, results indicated that 25μL was insufficient for capturing ABP-labelled species. Comparable DUB signal was observed in the first and second pulldowns. In contrast, 85% of total DUB signal was observed in the first round of pulldowns with 125μL of magnetic streptavidin bead slurry, further supporting this is an issue with bead capacity. The third pulldown, occurring after a second round of ABP treatment, failed to capture DUBs to any significant extent. Together, the results indicate that 2μM of ABP was indeed sufficient for quantitative DUB labelling, and either two sequential pulldowns with 125μL of magnetic streptavidin bead slurry or one single pulldown with 25μL magnetic streptavidin bead slurry. Results were consistent across two replicates of this sequential pulldown experiment.

Table 2.5. Results for the sequential pulldown experiment, 25μL beads.

	25mL beads, 1st pulldown (127C)	25mL beads, 2nd pulldown (128C)	25mL beads, 2nd ABP treatment (129C)
Total Signal	3.8E+09	2.8E+09	3.7E+09
DUB Signal	2.1E+09	1.6E+09	3.1E+08
% of total DUB Signal	51%	41%	8%

Table 2.6. Results for the sequential pulldown experiment, 125μL beads.

	125mL beads, 1st pulldown (126)	125mL beads, 2nd pulldown (128N)	125mL beads, 2nd ABP treatment (130N)
Total Signal	1.1E+10	2.7E+09	3.7E+09
DUB Signal	4.8E+09	3.7E+08	3.1E+08
% of total DUB Signal	84%	6%	10%

2.5.2 Scaling down the assay

The condition with half the HEK293T input (127N, 129N) was designed to investigate whether enrichment efficiency would stay consistent when input was scaled down. If enrichment efficiency stays constant, signal is expected to decrease by two-fold from the conditions where the regular HEK293T input was used (126, 128N, 130N), in proportion with HEK293T input. Surprisingly, reducing HEK293T input by half results in a 10-fold diminishment in total DUB signal. [Table 2.7] Hence, we came to conclude that we should not attempt to scale down the input of the assay.

Table 2.7. Results for the sequential pulldown experiment, half HEK293T input.

	Full input, 25mL beads, 1st pulldown (127C)	Half input, 25mL beads, 1st pulldown (127N)	Full input, 25mL beads, 1st pulldown (128C)	Half input, 25mL beads, 2nd pulldown (129N)
Total Signal	3.8E+09	4.9E+08	2.8E+09	7.1E+08
DUB Signal	2.1E+09	2.0E+08	1.6E+09	2.1E+08
% of total DUB Signal	51%	43%	41%	47%

2.5.3 Incubation times for streptavidin pulldown

The conditions with reduced streptavidin incubation time (130C, 131) were designed to investigate whether enrichment efficiency would stay consistent with reduced incubation times. If enrichment efficiency stays constant, signal was expected to be the same as the condition where 2 hours was used as the incubation time (126). Surprisingly, reducing incubation time with the streptavidin beads led to an improvement in DUB signal. [Table 2.8]

Table 2.8. Results for testing different incubation times for streptavidin pulldown.

	2 hours (126)	30 minutes (130C)	1 hour (131)
Total Signal	3.8E+09	3.6E+09	5.2E+09
DUB Signal	2.1E+09	2.3E+09	2.6E+09
% DUB in total signal	53%	63%	51%

2.5.4 Titrating down TMT

Finally, to save costs in the final large-scale screen, I titrated down the volume of TMT reagents used to label the DUB peptides. As shown in Table 2.9, I identified the minimum amount of TMT reagent needed for >95% labelling of my DUB peptides to be between 4 and 8μL (20μg/μL). To allow for a comfortable margin for sample-to-sample variation, I determined 10μL of TMT reagent, corresponding to ¼ unit as sold by Thermo Fisher, to be an appropriate amount for labelling in the screen. Compared to the previous protocol, this saved TMT reagent expenditure by 33%, to approximately 300 USD per run.

Table 2.9. Results for the sequential pulldown experiment, half HEK293T input.

TMT reagent used (μL)	% Labelled
Acetonitrile only	0
2	48
4	96
8	98
12	99
16	99
20	99

2.6. The final screening assay

Taking the optimization results together, the final screening assay would be carried out as follows. 200uL of 10mg/mL HEK293T lysates would be treated with inhibitor or DMSO, then 2uM of ABP. No denaturation step would follow, and the assay would proceed directly to two

cycles of sequential streptavidin pulldown with 25μL bead slurry and 30 minutes incubation time each.

2.7. Conclusions

In this chapter, I discussed the vision, first pilot run, and subsequent optimization for our library versus library screening assay. This worked put down the foundation for our full library versus library screen.

Following our vision to revolutionize DUB inhibitor discovery, our approach represented a marked deviation from the standard paradigms of small molecule drug discovery. Our target-class approach is embodied by the target class tailored library and target class-wide primary screening assay. Similarly, our assay optimization process focused on width of scope for the primary screening assay in terms of the number of DUBs detected. As mass spectrometry ABPP had never been applied for small molecule primary screening, we also set out to ascertain quantitative labelling and extraction of all DUBs in our inputs.

We found carrying out the pilot screen to be an extremely informative exercise. First, preliminary results on compound activity validated our library design strategy and indicated that screening the full library would be a fruitful exercise. The multiple runs involved in the pilot screen also revealed vulnerabilities in the assay; subsequent optimization improves assay performance and hence the quality of the output dataset from the full screen. Hence, we would recommend any researchers interested in running similar campaigns in the future to start with a small-scale pilot run to inform on both library design and assay performance.

3. Library versus library screen

In this chapter, I discuss efforts in leveraging our first-in-kind DUB-targeted covalent compound library and integrated platform of novel and established assays to achieve target-class inhibitor discovery against the DUBs. We carried out a data-dense primary screen with 178 DUB-focused library compounds against 54 native cellular DUBs by the ABPP screening assay described in the previous chapter. Hits were identified against 38 cellular DUBs spanning four subfamilies, then comprehensively triaged for biochemical inhibition, target engagement, hyperreactivity, and proteomic selectivity against a pipeline of biochemical, cellular and chemoproteomic assays.

3.1. Primary Screening

Primary screening was carried out over a series of twenty-five 11-plex TMT runs. [Figure 3.1] On average, 59 DUBs were detected in each run, 79% of cysteine proteases DUBs expressed in HEK293T cells.[81] [Figure 3.2a] We observed that 60% (101 of 178) of the compounds were active, defined as blocking ≥ 50% of ABP binding for at least a single DUB. [Figure 3.2b] 50% was selected for our cutoff to capture compounds with in-lysate IC50 values below 50μM.

Encouragingly, both chemical diversity and target coverage of the 101 active compounds were broad; they spanned 12 electrophiles, with 38 DUBs across five subfamilies bound by at least one compound [Figure 3.2b-c]. No hits were identified for the MINDY subfamily, suggesting that family might have different requirements for compound design. As anticipated for a first-generation library, we observed a range of selectivity for our compounds, from selective for one DUB to targeting 20% of all detected DUBs. [Figure 3.2b]

We observed hits against 69% (45 out of 65) of DUBs detected, with consistent hit rates across five subfamilies (USP, UCH, OTU, MJD, ZUP1). [Figure 3.2d] Of the hit DUBs, 84% (38 out of 45) have no validated inhibitors. [Figure 3.2c] As anticipated, targets of foundational chemotype F22 (JOSD1, UCHL1, USP30) appeared more prone to inhibition by our library compounds, supporting even more radical diversification in the next generation. [Figure 3.2a]

Preliminary examination of hits revealed series of compounds where changes to building block, linker, and warhead led to diversification of target landscape. [Figure 3.2e] Here, we discuss one such series. F22 is a multitargeted compound that targets UCHL1, USP32, and a few more targets at 50μM. Changing F22's imidazole linker to a more flexible alkyl linker leads to USP28 inhibitor WH188. By holding the N-cyano electrophile constant but shrinking the pyrrolidine ring to an azetidine, XL005C with improved UCHL1 selectivity results. From XL005C, shortening the linker and swapping to an alpha-chloroacetamide warhead yields WH103C selective for VCPIP1. Expanding the electrophile-bearing azetidine ring in WH103C to an indoline then leads to WH119C, which hits USP48. These series of compounds diversified by structure and activity supports the fruitfulness of our library design approach.

As initial validation, we tested roughly a quarter (24) of active compounds in 3-point dose response. [Figure 3.2f] 20 out of 60 selective (<3 targets) hits representative of most chemical and target diversity were supplemented with 4 less potent hits to diversify target scope. 75% (18 out of 24) hits had at least one DUB target confirmed in dose response, encompassing 15 DUBs in target scope, including 9 DUBs for which there are no known inhibitors (JOSD1, JOSD2, BAP1, UCHL3, USP3, USP16, USP22, USP48, VCPIP1).

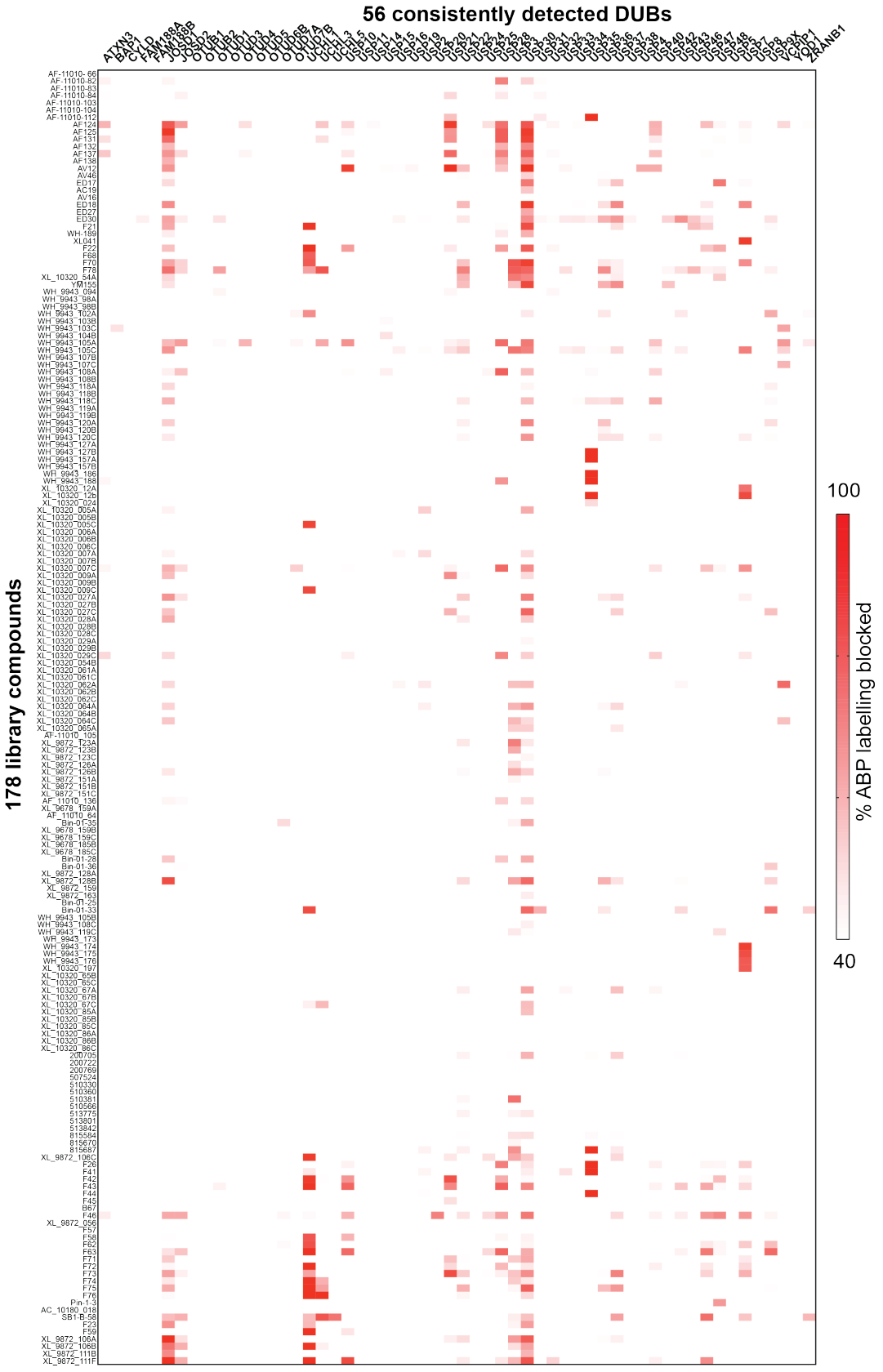


Figure 3.1. Primary screen results.

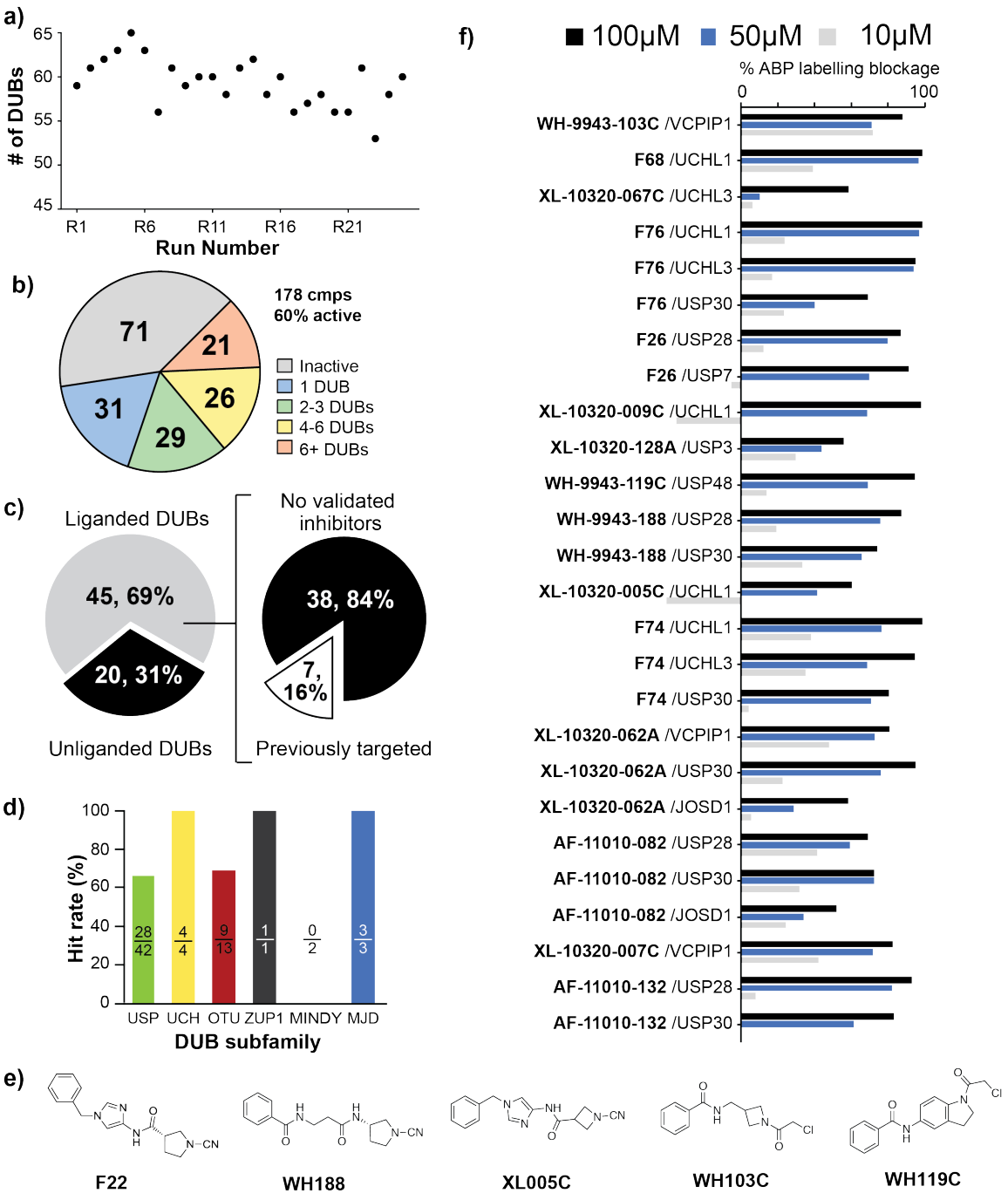


Figure 3.2. Summary of the primary screen. a) On average, 59 DUBs were detected in each run of the MS-ABPP assay. b) Of the 178 library compounds, 60% were active and displayed a wide range of DUB selectivities. c) Of the 65 detected DUBs, 45 were hit by at least one compound, while 38 out of 45 lack a validated reported inhibitor. d) Hit rates were relatively consistent

(continued from previous page) over the DUB subfamilies, expect for the MINDY family against which we had no hits. e) Hits showed clear SAR trends, with each part of the molecule impacting DUB inhibitory scope. f) Selected hits were re-screened over a three-point dose curve for initial validation.

3.2. Hit Validation/Characterization

We comprehensively validated hit compounds through a series of biochemical and cellular assays, for biochemical inhibition, covalent bond formation, site of covalent modification, and proteome-wide selectivity. In total, we have validated 24 compounds targeting 10 DUBs across the four most well-studied subfamilies.

First, we examined whether probe labelling blockage (or lack thereof) observed in ABPP is consistent with biochemical inhibition using recombinant DUBs. 24 compounds corresponding to those in dose response ABPP were profiled against 7 purified DUBs across 4 subfamilies (UCHL1, UCHL3, USP7, USP28, USP48, JOSD1, VCPIP1), selected to match hit target landscapes and maximize DUB diversity (Table 3.1). For 85% (142 out of 168) of DUB/compound pairs, biochemical inhibition or lack thereof was consistent with ABPP primary screening, using a cutoff of IC_{50}=20μM for biochemical inhibition. Potency of the most selective hits are in the high nanomolar to single-digit micromolar range, on par with most currently published DUB inhibitors. For example, WH-9943-103C inhibited VCPIP1 at IC_{50} of 500nM. Biochemical selectivity of hits was further examined against a panel of 41 purified DUBs in single-point dose, using Ubiquigent's DUBprofiler service.

Table 3.1. Biochemical validation of ABPP screen hits against 6 DUBs

Hit	ABPP Target	Biochemistry IC50 (μM)						
		VCPIP	UCHL1	UCHL3	USP7	USP28	USP48	JOSD1
WH_9943_107C	VCPIP1	1.2	>100	>100	>100	>100	>100	45.8
WH_9943_103C	VCPIP1	2.0	>100	>100	>100	>100	>100	44.1
XL_10320_062A	VCPIP1	3.0	>100	>100	>100	>100	>100	81.0
XL_10320_009C	UCHL1	>100	6.6	>100	>100	62.0	87.2	78.7
XL_10320_005C	UCHL1	33.7	3.5	42.6	>100	>100	>100	>100
F68	UCHL1	>100	3.4	>100	>100	69.1	78.0	38.6
F59	UCHL1/USP10	>100	0.5	>100	>100	>100	>100	>100
XL_10320_67C	UCHL3	14.9	18.1	6.3	>100	>100	>100	>100
F76	UCHL3/1	>100	0.1	1.0	>100	41.2	81.6	31.5
WH_9943_119C	USP48	>100	>100	>100	>100	>100	9.2	29.0
AC19	USP27X	>100	>100	58.4	29.5	40.1	8.6	15.2
F26	USP7/USP28	>100	>100	92.0	66.9	18.5	>100	9.2
WH_9943_188	USP28	>100	>100	>100	>100	17.5	>100	12.3
AF132	USP28	>100	>100	59.0	>100	27.9	>100	3.7
815687	USP3	>100	>100	>100	>100	59.8	>100	75.2
510381	USP3	>100	>100	>100	>100	>100	>100	>100
XL_9872_163	USP3	>100	>100	>100	>100	>100	>100	>100
XL_9872_126A	USP3	41.8	>100	>100	59.4	80.8	>100	43.5
XL_9872_123A	USP3	>100	>100	>100	>100	>100	94.5	48.4
AF_11010_112	USP21/31	>100	>100	>100	>100	>100	>100	>100
F45	USP21	>100	81.0	>100	>100	>100	>100	>100
F46	Multitargeted	>100	9.9	20.1	25.8	8.1	41.5	4.9
AF124	Multitargeted	12.3	10.0	8.6	22.5	4.4	>100	0.9
XL-9872-111B	JOSD1, USP30	14.5	0.3	7.8	5.7	5.3	21.1	2.6

We further confirmed target engagement of each hit in HEK lysate by ABPP coupled to an orthogonal Western blot readout, probed for six of the DUBs from the biochemical panel. JOSD1 could not be probed due to lack of a high-fidelity antibody. Inhibition/lack thereof for

95% (137 out of 144) of DUB/compound pairs were consistent across WB-ABPP and MS-ABPP, with 50% probe blockage as a cutoff, judged from the intensity of the +10kDa band on ImageJ.

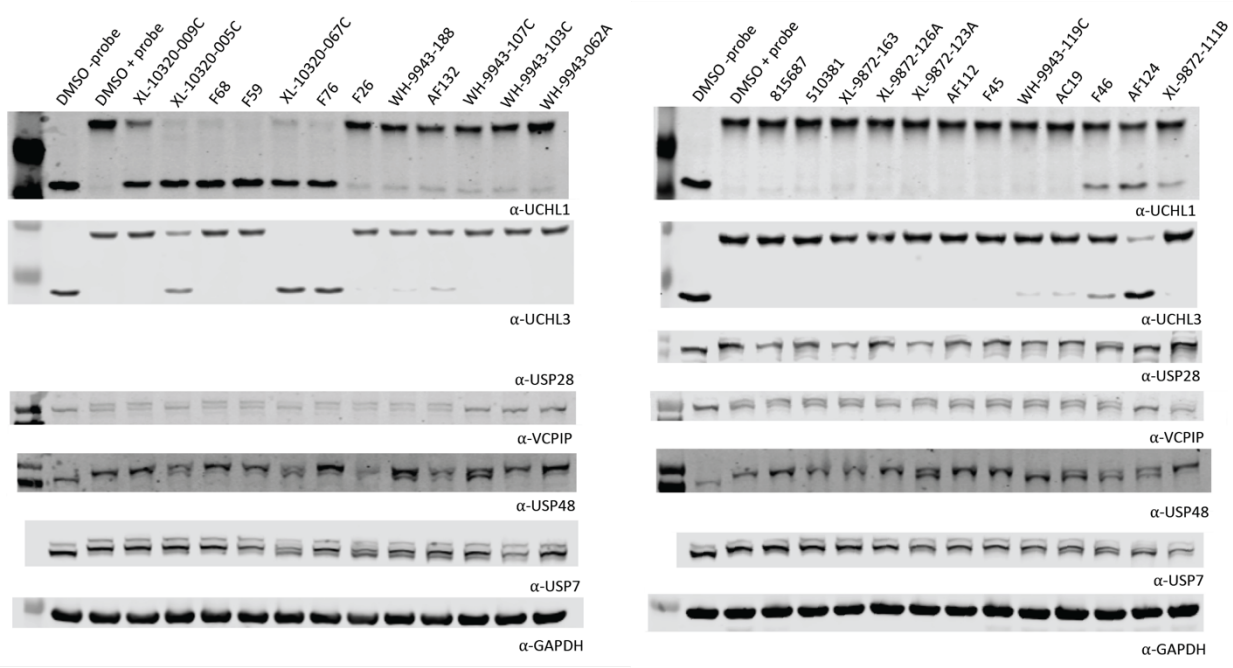


Figure 3.3. Western blot target engagement. Live HEK293T cells were treated with 50μM compound, then incubated with DUB ABP. ABP binding to the target DUB leads to a 10kDa shift in molecular weight, which was visualized on a Western blot. Inhibitor target engagement to the DUB catalytic cysteine block ABP binding, leading to disappearance of the +10kDa band. GAPDH was blotted for a loading standard.

Hits successfully validated in biochemistry were carried forward for further validation in intact protein mass spectrometry. We were able to confirm 100% 1:1 covalent labelling for the majority of examined hits, although some compounds showed partial labelling. [Figure 3.4a-c] This compares favorably to previous studies whose hits mostly only partially labelled protein in vitro.[61] Importantly, one compound displayed hyperreactive overlabelling of the target DUB,

allowing agile triaging of the hit which displays undesirable liabilities. We then took the compound-DUB reaction mixtures and digested to identify labelling site by LC-MS: as expected, the compounds targeted the catalytic cysteine residues in each DUB. [Table 3.2]

Overall, 16 hits against 11 DUBs were successfully validated in at least two orthogonal assays across biochemical inhibition, 1:1 covalent bond formation, and in-cell target engagement. Exemplar data for full hit validation is shown for UCHL1 hit F68. [Figure 3.5a-e] There were very few outliers: a notable example is XL-9872-111B, selective against JOSD1 in ABPP, showed multitargeted behavior in biochemistry, pointing at assay interference. Below, I will discuss hit trends surrounding some target DUBs of interest.

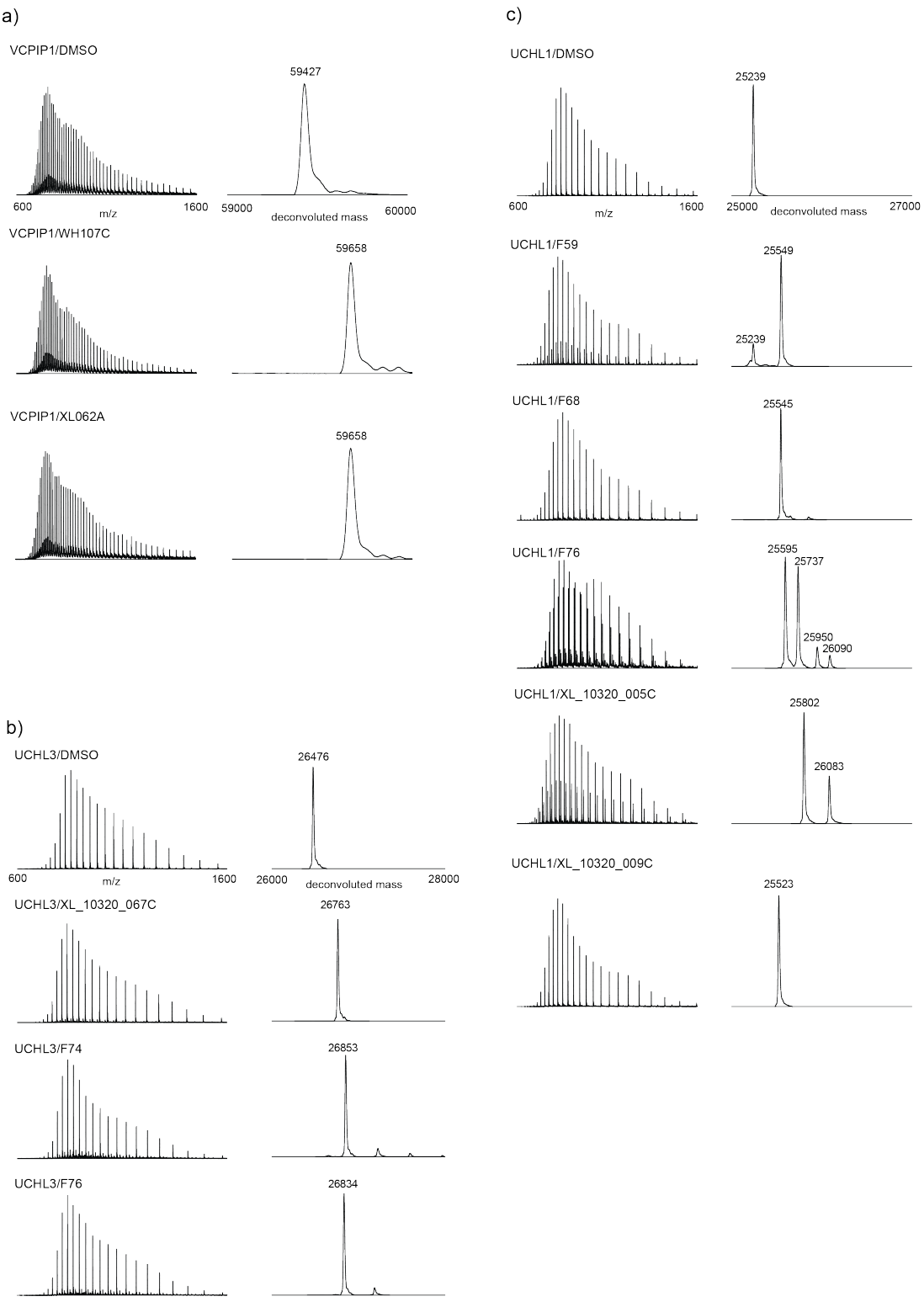


Figure 3.4. Intact protein MS validation. Purified a) VCPIP1, b) UCHL3, and c) UCHL1 protein were incubated with a 10x molar excess of compound and analyzed by LC-MS.

Table 3.2. Peptide-level CE-MS for selected DUB/compound pairs.

DUB	Compound	Cysteine(s) labelled
UCHL1	F59	**C90**
	F68	**C90**
	F76	**C90**, C132
	XL-10320-005C	**C90**, in addition to 5 other Cys sites
	XL-10320-009C	**C90**
UCHL3	XL-10320-067C	**C95**, C209
	F74	**C95**, C209
	F76	**C95**, C209
VCPIP1	WH-9943-107C	**C219** in addition to 9 other Cys sites
	XL-10320-062A	**C219** in addition to 7 other Cys sites
USP48	WH-9943-119C	**C98**

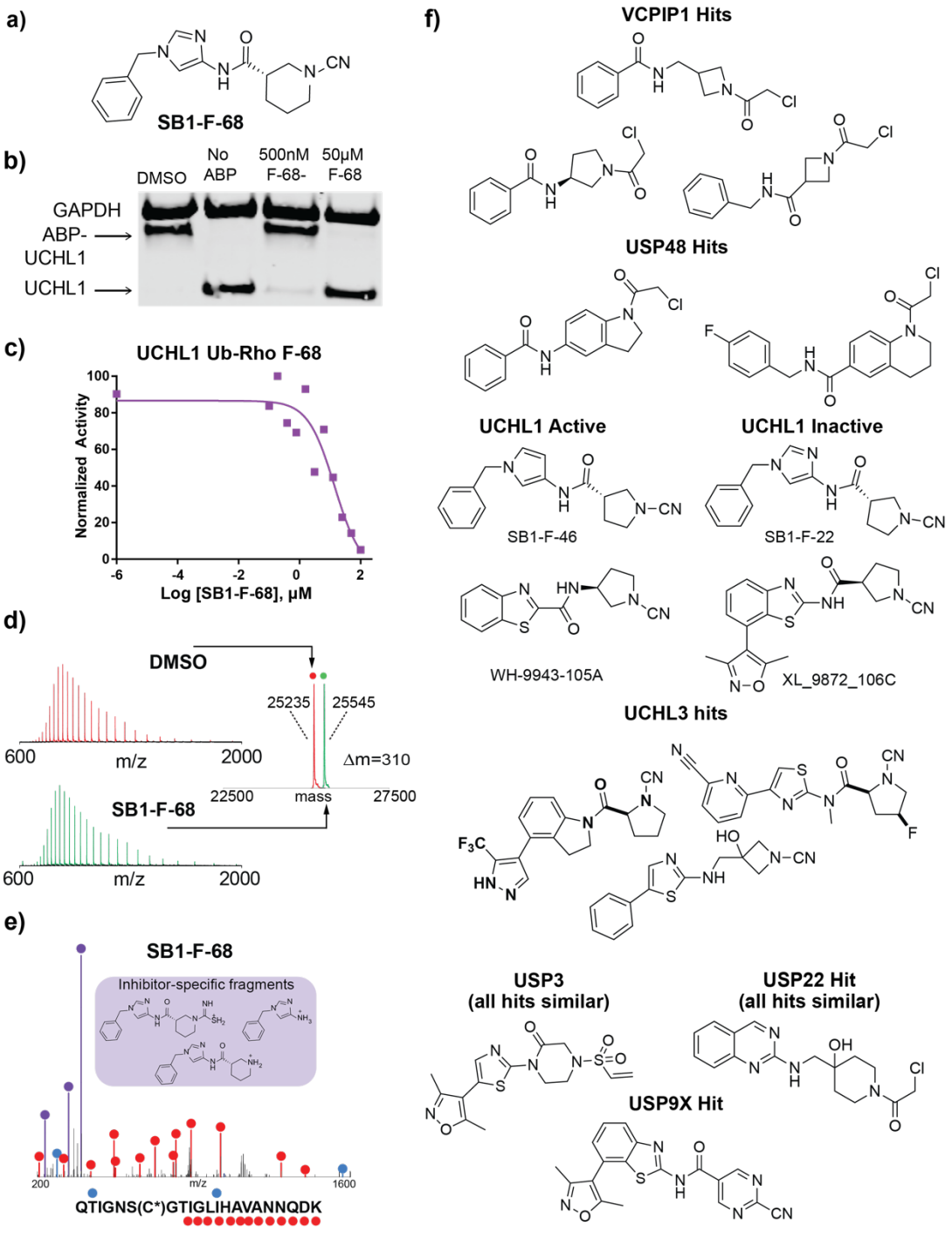


Figure 3.5. Hit validation and trends. a) Structure of UCHL1 hit F68. b) In-cell target engagement with F68 against UCHL1. c) F68 inhibited the catalytic activity of purified UCHL1 against fluorescent substrate ubiquitin-rhodamine with an IC_{50} of approximate 1μM. d-e) Intact

(continued from previous page) protein mass analysis revealed 1:1 labelling of purified UCHL1 protein; subsequent digest and CE-MS indicated modification occurred at the catalytic cysteine residue. f) Hits against the same DUB showed structural similarity, increasing confidence.

3.2.1 VCPIP1 Hits

One striking cluster of compounds hit VCPIP1: WH-9943-103C, WH-9943-107C, and XL-10320-062A. [Figure 3.5f] The three all possess an N-azetidinyl chloroacetamide warhead, and only differ by placement of a methylene group along the compound backbone. From primary screening, WH-9943-103C appeared to be slightly more selective than WH-9943-107C and XL-10320-062A.

WH-9943-103C, WH-9943-107C, and XL-10320-062A inhibited the catalytic activity of target VCPIP1 in a dose-dependent fashion with IC_{50} values in the low single-digit micromolar range. [data not shown] By incubating purified enzyme with each compound and LC-MS analysis, we observed complete 1:1 labelling of VCPIP1 with WH-9943-103C. We proceeded to identify the labelled cysteine residue by trypsin digestion and capillary electrophoresis mass spectrometry (CE-MS). WH-9943-103C was identified to react with the catalytic C219 on VCPIP1. As reported in a previous publication, we were also able to identify the thiolated ions formed from dissociation of the cysteine-inhibitor adduct in both cases, increasing confidence in labelled peptide identification.

Interestingly, both WH-9943-107C and XL-10320-062A reacted with multiple cysteines indiscriminately over purified VCPIP1, suggesting possible hyperreactivity. This is consistent

with multitargeted behavior of the two hits in the primary screening assay. This motivated us to drop the compounds from further consideration. In a confirmation for live cell target engagement, WH-9943-103C successfully blocked ABP labelling of the target DUB in live cells with an orthogonal Western blot readout. Consistent with ABPP primary screening data, WH-9943-103C continues to display exquisite selectivity for VCPIP1 in the Ubiquigent assay panel, with little to no inhibition of any DUBs other than VCPIP1.

Cognizant of off-target liabilities, we investigated the proteome-wide selectivity of hits WH-9943-103C. WH-9943-103C significantly labelled (compound-treated/DMSO ratio > 4) only 2 out of 28034 cysteine residues captured in a cysteine-profiling experiment (Figure 3.6a). While we did not observe the VCPIP1 catalytic cysteine residue, the top ranking hits were previously reported as hyperreactive cysteine residues.

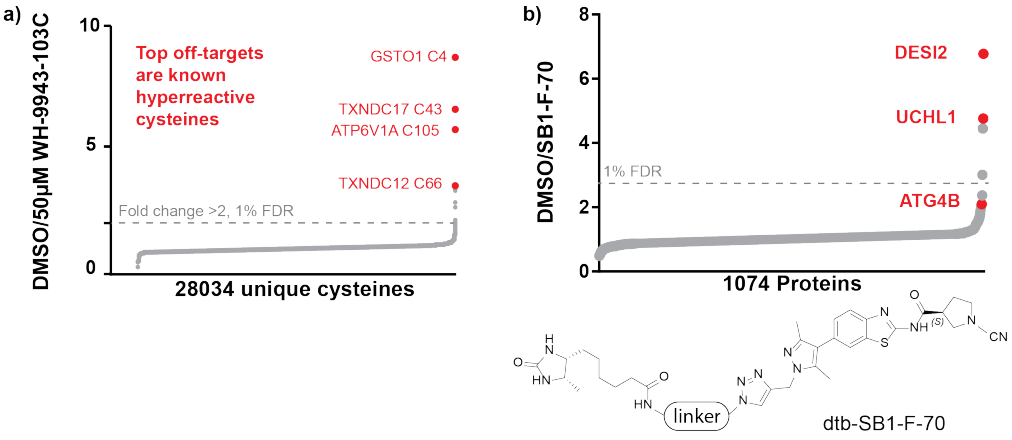


Figure 3.6. Proteome-wide hit validation. a) A proteome-wide cysteine profiling experiment revealed few off-targets for the DUBs. b) Competition with a biotinylated analog revealed selective binding to DUBs and ubiquitin-like proteases by N-cyanopyrrolidine compound F-70.

3.2.2 USP48 Hits

Structurally similar hits against USP48 were also identified. [Figure 3.5f] They all possessed chloroacetamide electrophiles with bulky indoline or tetrahydroquinoline pendant rings, again increasing our confidence that these are indeed real hits. Hit WH-9943-119C was successfully validated for biochemical inhibition and labelling at the catalytic cysteine residue.

A separate USP48 hit was identified in Pin-1-3, a covalent inhibitor for the prolyl isomerase Pin1. Another ABPP run using an analog with the chloroacetamide replaced by a methylamide confirmed the electrophilic warhead to be necessary for activity, increasing confidence in the hit. Further, structurally analogous fragments were reported to bind to C48 on USP48 in a electrophilic fragment screening study from the Gygi Lab, further corroborating this compound's activity against USP48. USP48 hits are being actively optimized by Isabella Jaen Maisonet in the Buhrlage Lab.

3.2.3 UCHL1 Hits

Foundational chemotype F22 had UCHL1 among one of its top targets; hence, it came as little surprise that we obtained a cluster of UCHL1 hits, each containing N-cyano warheads. In this screening effort, we observed additional SAR which informs future inhibitor development against UCHL1 or other DUBs. [Figure 3.5f]

One of the nitrogen atoms on the imidazole linker was critical for UCHL1 inhibition. F22 and F46, which differed only in the presence/absence of this nitrogen atom, differed 400-fold in biochemical IC_{50} against UCHL1. Another determinant for UCHL1 inhibitory activity was amide

orientation at the warhead. A series of closely matched equivalents with the amide bonds at opposite orientations clearly showed that UCHL1 favored compounds with the carbonyl on the side of the warhead. It was also clear that UCHL1 inhibition is indiscriminate of stereochemistry at the beta-proline warhead. Various matched pairs of compound with opposite stereochemistry displayed equivalent inhibition.

Noting the commonality of the N-cyano pyrrolidine/azetidine warheads in the screen hits, we further investigated the proteome-wide selectivity of the scaffold. We synthesized a desthiobiotinylated analog of the N-cyanopyrrolidine compound F-70 and competed it against parent compound for pulldown in HEK lysate. Analysis confirmed UCHL1 as one of only 6 bound targets throughout the proteome. [Figure 3.6b] Interestingly, two off-target enzymes (DESI-2 and ATG4B) were recently reported to deconjugate ubiquitin-like proteins.[82,83] In aggregate with the primary screening data, we would like to nominate the N-cyanopyrrolidines as a potential privileged scaffold for DUBs and ubiquitin-like proteases.

Several selective UCHL1-targeting compounds with N-cyanopyrrolidine warheads were reported in the literature since the start of the study. There is also intense competition on UCHL1 inhibitor development from the biotech company Mission Therapeutics. As a result, we decided against further pursuing UCHL1 inhibitors.

3.2.4 UCHL3 Hits

Another cluster of two hits targeted UCHL3 in addition to UCHL1. [Figure 3.5f] Both possess a N-cyano warhead and a thiazole linker, suggesting that these are important determinants for targeting UCHL3. Outside of this commonality, the two had vastly different pendant rings: XL-

10320-067C had a hydroxyazetidine ring with a methylene spacer, while F76 had an fluoropyrrolidine ring attached the the nitrile electrophile.

While there are no compounds that selectively targeted UCHL3 over UCHL1, these are nonetheless significant results. One can imagine reversing the determinants for UCHL1 targeting above to engineer compounds that are inactive against UCHL1. An example would be to reverse the amide bond orientation at the warhead: as demonstrated by multitargeted compound WH-9943-105A, while having the carbonyl on the side of the warhead is strictly necessary for UCHL1 targeting, it is not necessary for UCHL3 targeting. Hence, we can imagine flipping the amide bond orientation in F76 in an attempt to remove UCHL1 inhibition.

3.2.5 USP10 Hits

Our group has previously identified USP10 as a potential therapeutic target in FLT3 mutant malignancies, due to its ability to selectively degrade the oncogenic mutant over wild-type protein.[21] USP10 also regulates various proteins along the nutrient sensing and Beclin-1 signaling, suggesting a central role in autophagy.[84,85] As such, USP10 is a DUB for which we are highly interested in developing inhibitors against.

While no library compounds selectively hit USP10 preferentially over other target, we had a wealth of multitargeted compounds that hit USP10, and selective compound which hit USP10 as a secondary target. [Figure 3.5f] Across the two categories, we observed paired inhibition of UCHL1 and USP10 in many compounds. In fact, the cleanest compound against USP10, F59, hit both UCHL1 and USP10. Nonetheless, we were encouraged by how USP10 does not have any of the structural requirements for inhibition as listed above for UCHL1; compounds WH-10417-

046A and F-46 both inhibit USP10 despite having a flipped amide orientation and lacking the necessary nitrogen atom on the linker respectively. This suggests that we might be able to disentangle USP10/UCHL1 inhibition in the future.

3.2.6 USP3, USP9X, USP22

Structurally related hit clusters were identified against USP22 and USP3. One further hit was identified against USP9X. [Figure 3.5f]

We encountered significant difficulty when attempting to validate these hits. While USP22 protein could be expressed and purified, it had no measurable catalytic activity. This finding is consistent with how USP22 has been reported to be activated by binding to the SAGA complex, which it is usually in complex with in the cell. The lack of activity against the Ub-Rho substrate precluded biochemical hit validation. The large size of the protein precluded intact protein mass analysis, while digestion experiments failed to identify the catalytic cysteine containing residue. The USP3 protein could not be expressed/purified at all, and no orthogonal validation could take place. Commercial USP9X protein was inactive in Ub-Rho biochemical assay, and its large size precluded target engagement validation. As a result of the inability to assay these proteins, we declined to follow-up on USP3, USP9X, and USP22 hits.

3.3. Target/target-class optimization of VCPIP1 hit

Following validation, the azetidinyl chloroacetamide scaffold became of interest for several reasons: (1) two azetidyl chloroacetamide compounds selectively targeted VCPIP1, a member of the hitherto untargeted OTU subfamily; (2) VCPIP1 was of interest for its reported roles in regulating ER structure and botulism neurotoxin stability;[86,87] (3) Our dataset nominated a small

set of additional hitherto untargeted DUBs, including BAP1, USP16 and USP40, as potentially selectively targetable by the chemotype [Figure 3.1]; (4) only six compounds from this series were contained in our initial library. We sought to further explore the chemotype by synthesizing and screening a novel 30-compound mini-library.

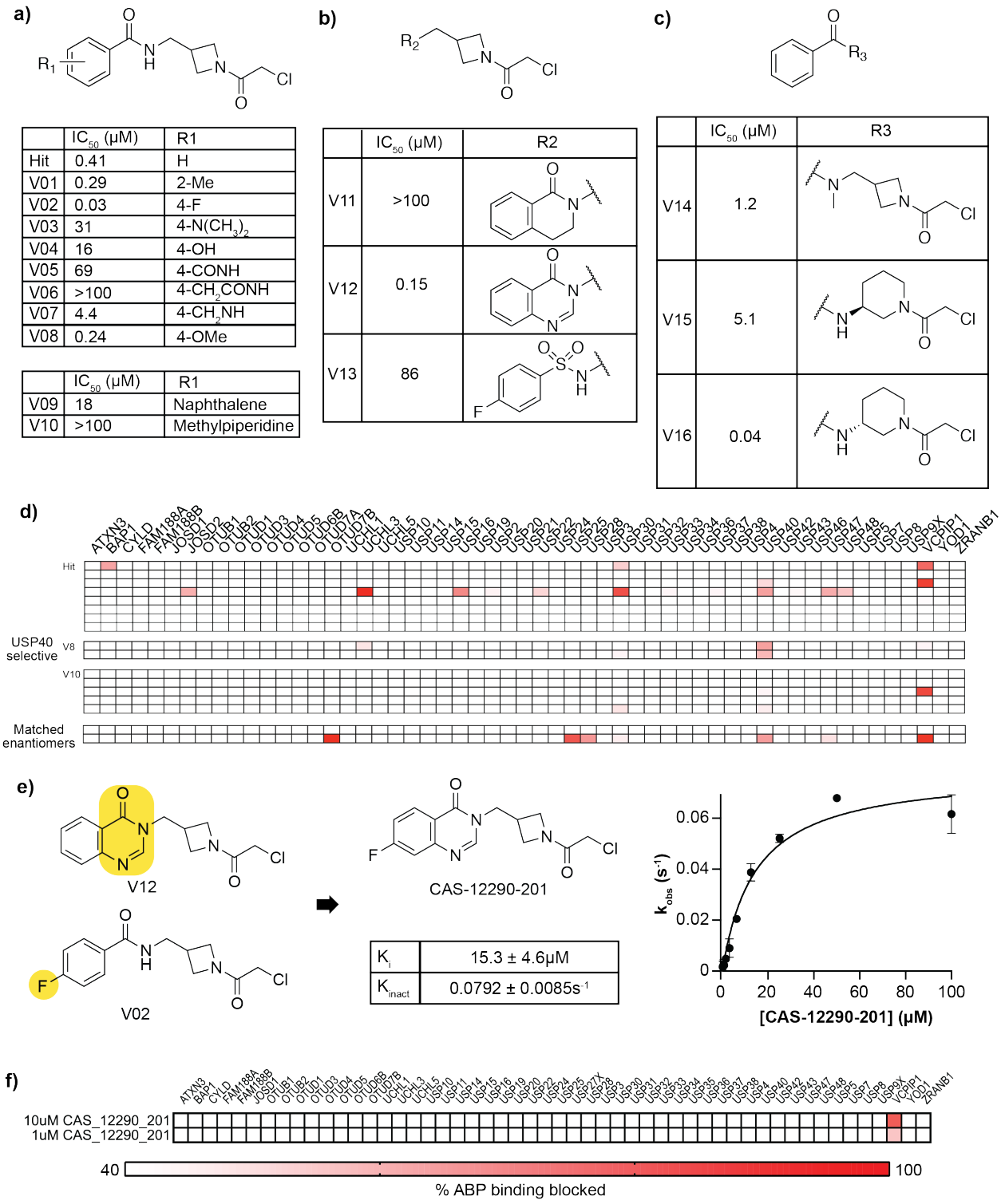
Our goals for the library were two-fold: develop a first-in-class probe for VCPIP1 and further establish the potential of the scaffold across the target class. With these in mind, we varied the noncovalent building block, linker, and electrophilic warhead. We varied the benzoic acid starting material in electronics and bulk by adding substituents or swapping to different ring systems. Rigidity and bulk of the methylene amide linker was variously incorporated into fused heterocycles or with the amide swapped for a sulfonamide. SAR insight from the primary screen informed design: the benzothiazole building block on multitargeted VCPIP1 hit WH-9943-105A inspired us to fuse the amide into a closed quinazoline. Ring size around the electrophilic warhead was changed as well. All library compounds were evaluated using a hybrid target/target-class approach that employs parallel evaluation of VCPIP1 inhibition in a biochemical assay and profiling using our ABPP-MS assay. Compounds with VCPIP1 biochemical IC_{50} values below 250nM were screened in ABPP at 10μM, otherwise 50μM.

3.3.1 A potent and selective first-in-class VCPIP1 inhibitor

On the benzene ring, small substituents such as methyl or halogen groups were well-tolerated, but larger substituents such as dimethylamines or methylamides obliterated activity. [Figure 3.7a-d] This is consistent with how swapping the benzene with naphthalene led to loss of

activity. We found the methylene spacer on the linker to be essential for activity, as changes to incorporate the methylene into more rigid ring systems led to dramatic loss in activity.

Two specific changes on different parts of the molecule led to improvements in potency towards VCPIP1. First, addition of halogen substituents to the benzoic acid noncovalent building block, such as in the fluorophenyl derivative CAS-11478-188, or difluorophenyl derivative XL-11478-092D greatly improved potency. Integrating the amide into a closed fused bicyclic system, as in the case of quinazolinone compounds CAS-12290-201 or indolinone AYA-01-045 improved both potency and selectivity.



	IC_{50} (μM)	R1
Hit	0.41	H
V01	0.29	2-Me
V02	0.03	4-F
V03	31	4-N(CH_3)$_2$
V04	16	4-OH
V05	69	4-CONH
V06	>100	4-CH_2CONH
V07	4.4	4-CH_2NH
V08	0.24	4-OMe

	IC_{50} (μM)	R1
V09	18	Naphthalene
V10	>100	Methylpiperidine

	IC_{50} (μM)	R2
V11	>100	
V12	0.15	
V13	86	

	IC_{50} (μM)	R3
V14	1.2	
V15	5.1	
V16	0.04	

K_i	15.3 ± 4.6μM
K_{inact}	0.0792 ± 0.0085s^{-1}

Figure 3.7. Target/target-class VCPIP1 inhibitor development. a-c) Biochemical IC_{50} values and structures for the azetidinyl chloroacetamide focused mini-library. d) ABPP results for the mini-library compounds. e) Productive changes were combined to yield CAS-12290-201, novel potent

(continued from previous page) and selective VCPIP1 inhibitor with kinetic characterization. f) CAS-12290-201 selectively inhibits VCPIP1 in the MS-ABPP assay.

We combined the two productive changes to yield fluoro-quinazolinone compound CAS-12290-201. CAS-12290-201 inhibited VCPIP1 with an IC_{50} of 70nM with a preincubation time of 6 hours [Data not shown]. Kinetic analysis revealed Ki = 15.3 ±4.6μM and kinact = 0.0792 ± 0.0085s^{-1}. [Figure 3.7e] In the ABPP-MS assay, CAS-12290-201 strongly blocked ABP labeling of VCPIP1 at two concentrations (10μM, 1 μM) tested while exhibiting little to no activity toward any other DUBs [Figure 3.7f]. Taken together, these data credential CAS-12290-201 as a first-in-class chemical probe for VCPIP1.

3.3.2 Additional targets of the scaffold

The library compounds confirmed BAP1, UCHL3, USP16, USP40, and OTUD7B as additional DUB targets of the chemical series, suggesting the series may be specific to a cluster of family members. [Figure 3.7d]

Change to the phenyl ring were found to be a key determinant of target spectrum, underscoring the importance of noncovalent interactions in DUB binding. For example, replacement of the phenyl ring with naphthalene led to an improvement in potency toward USP40 and dialing out VCPIP1 activity. While we could not validate this activity because of inability to recombinantly express and purify USP40, this would warrant future investigation.

Altering the methylene linker and pendant azetidine ring into a piperidine leads to targeting of OTUD7B in addition BAP1, USP28, and VCPIP1. This is particularly striking as this is the only

OTUD7B-targeting compound we have obtained thus far in screening. The tetrahedral carbon where the amide meets the piperidine rings is a stereocenter, providing opportunities for matched enantiomeric control compounds. Upon generating each of the enantiomers, we observe that the two differ 100-fold in biochemical potency against VCPIP1. [Figure 3.7c] ABPP confirmed the difference against VCPIP1 and for the rest of the targets. [Figure 3.7d] We anticipate that this piperidinyl chloroacetamide scaffold can be further diversified as a starting point for OTUD7B inhibitors.

In all, through our hybrid target/target-class approach, we were able to obtain a highly potent and selective VCPIP1 inhibitor, while gleaning insight for future projects based on the same scaffold.

3.4. Discussion

The novel application of mass spectrometry ABPP as a primary screening assay enables multipronged quantitative hit identification against multiple DUBs in one screen. While compound throughput is lower than a conventional target-based biochemical high-throughput screen, our advantage comes in early availability of biologically relevant selectivity information. In aggregate with the validation assays, both potency and selectivity information for 54 native DUBs can be identified in one screen.

This initial endeavor has already revealed a wealth of hits, valuable structure-activity relationships, and DUB inhibitor design clues. Even with a very limited number of 178 compounds, we already expanded the set of DUBs with covalent inhibitor starting points by 15.

This is especially impressive when compared to 5 DUBs with well inhibitors reported with both biochemical and cellular selectivity characterization data. Potency of the most selective hits are in the high nanomolar to single-digit micromolar range, which is on par with most currently published DUB inhibitors, even ones that have been put through extensive optimization.

Screening results confirmed the efficacy of a customized library focused on the DUB catalytic site, with the noncovalent building block, linker, and warhead all contributing to selectivity. Acrylamides, common in existing commercial fragment libraries, were ineffective against DUBs, exemplifying the importance of tailoring molecules to the target class. Sets of DUBs inhibited by each hit did not cluster by sequence phylogeny, supporting our intensive scaffold hopping to expand DUB inhibition scope, as exemplified by going from the foundational N-cyanopyrrolidines to the azetindine chloroacetamides.

During hit validation, it came to our attention that the N-cyano pyrrolidine warhead exhibits a very wide range of selectivities and DUB subfamilies depending on the noncovalent binding piece. The aforementioned XL005C and WH188 are selective inhibitors for a UCH subfamily and USP subfamily DUB respectively. In contrast, AF124 is a multitargeted compound which hits 13 DUBs across the UCH, USP, OTU, and MJD subfamilies. Myriad compounds occupy the selectivity spectrum between these two extremes, while our proteome-wide IBP-based exploration of F-70 target landscaped revealed activity against ubiquitin-like proteases DESI2 and ATG7B. Together, these data nominate the N-cyano pyrrolidine as a potential privileged warhead for targeting DUBs and other cysteine proteases with specificity for ubiquitin-like proteins.

This endeavor highlighted difficulties in studying particular DUBs using currently available assays. This is exemplified in our struggles to characterize compounds against USP3 and USP9X. USP3 did not express in *E.Coli* in our hands, could not be detected robustly in ABPP, was excluded in commercial DUB testing panels, and also lacked high fidelity antibodies for Western blot target engagement. This stresses the importance of future reagent and assay development to expand accessiblity of DUB targets for pharmacological discovery.

In the DUB-ABPP primary screening assay, some DUBs were more prone to inhibition than others as read out by the ABPP assay, namely USP30 and JOSD1. In the case of JOSD1, all compounds that targeted JOSD1 hold the N-cyanopyrrolidine warhead, suggesting that the DUB itself might possess a preference for the warhead. However, in the cases of USP30, no such patterns in hits were identified. Nonetheless, USP30 biochemical inhibition was indeed observed for compounds identified as "hits" in DUB ABPP, suggesting that this susceptibility was not an assay artifact. We speculate that perhaps USP30 is especially prone to covalent inhibition due to reactivity or placement of the catalytic cysteine residue.

As with all screens, there were some hits which failed to validate in orthogonal assays. Specifically, we encountered several discrepancies across the DUB ABPP screening assay and the purified enzyme biochemical assay. Instead of viewing this as a simple failure to validate, we speculate this might be due to a difference between purified enzyme systems and the endogenous lysate environment in which DUB ABPP is carried out. For example, USP22 exists in complex with the SAGA proteins in the cell, but biochemical studies were carried out with purified USP22 only.[88] Since the vast majority of DUB screening efforts take place by biochemical assays at the moment, this discrepancy raises questions regarding false negatives

in biochemical screening. Namely, these compounds are only “hits” by reversing the paradigm of using chemoproteomic assays for late-stage validation, since such compounds that fail to show efficacy in biochemical assays would have been triaged in a more orthodox target-based pipeline. The experience and wisdom needed to answer these questions require another generation of screening and deeper physiological insight on how DUBs operate.

As the screening effort continues and the well-annotated library grows, chemoinformatic approaches would become feasible, to start extracting SAR information from the complex patterns exhibited by the multitargeted compounds. While it is currently not feasible due to the small size of the initial screening library, this would be an important avenue to keep in mind for the future. Perfectly matching paired analogs would be useful to both confirm SAR trends and increasing the statistical power of such computation models.

In summary, we report a target-class approach for DUB inhibitor discovery, leveraging a first-in-kind DUB-targeted covalent compound library and an integrated platform of novel and established assays. We showed covalent targetability for a large swathe of the target class, and comprehensively validated hits for biochemical inhibition, target engagement, hyperreactivity, and proteomic selectivity.

4. Statistical analysis of chemoproteomic screening data

In the previous chapters, I have discussed our efforts to develop a chemoproteomic screening platform with an ABPP-based primary screen, and subsequent efforts to optimize resulting hits. While driving these efforts, questions relating to data processing and analysis repeatedly emerged, motivating us to take a closer look at data analysis in the specific context of chemoproteomic primary screening.

In this chapter, I first discuss our observations and challenges encountered in data analysis/processing for our screening effort. As a direct result, we developed a new approach we term individualized protein baseline extraction (INPROBE), a novel method for protein-by-protein statistical hit identification.

4.1. Observed challenges in chemoproteomic data analysis

Chemoproteomic screening techniques leverage reactive probes which biochemically capture a defined set of cellular targets, including amino acid residues (e.g. cysteine, lysine, tyrosine) or mechanistically related enzymes (e.g. serine hydrolases, cysteine proteases) for affinity enrichment. Small molecules in a screening library compete against the probe for target binding, leading to reduced enrichment, which is quantified by mass spectrometry. Quantified protein abundances are then compared across null (e.g. DMSO) and treatment conditions to obtain “competition ratios”, a general measure for inhibitor blockage of probe binding to each target species. [Figure 4.1a] As each compound is conventionally tested at a single dose without replicates, the final output of a chemoproteomic experiment is often a matrix of competition

ratios with detected protein/peptides as rows and compounds as columns. Our dataset discussed in the previous chapters exemplifies these datasets and workflow.

The need for statistical control arises naturally when dealing with high-density data, and chemoproteomics is no exception. Even with the same conditions, quantified abundances of each protein species are detected with a range of baseline variation in the mass spectrometer. Besides biological variations in expression, additional noise comes from differences in probe binding, affinity enrichment, TMT-labelling, and sample handling. Thus, any differences in quantified protein abundances due to inhibitor treatment must be distinguished from this baseline variation, leading to a thresholding problem.

To our surprise, during our screen, we found that there were no methods for statistical distinction of baseline variance v.s. inhibitor-derived differences in signal for chemoproteomic primary screening data. This is mostly due to the single dose, no replicate format of chemoproteomic screening. Common replicate-reliant data analysis methods such as the popular R *limma* package or *MSstats* package requires multiple biological and technical replicates for variance modelling, precluding their use in chemoproteomic primary screening contexts.

As discussed in the introduction, the existing two thresholding approaches were plagued with various problems. [Figure 4.1b] Declaring arbitrary integers as significance thresholds is statistically meaningless and precludes a common methodology for the field. Concatenating competition ratios into a single distribution relies on dubious assumptions that baseline distribution for each detected protein is similar, and that competition ratios are statistically comparable across protein species.

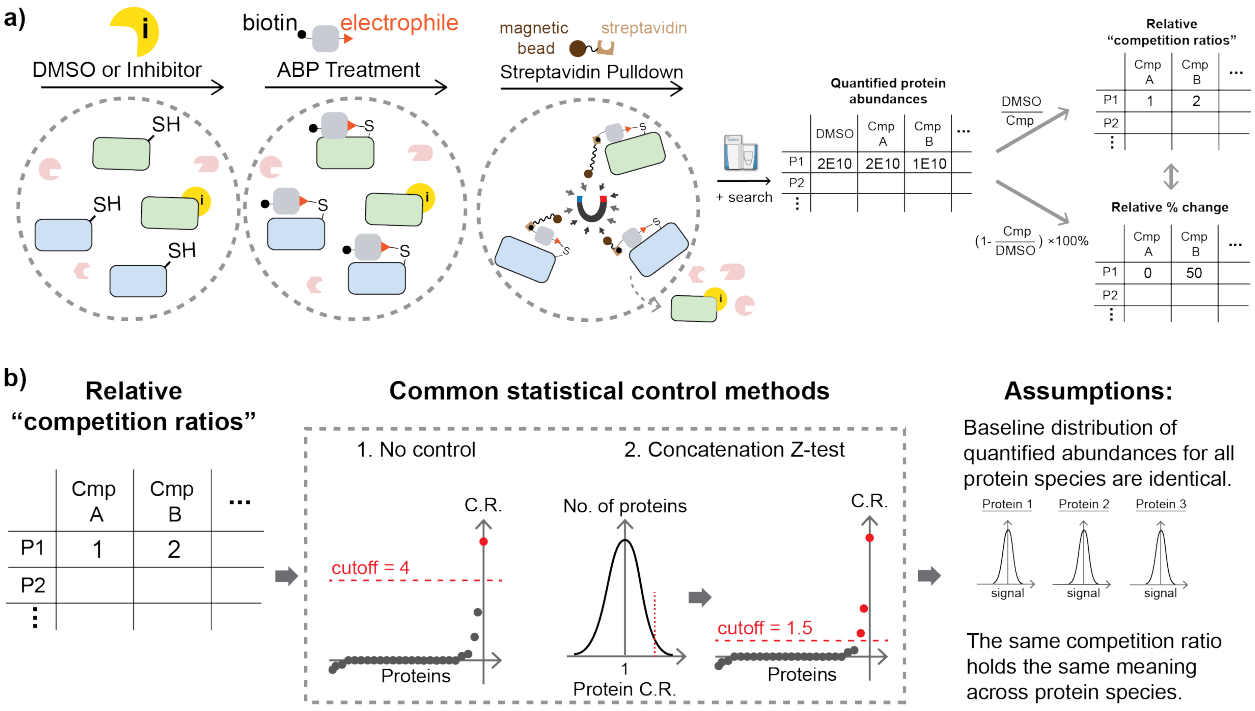


Figure 4.1. Existing methods for statistical analysis of chemoproteomic screening datasets. a) Data output of a chemoproteomic primary screen is in the form of a NxM matrix with N proteins (rows) and M conditions (columns), filled by values of quantified protein abundances for each protein in each condition. The abundances in each treatment condition can be compared to that in the DMSO condition by two measures, relative competition ratios and relative % change. The two are interconvertible. b) Currently, common statistical control measures for chemoproteomic screening start from relative competition ratios. However, both methods assume that the baseline distribution of quantified abundances of all protein species are identical, and that the same competition ratio holds the same meaning across protein species.

Limited by the accepted methodologies in the field, we selected the 50% competition threshold for hit identification in our DUB library versus library primary screen by a combination of these two approaches in initial screening described in Chapter 3. We first theorized that at our screening concentration of 50μM, a 50% blockage of ABP binding would translate to an in-lysate IC_{50} of 25μM, a reasonable starting point for medicinal chemistry efforts. Then, we validated this cutoff with a concatenation-based strategy, by confirming that a 50% cutoff does indeed correspond to an FDR of <1%.

While our hit validation rates were indeed impressive, suggesting a relatively well-performing threshold, we could not help but wonder about the cases in which DUB-inhibitor pairs failed to replicate/validate. These tend to be concentrated in DUBs with low signal in primary screening (JOSD2, USP2, USP20, USP21, USP3). On the opposite end of the spectrum, we also observed DUB-inhibitor pairs for which percentage blockage was below the 50% threshold in primary screening, but nonetheless displayed appreciable biochemical IC_{50} values in orthogonal validation. While it is tempting to brush these off as random irregularities in the data, we wondered if these problems were connected to the aforementioned deficiencies and unproven assumptions taken in the one-size-fit-all approach.

As such, I turned my attention to two questions about data analysis, namely: (1) are baseline variances indeed uniform and comparable across DUBs in primary screening as assumed in one-size-fit-all approaches? (2) Would considering differing variances across proteins improve hit identification in chemoproteomic primary screening?

To start answering these questions, we drew on our primary screening dataset of 25 11-plex ABPP replicates with 178 focused inhibitors against a collection of 54 cellular deubiquitinating enzymes (DUBs). Using the large number of null replicates in our screen, we challenged the assumption of statistical equivalence in current methods and propose an alternative method of statistical control. In this new approach we term individualized protein baseline extraction (INPROBE), baseline distributions are extracted protein-by-protein from null replicates, and signals in treatment conditions are compared against this null distribution for a more case-specific control. We compared screening results analyzed by INPROBE with that from conventional competition-ratio based methods and concluded that INPROBE gives more consistent results across replicates, higher sensitivity for high-abundance species, and fewer false positives for low-abundance species.

4.2. Statistical modelling of baseline variance

Instead of carrying out statistical control at the level of competition ratios as conventional methods do, we reasoned that TMT reporter ion intensities, a measure of detected peptide abundance, is a better point. To obtain baseline variation of the enrichment process without inhibitor competition, we focused our analysis on DMSO replicates. As each 11-plex run of the screen included DMSO runs, the basis of our analysis was 2 x 25 = 50 DMSO runs, across a set of 50 DUBs detected in >80% of runs.

To examine baseline variance without competition, we reasoned that we could model baseline distributions in quantified protein abundance for each enriched species as an individual normal distribution. This model could be used to estimate the probability of observing a given

competition ratio in a protein-inhibitor pair and evaluate the statistical significance. Unlike concatenation methods which treat all proteins as equivalent in baseline distribution, this analysis could be applied in a protein-by-protein fashion, allowing statistical significance thresholds to be customized to the distribution of each protein detected.

4.2.1 INPROBE procedure

We applied the following python workflow to our recently reported dataset of 25 11-plex ABPP replicates in the following steps. [Figure 4.2a] Raw TMT reporter ion intensities were summed across peptides for each protein, and DMSO negative control channels were aggregated into one dataframe. As each 11-plex experiment contained two DMSO null replicates, this resulted in a collection of 50 DMSO null runs. Next, TMT reporter ion intensities were normalized across the 50 runs by total summed DUB signal, and normalized abundances for each protein was modelled after a normal distribution in a protein-by-protein fashion.

In line with the field's established methodology of using "% thresholds" to identify chemical hits, we went off to translate the modelled probability information into competition thresholds. By the definition of the normal distribution, approximately 95% of the data lies within 2 standard deviations from the mean ($\pm 2\sigma$). Since we were only concerned with signal decreases in this case, -2σ was the threshold for which there was a <5% probability that the given competition ratio was observed due to random chance ($p<0.05$ by Z-test). We translated this -2σ threshold into competition ratio/percentage competition space by redefining the standard deviation in terms of the mean, to obtain a percentage competition statistical significance threshold that was customized for each protein. In other words, the hit identification threshold

is obtained by taking the standard deviation as a fraction of the mean, then multiplying that by two, which corresponds to Z-test 5% significance threshold. We named this process individualized protein baseline extraction (INPROBE).

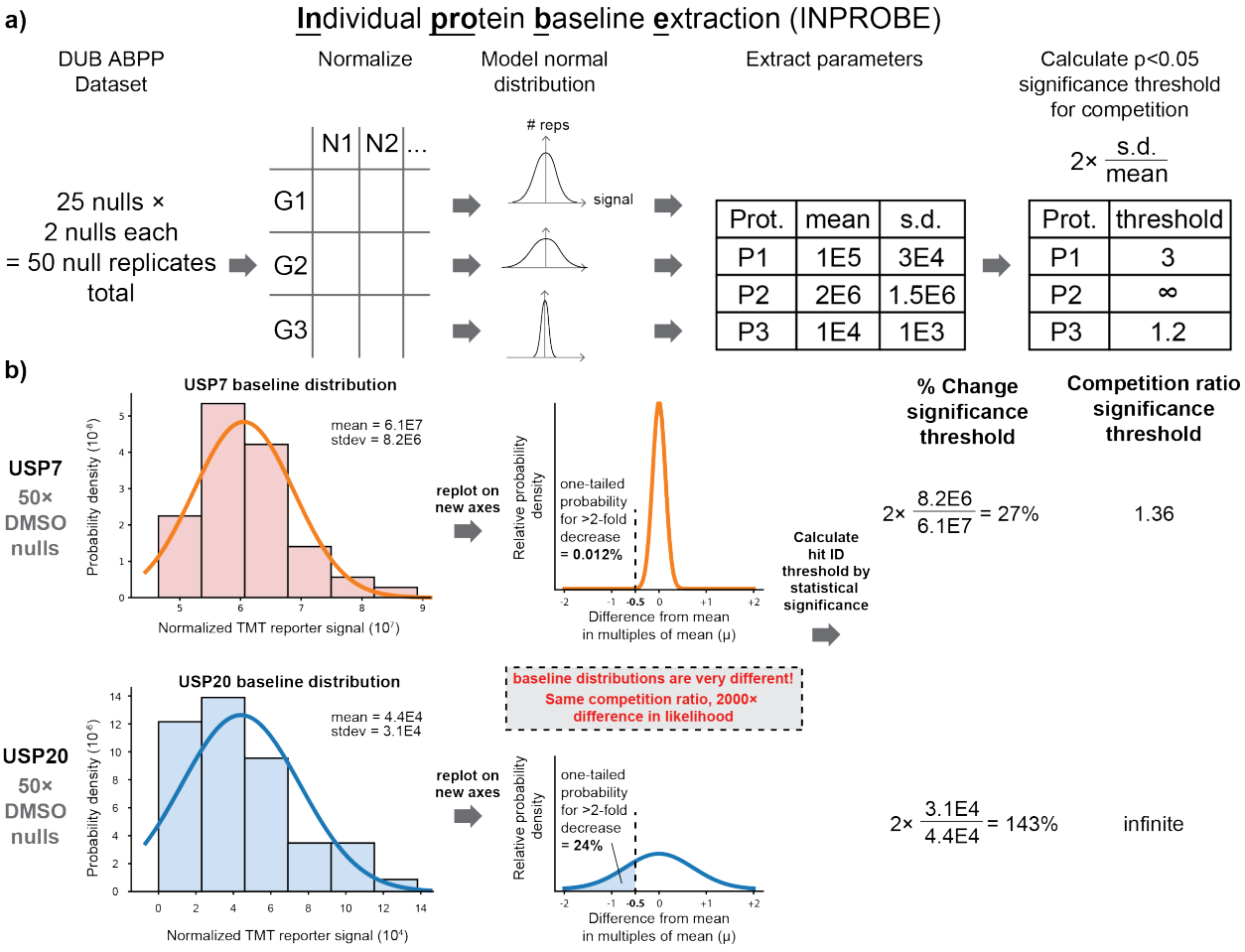


Figure 4.2. Procedure and initial application of INPROBE. a) Process of individual protein baseline extraction (INPROBE), a novel strategy for chemoproteomic screening data analysis. DMSO null replicates from the ABPP primary screen are extracted, normalized by total DUB signal across channels, then fitted to normal distributions in a protein-by-protein fashion. Under INPROBE, the hit ID threshold is obtained by taking the standard deviation as a fraction

(continued from previous page) of the mean, then multiplying that by two, which corresponds to Z-test 5% significance threshold. b) Example distributions for USP7 and USP20.

4.2.2 Preliminary analysis of baseline variance

Our analysis showed divergent magnitudes of signal variance across DUBs. [Figure 4.2b] Baseline signal variation for USP7 and USP20 modelled well to a normal distribution, but the magnitude of the standard deviation as a fraction of the mean was much larger for USP20 than for USP7. When replotted on the same set of axes, it became apparent that the baseline distribution/variance for USP20 in the ABPP-based primary screen was much larger than that for USP7.

This difference in baseline distribution holds important implications in hit identification. Based on the modelled probability distributions, the one-tail end probability corresponding to the 2-fold change used as a threshold in original hit identification were dramatically different for the two species: 0.012% for USP7, and 24% for USP20. In other words, the probability that a 2-fold change could occur due to random noise in the baseline distribution for USP20 was 2000-fold higher than for USP7. This challenged the conventional “one-size-fits-all” approach for arbitrary thresholding or concatenation and supported the application of protein-by-protein threshold determination.

Using the INPROBE procedure, we recalculated the hit identification cutoffs for USP7 and USP20. Calculated competition ratio significance threshold for USP7 was 27%, while the significance cutoff for USP20 was 143%, a five-fold difference. Furthermore, as the theoretical

maximum for a percentage decrease in protein abundance is 100% (quantified protein abundance signal cannot be negative), this means that it was impossible to obtain meaningful data for USP20 from one single run of the DUB ABPP assay in the screen.

At this point, preliminary application of INPROBE to two sample DUBs have challenged common assumptions in chemoproteomic primary screening data analysis. Contrary to conventional assumptions, the quantified abundances for USP7 and USP20 displayed very different baseline distributions. In this case, it was clear that the same competition ratio did not hold the same meaning across protein species, and calculations suggested there is a five-fold difference in the significance threshold (27% v.s. 143%) for a $p<0.05$ cutoff.

4.3. Impact on hit identification

To investigate the differences of significance thresholds across DUBs detected in the primary ABPP assay, we applied the INPROBE procedure to each DUB. **[Figure 4.3a]** Correlation with DUB mean signal magnitude suggested significance thresholds decreased with DUB signal, consistent with the abundance-biased nature of data-dependent acquisition and signal suppression. Two dummy non-DUBs were included in the analysis as validation; as expected, both gave a significance threshold of >100%, confirming these were not specifically enriched by the probe. A cluster of ~10 DUBs had significance thresholds that lie above the theoretical maximum of 100%, suggesting that it would be impossible to distinguish any signal from noise for these DUBs in one single run. 30-some DUBs had cutoffs below 50% which was used for the original hit identification effort, indicating that we were losing out on valuable sensitivity for this group.

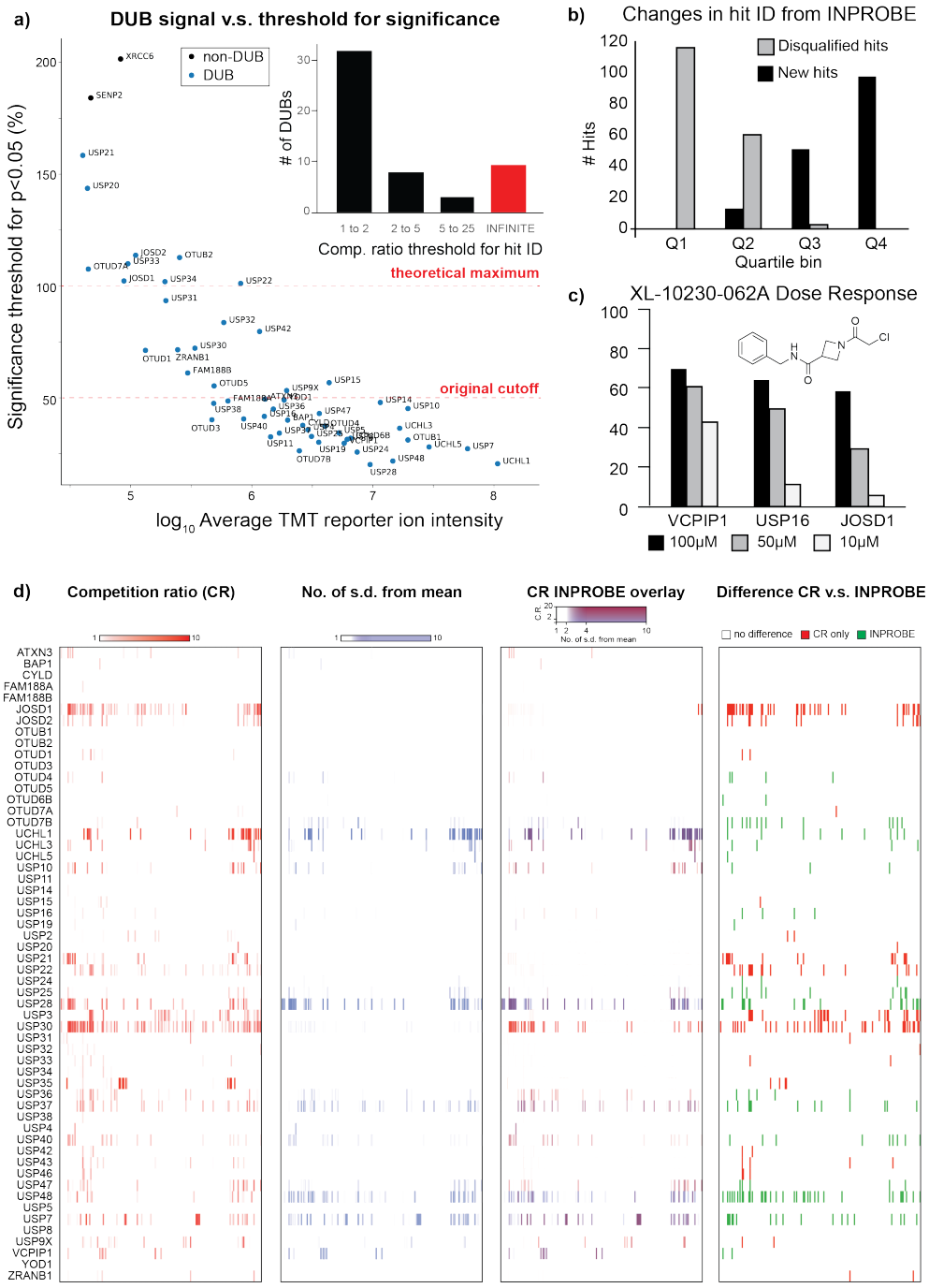


Figure 4.3. Impact of INPROBE on screening hit identification. a) INPROBE significance thresholds for each DUB plotted against log10 of its average TMT reporter ion intensity. *Inset*: bar chart showing number of DUBs in each significance threshold bin. b) Number of hits identified uniquely or disqualified by INPROBE when compared to taking a uniform 50%

(continued from previous page) threshold as in original screening, plotted in signal-dependent quartile bins. c) Dose response ABPP revealed XL-10320-062A indeed targeted USP16. d) Heatmaps showing screen results from *left:* a uniform competition ratio-based cutoff of 2-fold (percentage change of 50%), *mid-left:* INPROBE statistical significance only, *mid-right*: an overlay of both effect size and statistical significance, and *right*: disagreements across the two methods.

We next sought to re-evaluate data from the original primary screen by taking statistical significance for hit identification. In this new attempt, we defined hits as any DUB-inhibitor pair that gave a competition percentage larger than the significance threshold. We compared hit designation by INPROBE statistical significance only to the original designation determined using the uniform 50% cutoff in the original screening campaign. Of the 9904 inhibitor-DUB data points in the primary assay, INPROBE agreed with the original analysis method at 9553 (96%) of the data points. 162 novel hits were identified, and 189 inhibitor-DUB pairs were disqualified as hits due to having a competition ratio below the customized statistical significance threshold for each DUB.

INPROBE showed enhanced sensitivity for inhibitor binding to high-signal DUBs, consistent with unexpected biochemical inhibition observed in orthogonal validation. An example lies in UCHL1, where compounds such as F46 with competition ratios below the 50% cutoff had appreciable IC50s in purified enzyme biochemistry. The same applied for USP48 and F26. This higher sensitivity was most apparent in the upper quantiles of DUB signal intensity where most of the identified false negatives lie. Interestingly, several new hits were identified against USP16 and

USP40, DUBs in the second-lowest significance quantiles. One such new hit, XL-10320-062A, was originally flagged as a VCPIP1 hit; subsequent dose response confirmed USP16 activity. **[Figure 4.3c]** This indicates that INPROBE's higher sensitivity was not only limited to proteins detected at the highest signals.

Hits disqualified by statistical significance through INPROBE concentrated in DUBs with low signal in primary screening. These DUBs were also ones that failed in orthogonal hit validation or did not replicate in repeat MS-ABPP runs (USP20, USP21, USP22, JOSD2). **[Figure 4.3c]** Put together, this suggested that considering statistical significance improves sensitivity for high-signal species and reduces false positives against low-signal species.

Of interest, while the original screen did not identify selectivity relationships that correspond to DUB catalytic domain sequence phylogeny, INPROBE analysis reveals such clustering. This group of compounds usually target one member of a closely related group with higher potency, and the other members with lower potency. This lower potency leads to this activity being missed in concatenation and is subsequently recovered by INPROBE. An example cluster existed with USP7, USP47, USP48, USP16, and USP40. The emergence of this relationship bolsters confidence in these INPROBE-identified activity.

4.4. Discussion

Overall, our results have significant implications for current data analysis workflows in chemoproteomic covalent fragment screening. While informative, current statistical analysis assume competition ratios/baseline distributions are comparable across protein species, and

fail to distinguish baseline variance from inhibitor-derived signal. INPROBE revealed that assumption to be untrue, and that customizing hit identification thresholds to the baseline signal variance for each protein improves sensitivity and reduces false positives. Looking ahead, we are hopeful that INPROBE will be a powerful tool to improve the efficiency and robustness of future chemoproteomic primary screens.

Chemoproteomic experiments with this number of null replicates are rarely reported, and there are no reported attempts to examine protein-by-protein signal baseline variance on chemoproteomic data. Here, we provided both null raw signal and competition ratios for the general research community as a sample dataset for use by the general community. We hope that this accessible dataset can enable the development of more advanced analytical approaches.

An in-depth analysis of our first-reported null replicate dataset revealed that protein variances could not be assumed to be similar across species. The fact that correction factor thresholds ranged from close to one to infinity is alarming: current studies often apply a uniform integer threshold (usually between 2-5) for hit identification. As demonstrated with our library versus library screening dataset, a one-size-fit-all approach can lead to reduced sensitivity for high-signal species and false discoveries for low-signal species. Future chemoproteomic screens would do well to proceed with protein-by-protein statistical significance in mind.

That said, we would also like to caution the reader against the temptation to take a fatalistic attitude and declare all hits against proteins with threshold >100% to be false positives. INPROBE was designed for single-run primary screening data; hence, confidence with these hits can be strengthened if consistent results were obtained across multiple replicates. We would

also like to encourage researchers to validate using orthogonal assays in the chemical biology experimental design toolbox (biophysical binding assays, MS intact protein mass spectrometry, etc.) One example from our experience was the DUB JOSD1, whose significance threshold was 102% as determined in this study. In our experience, while many of the JOSD1 hits are indeed irreplicable, there is a handful of N-cyano compounds which has SAR, consistently generate large magnitude competition in ABPP, and can be validated in orthogonal assays.[20]

In addition to screen hit identification and prioritization, INPROBE also informs on priority for assay development. Reports on novel chemoproteomic workflow and assay development often focuses on depth of coverage and neglect consistency. That is indeed important in an identification context. Our results have revealed that further parameters such as consistency and variance are important to consider and motivate development in targeted PRM methods and associated spectral libraries.

Outside of the direct context of chemoproteomic screening in this paper, this finding might also hold ramifications for statistical analyses of quantitative proteomics experiments comparing two or more conditions, even those with replicates. Empirical Bayesian models used in the popular R *limma* package borrows variances across protein species, and in doing so assumes variances are similar across protein species, an assumption challenged by our data. As *limma* was borrowed from the field of genetics for use in proteomic data, a careful examination of how the data model assumptions translate from genetics to proteomics would be appropriate.

An interesting next step would be to apply the pipeline outlined here to peptide-level datasets to see if they follow similar variance patterns. Analyses reported in this chapter were carried out on a protein-level dataset. Hence, there is built-in levelling of random noise from summing

TMT reporter ion intensities across peptides. Addressing noise at the peptide level can hold important implications in the design and interpretation of popular cysteine-directed peptide-level screens. As the multiplexing capabilities and scale of chemoproteomic screens increase, it might even become possible to apply Z' calculations from high-throughput screening assay development to chemoproteomic screening.

Ultimately, we view the work described herein as an important step towards common rigorous analytical workflows for chemoproteomic primary screening data. As chemoproteomics become increasingly applied in large-scale covalent inhibitor discovery, our method holds the potential to improve efficiency and robustness in hit identification, subsequent optimization, and ultimately the development of first-in-class covalent probes and inhibitors against hitherto inaccessible targets.

5. Future directions and conclusions

5.1. Future directions

In Chapters 2, 3, and 4, we have demonstrated the power of our chemoproteomic library versus library screen in identifying valuable covalent hits targeting the DUB active site. In this chapter, I discuss our preliminary efforts to expand our strategy, through (1) developing a second-generation covalent focused library targeting the active site, (2) expanding our focus to additional novel binding pockets in the target class, and (3) further refining the ABPP primary screening assay. These new focused libraries will lead to future hits against a diversified portfolio of DUBs, enabling future chemical genetics studies to elucidate DUB biological function and disease relevance.

5.1.1 Second-generation active-site focused library

The first iterative step forward will be to generate a second-generation covalent active-site directed DUB-focused library for screening using the same platform. Insights from the library versus library screen of the first-generation library will be used to inform design.

The second-generation library will include structural elements underrepresented in the first-generation library. In first-generation library screening, we saw that length and flexibility of alkyl linkers impacted DUB targeting (WH-9943-188, WH-9943-189, WH-9943-103C). However, few compounds included a flexible alkyl chain, and there was no systematic attempt to vary their length. The second-generation library will include compounds that explore this structural motif.

Another underexplored determinant of inhibitor selectivity was ring size, saturation, and substituents on the pendant ring at the warhead. Bulky warheads in particular were underrepresented in the first-generation library, but tetrahydroquinoline and indoline chloroacetamide warheads both yielded interesting USP48 hits. Also of note is the fact that XL177A, selective USP7 inhibitor which is also a foundational chemotype in first-generation library design, also possessed a rather bulky tetrahydroacridine warhead.[33] This evidence motivates further incorporation of bulky pendant rigns (tetrahydroacridines, quinolines, indolines) in the second-generation library. F59 was the sole example with a substitutent (a methyl group) on the pendant pyrrolidine ring: sole addition of the methyl group greatly improved selectivity for UCHL1 and USP10 relative to F22. This motivates addition of substituents such as methyl and hydroxy groups to the pendant rings.

The second generation will also feature novel structural elements unexplored in the first-generation library. For example, most molecules in the first-generation library were linear in structure, featuring little branching. Another possible direction will be the use of more complex leaving groups on acryl warheads, such as ethyl esters and benzene groups. As the electrophilic warhead is wedged in an oxanion hole meant to stabilize the substrate transition state, tuning of electronics and bulk might yield additional determinants of DUB-directed potency and selectivity. Building further to the right of the electrophilic warheads might offer access to pockets occupied by the substrate lysine side chain.

Finally, scaffolds shown to target a small set of DUBs in first-generation screening will be further diversified. The hope is to identify analogs with refined selectivity against hitherto untargeted DUBs. An example here would be the azetidinyl chloroacetamide scaffold, which

targeted VCPIP1, OTUD7B, BAP1, USP16, and USP40. While we were able to rapidly optimize hits into a potent and selective VCPIP1 probe, requisite changes for targeting OTUD7B, BAP1, and USP16 remain unclear. We plan to further diversify at the pendant ring and attempt to access further pockets by extending a flexible alkyl linker from the current scaffold.

With these considerations in mind, we are confident that the second-generation library will be a fruitful endeavor that will both extend and complement findings from first-generation library versus library screening. In the fullness of time, repeated iterations of this methodology will lead to first-in-class inhibitors and drugs targeted hitherto understudied DUBs and illuminate DUB biology as we have already begun with VCPIP1.

5.1.2 Beyond the active site: noncovalent library versus library screening

In the previous chapters, I have discussed our successes at targeting the DUB active site using catalytic cysteine directed covalent small molecules. However, DUBs possess myriad other domains outside of the active site, and that additional structural and sequence diversity can be leveraged to discover novel potent and selective inhibitors as well.

In the years since I have started my PhD, we and others have identified additional potential targetable pockets which might exist across DUB family members. A set of chemical scaffolds targets each pocket, and a few prototype compounds have emerged with validated activity against specific cognate DUBs. These include the (1) hydroxypiperidines and dimethylpyrroles targeting the S3-S5 pocket, and (2) oxadiazoles targeting a distal pocket in the ubiquitin-binding site on USP family DUBs.

The first such pocket is the S3-S5 pocket, named according to protease structural nomenclature. Two series of compound bind at that pocket: the hydroxypiperidines and the dimethylpyrrole series. Multiple USP7 inhibitors feature hydroxypiperidine moieties, and interactions with blocking loops 1 and 2 have been characterized by X-ray crystallography.[30] A USP19 inhibitor reported by the biotechnology company Almac also possesses a hydroxypiperidine core and has been successfully validated by our in-house pipeline.[39] Dimethylpyrrole USP14 inhibitors have been identified with crystallographic information as early as 2010; we became interested in the scaffold's potential to inhibit a wide swathe of DUBs when we identified another dimethylpyrrole compound in a high-throughput screen against USP8.[89,90]

The oxadiazole moiety became of interest due to our work with USP25/28.[90] Parallel high-throughput screening of 8 DUBs against a library of 50,000 noncovalent small molecules revealed multiple compounds with an oxadiazole core with preferential inhibitory activity against USP25/28. Medicinal chemistry optimization efforts led to AV-180, an oxadiazole compound with IC_{50} = 100nM against USP25 and USP28. De Lello et al. at Genentech identified oxadiazole hits in an NMR-based fragment screen against USP7, and found the series to bind to a pocket where ubiquitin binds to USP7, distal from the active site[32]. Several unvalidated USP47 inhibitors reported in the patent literature also contain an oxadiazole core.[91] Together, these studies nominate the oxadiazoles as a scaffold which might be able to bind to DUBs broadly across the USP subfamily.

With bona fide DUB binders targeting each pocket, the same library versus library setup of my thesis work can be repurposed to discover noncovalent chemical matter which binds to

unprecedented DUB targets at the S3-S5 and distal pocket. Modular synthesis can center around each of the hydroxypiperidines, dimethylpyrroles, and oxadiazoles cores, with systematic changes to enable SAR exploration. ABPP can be used to read out target scope against the DUB target class, and orthogonal validation can take the form of biochemical inhibition and biophysical binding assays such as ITC or SPR.

As a first step towards a noncovalent DUB-focused library versus library screen, I made preliminary efforts towards adapting the protocol for use with noncovalent molecules. Since we would be competing the irreversible covalent ABPs against reversible noncovalent molecules, ABP incubation time had to be calibrated to avoid outcompeting the noncovalent molecules entirely. A pilot run was carried out using a panel of inhibitors from a high-throughput screen, in which we were able to successfully demonstrate enrichment of 59 DUBs with competition readout against known DUB targets.[90]

We have obtained funding to start synthesizing a focused library based on the hydroxypiperidine, dimethylpyrrole, and oxadiazole scaffolds. Design and synthesis efforts have begun under the direction of Xiaoxi Liu, Mona Sharafi, Yanran Lu, and Cara Starnbach. This effort promises discovery of new scaffolds and chemical ligands targeting diverse pockets on the DUB target class. In addition, any ligands discovered in this effort can be repurposed for DUB-recruiting chimeric molecules or DUB-degraders, providing additional avenues for productive application.

5.1.3 Enhancing the DUB ABPP chemoproteomic primary screening assay

Both libraries outlined above would be screened with ABPP, creating an impetus for us to further improve the scope, robustness, and throughput of the assay.

Opportunity exists in further expanding the scope of robustly detected DUBs. As shown in the bioinformatic analyses in Chapter IV, most DUBs which are inconsistently detected in the ABPP assay are expressed at low levels in HEK293T cells (e.g. OTUD6A, OTUD7A, A20). A potential improvement would be to "source" additional native DUBs from additional cell lines. One would start by identifying cell lines with a complementary DUB expression profile to HEK293Ts, so that the new cell line would express currently inconsistently detected DUBs at high levels. This can be done by mining publicly available proteomic datasets such as ProteinAtlas.org. After that, ABPP can be performed on mixtures of lysates from both HEK293Ts and the new cell line, and the number of DUBs consistently detected can be assessed as in Chapter 2.

In terms of enzyme scope outside of the DUBs, we showed in Chapter 3 that some DUB-reactive compounds such as F70 might have reactivity against proteases that target ubiquitin-like PTMs such as Nedd8, ISG15, and SUMO. Activity against Ubl cysteine proteases can be gauged by including their cognate ABPs in the assay probe cocktail, so that competitive binding can be readout for DUBs and Ubls in one single workflow. This acts as a built-in off-target identification on one hand and offers opportunities to identify Ubl-proteases on the other. As there are few proteases targeting these ubiquitin-like modifications, the expanded assay scope should introduce minimal additional analytical difficulty.

Active efforts are underway by He (Eric) Zhu and Zhenze Jiang to innovate on the MS acquisition part of the screening assay. The first approach involves using a targeted PRM CE-MS method to accelerate MS analysis and improve consistency of peptide detection. As a more long-term endeavor, we are also developing a label-free data independent acquisition strategy, where we will leverage the SWATH-MS platform developed by the Aebersold and Mann Labs for quantification instead of TMT. Removing TMT from the procedure will speed up sample preparation and improve chromatographic performance through elimination of the hydrophobic TMT modification. However, since DUBs are underrepresented existing spectral libraries required for SWATH-MS, the Marto lab is currently focusing on generation of DUB entries in SWATH-MS spectral libraries.

Members across the Burhlage and Marto groups are actively contributing to the development of these upcoming screens. Xiaoxi Liu, Mona Sharafi, Cara Starnbach, Yanran Lu, and Wei Pin Teh are synthesizing compounds for both libraries. Scott Ficarro, He (Eric) Zhu, and Nick Girardi are working on analytical methodology. We expect that these pipelines will yield multiple DUB-targeting compounds for DUB inhibitor development.

5.2. Overall conclusions

The work described in this thesis was fundamentally motivated by a need to expand the druggable proteome, and thereby expand the scope of diseases which can be treated with small molecule therapeutics. We have undertaken this challenge by rejecting the current

paradigms of well-trodden “druggable” enzyme targets and stale methods of noncovalent target-based inhibitor discovery.

Here, I discuss the overall conclusions from this study on two fronts: (1) contributions to DUB inhibitor discovery and (2) application of chemoproteomic target-class approaches to covalent drug discovery.

5.2.1 Contributions to DUB inhibitor discovery

Ubiquitin signaling and DUBs are both relatively novel discoveries in the biological sciences. Ubiquitin was discovered in 1975, while the first DUB was discovered in the early 1980s.[10,12] As such, we view our contributions to DUB inhibitor discovery in the lens of a field in its adolescence, on the verge of immense progress and transformation.

In the past five years, we have seen extensive progress DUB inhibitor discovery. The first *bona fide* potent and selective DUB inhibitors were published in 2017 against USP7, followed by inhibitors targeting other well-studied DUBs of known therapeutic interest such as UCHL1, USP19, USP30, and CSN5.[23,39,43] At the start of this work, main questions surrounded less-studied members of the target class: is the rest of the target class druggable? Are there privileged scaffolds or conserved structural motifs which will enable target-class inhibitor discovery as in the case of kinases and ATP mimetics? These questions arbitrate whether the DUBs consist of only a few enzymes which just so happened to be druggable, or if the entire family holds promise as a trove of actionable potential disease targets.

Our results have answered these questions vital to the future of DUB research. As the first study utilizing a DUB-focused compound library to systematically explore chemical space for

DUB inhibition, our results provide high quality hits against multiple DUBs, yield design principles for DUB inhibition, and demonstrate opportunities in targeting DUBs as a broader target class. We have credentialed covalent targeting of the catalytic cysteine residue to be a generalizable method for selective inhibition against diverse DUBs. We identified multiple covalent warheads as potential privileged scaffolds for DUB targeting. Our discovery of a potent and selective VCPIP1 inhibitor expands targeting to yet another DUB subfamily (OTUs), bringing the number of DUB subfamilies with validated inhibitors from 3 to 4. Taken together, these results validate the DUB family as generally druggable, and demystifies the vast majority of the understudied target class.

Our hits and platform are both well-positioned to answer future needs for DUB inhibitors. As our work with USP7 and others' have effectively demonstrated, DUB inhibition holds immense potential for both targeted protein degradation in cancer and to modulate other diseases. It is hence likely that future research will nominate DUBs as novel proteins of therapeutic interest with the discovery of as-yet unknown clinical phenotypes. The myriad hits identified in this thesis lay the groundwork for accelerating targeted inhibitor discovery when such discoveries call for chemical starting points. The platform described in this thesis can then be re-purposed for testing, validating, and optimizing such inhibitors.

As potent and selective DUB inhibitors become available, the next questions surround the viability of DUB inhibition as an approach for disease therapy. These include functional redundancies across DUBs, impact of DUB inhibition on normal physiology, and deciphering degradative v.s. non-degradative regulatory functions. Well-characterized inhibitors will be invaluable tools for answering these questions and determining the future of the DUBs as a

therapeutic target class. Overall, we are optimistic that DUB inhibition is well on its way to becoming the next frontier in drug discovery against a wide variety of indications.

5.2.2 MS chemoproteomic primary screening

As highlighted throughout this report, it is the target-class nature of our approach that enabled our success: the DUB-wide scope of our study allowed us to simultaneously deconvolve hits as well as DUB-family SAR. Modular covalent design based on knowledge of the DUB active site led to selective targeting of diverse DUBs over other enzyme classes. Our expertise in DUB target-class biochemical and target engagement assays facilitated rapid, rigorous hit validation without *a priori* knowledge about the specific DUB target. This has culminated in a potent and selective inhibitor against the unprecedented DUB VCPIP1, validating that our novel approach does indeed lead to expanded access to unprecedented DUB targets.

One lesson that we have learned regarding chemoproteomic target-class inhibitor discovery approaches is that broad, consistent, and quantitative coverage of endogenous proteins is key to success. These attributes enable inhibitors to be discovered against the target class in general, and also enable deciphering of SAR relationships. In statistical analyses described in Chapter 4, we undertook in-depth analysis of assay variance protein-by-protein, which further emphasized consistency in enrichment and quantification as a critical parameter to optimize during assay development. We also found that considering fluctuations in variance across protein species improves hit identification. This informs best practices in future chemoproteomic screening campaigns.

We anticipate that the general methodology of our target class inhibitor discovery approach can be expanded to other enzymes classes facing the challenges of selectivity and focused interest in inhibitor development. One example is the protein tyrosine phosphatases (PTPs) family: while the past efforts for selective PTP inhibitors have been largely unsuccessful, proof-of-concept compounds do exist.[92] Knowledge and technology such as PTP-specific activity-based probes, fluorogenic substrate assays, and PTP crystal structures are available as well, supporting the use of a target-class approach as in our study. Since the chemical, biochemical, and chemoproteomic methods utilized herein are tractable in a wide range of research environments, the blueprint outlined in this work should be broadly accessible to the research community.

While there are indeed many advantages in using a chemoproteomic library versus library format, limitations remain. First, there is no control over which target class member we will obtain chemical matter against. At the start of the screen, we had no idea which DUBs we will obtain hits against, and we were certainly not interested in VCPIP1. Thus, the unbiased nature of the screen could be a double-edged sword: while it enables general inquiry into the target class, we also lose control over focusing our interest on particular target class members. Such a screen is also very resource-intensive: customized libraries, probes, and dedicated time on a mass spectrometer are all necessary.

Ultimately, this comes down to a question of fit. The approach described in this thesis is certainly not a one-size-fit-all: the chemoproteomic target-class approach is best applied to underexplored protein families where any chemical matter against any unspecified family member can be valuable. A fruitful target class for such an approach should resemble the DUBs

at the beginning of this thesis. Chemically, the lack of selective inhibitors against most family members, combined with a lack of inhibitor design principles motivate such a screen. Biologically, as long as little is known about the substrates and pathways for most family members, any hit chemical matter is likely to be useful as a tool molecule, even in the absence of immediate disease relevance.

5.3. Towards new paradigms of therapeutic discovery

Over the years, novel therapeutic modalities and approaches have become part of the therapeutic paradigm, expanding the realm of treatable diseases. Targeted cancer therapies, immunotherapy, and gene therapy are a few examples that have undergone this transformation from novel to common in the past two decades. The work in this thesis embodies an early step in defining new paradigms: our success at discovering covalent inhibitors against the deubiquitinating enzymes by a chemoproteomic target-class approach motivates continued effort. I hope this story will become a good case study for the potential benefits that can come from taking the path less travelled in future therapeutic discovery. Ultimately, we are optimistic that chemoproteomic inhibitor discovery against DUBs and other novel target classes will allow novel molecules to enter the clinic for improved therapeutics against currently unmet medical needs.

6. Methods and materials

6.1. Methods and regents

6.1.1 Synthetic procedures for library compounds

Abbreviation

Et_3N: Triethylamine

HATU: 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate

EtOAc: Ethyl acetate

DMF: Dimethylformamide

MeOH: Methanol

TFA: Trifluoroacetic acid

HPLC: High performance liquid chromatography

$Pd_2(dba)_3$: Tri(dibenzylideneacetone)dipalladium (0)

Xantphos: 4,5-Bis(diphenylphosphino)9,9-dimethylxanthene

DIAD: Diisopropyl azodicarboxylate

$Pd(dppf)Cl_2$: [1,1'-Bis(diphenylphosphino)ferrocene]dichloropalladium(II)

General Procedure 1:

Step 1: Amines (1.0 eq.), carboxylic acids (1.2 eq.) Et_3N (5.0 eq.) and HATU (1.5 eq.) were added into DMF (3-5mL). The mixture was stirred at room temperature overnight. If necessary, the mixture was diluted with EtOAc (50mL), and washed with brine (30mL×2) to remove excess DMF.

Organic layer was dried over anhydrous sodium sulfate (Na_2SO_4), filtered, and concentrated under reduced pressure. The crude material was then purified by flash column chromatography (hexanes/EtOAc/MeOH).

Step 2: Products from last step were dissolved in DCM (2-3mL) and treated with TFA (2-3mL). The mixtures were stirred at room temperature until the tert-butyloxycarbonyl protecting group was cleaved tracking by UPLC-MS. The mixture was concentrated and flushed by flash column chromatography (EtOAc/MeOH/0.5%Et_3N).

Step 3: Products from last step were dissolved in DCM (2-3mL) with Et_3N (2 eq.) at 0°C. Chloroacetyl chloride (1.2 eq.), or acryloyl chloride (1.2 eq.), or cyanogen bromide (1.2eq) was added dropwisely. The mixture was then stirred at 0°C for 1 hour, and directly purified by flash chromatography (hexanes/EtOAc/MeOH) followed by preparative HPLC (MeOH or CH_3CN/H_2O with 0.0425% TFA) to afford the target products.

General Procedure 2:

Step 1: amines (1.0 eq.), epoxides (1.0 eq.) and cesium carbonate (3.0 eq.) were added into anhydrous DMF (10-15mL). The mixture was heated at 60-80 °C overnight, then cooled down to room temperature before dilution with EtOAc (~50mL). The organic layer was washed with brine (~30mL×2). Combined organic layer was dried over anhydrous sodium sulfate (Na_2SO_4), filtered, and concentrated under reduced pressure. The crude material was then purified by flash column chromatography (hexanes/EtOAc/MeOH).

Step 2: Products from last step were dissolved in DCM (2-3mL) and treated with TFA (2-3mL). The mixtures were stirred at room temperature until the tert-butyloxycarbonyl protecting group was cleaved tracking by UPLC-MS. The mixture was concentrated and flushed by flash column chromatography (EtOAc/MeOH/0.5%Et_3N).

Step 3: Products from the last step (1.0 eq.) were dissolved in DCM (2-3mL) with Et_3N (2.0-5.0 eq.) at 0°C. Chloroacetyl chloride (1.2 eq.), or acryloyl chloride (1.2 eq.), or cyanogen bromide (1.2eq) was added dropwisely. The mixture was then stirred at 0°C for 1 hour, and directly purified by flash chromatography (hexanes/EtOAc/MeOH) followed by preparative HPLC (MeOH or CH_3CN/H_2O with 0.0425% TFA) to afford the target products.

General Procedure 3:

Step 1: bromo-substituted benzo[d]thiazol-2-amine (1.0 eq.) carboxylic acids (1.2 eq.), Et3N (5.0-10.0 eq.) and HATU (1.5-2.0 eq.) were added sequentially in anhydrous DMF (5-10mL). The mixture was stirred at room temperature overnight. If necessary, the mixture was diluted with EtOAc (50mL), and washed with brine (30mL×2) to remove excess DMF. Organic layer was dried over anhydrous sodium sulfate (Na_2SO_4), filtered, and concentrated under reduced pressure. The crude material was then purified by flash column chromatography (hexanes/EtOAc/MeOH).

Step 2: The isolated products from step 1 (1.0 eq.) was dissolved in 1,4-dioxane and H_2O (3:1). Into the solution were added boronic acids or boronate ester (3.0 eq.), potassium carbonate (3.0 eq.) and $Pd(PPh_3)_4$ (0.2 eq.). The mixture was degassed by bubbling through N2 for 10min before heating up to 95°C and stirred at this temperature for 2-8 hours. The reaction was then cooled down to room temperature and diluted with EtOAc (50mL). The organic phase was washed with saturated ammonium chloride (30mL×2). Aqueous layer was then extracted with more EtOAc (50mL). Combined organic layers were washed with brine, dried over anhydrous sodium sulfate (Na_2SO_4), filtered, and concentrated under reduced pressure to afford crude material, which was then purified by flash chromatography (hexanes/EtOAc/MeOH).

Step 3: Products from last step were dissolved in DCM (2-3mL) and treated with TFA (2-3mL). The mixtures were stirred at room temperature until the tert-butyloxycarbonyl protecting group was cleaved tracking by UPLC-MS. The mixture was concentrated and flushed by flash column chromatography (EtOAc/MeOH/0.5%Et_3N).

Step 4: Products from the last step (1.0 eq.) were dissolved in DCM (2-3mL) with Et_3N (2.0-5.0 eq.) at 0°C. Cyanogen bromide (1.2eq) was then added. The mixture was then stirred at 0°C for 1 hour, and directly purified by flash chromatography (hexanes/EtOAc/MeOH) followed by preparative HPLC (MeOH or CH_3CN/H_2O with 0.0425% TFA) to afford the target products.

General Procedure 4:

Step 1: The mixture of bromobenzo[d]thiazol-2-amine (1.0 eq.), 3,5-dimethylisoxazole-4-boronic acid (1.3 eq.), sodium carbonate (2.0 eq.) were mixed in 1,4-dioxane, EtOH and H_2O (8:2:1). N_2 was bubbled through the suspension for 10 to 15min, followed by addition of

tetrakis(triphenylphosphine palladium (0) (0.1 eq.) The mixture was purged with N_2 for another 5 min before stirring at 95°C overnight under N_2. Then the mixture was concentrated under reduced pressure, diluted with EtOAc, and washed with saturated NH_4Cl. Combined aqueous layer was extracted with EtOAc. Combined organic layer was washed once with brine, dried over anhydrous Na_2SO_4, filtered, and evaporated under reduced pressure to afford crude material, which was then purified by flash chromatography (hexanes/EtOAc/MeOH).

Step 2: The products isolated from last step (1.0 eq.) and (*S*)-1-Boc-pyrrolidine-3-carboxylic acid (1.2 eq.), Et_3N (5.0 eq.) and HATU (1.5 eq.) were added into DCM/DMF. The solution was stirred at room temperature overnight. The crude was then directly purified by flash chromatography (hexanes/EtOAc/MeOH) to afford desired product.

Step 3: Products from last step were dissolved in DCM (2-3mL) and treated with TFA or 4M HCl in 1,4-dioxane (2-3mL). The mixtures were stirred at room temperature until the tert-butyloxycarbonyl protecting group was cleaved tracking by UPLC-MS. The mixture was concentrated and flushed by flash column chromatography (EtOAc/MeOH/0.5%Et_3N).

Step 4: Products from the last step (1.0 eq.) were dissolved in DCM (2-3mL) with Et_3N (2.0-5.0 eq.) at 0°C. Cyanogen bromide (1.2eq) was then added. The mixture was then stirred at 0°C for 1 hour, and directly purified by flash chromatography (hexanes/EtOAc/MeOH) followed by preparative HPLC (MeOH or CH_3CN/H_2O with 0.0425% TFA) to afford the target products.

General Procedure 5:

Step 1: 5-(3,5-dimethylisoxazol-4-yl)benzo[d]thiazol-2-amine, which was synthesized in step 1 of General Procedure 4 (0.075g, 0.3mmol) was added in 3mL anhydrous MeCN. Into the solution was added $CuBr_2$ (0.065g, 0.45mmol) and t-butyl nitrite (0.046g, 0.45mmol) at 0°C. The mixture was then warmed up to room temperature then 65°C, and stirred for 4 hours. The reaction was cooled to room temperature, and diluted with water (30mL). The mixture was acidified with 12M HCl to pH=2 and extracted with EtOAc (30mL×2). The combined organic layer was washed with brine, dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure to afford the crude material. The material was purified by flash chromatography (hexanes/EtOAc/MeOH) to afford a mixture of desired product and chloride-substituted analogue, which did not undergo

further purification and used directly in the next step. LC/MS (ESI) *m/z* 264.77; $[M+H]^+$; calcd for $C_{12}H_{10}ClN_2OS^+$: 265.02

Step 2: Products from the last step (0.14g, 0.5mmol), 1-Boc-3-oxopiperazine (0.2g, 1.0mmol), cesium carbonate (0.65g, 2.0mmol), $Pd_2(dba)_3$ (0.046g, 0.05mmol), and Xantphos (0.058g, 0.1mmol) were added into 5mL 1,4-dioxane. The mixture was degassed by bubbling in N_2 for 10-15min before heated at 95°C overnight. Then the mixture was cooled to room temperature before diluted with EtOAc (30mL). Organic layer was washed with 20% citric acid (20mL×2). Combined aqueous layer was extracted with EtOAc (30mL). Combined organic layer was washed with brine, dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure to afford crude material. The crude material was purified by flash chromatography (hexanes/EtOAc/MeOH) to afford desired product (0.12g,) LC/MS (ESI) *m/z* 428.87; $[M+H]^+$; calcd for $C_{21}H_{25}N_4O_4S^+$: 429.16

Step 3: Products from last step were dissolved in DCM (2-3mL) and treated with TFA (2-3mL). The mixtures were stirred at room temperature until the tert-butyloxycarbonyl protecting group was cleaved tracking by UPLC-MS. The mixture was concentrated and flushed by flash column chromatography (EtOAc/MeOH/0.5%Et_3N).

Step 4: Products from the last step (0.04g, 0.1mmol, 1.0 eq.) were dissolved in DCM (3mL) with Et_3N (0.07mL, 0.5mmol, 5.0 eq.) at 0°C. 2-chloroethane-1-sulfonyl chloride (16µL, 0.15mmol, 1.5 eq.), or acryloyl chloride (13µL, 0.15mmol, 1.5 eq.), or cyanogen bromide (3M) (50µL, 0.15mmol, 1.5 eq) was added dropwisely. The mixture was then stirred at 0°C for 1 hour, and directly purified by flash chromatography (hexanes/EtOAc/MeOH) followed by preparative HPLC (MeOH or CH_3CN/H_2O with 0.0425% TFA) to afford the target products.

General Procedure 6:

Step 1: 6-(3,5-dimethylisoxazol-4-yl)benzo[d]thiazol-2-amine, which was synthesized in step 1 of General Procedure 4 (0.25g, 1.0mmol) was added in 10mL anhydrous MeCN. Into the solution was added $CuBr_2$ (0.22g, 1.5mmol) and t-butyl nitrite (0.16g, 1.5mmol) at 0°C. The mixture was then warmed up to room temperature then 65°C, and stirred for 4 hours. The reaction was cooled

to room temperature, and diluted with water (30mL). The mixture was acidified with 12M HCl to pH=2 and extracted with EtOAc (30mL×2). The combined organic layer was washed with brine, dried over anhydrous Na2SO4, filtered, and concentrated under reduced pressure to afford the crude material. The material was purified by flash chromatography (hexanes/EtOAc/MeOH) to afford 0.26g mixture of desired product and chloride-substituted analogue, which did not undergo further purification and used directly in the next step. LC/MS (ESI) *m/z* 308.87; $[M+H]^+$; calcd for $C_{12}H_{10}BrN_2OS^+$: 308.97

Step 2: The product isolated from last step (0.08g, 0.26mmol), 3 or 4-aminophenylboronic acid (0.05g, 0.4mmol), potassium carbonate (0.07g, 0.52mmol), and Pd(dppf)Cl_2 (0.022g, 0.03mmol) were added into 1,4-dioxane/H_2O (4mL, 3:1). The mixture was degassed by bubbling in N_2 for 10-15min before heated at 95°C overnight. Then the mixture was cooled to room temperature before diluted with EtOAc (30mL). Organic layer was washed with saturated ammonium chloride. Combined aqueous layer was extracted with EtOAc (30mL). Combined organic layer was washed with brine, dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure to afford crude material. The crude material was purified by flash chromatography (hexanes/EtOAc/MeOH) to afford desired products (4-amino: 0.084g, LC/MS (ESI) *m/z* 322.04; $[M+H]^+$; calcd for $C_{18}H_{16}N_3OS^+$: 322.10; 3-amino:0.073g LC/MS (ESI) *m/z* 322.04; $[M+H]^+$; calcd for $C_{18}H_{16}N_3OS^+$: 322.10)

Step 3a: Products from the last step (4-(6-(3,5-dimethylisoxazol-4-yl)benzo[d]thiazol-2-yl)aniline (0.042g, 0.13mmol, 1.0 eq.)) were dissolved in DCM (3mL) with Et_3N (0.056mL, 0.4mmol, 3.0 eq.) at 0°C. 2-chloroethane-1-sulfonyl chloride (22µL, 0.2mmol, 1.5 eq.), or acryloyl chloride (16µL, 0.2mmol, 1.5 eq) was added dropwisely. The mixture was then stirred at 0°C for 1 hour, and directly purified by flash chromatography (hexanes/EtOAc/MeOH) followed by preparative HPLC (MeOH or CH_3CN/H_2O with 0.0425% TFA) to afford the target products.

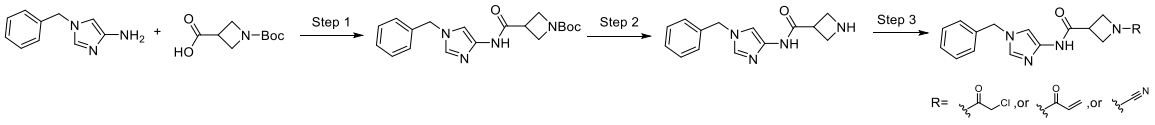
Step 3b: Products from the last step (3-(6-(3,5-dimethylisoxazol-4-yl)benzo[d]thiazol-2-yl)aniline (0.036g, 0.11mmol, 1.0 eq.)) were dissolved in DCM (3mL) with Et_3N (0.07mL, 0.55mmol, 5.0 eq.) at 0°C. 2-chloroethane-1-sulfonyl chloride (14µL, 0.12mmol, 1.1 eq.), or acryloyl chloride (11µL, 0.12mmol, 1.1 eq) was added dropwisely. The mixture was then stirred at 0°C for 1 hour, and

directly purified by flash chromatography (hexanes/EtOAc/MeOH) followed by preparative HPLC (MeOH or CH_3CN/H_2O with 0.0425% TFA) to afford the target products.

General Procedure 7:

Step 1: bromo-substituted heterocyclic carboxylic acids (1.0 eq.) were added into 5mL anhydrous DCM under N_2. Into the mixture was added benzylamine (1.0 eq.), Et_3N (10.0 eq.) and T3P (5.0 eq.). The reaction mixture was stirred at room temperature overnight, and directly purified by flash chromatography (hexanes/EtOAc/MeOH) followed by preparative HPLC (MeOH or CH_3CN/H_2O with 0.0425% TFA) to afford the target products.

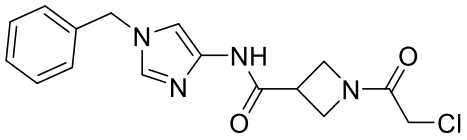
Synthesis of XL_10320_005A, 005B, 005C:



Step 1: The synthesis was preformed according to General Procedure 1 with 1-benzyl-1H-imidazol-4-amine (0.25g, 1.5mmol) and 1-(tert-butoxycarbonyl)azetidine-3-carboxylic acid (0.36g, 1.8mmol). 0.13g desired compound (tert-butyl 3-((1-benzyl-1H-imidazol-4-yl)carbamoyl)azetidine-1-carboxylate) was obtained (24%). LC/MS (ESI) *m/z* 357.07; $[M+H]^+$ calcd for $C_{19}H_{25}N_4O_3^+$: 357.19.

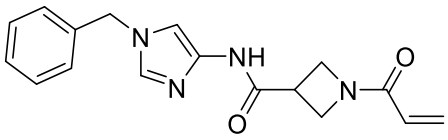
Step 2: The synthesis was performed according to the General Procedure 1 tert-butyl 3-((1-benzyl-1H-imidazol-4-yl)carbamoyl)azetidine-1-carboxylate (0.13g, 0.37mmol). 0.08g N-(1-benzyl-1H-imidazol-4-yl)azetidine-3-carboxamide (84%)

Step 3: The synthesis was performed according to the General Procedure 1 with N-(1-benzyl-1H-imidazol-4-yl)azetidine-3-carboxamide (0.016g, 0.06mmol) and chloroacetyl chloride (10.0 µL, 0.12mmol), or acryloyl chloride (10.0 µL, 0.12mmol), or cyanogen bromide (0.013g, 0.12mmol).

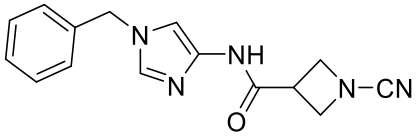


XL_10320_005A (5mg,25%) ^{1}H NMR (500 MHz, DMSO) δ 10.70 (s, 1H), 8.09 (s, 1H), 7.54 – 7.08 (m, 6H), 5.22 (s, 2H), 4.34 (t, *J* = 8.7 Hz, 1H), 4.25 (dd, *J* = 8.5, 5.9 Hz, 1H), 4.17 – 4.08 (d, *J* = 2.4

Hz , 2H), 4.03 (t, *J* = 9.3 Hz, 1H), 3.95 (dd, *J* = 9.6, 5.9 Hz, 1H). (3-H on the azetidine ring may overlap with water peak) LC/MS (ESI) *m/z* 332.87; $[M+H]^+$ calcd for $C_{16}H_{18}ClN_4O_2^+$: 333.11

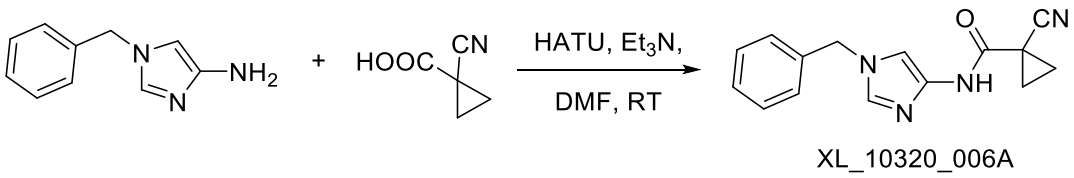


XL_10320_005B (4mg,22%) ^{1}H NMR (500 MHz, DMSO) δ 10.70 (s, 1H), 8.05 (s, 1H), 7.46 – 7.24 (m, 6H), 6.30 (dd, *J* = 17.0, 10.3 Hz, 1H), 6.10 (dd, *J* = 17.0, 2.2 Hz, 1H), 5.67 (dd, *J* = 10.3, 2.2 Hz, 1H), 5.22 (s, 2H), 4.35 (t, *J* = 8.7 Hz, 1H), 4.24 (dd, *J* = 8.4, 5.8 Hz, 1H), 4.05 (t, *J* = 9.4 Hz, 1H), 3.95 (dd, *J* = 9.8, 5.9 Hz, 1H). (3-H on the azetidine ring may overlap with water peak) LC/MS (ESI) *m/z* 310.97; $[M+H]^+$ calcd for $C_{17}H_{19}N_4O_2^+$: 311.15



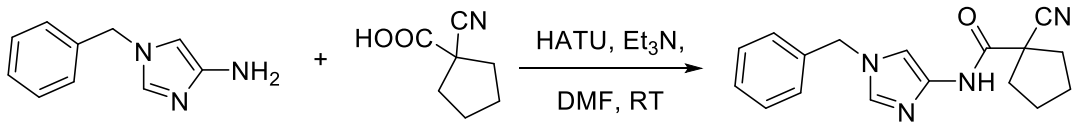
XL_10320_005C (8mg, 47%) 1H NMR (500 MHz, DMSO) δ 10.46 (s, 1H), 7.60 (s, 1H), 7.41 – 7.34 (m, 2H), 7.34 – 7.23 (m, 4H), 5.16 (s, 2H), 4.24 (t, J = 8.1 Hz, 2H), 4.18 (t, J = 6.9 Hz, 2H), 3.62 (ddd, J = 15.2, 8.6, 6.6 Hz, 1H). LC/MS (ESI) m/z 282.08; $[M+H]^+$ calcd for $C_{15}H_{16}N_5O^+$: 282.13

<u>Synthesis of XL_10320_006A, 006B, 006C:</u>



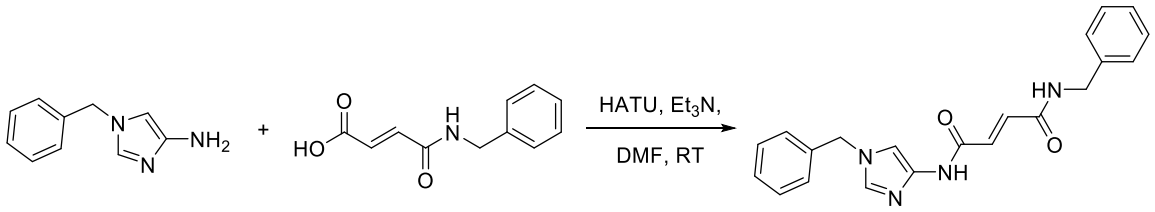
XL_10320_006A

1-benzyl-1H-imidazol-4-amine (0.06g, 0.3mmol), 1-cyanocyclopropane-1-carboxylic acid (0.05g, 0.5mmol) Et_3N (0.48mL, 3.0mmol), and HATU (0.24g, 0.64mmol) were added sequentially to 3mL anhydrous DMF. The reaction mixture was stirred overnight at room temperature, and purified by flash chromatography (hexanes/EtOAc/MeOH) and preparative HPLC (MeOH/H_2O w/ 0.0425% TFA) to afford XL_10320_006A (53mg, 67%) ^{1}H NMR (500 MHz, DMSO) δ 10.59 (s, 1H), 8.17 (s, 1H), 7.39 (m, 2H), 7.36 – 7.30 (m, 4H), 5.25 (s, 2H), 1.73 – 1.59 (m, 4H). LC/MS (ESI) *m/z* 266.97; $[M+H]^+$ calcd for $C_{15}H_{15}N_4O^+$: 267.12



XL_10320_006B

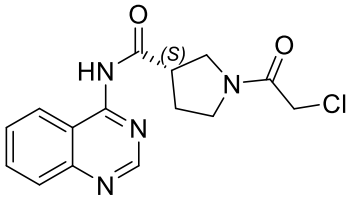
1-benzyl-1H-imidazol-4-amine (0.06g, 0.3mmol), 1-cyanocyclopentane-1-carboxylic acid (0.07g, 0.5mmol) Et_3N (0.48mL, 3.0mmol), and HATU (0.24g, 0.64mmol) were added sequentially to 3mL anhydrous DMF. The reaction mixture was stirred overnight at room temperature, and purified by flash chromatography (hexanes/EtOAc/MeOH) and preparative HPLC (MeOH/H_2O w/ 0.0425% TFA) to afford XL_10320_006B (61mg, 69%) ^1H NMR (500 MHz, DMSO) δ 10.79 (s, 1H), 7.65 (d, *J* = 1.5 Hz, 1H), 7.38 (dd, *J* = 9.9, 4.5 Hz, 2H), 7.31 (ddd, *J* = 13.2, 6.7, 4.1 Hz, 3H), 7.23 (d, *J* = 1.5 Hz, 1H), 5.16 (s, 2H), 2.33 – 2.13 (m, 4H), 1.81 – 1.60 (m, 4H). LC/MS (ESI) *m/z* 295.07 $[M+H]^+$; calcd for $C_{17}H_{19}N_4O^+$: 295.15



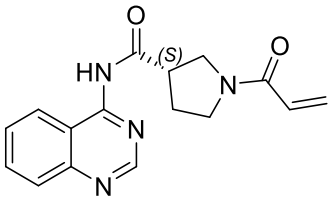
XL_10320_006C

1-benzyl-1H-imidazol-4-amine (0.06g, 0.3mmol), (E)-4-(benzylamino)-4-oxobut-2-enoic acid (0.10g, 0.5mmol) Et_3N (0.48mL, 3.0mmol), and HATU (0.24g, 0.64mmol) were added sequentially to 3mL anhydrous DMF. The reaction mixture was stirred overnight at room temperature, and purified by flash chromatography (hexanes/EtOAc/MeOH) and preparative HPLC (MeOH/H_2O w/ 0.0425% TFA) to afford XL_10320_006C (6mg, 6%) ^1H NMR (500 MHz, DMSO) δ 10.88 (s, 1H), 8.93 (t, *J* = 6.0 Hz, 1H), 7.65 (d, *J* = 1.4 Hz, 1H), 7.45 – 7.18 (m, 11H), 7.08 (d, *J* = 15.1 Hz, 1H), 6.95 (d, *J* = 15.1 Hz, 1H), 5.18 (s, 2H), 4.38 (d, *J* = 6.0 Hz, 2H). LC/MS (ESI) *m/z* 360.97; $[M+H]^+$ calcd for $C_{21}H_{21}N_4O_2^+$: 361.17

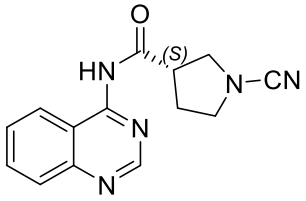
Synthesis of XL_10320_007A, 007B, 007C:



XL_10320_007A (10mg, 11%) 1H NMR (500 MHz, DMSO, mixture of rotamers) δ 8.94 (d, J = 4.8 Hz, 1H), 8.22 (d, J = 8.4 Hz, 1H), 8.01 – 7.89 (m, 2H), 7.73 – 7.64 (m, 1H), 4.34 (d, J = 2.2 Hz, 1H), 3.86 (dd, J = 10.0, 7.7 Hz, 0.6H), 3.79 – 3.63 (m, 2H), 3.63 – 3.50 (m, 2H), 3.40 (dt, J = 11.8, 7.6 Hz, 0.7H), 2.35 – 2.28 (m, 0.6H), 2.28 – 2.18 (m, 1H), 2.17 – 2.06 (m, 0.7H). LC/MS (ESI) m/z 318.97; [M+H]+ calcd for $C_{15}H_{16}ClN_4O_2^+$: 319.10

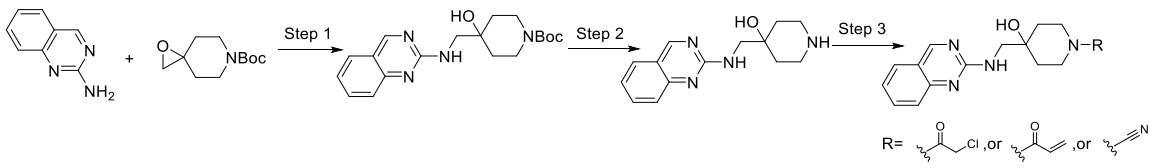


XL_10320_007B (3mg, 3%) 1H NMR (500 MHz, DMSO, mixture of rotamers) δ 8.94 (d, J = 6.0 Hz, 1H), 8.22 (d, J = 8.3 Hz, 1H), 7.94 (m, 2H), 7.69 (m, 1H), 6.62 (ddd, J = 16.8, 12.0, 10.3 Hz, 1H), 6.22 – 6.07 (m, 1H), 5.68 (dt, J = 10.3, 2.4 Hz, 1H), 3.91 (dd, J = 10.3, 7.8 Hz, 0.6H), 3.86 – 3.77 (m, 0.6H), 3.77 – 3.52 (m, 3H), 3.47 – 3.39 (m, 0.6H), 2.26 (m, 1.6H), 2.12 (m, 0.8H) LC/MS (ESI) m/z 296.87; [M+H]+ calcd for $C_{16}H_{17}N_4O_2^+$: 297.13



XL_10320_007C (4mg, 5%) 1H NMR (500 MHz, DMSO, mixture of rotamers) δ 8.91 (s, 1H), 8.23 (d, J = 8.3 Hz, 1H), 8.00 – 7.88 (m, 2H), 7.72 – 7.62 (m, 1H), 3.68 (dd, J = 11.2, 10.0 Hz, 1H), 3.65 – 3.59 (m, 2H), 3.53 – 3.40 (m, 2H), 2.30 – 2.12 (m, 2H). LC/MS (ESI) m/z 267.97; [M+H]+ calcd for $C_{14}H_{14}N_5O^+$: 268.12

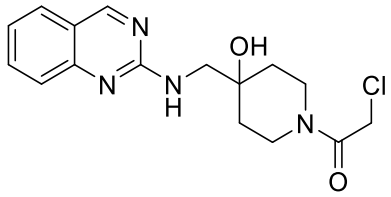
Synthesis of XL_10320_009A, 009B, 009C:



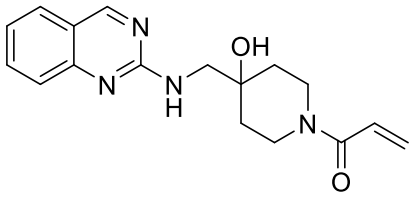
Step 1: The synthesis was performed according to General Procedure 2 with quinazolin-2-amine (0.1g, 0.7mmol) and tert-butyl 1-oxa-6-azaspiro[2.5]octane-6-carboxylate (0.15g, 0.7mmol) with exception of using NaH (60% in mineral oil) (0.03g, 0.75mmol) as base instead of cesium carbonate. 0.16g desired product (tert-butyl 4-hydroxy-4-((quinazolin-2-ylamino)methyl)piperidine-1-carboxylate) was obtained (64%). LC/MS (ESI) *m/z* 358.97; $[M+H]^+$ calcd for $C_{19}H_{27}N_4O_3^+$: 359.21

Step 2: The synthesis was performed according to the General Procedure 2 tert-butyl 4-hydroxy-4-((quinazolin-2-ylamino)methyl)piperidine-1-carboxylate (0.16g, 0.45mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.12g 4-((quinazolin-2-ylamino)methyl)piperidin-4-ol was obtained (quant.)

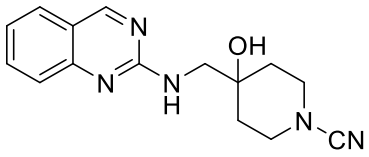
Step 3: The synthesis was performed according to the General Procedure 2 with 4-((quinazolin-2-ylamino)methyl)piperidin-4-ol (0.04g, 0.16mmol), Et_3N (0.13mL, 0.9mmol) and chloroacetyl chloride (14.0µL, 0.18mmol), or acryloyl chloride (15.0µL, 0.18mmol), or cyanogen bromide (0.019g, 0.18mmol).



XL_10320_009A (5mg, 9%)^{1}H NMR (500 MHz, DMSO) δ 9.38 (m, 1H), 8.53 (d, *J* = 234.4 Hz, 1H), 7.88 (d, *J* = 51.8 Hz, 2H), 7.60 (d, *J* = 8.4 Hz, 1H), 7.39 (s, 1H), 4.44 – 4.27 (m, 2H), 4.06 (d, *J* = 13.0 Hz, 1H), 3.63 (d, *J* = 13.5 Hz, 2H), 3.35 (dd, *J* = 18.4, 9.4 Hz, 1H), 3.02 (t, *J* = 11.5 Hz, 1H), 1.55 (dd, *J* = 45.8, 17.5 Hz, 4H). LC/MS (ESI) *m/z* 334.87; $[M+H]^+$ calcd for $C_{16}H_{20}ClN_4O_2^+$: 335.13

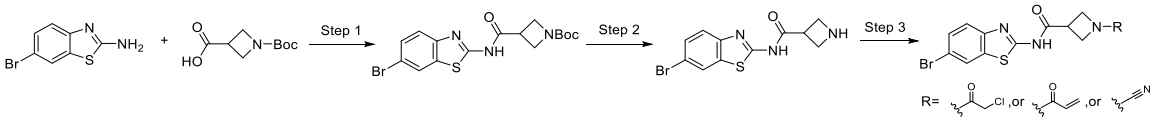


XL_10320_009B (2mg, 1%) 1H NMR (500 MHz, DMSO) δ 9.12 (s, 1H), 7.80 (m, 1H), 7.75 – 7.58 (m, 1H), 7.44 (m, 1H), 7.31 – 7.09 (m, 2H), 6.92 – 6.69 (m, 1H), 6.06 (dd, J = 16.6, 6.3 Hz, 1H), 5.77 – 5.50 (m, 1H), 5.10 (br, 1H), 4.12 (m, 1H), 3.82 (d, J = 11.4 Hz, 1H), 3.45 (m, 2H), 3.04 (m, 1H), 1.53 (m, 4H). LC/MS (ESI) m/z 312.87; $[M+H]^+$ calcd for $C_{15}H_{18}N_5O^+$: 313.17



XL_10320_009C (10mg, 8%). 1H NMR (500 MHz, DMSO) δ 9.23 (s, 1H), 7.89 (d, J = 7.9 Hz, 1H), 7.77 (dd, J = 20.4, 12.9 Hz, 1H), 7.53 (t, J = 10.9 Hz, 1H), 7.33 (dd, J = 17.2, 9.7 Hz, 1H), 3.54 (s, 2H), 3.34 – 3.24 (m, 2H), 3.21 (dt, J = 12.7, 4.1 Hz, 2H), 1.77 – 1.64 (m, 2H), 1.60 (d, J = 13.2 Hz, 2H). LC/MS (ESI) m/z 283.87; $[M+H]^+$ calcd for $C_{15}H_{18}N_5O^+$: 284.15

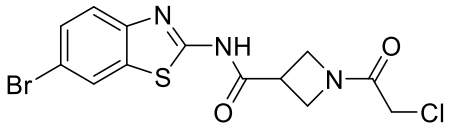
Synthesis of Xl_10320_027A, 027B, 027C



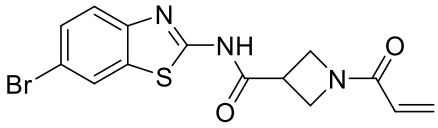
Step 1: The synthesis was preformed according to General Procedure 1 with 6-bromobenzo[d]thiazol-2-amine (0.39g, 1.7mmol) and 1-(tert-butoxycarbonyl)azetidine-3-carboxylic acid (0.42g, 2.1mmol). 0.21g desired compound (tert-butyl 3-((6-bromobenzo[d]thiazol-2-yl)carbamoyl)azetidine-1-carboxylate) was obtained (30%). LC/MS (ESI) *m/z* 356.07 (M+H−*t*-butyl); $[M+H]^+$ calcd for $C_{16}H_{19}BrN_3O_3S^+$: 412.03

Step 2: The synthesis was performed according to the General Procedure 1 with tert-butyl 3-((6-bromobenzo[d]thiazol-2-yl)carbamoyl)azetidine-1-carboxylate (0.21g, 0.5mmol). 0.18g N-(6-bromobenzo[d]thiazol-2-yl)azetidine-3-carboxamide was obtained (quant.)

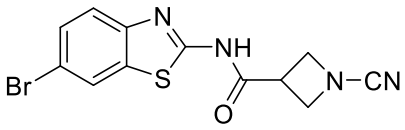
Step 3: The synthesis was performed according to the General Procedure 1 with N-(4-phenylthiazol-2-yl)azetidine-3-carboxamide (0.06g, 0.17mmol) and chloroacetyl chloride (0.017mL, 0.2mmol), or acryloyl chloride (0.017mL, 0.2mmol), or cyanogen bromide (0.02g, 0.2mmol).



XL_10320_027A (43mg, 65%) ^{1}H NMR (500 MHz, DMSO) δ 12.62 (s, 1H), 8.28 (d, *J* = 1.9 Hz, 1H), 7.69 (d, *J* = 8.6 Hz, 1H), 7.58 (dd, *J* = 8.6, 2.0 Hz, 1H), 4.47 – 4.33 (m, 2H), 4.17 (s, 2H), 4.15 – 4.03 (m, 2H), 3.77 (tt, *J* = 8.9, 5.9 Hz, 1H), 3.18 (d, *J* = 4.7 Hz, 1H). LC/MS (ESI) *m/z* 387.77; $[M+H]^+$ calcd for $C_{13}H_{12}BrClN_3O_2S^+$: 387.95

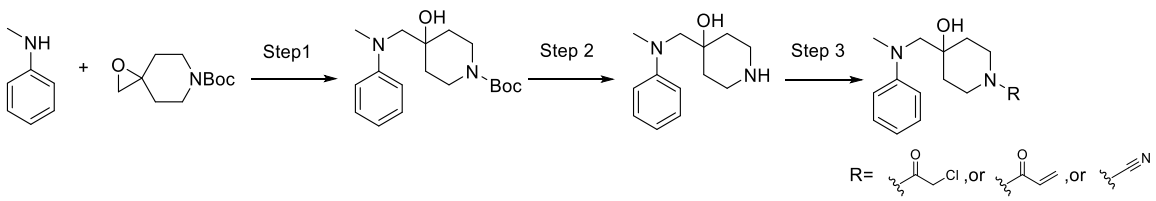


XL_10320_027B (31mg, 50%) ^{1}H NMR (500 MHz, DMSO) δ 12.62 (s, 1H), 8.28 (d, *J* = 2.0 Hz, 1H), 7.69 (d, *J* = 8.6 Hz, 1H), 7.58 (dd, *J* = 8.6, 2.0 Hz, 1H), 6.33 (dd, *J* = 17.0, 10.3 Hz, 1H), 6.12 (dd, *J* = 17.0, 2.2 Hz, 1H), 5.69 (dd, *J* = 10.3, 2.2 Hz, 1H), 4.48 – 4.33 (m, 2H), 4.17 – 4.03 (m, 2H), 3.76 (tt, *J* = 8.9, 5.8 Hz, 1H). LC/MS (ESI) *m/z* 365.77; $[M+H]^+$ calcd for $C_{14}H_{13}BrN_3O_2S^+$: 365.99



XL_10320_027C (36mg, 63%) ^{1}H NMR (500 MHz, DMSO) δ 12.56 (s, 1H), 8.27 (t, *J* = 4.2 Hz, 1H), 7.68 (d, *J* = 8.6 Hz, 1H), 7.58 (dd, *J* = 8.6, 2.1 Hz, 1H), 4.34 (dt, *J* = 14.0, 7.8 Hz, 4H), 3.82 (tt, *J* = 8.9, 6.3 Hz, 1H). LC/MS (ESI) *m/z* 336.77; $[M+H]^+$; calcd for $C_{12}H_{10}BrN_4OS^+$: 336.98

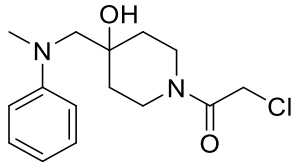
<u>Synthesis of XL_10320_028A, 028B, 028C:</u>



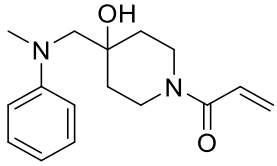
Step 1: The synthesis was performed according to General Procedure 2 with *N*-methylaniline (0.086mL, 0.8mmol) and tert-butyl 1-oxa-6-azaspiro[2.5]octane-6-carboxylate (0.17g, 0.8mmol). LiHMDS (1M) (0.8mL, 0.8mmol) was used instead of cesium carbonate. LiHMDS was dropwisely added to the solution of *N*-Methylaniline in 3mL anhydrous THF at 0°C. The mixture was stirred for 0.5h at 0°C before introducing solution of tert-butyl 1-oxa-6-azaspiro[2.5]octane-6-carboxylate in 2mL THF. The reaction was stirred at room temperature overnight before the general work-up procedure was followed. 0.25g desired product (tert-butyl 4-hydroxy-4-((methyl(phenyl)amino)methyl)piperidine-1-carboxylate) was obtained (96%). LC/MS (ESI) *m/z* 320.97; $[M+H]^+$; calcd for $C_{18}H_{29}N_2O_3^+$: 321.22

Step 2: The synthesis was performed according to the General Procedure 2 tert-butyl 4-hydroxy-4-((methyl(phenyl)amino)methyl)piperidine-1-carboxylate (0.25g, 0.78mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.15g 4-((methyl(phenyl)amino)methyl)piperidin-4-ol was obtained (quant.)

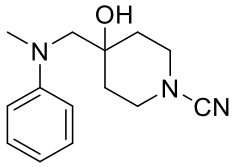
Step 3: The synthesis was performed according to the General Procedure 2 with 4-((methyl(phenyl)amino)methyl)piperidin-4-ol (0.05g, 0.2mmol), Et_3N (0.14mL, 1.0mmol) and chloroacetyl chloride (19.0µL, 0.24mmol), or acryloyl chloride (20.0µL, 0.24mmol), or cyanogen bromide (0.025g, 0.24mmol).



XL_10320_028A (33mg, 55%) ^{1}H NMR (500 MHz, DMSO) δ 7.13 (dd, *J* = 8.7, 7.3 Hz, 2H), 6.78 (d, *J* = 8.2 Hz, 2H), 6.58 (t, *J* = 7.2 Hz, 1H), 4.61 (s, 1H), 4.35 (s, 2H), 4.16 (t, *J* = 15.8 Hz, 1H), 3.63 (d, *J* = 13.3 Hz, 1H), 3.27 (m, 2H), 2.96 (s, 3H), 2.92 – 2.83 (m, 1H), 1.58 (m, 3H), 1.43 (td, *J* = 12.9, 4.6 Hz, 1H). LC/MS (ESI) *m/z* 296.97; $[M+H]^+$ calcd for $C_{15}H_{22}ClN_2O_2^+$: 297.14

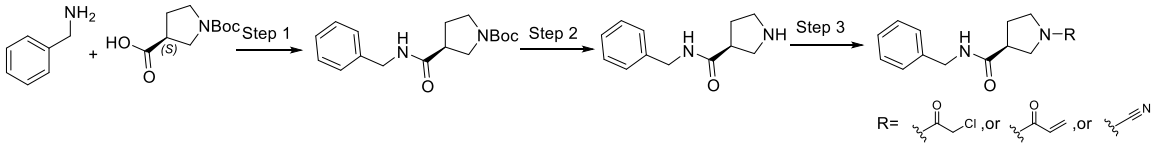


XL_10320_028B (24mg, 44%) ^{1}H NMR (500 MHz, DMSO) δ 7.12 (t, *J* = 7.9 Hz, 2H), 6.80 (m, 3H), 6.57 (t, *J* = 7.2 Hz, 1H), 6.07 (dd, *J* = 16.7, 2.4 Hz, 1H), 5.63 (dd, *J* = 10.4, 2.4 Hz, 1H), 4.23 (d, *J* = 11.0 Hz, 1H), 3.85 (d, *J* = 13.5 Hz, 1H), 3.30 (m, 3H), 2.94 (s, 3H), 2.93 – 2.83 (m, 1H), 1.46 (m, 4H). LC/MS (ESI) *m/z* 274.97; $[M+H]^+$ calcd for $C_{16}H_{23}N_2O_2^+$: 275.18



XL_10320_028C (32mg, 65%) ^{1}H NMR (500 MHz, DMSO) δ 7.14 (t, *J* = 8.0 Hz, 2H), 6.80 (d, *J* = 8.1 Hz, 2H), 6.64 – 6.48 (m, 1H), 3.68 (d, *J* = 12.7 Hz, 1H), 3.30 (s, 1H), 3.27 (s, 1H), 3.23 (dd, *J* = 12.3, 2.7 Hz, 1H), 3.17 (dd, *J* = 12.4, 2.8 Hz, 1H), 2.99 – 2.91 (m, 4H), 1.64 (td, *J* = 13.0, 5.1 Hz, 1H), 1.57 – 1.45 (m, 1H), 1.45 – 1.32 (m, 2H). LC/MS (ESI) *m/z* 245.88; $[M+H]^+$ calcd for $C_{14}H_{20}N_3O^+$: 246.16

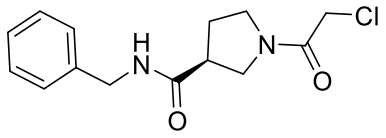
<u>Synthesis of 029A, 029B, 029C</u>:



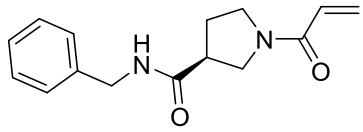
Step 1: The synthesis was preformed according to General Procedure 1 with benzylamine (0.18mL, 1.64mmol) and (*S*)-1-(tert-butoxycarbonyl)pyrrolidine-3-carboxylic acid (0.43g, 1.9mmol). 0.38g desired compound (tert-butyl (*S*)-3-(benzylcarbamoyl)pyrrolidine-1-carboxylate) was obtained (76%). LC/MS (ESI) *m/z* 249.17 (M+H−*t*-butyl); $[M+H]^+$; calcd for $C_{17}H_{25}N_2O_3^+$: 305.19

Step 2: The synthesis was performed according to the General Procedure 1 tert-butyl (*S*)-3-(benzylcarbamoyl)pyrrolidine-1-carboxylate (0.38g, 1.2mmol). 0.30g (*S*)-N-benzylpyrrolidine-3-carboxamide (quant.)

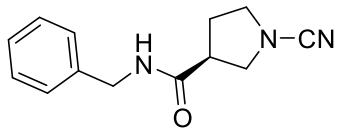
Step 3: The synthesis was performed according to the General Procedure 1 with (S)-N-benzylpyrrolidine-3-carboxamide (0.1g, 0.4mmol) and chloroacetyl chloride (40.0 µL, 0.5mmol), or acryloyl chloride (40.0 µL, 0.5mmol), or cyanogen bromide (0.53g, 0.5mmol).



XL_10320_029A (45mg, 40%) ^{1}H NMR (500 MHz, DMSO) δ 8.55 (dt, *J* = 11.8, 5.8 Hz, 1H), 7.38 – 7.28 (m, 2H), 7.25 (m, Hz, 3H), 4.32 – 4.28 (m, 4H), 3.71 (dd, *J* = 10.1, 7.9 Hz, 0.5H), 3.67 – 3.53 (m, 1.5H), 3.54 – 3.40 (m, 1H), 3.42 – 3.36 (m, 0.5H), 3.32 – 3.25 (m, 0.5H), 3.15 – 3.03 (m, 0.5H), 2.99 (p, *J* = 7.7 Hz, 0.5H), 2.22 – 1.85 (m, 2H). (mixture of rotamers) LC/MS (ESI) *m/z* 280.87; $[M+H]^+$ calcd for $C_{14}H_{18}ClN_2O_2^+$: 281.11

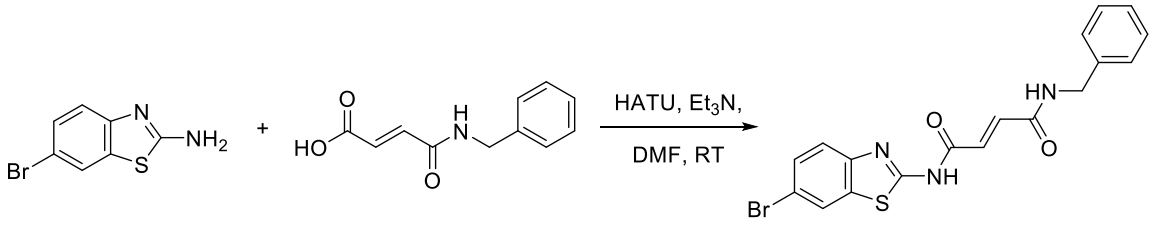


XL_10320_029B (61mg, 59%) ^{1}H NMR (500 MHz, DMSO) δ 8.54 (d, *J* = 5.2 Hz, 1H), 7.37 – 7.30 (m, 2H), 7.25 (d, *J* = 6.8 Hz, 3H), 6.57 (ddd, *J* = 16.7, 10.3, 6.4 Hz, 1H), 6.13 (dt, *J* = 16.8, 2.0 Hz, 1H), 5.78 – 5.52 (m, 1H), 4.29 (d, *J* = 7.3 Hz, 2H), 3.72 – 3.65 (m, 0.5H), 3.65 – 3.57 (m, 1H), 3.53 (m, 1H), 3.43 (dd, *J* = 12.1, 7.2 Hz, 0.5H), 3.38 – 3.30 (m, 0.5H), 3.09 (dt, *J* = 15.3, 7.6 Hz, 0.5H), 2.99 (dt, *J* = 15.1, 7.6 Hz, 0.5H), 2.17 – 2.02 (m, 1.5H), 1.94 (m, 0.5H). (mixture of rotamers) LC/MS (ESI) *m/z* 258.87; $[M+H]^+$ calcd for $C_{15}H_{19}N_2O_2^+$: 259.14



XL_10320_029C (88mg, 96%) ^{1}H NMR (500 MHz, DMSO) δ 8.52 (dt, *J* = 32.7, 5.7 Hz, 1H), 7.38 – 7.29 (m, 2H), 7.29 – 7.21 (m, 3H), 4.39 – 4.19 (m, 2H), 3.55 (dd, *J* = 9.2, 7.9 Hz, 2H), 3.47 – 3.39 (m, 2H), 3.39 – 3.32 (m, 1H), 3.10 – 2.98 (m, 1H), 2.17 – 2.02 (m, 1H), 2.03 – 1.92 (m, 1H). LC/MS (ESI) *m/z* 230.08; $[M+H]^+$ calcd for $C_{13}H_{16}N_3O^+$: 230.13

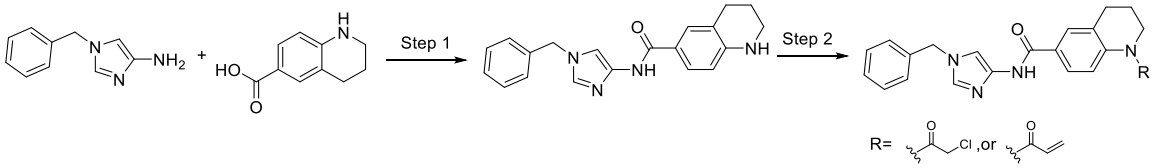
Synthesis of XL_10320_061A:



XL_10320_061

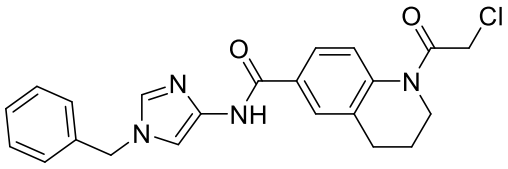
6-bromobenzo[d]thiazol-2-amine (0.046g, 0.2mmol), (E)-4-(benzylamino)-4-oxobut-2-enoic acid (0.05g, 0.24mmol) Et_3N (0.28mL, 2.0mmol), and HATU (0.15g, 0.4mmol) were added sequentially to 3mL anhydrous DMF. The reaction mixture was stirred overnight at room temperature, and purified by flash chromatography (hexanes/EtOAc/MeOH) and preparative HPLC (MeOH/H_2O w/ 0.0425% TFA) to afford XL_10320_061A (14mg, 17%): ^1H NMR (500 MHz, DMSO) δ 12.91 (s, 1H), 9.10 (t, *J* = 5.9 Hz, 1H), 8.30 (d, *J* = 2.0 Hz, 1H), 7.73 (d, *J* = 8.6 Hz, 1H), 7.60 (dd, *J* = 8.6, 2.1 Hz, 1H), 7.35 (dd, *J* = 10.1, 4.6 Hz, 1H), 7.28 (dd, *J* = 16.4, 7.7 Hz, 2H), 7.25 – 7.15 (m, 2H), 6.54 (s, 2H), 4.42 (d, *J* = 5.9 Hz, 2H). LC/MS (ESI) *m/z* 415.77; $[M+H]^+$ calcd for $C_{18}H_{15}BrN_3O_2S^+$: 416.01

Synthesis of XL_10320_054A, 054B:

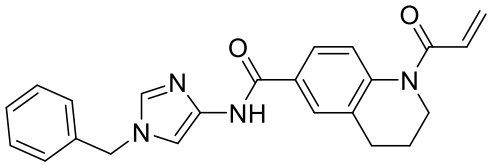


Step 1: The synthesis was preformed according to General Procedure 1 with 1-benzyl-1H-imidazol-4-amine (0.25g, 1.5mmol) and 1,2,3,4-tetrahydroquinoline-6-carboxylic acid (0.38g, 1.8mmol). 0.17g desired compound (*N*-(1-benzyl-1H-imidazol-4-yl)-1,2,3,4-tetrahydroquinoline-6-carboxamide) was obtained (34%). LC/MS (ESI) *m/z* 332.87; $[M+H]^+$ calcd for $C_{20}H_{21}N_4O^+$: 333.17

Step 2: The synthesis was performed according to the General Procedure 1 with N-(1-benzyl-1H-imidazol-4-yl)-1,2,3,4-tetrahydroquinoline-6-carboxamide (0.087g, 0.26mmol) and chloroacetyl chloride (25.0 µL, 0.32mmol), or acryloyl chloride (25.0 µL, 0.32mmol).

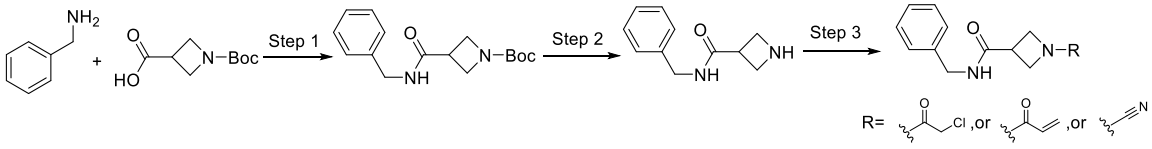


XL_10320_054A (42mg, 39%) ^{1}H NMR (500 MHz, DMSO) δ 10.98 (s, 1H), 8.33 (s, 1H), 7.83 (s, 1H), 7.80 (d, *J* = 8.5 Hz, 1H), 7.75 (s, 1H), 7.51 (s, 1H), 7.45 – 7.33 (m, 5H), 5.30 (s, 2H), 4.62 (s, 2H), 3.74 (dd, *J* = 14.8, 8.6 Hz, 2H), 2.78 (dd, *J* = 22.6, 16.0 Hz, 2H), 2.01 – 1.87 (m, 2H). LC/MS (ESI) *m/z* 409.37; $[M+H]^+$ calcd for $C_{22}H_{22}ClN_4O_2^+$: 409.14



XL_10320_054B (41mg, 41%) ^{1}H NMR (500 MHz, DMSO) δ 10.95 (s, 1H), 8.30 (s, 1H), 7.85 (d, *J* = 1.7 Hz, 1H), 7.80 (dd, *J* = 8.5, 2.0 Hz, 1H), 7.50 (d, *J* = 1.3 Hz, 1H), 7.46 – 7.26 (m, 6H), 6.59 (dd, *J* = 16.7, 10.3 Hz, 1H), 6.26 (dd, *J* = 16.8, 2.1 Hz, 1H), 5.77 (dd, *J* = 10.3, 2.1 Hz, 1H), 5.30 (s, 2H), 3.77 (t, *J* = 6.4 Hz, 2H), 2.79 (t, *J* = 6.5 Hz, 2H), 1.91 (p, *J* = 6.5 Hz, 2H). LC/MS (ESI) *m/z* 387.37; $[M+H]^+$ calcd for $C_{23}H_{23}N_4O_2^+$: 387.18

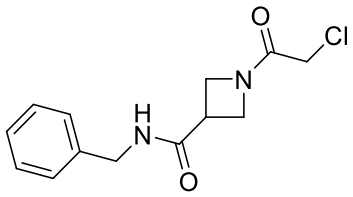
<u>Synthesis of XL_10320_062A, 062B, 062C</u>:



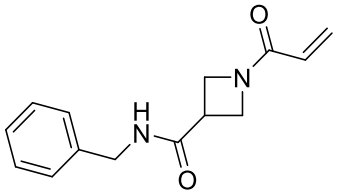
Step 1: The synthesis was preformed according to General Procedure 1 with benzylamine (0.22mL, 2.0mmol) and 1-(tert-butoxycarbonyl)azetidine-3-carboxylic acid (0.6g, 3.0mmol). 0.37g desired compound (tert-butyl 3-(benzylcarbamoyl)azetidine-1-carboxylate) was obtained (64%). LC/MS (ESI) *m/z* 291.17; $[M+H]^+$ calcd for $C_{16}H_{23}N_2O_3^+$: 291.17

Step 2: The synthesis was performed according to the General Procedure 1 tert-butyl 3-(benzylcarbamoyl)azetidine-1-carboxylate (0.37g, 1.2mmol). 0.23g *N*-benzylazetidine-3-carboxamide (quant.)

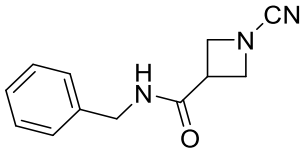
Step 3: The synthesis was performed according to the General Procedure 1 *N*-benzylazetidine-3-carboxamide (0.08g, 0.4mmol) and chloroacetyl chloride (40.0 µL, 0.5mmol), or acryloyl chloride (40.0 µL, 0.5mmol), or cyanogen bromide (0.53g, 0.5mmol).



XL_10320_062A (31mg, 29%) ^{1}H NMR (500 MHz, DMSO) δ 8.54 (t, *J* = 5.4 Hz, 1H), 7.38 – 7.29 (m, 2H), 7.29 – 7.20 (m, 3H), 4.32 (dd, *J* = 10.5, 7.3 Hz, 3H), 4.27 – 4.21 (m, 1H), 4.18 – 4.09 (m, 2H), 4.03 (t, *J* = 9.2 Hz, 1H), 3.93 (dd, *J* = 9.4, 6.0 Hz, 1H), 3.45 – 3.38 (m, 1H). LC/MS (ESI) *m/z* 267.17; $[M+H]^+$ calcd for $C_{13}H_{16}ClN_2O_2^+$: 267.09

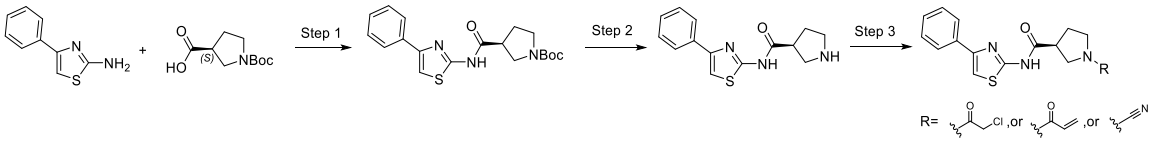


XL_10320_062B (36mg, 37%) ^{1}H NMR (500 MHz, DMSO) δ 8.55 (t, *J* = 5.7 Hz, 1H), 7.38 – 7.29 (m, 2H), 7.29 – 7.20 (m, 3H), 6.31 (dd, *J* = 17.0, 10.3 Hz, 1H), 6.10 (dd, *J* = 17.0, 2.2 Hz, 1H), 5.67 (dd, *J* = 10.3, 2.2 Hz, 1H), 4.39 – 4.28 (m, 3H), 4.23 (dd, *J* = 8.3, 5.9 Hz, 1H), 4.04 (t, *J* = 9.3 Hz, 1H), 3.93 (dd, *J* = 9.7, 5.9 Hz, 1H). (one proton overlaps with HDO peak) LC/MS (ESI) *m/z* 245.28; $[M+H]^+$ calcd for $C_{14}H_{17}N_2O_2^+$: 245.13



XL_10320_062C (51mg, 56%) ^{1}H NMR (500 MHz, DMSO) δ 8.50 (t, *J* = 5.4 Hz, 1H), 7.37 – 7.29 (m, 2H), 7.29 – 7.21 (m, 3H), 4.30 (d, *J* = 5.9 Hz, 2H), 4.26 (dd, *J* = 8.8, 7.5 Hz, 2H), 4.21 – 4.16 (m, 2H), 3.49 (tt, *J* = 8.9, 6.4 Hz, 1H). LC/MS (ESI) *m/z* 216.18; $[M+H]^+$ calcd for $C_{12}H_{14}N_3O^+$: 216.11

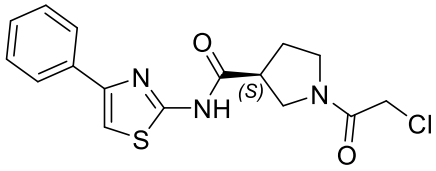
<u>Synthesis of XL_10320_064A, 064B, 064C:</u>



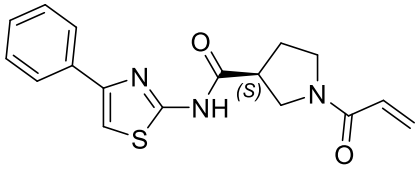
Step 1: The synthesis was preformed according to General Procedure 1 with 4-phenylthiazol-2-amine (0.26g, 1.5mmol) and (S)-1-(tert-butoxycarbonyl)pyrrolidine-3-carboxylic acid (0.38g, 1.8mmol). 0.6g desired compound (tert-butyl (*S*)-3-((4-phenylthiazol-2-yl)carbamoyl)pyrrolidine-1-carboxylate) was obtained (quant.). LC/MS (ESI) *m/z* 274.17 (M+H−Boc); $[M+H]^+$ calcd for $C_{19}H_{24}N_3O_3S^+$: 374.15

Step 2: The synthesis was performed according to the General Procedure 1 with tert-butyl (*S*)-3-((4-phenylthiazol-2-yl)carbamoyl)pyrrolidine-1-carboxylate (0.6g, 1.5mmol). 0.40g (*S*)-N-(4-phenylthiazol-2-yl)pyrrolidine-3-carboxamide (quant.)

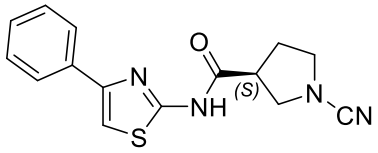
Step 3: The synthesis was performed according to the General Procedure 1 with (*S*)-N-(4-phenylthiazol-2-yl)pyrrolidine-3-carboxamide (0.13g, 0.5mmol) and chloroacetyl chloride (0.08mL, 1.0mmol), or acryloyl chloride (0.08mL, 1.0mmol), or cyanogen bromide (0.11g, 1.0mmol).



XL_10320_064A (12mg, 7%) ^{1}H NMR (500 MHz, DMSO) δ 12.48 (d, *J* = 7.6 Hz, 1H), 7.90 (d, *J* = 7.4 Hz, 2H), 7.65 (d, *J* = 2.5 Hz, 1H), 7.44 (t, *J* = 7.7 Hz, 2H), 7.33 (t, *J* = 7.3 Hz, 1H), 4.33 (d, *J* = 2.2 Hz, 2H), 3.80 (dd, *J* = 10.4, 7.8 Hz, 1H), 3.74 – 3.60 (m, 2H), 3.53 (ddd, *J* = 17.8, 11.3, 6.5 Hz, 3H), 2.33 – 2.00 (m, 3H).(Conformational isomers were observed. Protons overlaps with HDO peak) LC/MS (ESI) *m/z* 350.17; $[M+H]^+$ calcd for $C_{16}H_{17}ClN_3O_2S^+$: 350.07

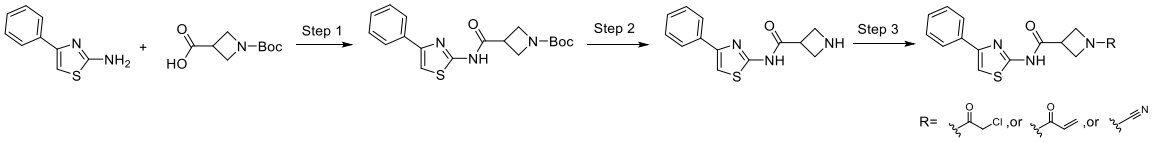


XL_10320_064B (8mg, 5%) ^{1}H NMR (500 MHz, DMSO) δ 12.47 (s, 1H), 7.95 – 7.85 (m, 2H), 7.65 (d, *J* = 1.8 Hz, 1H), 7.44 (t, *J* = 7.7 Hz, 2H), 7.33 (t, *J* = 7.4 Hz, 1H), 6.60 (ddd, *J* = 16.8, 10.3, 3.5 Hz, 1H), 6.15 (dd, *J* = 16.8, 2.4 Hz, 1H), 5.69 (dd, *J* = 10.3, 2.4 Hz, 1H), 3.87 (dd, *J* = 10.3, 7.8 Hz, 1H), 3.72 (ddd, *J* = 12.2, 9.2, 5.4 Hz, 2H), 3.41 (dd, *J* = 13.4, 6.4 Hz, 2H), 3.32 (dt, *J* = 14.0, 7.1 Hz, 1H), 2.33 – 1.99 (m, 3H). (Conformational isomvers were observed) LC/MS (ESI) *m/z* 328.17; $[M+H]^+$ calcd for $C_{17}H_{18}N_3O_2S^+$: 328.11



XL_10320_064C (15mg, 10%) ^{1}H NMR (500 MHz, DMSO) δ 12.45 (d, *J* = 25.3 Hz, 1H), 7.97 – 7.84 (m, 2H), 7.65 (d, *J* = 7.8 Hz, 1H), 7.44 (t, *J* = 7.7 Hz, 2H), 7.33 (t, *J* = 7.3 Hz, 1H), 3.64 (dd, *J* = 9.6, 7.8 Hz, 1H), 3.61 – 3.52 (m, 1H), 3.52 – 3.40 (m, 2H), 3.40 – 3.31 (m, 1H), 2.22 (td, *J* = 13.2, 7.4 Hz, 1H), 2.10 (dt, *J* = 19.7, 7.0 Hz, 1H). LC/MS (ESI) *m/z* 298.97; $[M+H]^+$ calcd for $C_{15}H_{15}N_4OS^+$: 299.10

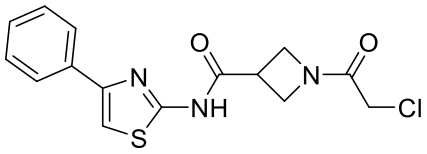
Synthesis of XL_10320_065A, 065B, and 065C:



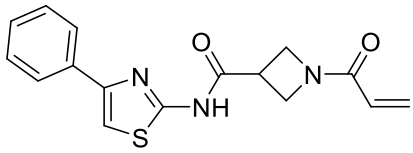
Step 1: The synthesis was preformed according to General Procedure 1 with 4-phenylthiazol-2-amine (0.26g, 1.5mmol) and 1-(tert-butoxycarbonyl)azetidine-3-carboxylic acid (0.36g, 1.8mmol). 0.52g desired compound (tert-butyl 3-((4-phenylthiazol-2-yl)carbamoyl)azetidine-1-carboxylate) was obtained (96%). LC/MS (ESI) *m/z* 260.17 (M+H−Boc); $[M+H]^+$ calcd for $C_{18}H_{22}N_3O_3S^+$: 360.14

Step 2: The synthesis was performed according to the General Procedure 1 with tert-butyl 3-((4-phenylthiazol-2-yl)carbamoyl)azetidine-1-carboxylate (0.52g, 1.5mmol). 0.46g N-(4-phenylthiazol-2-yl)azetidine-3-carboxamide was obtained (quant.)

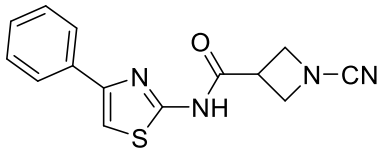
Step 3: The synthesis was performed according to the General Procedure 1 with N-(4-phenylthiazol-2-yl)azetidine-3-carboxamide (0.15g, 0.6mmol) and chloroacetyl chloride (0.08mL, 1.0mmol), or acryloyl chloride (0.08mL, 1.0mmol), or cyanogen bromide (0.11g, 1.0mmol).



XL_10320_065A (33mg, 16%) ^{1}H NMR (500 MHz, DMSO) δ (ppm): 12.43 (s, 1H), 7.94 – 7.84 (m, 2H), 7.66 (s, 1H), 7.43 (dd, *J* = 10.6, 4.8 Hz, 2H), 7.36 – 7.27 (m, 1H), 4.45 – 4.32 (m, 2H), 4.17 (s, 2H), 4.13 – 4.00 (m, 2H), 3.76 – 3.70 (m, 1H). LC/MS (ESI) *m/z* 336.17; $[M+H]^+$ calcd for $C_{15}H_{15}ClN_3O_2S^+$: 336.06

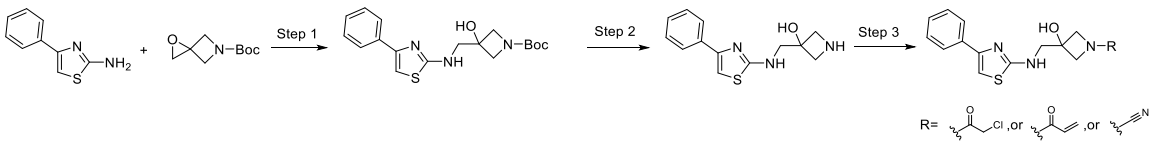


XL_10320_065B (18mg, 10%) ^{1}H NMR (500 MHz, DMSO) δ (ppm): 12.44 (s, 1H), 7.92 – 7.86 (m, 2H), 7.66 (s, 1H), 7.43 (t, *J* = 10.0Hz, 2H), 7.32 (t, *J* = 10.0 Hz, 1H), 6.32 (dd, *J* = 17.0, 10.3 Hz, 1H), 6.11 (dd, *J* = 17.0, 2.2 Hz, 1H), 5.68 (dd, *J* = 10.3, 2.2 Hz, 1H), 4.41 (t, *J* = 10.0 Hz, 1H), 4.35 (dd, *J* = 8.5, 5.8 Hz, 2H), 4.12 (t, *J* = 9.5 Hz, 2H), 4.04 (dd, *J* = 9.9, 5.8 Hz, 2H), 3.72 (tt, *J* = 8.9, 5.7 Hz, 1H). LC/MS (ESI) *m/z* 314.17; $[M+H]^+$ calcd for $C_{16}H_{16}N_3O_2S^+$: 314.10



XL_10320_065C (3mg, 2%) ^{1}H NMR (500 MHz, DMSO) δ (ppm): 12.38 (s, 1H), 7.89 (d, *J* = 7.5 Hz, 2H), 7.66 (s, 1H), 7.43 (t, *J* = 7.6 Hz, 2H), 7.32 (t, *J* = 7.4 Hz, 1H), 4.32 (dt, *J* = 14.0, 7.7 Hz, 5H), 3.85 – 3.70 (m, 1H). LC/MS (ESI) *m/z* 285.07; $[M+H]^+$ calcd for $C_{14}H_{13}N_4OS^+$: 285.08

<u>Synthesis of XL_10320_067A, 067B, 067C:</u>

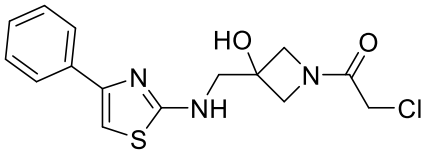


Step 1: The synthesis was performed according to General Procedure 2 with 4-phenylthiazol-2-amine (0.24g, 1.4mmol) and tert-butyl 1-oxa-5-azaspiro[2.3]hexane-5-carboxylate (0.25g, 1.4mmol). 0.17g desired product (tert-butyl 3-hydroxy-3-(((4-phenylthiazol-2-

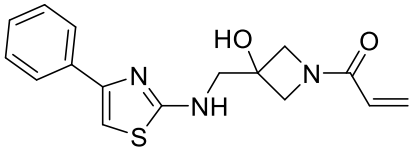
yl)amino)methyl)azetidine-1-carboxylate) was obtained (36%). LC/MS (ESI) *m/z* 361.97; $[M+H]^+$ calcd for $C_{18}H_{24}N_3O_3S^+$: 362.15

Step 2: The synthesis was performed according to the General Procedure 2 with tert-butyl 3-hydroxy-3-(((4-phenylthiazol-2-yl)amino)methyl)azetidine-1-carboxylate (0.12g, 0.3mmol). 0.084g 3-(((4-phenylthiazol-2-yl)amino)methyl)azetidin-3-ol was obtained (quant.)

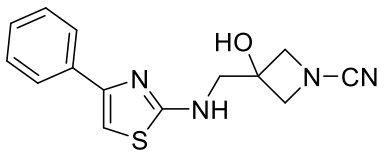
Step 3: The synthesis was performed according to the General Procedure 2 with 3-(((4-phenylthiazol-2-yl)amino)methyl)azetidin-3-ol (0.028g, 0.1mmol), Et_3N (0.077mL, 0.55mmol) and chloroacetyl chloride (9.2µL, 0.11mmol), or acryloyl chloride (9.2µL, 0.11mmol), or cyanogen bromide (0.012g, 0.11mmol).



XL_10320_067A (8mg, 24%) ^{1}H NMR (500 MHz, DMSO) δ (ppm): 8.06 (s, 1H), 7.83 – 7.77 (m, 2H), 7.38 (t, *J* = 7.6 Hz, 2H), 7.29 (t, *J* = 7.3 Hz, 1H), 7.08 (s, 1H), 4.27 (d, *J* = 9.2 Hz, 1H), 4.19 – 4.08 (dd, *J* = 15.0, 5.0 Hz, 2H), 4.02 (m, 2H), 3.72 (d, *J* = 10.2 Hz, 1H), 3.61 (s, 2H). LC/MS (ESI) *m/z* 338.07; $[M+H]^+$ calcd for $C_{15}H_{17}ClN_3O_2S^+$: 338.07

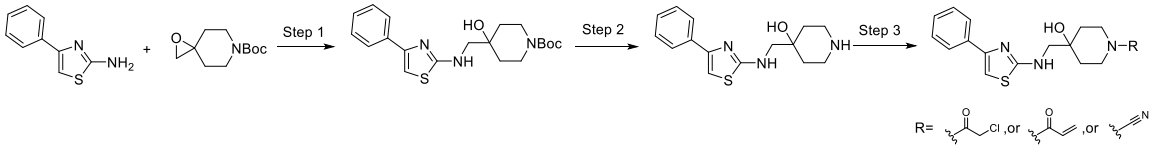


XL_10320_067B (8mg, 25%) ^{1}H NMR (500 MHz, DMSO) δ (ppm): 8.14 (s, 1H), 7.79 (d, *J* = 7.5 Hz, 2H), 7.38 (t, *J* = 7.7 Hz, 2H), 7.29 (t, *J* = 7.3 Hz, 1H), 7.08 (s, 1H), 6.32 (dd, *J* = 16.9, 10.3 Hz, 1H), 6.10 (dd, *J* = 17.0, 1.6 Hz, 1H), 5.65 (dd, *J* = 10.3, 1.6 Hz, 2H), 4.31 (d, *J* = 9.0 Hz, 1H), 4.01 (d, *J* = 9.8 Hz, 2H), 3.74 (d, *J* = 10.4 Hz, 1H), 3.62 (q, *J* = 13.8 Hz, 2H). LC/MS (ESI) *m/z* 315.87; $[M+H]^+$ calcd for $C_{16}H_{18}N_3O_2S^+$: 316.11



XL_10320_067C (15mg, 53%, mixture of rotamers) ^{1}H NMR (500 MHz, DMSO) δ (ppm): 7.84 – 7.77 (m, 2H), 7.43 – 7.36 (m, 2H), 7.29 (m, 1H), 7.11 – 7.04 (m, 1H), 4.19 (d, *J* = 8.5 Hz, 1H), 4.01 (d, *J* = 8.4 Hz, 1H), 3.83 (d, *J* = 9.7 Hz, 2H), 3.64-3.59 (m, 3H). LC/MS (ESI) *m/z* 286.97; $[M+H]^+$ calcd for $C_{14}H_{15}N_4OS^+$: 287.10

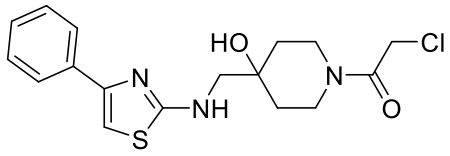
Synthesis of XL_10320_085A, 085B, 085C:



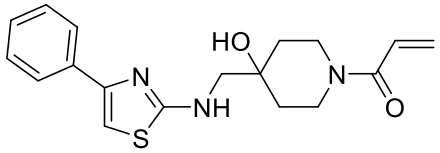
Step 1: The synthesis was performed according to General Procedure 2 with 4-phenylthiazol-2-amine (0.23g, 1.3mmol) and tert-butyl 1-oxa-6-azaspiro[2.5]octane-6-carboxylate (0.28g, 1.3mmol). 0.12g desired product (tert-butyl 4-hydroxy-4-(((4-phenylthiazol-2-yl)amino)methyl)piperidine-1-carboxylate) was obtained (24%). LC/MS (ESI) *m/z* 390.37; $[M+H]^+$ calcd for $C_{20}H_{28}N_3O_3S^+$: 390.18

Step 2: The synthesis was performed according to the General Procedure 2 tert-butyl 4-hydroxy-4-(((4-phenylthiazol-2-yl)amino)methyl)piperidine-1-carboxylate (0.12g, 0.3mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.08g 4-(((4-phenylthiazol-2-yl)amino)methyl)piperidin-4-ol was obtained (90%)

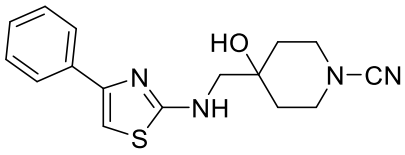
Step 3: The synthesis was performed according to the General Procedure 2 with 4-(((4-phenylthiazol-2-yl)amino)methyl)piperidin-4-ol (0.015g, 0.05mmol), Et_3N (0.077mL, 0.55mmol) and chloroacetyl chloride (5.0µL, 0.06mmol), or acryloyl chloride (5.0µL, 0.06mmol), or cyanogen bromide (0.007g, 0.06mmol).



XL_10320_085A (4mg, 22%) 1H NMR (500 MHz, DMSO) δ 7.80 (m, J = 7.4 Hz, 2H), 7.38 (t, J = 7.6 Hz, 2H), 7.28 (t, J = 7.3 Hz, 1H), 7.05 (s, 1H), 4.42 – 4.31 (m, 2H), 4.07 (d, J = 12.9 Hz, 1H), 3.64 (d, J = 13.3 Hz, 1H), 3.42 – 3.28 (m, 3H), 3.02 (t, J = 10.5 Hz, 1H), 1.56 (tdd, J = 31.9, 20.6, 11.0 Hz, 4H). LC/MS (ESI) m/z 365.87; [M+H]+ calcd for $C_{17}H_{21}N_3O_2S^+$: 366.10

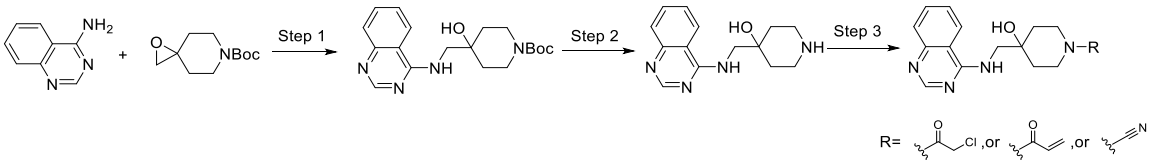


XL_10320_085B (4mg, 23%) 1H NMR (500 MHz, DMSO) δ 7.87 – 7.68 (m, 3H), 7.38 (t, J = 7.7 Hz, 2H), 7.27 (t, J = 7.4 Hz, 1H), 7.04 (s, 1H), 6.81 (dd, J = 16.7, 10.5 Hz, 1H), 6.08 (dd, J = 16.7, 2.5 Hz, 1H), 5.64 (dd, J = 10.4, 2.5 Hz, 1H), 4.15 (d, J = 13.3 Hz, 1H), 3.42 – 3.32 (m, 3H), 3.04 (t, J = 9.9 Hz, 1H), 1.53 (t, J = 10.7 Hz, 4H).). LC/MS (ESI) m/z 343.67; $[M+H]^+$ calcd for $C_{18}H_{22}N_3O_2S$+: 344.14



XL_10320_085C (6mg, 38%) ^{1}H NMR (500 MHz, DMSO) δ (ppm): 7.80 (m, 2H), 7.39 (m, 2H), 7.29 (m, 1H), 7.05 (s, 1H), 3.37 (m, 2H), 3.32 – 3.19 (m, 4H), 1.73 – 1.62 (m, 2H), 1.61 – 1.40 (m, 2H). LC/MS (ESI) *m/z* 314.67; $[M+H]^+$ calcd for $C_{16}H_{19}N_4OS^+$: 315.13

<u>Synthesis of XL_10320_086A, 086B, 086C:</u>

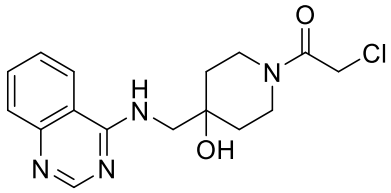


Step 1: The synthesis was performed according to General Procedure 2 with quinazolin-4-amine (0.29g, 2.0mmol) and tert-butyl 1-oxa-6-azaspiro[2.5]octane-6-carboxylate (0.42g, 2.0mmol). 0.35g desired product (tert-butyl 4-hydroxy-4-((quinazolin-4-ylamino)methyl)piperidine-1-carboxylate) was obtained (49%). LC/MS (ESI) *m/z* 359.27; $[M+H]^+$ calcd for $C_{19}H_{27}N_4O_3^+$: 359.21

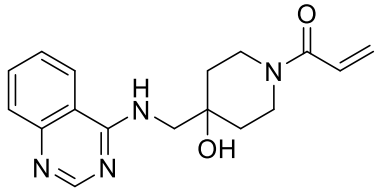
Step 2: The synthesis was performed according to the General Procedure 2 with tert-butyl 4-hydroxy-4-((quinazolin-4-ylamino)methyl)piperidine-1-carboxylate (0.35g, 1.0mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.26g 4-((quinazolin-4-ylamino)methyl)piperidin-4-ol was obtained (quant.)

Step 3: The synthesis was performed according to the General Procedure 2 with 4-((quinazolin-4-ylamino)methyl)piperidin-4-ol (0.086g, 0.3mmol), Et_3N (0.077mL, 0.55mmol) and chloroacetyl

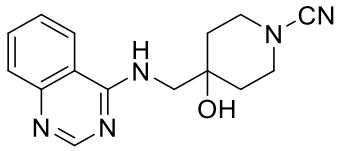
chloride (43.0µL, 0.5mmol), or acryloyl chloride (43.0µL, 0.5mmol), or cyanogen bromide (0.57g, 0.5mmol).



XL_10320_086A (31mg, 31%) ^{1}H NMR (500 MHz, DMSO) δ (ppm): 10.00 (t, *J* = 5.0 Hz, 1H), 8.89 (s, 1H), 8.59 (d, *J* = 8.0 Hz, 1H), 8.09 – 8.01 (m, 1H), 7.80 (m, 2H), 4.39 – 4.29 (d, *J* = 15.0, 20.0 Hz, 2H), 4.07 (d, *J* = 13.1 Hz, 1H), 3.81 (ddd, *J* = 30.7, 13.1, 6.2 Hz, 3H), 3.63 (d, *J* = 13.5 Hz, 1H), 3.41 – 3.26 (m, 1H), 2.99 (m, 1H), 1.72 – 1.43 (m, 4H). LC/MS (ESI) *m/z* 334.87; $[M+H]^+$ calcd for $C_{16}H_{20}ClN_4O_2^+$: 335.13

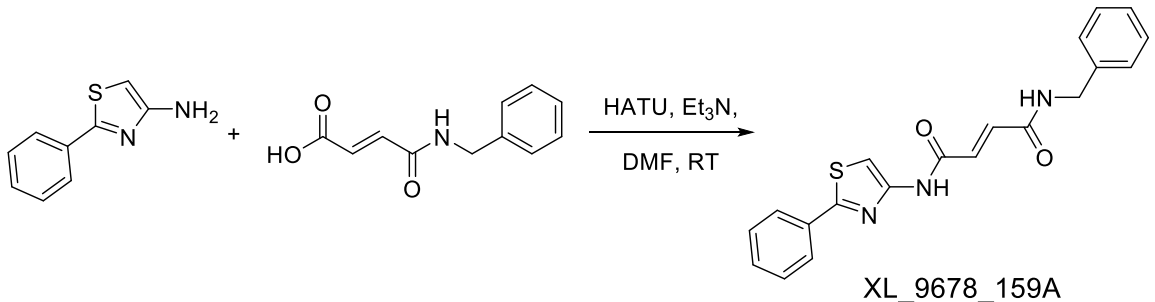


XL_10320_086B (53mg, 57%) ^{1}H NMR (500 MHz, DMSO) δ (ppm): 10.03 (t, *J* = 6.0 Hz, 1H), 8.89 (s, 1H), 8.59 (d, *J* = 7.9 Hz, 1H), 8.05 (t, *J* = 7.8 Hz, 1H), 7.80 (m, 2H), 6.79 (dd, *J* = 16.7, 10.5 Hz, 1H), 6.07 (dd, *J* = 16.7, 2.4 Hz, 1H), 5.64 (dd, *J* = 10.5, 2.4 Hz, 1H), 4.15 (d, *J* = 12.8 Hz, 1H), 3.89 – 3.74 (m, 3H), 3.34 (m, 1H), 3.02 (m, 1H), 1.54 (m, 4H). LC/MS (ESI) *m/z* 312.57; $[M+H]^+$ calcd for $C_{17}H_{21}N_4O_2^+$: 313.17



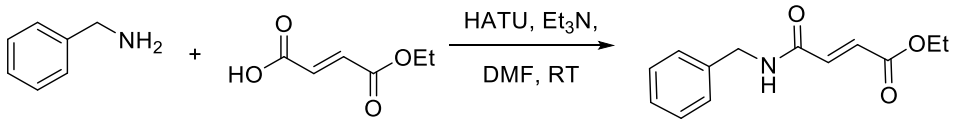
XL_10320_086C (28mg, 33%) ^{1}H NMR (500 MHz, DMSO) δ (ppm): 9.83 (s, 1H), 8.86 (s, 1H), 8.56 (d, *J* = 8.1 Hz, 1H), 8.03 (t, *J* = 7.3 Hz, 1H), 7.78 (m, 2H), 3.80 (d, *J* = 6.2 Hz, 2H), 3.29 – 3.19 (m, 5H), 1.77 – 1.64 (m, 2H), 1.58 (d, *J* = 13.4 Hz, 2H). LC/MS (ESI) *m/z* 283.77; $[M+H]^+$ calcd for $C_{15}H_{18}N_5O^+$: 284.15

Synthesis of XL_9678_159A, 159B, 159C, XL_9872_056, and AC_10180-18:



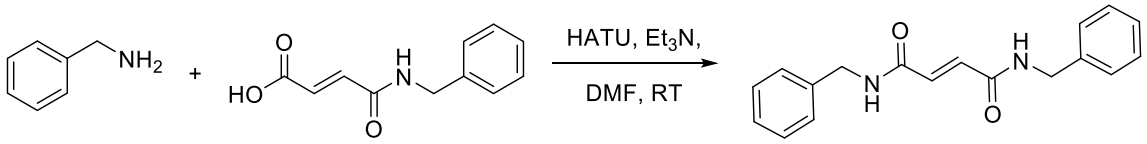
XL_9678_159A

2-phenylthiazol-4-amine (0.08g, 0.5mmol), (E)-4-(benzylamino)-4-oxobut-2-enoic acid (0.12g, 0.6mmol) Et_3N (0.35mL, 2.5mmol), and HATU (0.38g, 1.0mmol) were added sequentially to 3mL anhydrous DMF. The reaction mixture was stirred overnight at room temperature, and purified by flash chromatography (hexanes/EtOAc/MeOH) and preparative HPLC (MeOH/H_2O w/ 0.0425% TFA) to afford **XL_9678_159A** (20mg, 11%) ^1H NMR (500 MHz, DMSO) δ 12.79 (s, 1H), 9.10 (t, *J* = 5.8 Hz, 1H), 7.92 (d, *J* = 7.4 Hz, 2H), 7.71 (s, 1H), 7.44 (t, *J* = 7.6 Hz, 2H), 7.31 (m, 6H), 7.19 (d, *J* = 11.0 Hz, 2H), 4.42 (d, *J* = 5.8 Hz, 2H). LC/MS (ESI) *m/z* 363.87; $[M+H]^+$ calcd for $C_{20}H_{18}N_3O_2S^+$: 364.11



XL_9872_056

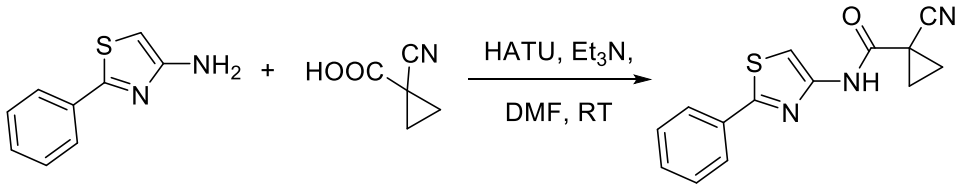
Benzylamine (0.11g, 1.0mmol), monoethyl fumarate (0.14g, 1.0mmol) Et3N (0.7mL, 5.0mmol), and HATU (0.57g, 1.5mmol) were added sequentially to 3mL anhydrous DMF. The reaction mixture was stirred overnight at room temperature, and purified by flash chromatography (hexanes/EtOAc/MeOH) and preparative HPLC (MeOH/H2O w/ 0.0425% TFA) to afford **XL_9872_056** (0.19g, 81%) 1H NMR (500 MHz, DMSO) δ 9.03 (t, J = 5.7 Hz, 1H), 7.37 – 7.31 (m, 2H), 7.31 – 7.24 (m, 3H), 7.07 (d, J = 15.5 Hz, 1H), 6.62 (d, J = 15.5 Hz, 1H), 4.40 (d, J = 5.9 Hz, 2H), 4.19 (q, J = 7.1 Hz, 2H), 1.25 (t, J = 7.1 Hz, 3H). LC/MS (ESI) m/z 234.08; [M+H]+ calcd for $C_{13}H_{16}NO_3^+$: 234.11



AC_10180_18

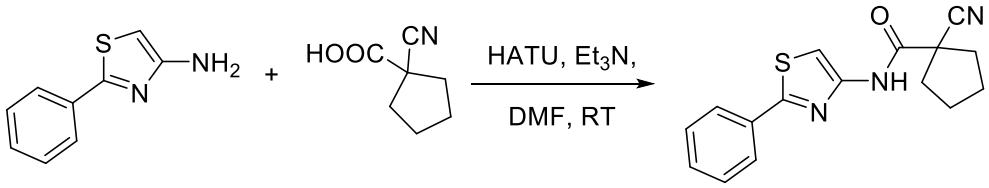
Benzylamine (0.026mL, 0.24mmol), (*E*)-4-(benzylamino)-4-oxobut-2-enoic acid (0.06g, 0.29mmol) Et_3N (0.17mL, 1.2mmol), and HATU (0.14g, 0.36mmol) were added sequentially to 1mL anhydrous DMF. The reaction mixture was stirred overnight at room temperature, and purified by flash chromatography (hexanes/EtOAc/MeOH) and trituation from diethyl ether to afford **AC-10180-18** (22mg, 31%) ^{1}H NMR (500 MHz, DMSO) δ 8.92 (t, *J* = 5.9 Hz, 2H), 7.39 – 7.30 (m, 4H), 7.30 – 7.20 (m, 6H), 6.93 (s, 2H), 4.38 (d, *J* = 6.0 Hz, 4H). LC/MS (ESI) *m/z* 294.80; $[M+H]^+$ calcd for $C_{18}H_{19}N_2O_2^+$: 295.14

2-



XL_9678_159B

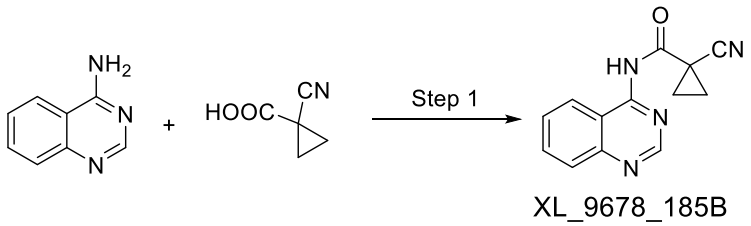
phenylthiazol-4-amine (0.08g, 0.5mmol), 1-cyanocyclopropane-1-carboxylic acid (0.07g, 0.6mmol) Et_3N (0.35mL, 2.5mmol), and HATU (0.38g, 1.0mmol) were added sequentially to 3mL anhydrous DMF. The reaction mixture was stirred overnight at room temperature, and purified by flash chromatography (hexanes/EtOAc/MeOH) and preparative HPLC (MeOH/H_2O w/ 0.0425% TFA) to afford **XL_9678_159B** (40mg, 30%). ^{1}H NMR (500 MHz, DMSO) δ 12.43 (s, 1H), 7.91 (d, *J* = 7.6 Hz, 2H), 7.68 (s, 1H), 7.45 (t, *J* = 7.7 Hz, 2H), 7.35 (t, *J* = 7.3 Hz, 1H), 1.87 (br, 2H), 1.79 (br, 2H). LC/MS (ESI) *m/z* 269.87; $[M+H]^+$ calcd for $C_{14}H_{12}N_3OS^+$: 270.07



XL_9678_159C

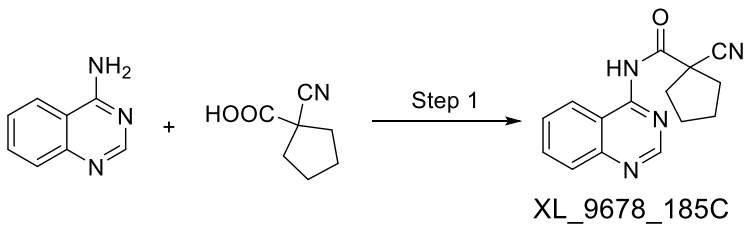
2-phenylthiazol-4-amine (0.08g, 0.5mmol), 1-cyanocyclopentane-1-carboxylic acid (0.08g, 0.6mmol), Et_3N (0.35mL, 2.5mmol), and HATU (0.38g, 1.0mmol) were added sequentially to 3mL anhydrous DMF. The reaction mixture was stirred overnight at room temperature, and purified by flash chromatography (hexanes/EtOAc/MeOH) and preparative HPLC (MeOH/H_2O w/ 0.0425% TFA) to afford **XL_9678_159C** (56mg, 38%) ^{1}H NMR (500 MHz, DMSO) δ 12.92 (s, 1H), 7.93 (d, *J* = 7.5 Hz, 2H), 7.73 (s, 1H), 7.45 (t, *J* = 7.6 Hz, 2H), 7.35 (t, *J* = 7.3 Hz, 1H), 2.46 – 2.26 (m, 4H), 1.80 (m, 4H). LC/MS (ESI) *m/z* 297.87; $[M+H]^+$ calcd for $C_{16}H_{16}N_3OS^+$: 298.10

Synthesis of XL_9678_185B, 185C:



XL_9678_185B

Step 1: quinazolin-4-amine (0.043g, 0.3mmol), 1-cyanocyclopropane-1-carboxylic acid (0.044g, 0.4mmol), Et_3N (0.21mL, 1.5mL), and HATU (0.23g, 0.6mmol) were added into 3mL DMF. The reaction was stirred at room temperature overnight. The reaction mixture was stirred overnight at room temperature, and purified by flash chromatography (hexanes/EtOAc/MeOH) and preparative HPLC (MeOH/H_2O w/ 0.0425% TFA) to afford **XL_9678_185B**. ^{1}H NMR (500 MHz, DMSO) δ 13.60 (s, 1H), 8.48 (d, *J* = 8.0 Hz, 1H), 8.41 (s, 1H), 8.03 – 7.94 (m, 1H), 7.81 (d, *J* = 8.2 Hz, 1H), 7.75 – 7.64 (m, 1H), 1.78 – 1.67 (m, 4H). LC/MS (ESI) *m/z* 238.98; $[M+H]^+$ calcd for $C_{13}H_{11}N_4O^+$: 239.09.

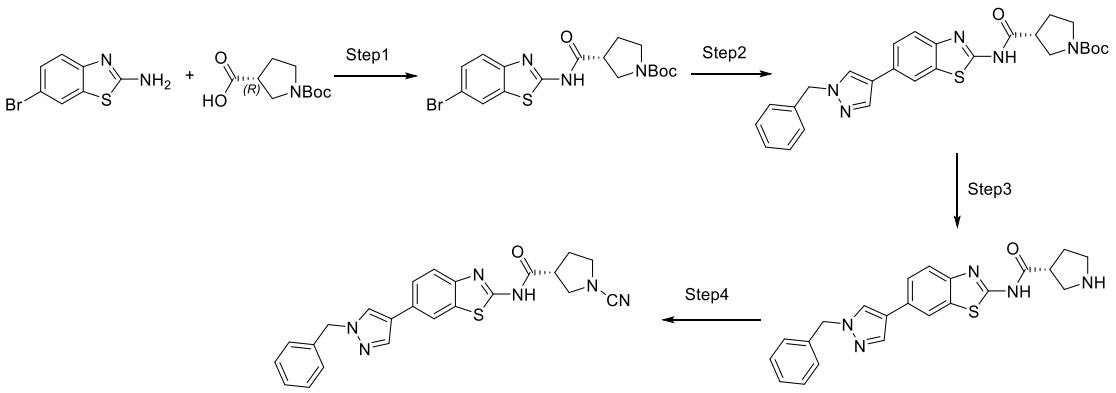


XL_9678_185C

Step 1: quinazolin-4-amine (0.043g, 0.3mmol), 1-cyanocyclopentane-1-carboxylic acid (0.055g, 0.4mmol), Et3N (0.21mL, 1.5mL), and HATU (0.23g, 0.6mmol) were added into 3mL DMF. The reaction was stirred at room temperature overnight. The reaction mixture was stirred overnight

at room temperature, and purified by flash chromatography (hexanes/EtOAc/MeOH) and preparative HPLC (MeOH/H2O w/ 0.0425% TFA) to afford **XL_9678_185C** (2mg, 3%). 1H NMR (500 MHz, DMSO) δ 8.42 (br, 2H), 8.00 (t, J = 7.2 Hz, 1H), 7.85 (d, J = 6.7 Hz, 1H), 7.72 (t, J = 7.6 Hz, 1H), 2.37 (m, 2H), 2.25 (m, 2H), 1.92 – 1.78 (m, 4H). LC/MS (ESI) *m/z* 267.10; $[M+H]^+$ calcd for $C_{15}H_{15}N_4O^+$: 267.12.

Synthesis of XL_9872_111B:



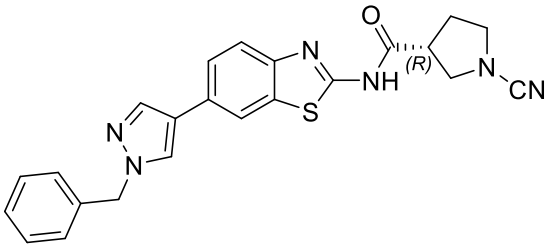
Step 1: 6-bromobenzo[d]thiazol-2-amine (0.41g, 1.8mmol), (R)-1-(tert-butoxycarbonyl)pyrrolidine-3-carboxylic acid (0.47g, 2.2mmol), Et_3N (1.2mL, 9.0mmol) and HATU (1.03g, 2.7mmol) were added sequentially in anhydrous DMF (5mL). The mixture was stirred at room temperature overnight. The mixture was then diluted with EtOAc (50mL), and washed with brine (30mL×2) to remove excess DMF. Organic layer was dried over anhydrous sodium sulfate (Na_2SO_4), filtered, and concentrated under reduced pressure. The crude material was then purified by flash column chromatography (hexanes/EtOAc/MeOH) to afford 0.72g (94%) material. LC/MS (ESI) *m/z* 426.27; $[M+H]^+$ calcd for $C_{17}H_{21}BrN_3O_3S^+$: 426.05.

Step 2: The isolated product tert-butyl (R)-3-((6-bromobenzo[d]thiazol-2-yl)carbamoyl)pyrrolidine-1-carboxylate from step 1 (0.064g, 0.15mmol) was dissolved in 1,4-dioxane and H_2O (4mL, 3:1). Into the solution were added 1-benzyl-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-1H-pyrazole (0.09g, 0.45mmol), potassium carbonate (0.062g, 0.45mmol) and $Pd(PPh_3)_4$ (0.035g, 0.03mmol). The mixture was degassed by bubbling through N_2 for 10min before heating up to 95°C and stirred at this temperature for 2-8 hours. The reaction was then cooled down to room temperature and diluted with EtOAc (50mL). The organic phase was

washed with saturated ammonium chloride (30mL×2). Aqueous layer was then extracted with more EtOAc (50mL). Combined organic layers were washed with brine, dried over anhydrous sodium sulfate (Na_2SO_4), filtered, and concentrated under reduced pressure to afford crude material, which was then purified by flash chromatography (hexanes/EtOAc/MeOH) to afford 0.01g product (13%) LC/MS (ESI) *m/z* 503.88; $[M+H]^+$ calcd for $C_{27}H_{30}N_5O_3S^+$: 504.21.

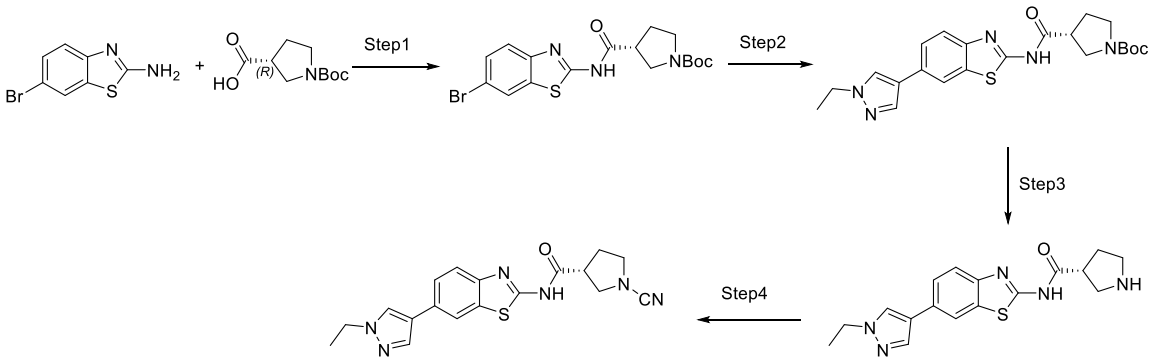
Step 3: Products from last step were dissolved in DCM (1mL) and treated with TFA (1mL). The mixtures were stirred at room temperature until the tert-butyloxycarbonyl protecting group was cleaved tracking by UPLC-MS. The mixture was concentrated and flushed by flash column chromatography (EtOAc/MeOH/0.5%Et_3N).

Step 4: Products from the last step (0.008g, 0.02mmol) were dissolved in DCM (2mL) with Et_3N (14µL, 0.1mmol) at 0°C. Cyanogen bromide 3M solution in DCM (13µL, 0.04mmol) was then added. The mixture was then stirred at 0°C for 1 hour, and directly purified by flash chromatography (hexanes/EtOAc/MeOH) followed by preparative HPLC (MeOH or CH_3CN/H_2O with 0.0425% TFA) to afford the target product.



XL_9872_111B (4mg, 47%) ^{1}H NMR (500 MHz, DMSO) δ 12.52 (s, 1H), 8.32 (s, 1H), 8.19 (s, 1H), 7.97 (s, 1H), 7.71 (d, *J* = 8.4 Hz, 1H), 7.65 (d, *J* = 8.4 Hz, 1H), 7.37 (m, 2H), 7.30 (m, 3H), 5.36 (s, 2H), 3.69 – 3.62 (m, 1H), 3.58 (dd, *J* = 9.6, 6.1 Hz, 1H), 3.46 (m, 2H), 3.41 – 3.34 (m, 1H), 2.23 (dt, *J* = 13.1, 7.3 Hz, 1H), 2.11 (dt, *J* = 19.7, 7.0 Hz, 1H). LC/MS (ESI) *m/z* 428.87; $[M+H]^+$ calcd for $C_{23}H_{21}N_6OS^+$: 429.15

Synthesis of XL_9872_111F:



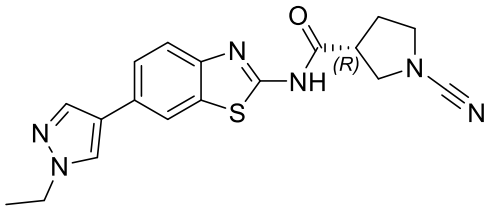
Step 1: 6-bromobenzo[d]thiazol-2-amine (0.41g, 1.8mmol), (R)-1-(tert-butoxycarbonyl)pyrrolidine-3-carboxylic acid (0.47g, 2.2mmol), Et3N (1.2mL, 9.0mmol) and HATU (1.03g, 2.7mmol) were added sequentially in anhydrous DMF (5mL). The mixture was stirred at room temperature overnight. The mixture was then diluted with EtOAc (50mL), and washed with brine (30mL×2) to remove excess DMF. Organic layer was dried over anhydrous sodium sulfate (Na_2SO_4), filtered, and concentrated under reduced pressure. The crude material was then purified by flash column chromatography (hexanes/EtOAc/MeOH) to afford 0.72g (94%) material. LC/MS (ESI) *m/z* 426.27; $[M+H]^+$ calcd for $C_{17}H_{21}BrN_3O_3S^+$: 426.05.

Step 2: The isolated product tert-butyl (R)-3-((6-bromobenzo[d]thiazol-2-yl)carbamoyl)pyrrolidine-1-carboxylate from step 1 (0.064g, 0.15mmol) was dissolved in 1,4-dioxane and H_2O (4mL, 3:1). Into the solution were added 1-ethyl-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-1H-pyrazole (0.1g, 0.45mmol), potassium carbonate (0.062g, 0.45mmol) and $Pd(PPh_3)_4$ (0.035g, 0.03mmol). The mixture was degassed by bubbling through N_2 for 10min before heating up to 95°C and stirred at this temperature for 2-8 hours. The reaction was then cooled down to room temperature and diluted with EtOAc (50mL). The organic phase was washed with saturated ammonium chloride (30mL×2). Aqueous layer was then extracted with more EtOAc (50mL). Combined organic layers were washed with brine, dried over anhydrous sodium sulfate (Na_2SO_4), filtered, and concentrated under reduced pressure to afford crude material, which was then purified by flash chromatography (hexanes/EtOAc/MeOH) to afford 0.05g product (76%). LC/MS (ESI) *m/z* 441.88; $[M+H]^+$ calcd for $C_{22}H_{28}N_5O_3S^+$: 442.19.

Step 3: Products from last step were dissolved in DCM (1mL) and treated with TFA (1mL). The mixtures were stirred at room temperature until the tert-butyloxycarbonyl protecting group was

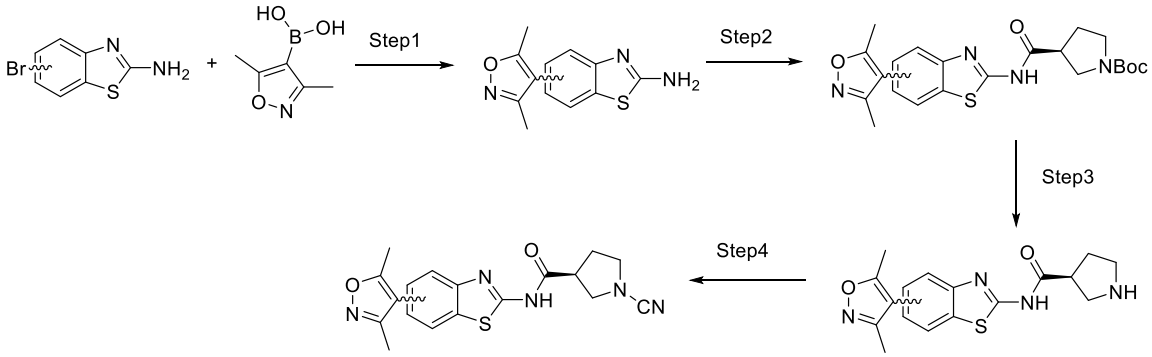
cleaved tracking by UPLC-MS. The mixture was concentrated and flushed by flash column chromatography (EtOAc/MeOH/0.5%Et_3N).

Step 4: Products from the last step (0.041g, 0.12mmol) were dissolved in DCM (2mL) with Et_3N (84µL, 0.6mmol) at 0°C. Cyanogen bromide 3M solution in DCM (60µL, 0.24mmol) was then added. The mixture was then stirred at 0°C for 1 hour, and directly purified by flash chromatography (hexanes/EtOAc/MeOH) followed by preparative HPLC (MeOH or CH_3CN/H_2O with 0.0425% TFA) to afford the target product.



XL_9872_111F (10mg, 23%) ^{1}H NMR (500 MHz, DMSO) δ 12.53 (s, 1H), 8.23 (s, 1H), 8.19 (d, *J* = 1.5 Hz, 1H), 7.92 (d, *J* = 0.6 Hz, 1H), 7.72 (d, *J* = 8.4 Hz, 1H), 7.68 – 7.64 (dd, *J* = 1.7, 8.5 Hz, 1H), 4.21 – 4.11 (q, *J* = 7.4 Hz, 2H), 3.65 (dd, *J* = 9.6, 7.7 Hz, 1H), 3.59 (dd, *J* = 9.6, 6.0 Hz, 1H), 3.52 – 3.42 (m, 2H), 3.42 – 3.35 (m, 1H), 2.24 (dt, *J* = 13.4, 7.5 Hz, 1H), 2.11 (dt, *J* = 14.3, 6.9 Hz, 1H), 1.42 (t, *J* = 7.3 Hz, 3H). LC/MS (ESI) *m/z* 366.87; $[M+H]^+$ calcd for $C_{18}H_{19}N_6OS^+$: 367.13

<u>Synthesis of XL_9872_106A, 106B, 106C</u>:



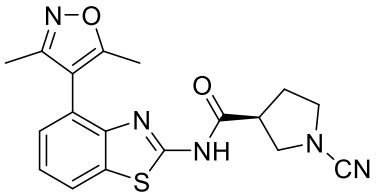
Step 1: 2-amino-4-bromobenzothiazole (0.69g, 3.0mmol), 2-amino-5-bromobenzothiazole (0.69g, 3.0mmol), or 2-amino-7-bromobenzothiazole (0.69g, 3.0mmol) were used as starting material in the synthesis described in Step 1 of General Procedure 4 to afford desired products (4-(3,5-dimethylisoxazol-4-yl)benzo[d]thiazol-2-amine: 0.6g (82%) LC/MS (ESI) *m/z* 245.98; $[M+H]^+$ calcd

for $C_{12}H_{12}N_3OS^+$: 246.07; 5-(3,5-dimethylisoxazol-4-yl)benzo[d]thiazol-2-amine: 0.58g (79%) LC/MS (ESI) *m/z* 245.98; $[M+H]^+$ calcd for $C_{12}H_{12}N_3OS^+$: 246.07; 7-(3,5-dimethylisoxazol-4-yl)benzo[d]thiazol-2-amine: 0.39g (53%) LC/MS (ESI) *m/z* 245.88; $[M+H]^+$ calcd for $C_{12}H_{12}N_3OS^+$: 246.07.

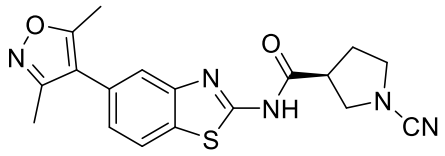
Step 2: The isolated products from Step 1 (0.05g, 0.2mmol) were used in Step 2 described in General Procedure 4 to afford desired products (tert-butyl (S)-3-((4-(3,5-dimethylisoxazol-4-yl)benzo[d]thiazol-2-yl)carbamoyl)pyrrolidine-1-carboxylate: 0.092g (84%); LC/MS (ESI) m/z 443.08; [M+H]+ calcd for $C_{22}H_{27}N_4O_4S^+$: 443.17; tert-butyl (S)-3-((5-(3,5-dimethylisoxazol-4-yl)benzo[d]thiazol-2-yl)carbamoyl)pyrrolidine-1-carboxylate: 0.19g (over 100%) LC/MS (ESI) m/z 443.08; [M+H]+ calcd for $C_{22}H_{27}N_4O_4S^+$: 443.17; tert-butyl (S)-3-((7-(3,5-dimethylisoxazol-4-yl)benzo[d]thiazol-2-yl)carbamoyl)pyrrolidine-1-carboxylate: 0.20g (over 100%) LC/MS (ESI) *m/z* 387.27 (M+H−*t*-Butyl); $[M+H]^+$ calcd for $C_{22}H_{27}N_4O_4S^+$: 443.54).

Step 3: The isolated products from Step 2 were used in Step 3 described in General Procedure 4 to afford desired products ((S)-N-(4-(3,5-dimethylisoxazol-4-yl)benzo[d]thiazol-2-yl)pyrrolidine-3-carboxamide: 0.08g (quant.); (S)-N-(5-(3,5-dimethylisoxazol-4-yl)benzo[d]thiazol-2-yl)pyrrolidine-3-carboxamide: 0.08g (quant.); (S)-N-(7-(3,5-dimethylisoxazol-4-yl)benzo[d]thiazol-2-yl)pyrrolidine-3-carboxamide: 0.07g (quant.)).

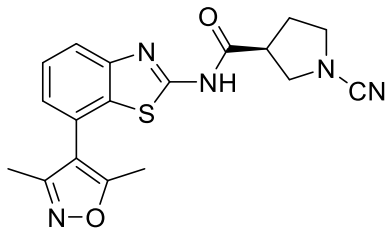
Step 4: The isolated products from Step 3 (0.072g, 0.21mmol) were used in Step 4 described in General Procedure 4 to afford desired products.



XL_9872_106A (0.026g, 34%) ^{1}H NMR (500 MHz, DMSO) δ 12.60 (s, 1H), 8.10 – 7.99 (m, 1H), 7.45 – 7.35 (m, 2H), 3.63 (dd, *J* = 9.6, 7.8 Hz, 1H), 3.57 (dd, *J* = 9.7, 6.0 Hz, 1H), 3.51 – 3.40 (m, 2H), 3.37 (dd, *J* = 13.8, 6.9 Hz, 1H), 2.31 (s, 3H), 2.27 – 2.16 (m, 1H), 2.16 – 2.03 (m, 4H). LC/MS (ESI) *m/z* 367.77; $[M+H]^+$ calcd for $C_{18}H_{18}N_5O_2S^+$: 368.12

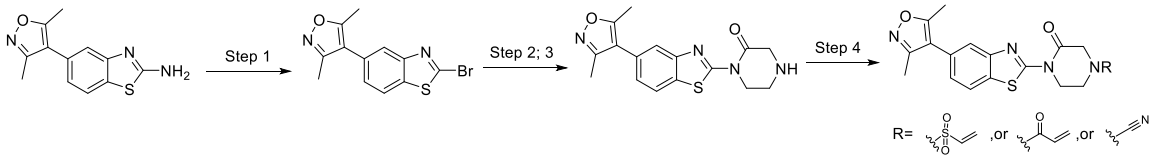


XL_9872_106B (0.022g, 29%) ^{1}H NMR (500 MHz, DMSO) δ 12.59 (s, 1H), 8.09 (d, *J* = 8.2 Hz, 1H), 7.74 (s, 1H), 7.31 (dd, *J* = 19.8, 9.9 Hz, 1H), 3.70 – 3.63 (m, 1H), 3.60 (dd, *J* = 9.6, 6.1 Hz, 1H), 3.52 – 3.37 (m, 3H), 2.43 (s, 3H), 2.28 – 2.21 (m, 4H), 2.12 (td, *J* = 14.1, 7.1 Hz, 1H). LC/MS (ESI) *m/z* 367.87; $[M+H]^+$ calcd for $C_{18}H_{18}N_5O_2S^+$: 368.12

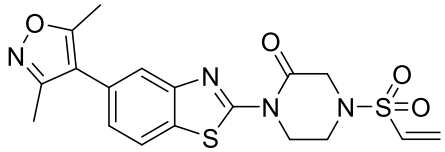


XL_9872_106C (0.029g, 40%) 1H NMR (500 MHz, DMSO) δ 7.82 (dd, J = 8.1, 0.9 Hz, 1H), 7.61 – 7.51 (m, 1H), 7.28 (dd, J = 7.4, 0.9 Hz, 1H), 3.68 – 3.60 (m, 1H), 3.56 (dd, J = 9.7, 5.9 Hz, 1H), 3.48 – 3.41 (m, 2H), 3.39 (dd, J = 13.5, 6.8 Hz, 1H), 2.31 (s, 3H), 2.21 (tt, J = 12.8, 6.4 Hz, 1H), 2.14 – 2.04 (m, 4H). LC/MS (ESI) m/z 367.97; [M+H]+ calcd for $C_{18}H_{18}N_5O_2S^+$: 368.12.

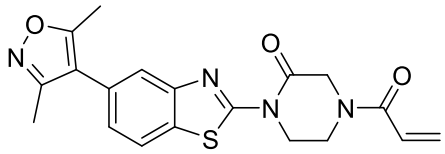
Synthesis of XL_9872_123A, 123B, 123C:



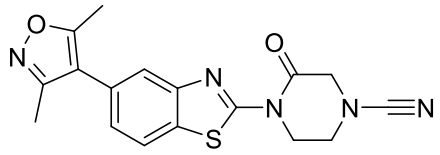
Synthesis were performed according to General Procedure 5.



XL_9872_123A (19mg, 45%) ^{1}H NMR (500 MHz, DMSO) δ 8.13 (d, *J* = 8.2 Hz, 1H), 7.86 (d, *J* = 1.4 Hz, 1H), 7.39 (dd, *J* = 8.2, 1.6 Hz, 1H), 7.00 (dd, *J* = 16.5, 10.0 Hz, 1H), 6.28 (dd, *J* = 13.2, 3.3 Hz, 2H), 4.38 – 4.29 (m, 2H), 4.18 (s, 2H), 3.73 – 3.61 (m, 2H), 2.44 (s, 3H), 2.27 (s, 3H). LC/MS (ESI) *m/z* 418.87; $[M+H]^+$ calcd for $C_{18}H_{19}N_4O_4S_2^+$: 419.08

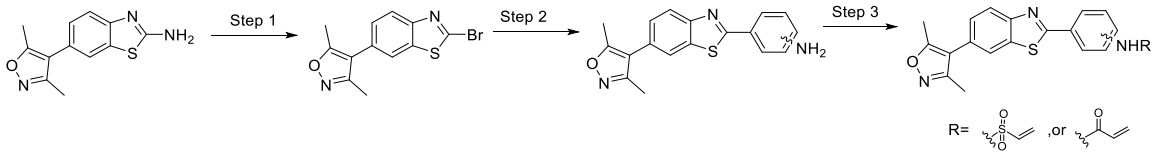


XL_9872_123B (16mg, 42%) ^{1}H NMR (500 MHz, DMSO) δ 8.12 (d, *J* = 8.2 Hz, 1H), 7.85 (s, 1H), 7.37 (dd, *J* = 8.2, 1.5 Hz, 1H), 6.98 – 6.78 (m, 1H), 6.22 (d, *J* = 16.6 Hz, 1H), 5.80 (d, *J* = 11.0 Hz, 1H), 4.68 (s, 1H), 4.50 (s, 1H), 4.32 (br, 2H), 4.04 (br, 2H), 2.44 (s, 3H), 2.27 (s, 3H). LC/MS (ESI) *m/z* 382.87; $[M+H]^+$ calcd for $C_{19}H_{19}N_4O_3S^+$: 383.12

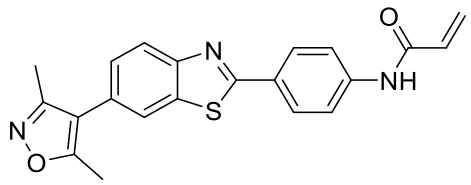


XL_9872_123C (16mg, 45%) ^{1}H NMR (500 MHz, DMSO) δ 8.13 (d, *J* = 8.2 Hz, 1H), 7.86 (s, 1H), 7.38 (d, *J* = 8.2 Hz, 1H), 4.36 (m, 4H), 3.85 – 3.72 (m, 2H), 2.44 (s, 3H), 2.27 (s, 3H). LC/MS (ESI) *m/z* 353.87; $[M+H]^+$ calcd for $C_{17}H_{16}N_5O_2S^+$: 354.10

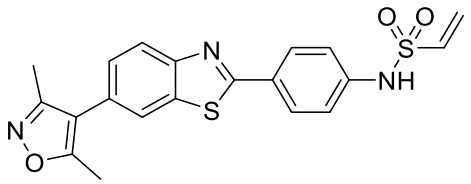
<u>Synthesis of XL_9872_126A, 126B, 128A, 128B</u>:



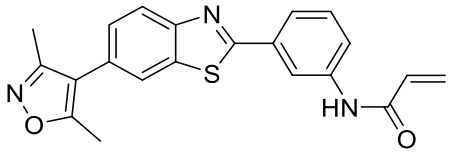
Synthesis was performed according to General Procedure 6.



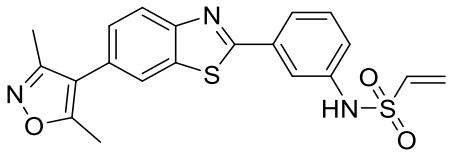
XL_9872_126A (22mg, 45%) ^{1}H NMR (500 MHz, DMSO) δ 10.49 (s, 1H), 8.19 (d, *J* = 1.6 Hz, 1H), 8.10 (dd, *J* = 7.1, 1.7 Hz, 3H), 7.90 (d, *J* = 8.7 Hz, 2H), 7.54 (dd, *J* = 8.4, 1.7 Hz, 1H), 6.56 – 6.44 (m, 1H), 6.33 (dd, *J* = 17.0, 1.8 Hz, 1H), 5.83 (dd, *J* = 10.1, 1.8 Hz, 1H), 2.47 (s, 3H), 2.29 (s, 3H). LC/MS (ESI) *m/z* 375.97; $[M+H]^+$ calcd for $C_{21}H_{18}N_3O_2S^+$: 376.11



XL_9872_126B (7mg, 13%) ^{1}H NMR (500 MHz, DMSO) δ 10.53 (s, 1H), 8.20 (d, *J* = 1.5 Hz, 1H), 8.08 (dd, *J* = 16.1, 8.5 Hz, 3H), 7.54 (dd, *J* = 8.4, 1.6 Hz, 1H), 7.34 (d, *J* = 8.7 Hz, 2H), 6.89 (dd, *J* = 16.4, 9.9 Hz, 1H), 6.25 (d, *J* = 16.4 Hz, 1H), 6.13 (d, *J* = 9.9 Hz, 1H), 2.46 (s, 3H), 2.29 (s, 3H). LC/MS (ESI) *m/z* 411.87; $[M+H]^+$ calcd for $C_{20}H_{18}N_3O_3S_2^+$: 412.08

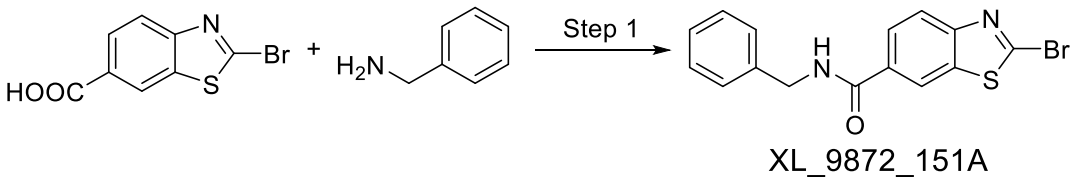


XL_9872_128A (25mg, 60%) ^{1}H NMR (500 MHz, DMSO) δ 10.44 (s, 1H), 8.59 (s, 1H), 8.23 (s, 1H), 8.15 (d, *J* = 8.4 Hz, 1H), 7.87 (d, *J* = 8.2 Hz, 1H), 7.81 (d, *J* = 7.7 Hz, 1H), 7.64 – 7.49 (m, 2H), 6.48 (dd, *J* = 17.0, 10.1 Hz, 1H), 6.33 (dd, *J* = 17.0, 1.6 Hz, 1H), 5.87 – 5.78 (m, 1H), 2.47 (s, 3H), 2.30 (s, 3H). LC/MS (ESI) *m/z* 375.97; $[M+H]^+$ calcd for $C_{21}H_{18}N_3O_2S^+$: 376.11



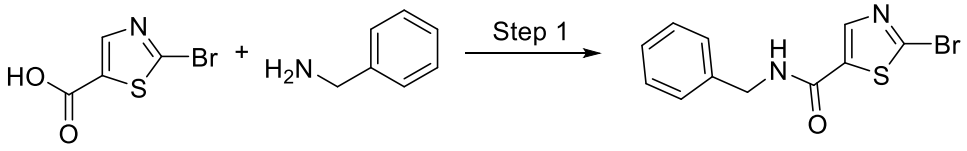
XL_9872_128B (17mg, 38%) ^{1}H NMR (500 MHz, DMSO) δ 10.35 (s, 1H), 8.23 (d, *J* = 1.5 Hz, 1H), 8.16 (d, *J* = 8.4 Hz, 1H), 7.95 (t, *J* = 1.8 Hz, 1H), 7.81 (d, *J* = 7.8 Hz, 1H), 7.65 – 7.45 (m, 2H), 7.38 (dd, *J* = 8.1, 1.6 Hz, 1H), 6.88 (dd, *J* = 16.4, 9.9 Hz, 1H), 6.20 (d, *J* = 16.4 Hz, 1H), 6.11 (d, *J* = 9.9 Hz, 1H), 2.47 (s, 3H), 2.30 (s, 3H). LC/MS (ESI) *m/z* 411.87; $[M+H]^+$ calcd for $C_{20}H_{18}N_3O_3S_2^+$: 412.08

<u>Synthesis of XL_9872_151A, 151B, 151C</u>:



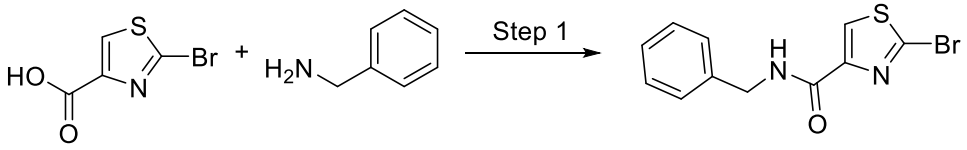
XL_9872_151A

XL_9872_151A (0.11g, 64%) ^{1}H NMR (500 MHz, DMSO) δ 9.21 (t, *J* = 5.8 Hz, 1H), 8.65 (s, 1H), 8.08 (d, *J* = 8.5 Hz, 1H), 8.04 (dd, *J* = 8.5, 1.4 Hz, 1H), 7.39 – 7.30 (m, 4H), 7.30 – 7.20 (m, 1H), 4.52 (d, *J* = 5.9 Hz, 2H). LC/MS (ESI) *m/z* 346.77; $[M+H]^+$ calcd for $C_{15}H_{12}BrN_2OS^+$: 346.98



XL_9872_151B

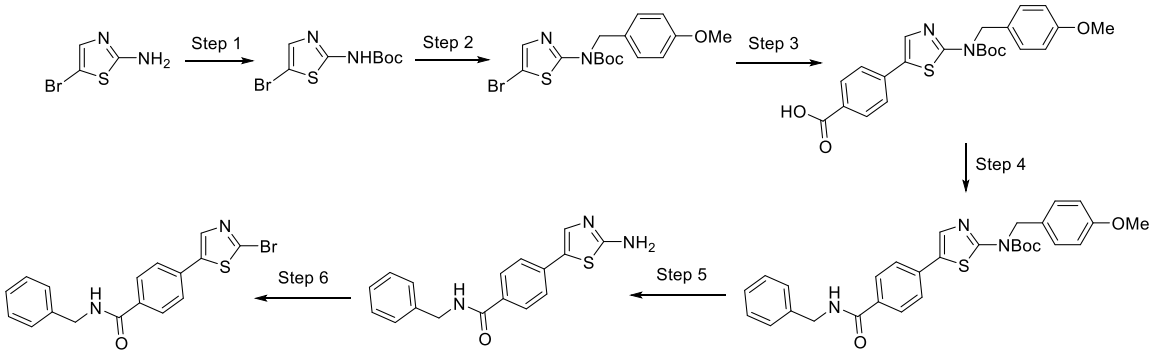
XL_9872_151B (0.075g, 51%) ^{1}H NMR (500 MHz, DMSO) δ 9.08 (t, *J* = 6.1 Hz, 1H), 8.31 (s, 1H), 7.37 – 7.27 (m, 4H), 7.27 – 7.19 (m, 1H), 4.43 (d, *J* = 6.3 Hz, 2H). LC/MS (ESI) *m/z* 296.77; $[M+H]^+$ calcd for $C_{11}H_{10}BrN_2OS^+$: 296.97



XL_9872_151C

XL_9872_151C ^{1}H NMR (500 MHz, DMSO) δ 9.32 (t, *J* = 5.8 Hz, 1H), 8.28 (d, *J* = 5.6 Hz, 1H), 7.39 – 7.30 (m, 4H), 7.30 – 7.21 (m, 1H), 4.46 (d, *J* = 5.9 Hz, 2H). LC/MS (ESI) *m/z* 296.87; $[M+H]^+$ calcd for $C_{11}H_{10}BrN_2OS^+$: 296.97

<u>Synthesis of XL_9872_159</u>:



Step 1: 2-amino-5-bromothiazole hydrobromide (1.3g, 5.0mmol), di-tert-butyl decarbonate (1.3g, 6.0mmol), Et3N (1.4mL, 10.0mmol), and DMAP (0.06g, 0.5mmol) were added into 5mL THF sequentially. The mixture was stirred at room temperature overnight, and diluted with EtOAc (30mL). The solution was washed with saturated sodium bicarbonate (30mL×2).

Combined aqueous layers was extracted with EtOAc (50mL). Combined organic layer was washed with brine, dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure to afford crude material. The crude was purified by flash chromatography (hexanes/EtOAc) to afford desired product. LC/MS (ESI) *m/z* 222.88 (M+H−*t*-butyl); $[M+H]^+$ calcd for $C_8H_{12}BrN_2O_2S^+$: 278.98

Step 2: tert-Butyl (5-bromothiazol-2-yl)carbamate (0.57g, 2.1mmol), PPh_3 (1.18g, 4.5mmol), and p-methoxybenzyl alcohol (0.57g, 4.1mmol) were mixed up in 10mL anhydrous THF. The mixture was stirred at 0°C while DIAD (0.91g, 4.5mmol) was added dropwisely. Upon completion, the mixture was warmed to room temperature, and stirred for 2 hours. The resulting reaction mixture was diluted with EtOAc (30mL). The solution was washed with saturated sodium bicarbonate (30mL×2). Combined aqueous layers was extracted with EtOAc (50mL). Combined organic layer was washed with brine, dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure to afford crude material. The crude was purified by flash chromatography (hexanes/EtOAc) to afford desired product (0.60g, 74%). LC/MS (ESI) *m/z* 342.24 (M+H−*t*-butyl); $[M+H]^+$ calcd for $C_{16}H_{20}BrN_2O_3S^+$: 399.04

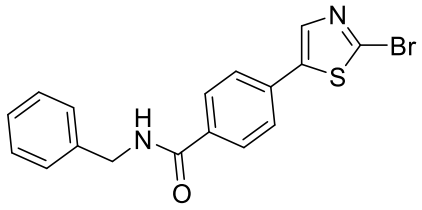
Step 3: tert-butyl (5-bromothiazol-2-yl)(4-methoxybenzyl)carbamate (0.2g, 0.5mmol), 4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzoic acid (0.26g, 1.0mmol), potassium carbonate (0.14g, 1.0mmol) and $Pd(dppf)Cl_2$ (0.04g, 0.05mmol) were added into 1,4-dioxane/H_2O (3mL/0.5mL). The mixture was purged with N_2 for 10min before stirring at 95°C overnight under N_2. Then the mixture was concentrated under reduced pressure, diluted with EtOAc (30mL), and washed with saturated NH_4Cl (30mL×2). Combined aqueous layer was extracted with EtOAc (50mL). Combined organic layer was washed once with brine, dried over anhydrous Na_2SO_4, filtered, and evaporated under reduced pressure to afford crude material, which was then purified by flash chromatography (hexanes/EtOAc/MeOH) to afford desired product (0.1g, 45%). LC/MS (ESI) *m/z* 384.77 (M+H−*t*-butyl); $[M+H]^+$ calcd for $C_{23}H_{25}N_2O_5S^+$: 441.15

Step 4: 4-(2-((tert-butoxycarbonyl)(4-methoxybenzyl)amino)thiazol-5-yl)benzoic acid (0.1g, 0.23mmol), benzylamine (0.03g, 0.28mmol), Et_3N (0.18mL, 1.25mmol), and HATU (0.13g, 0.35mmol) were added into 3mL DMF. The mixture was stirred at room temperature overnight,

and purified by flash chromatography (hexanes/EtOAc/MeOH) to afford desired product (0.03g, 25%). LC/MS (ESI) *m/z* 473.88 (M+H−*t*-butyl); $[M+H]^+$ calcd for $C_{30}H_{32}N_3O_4S^+$: 530.21

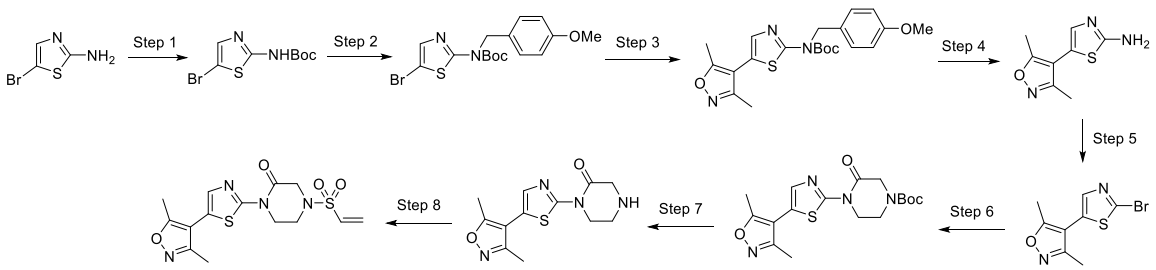
Step 5: tert-butyl (5-(4-(benzylcarbamoyl)phenyl)thiazol-2-yl)(4-methoxybenzyl)carbamate (0.03g, 0.06mmol) were dissolved in DCM (2-3mL) and treated with TFA (2-3mL). The mixtures were stirred at room temperature until the tert-butyloxycarbonyl protecting group was cleaved tracking by UPLC-MS. The mixture was concentrated and flushed by flash column chromatography (EtOAc/MeOH/0.5%Et_3N). LC/MS (ESI) *m/z* 309.87; $[M+H]^+$ calcd for $C_{17}H_{16}N_3OS^+$: 310.10

Step 6: 4-(2-aminothiazol-5-yl)-N-benzylbenzamide (0.028g, 0.09mmol), $CuBr_2$ (0.025g, 0.18mmol), and t-butyl nitrite (0.02g, 0.18mmol) were added into 2mL anhydrous MeCN at 0°C. The mixture was then warmed up to room temperature then 65°C, and stirred for 4 hours. The reaction was cooled to room temperature, and diluted with water (30mL). The mixture was acidified with HBr (48%wt in H_2O) to pH=2 and extracted with EtOAc (30mL×2). The combined organic layer was washed with brine, dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure to afford the crude material. The material was purified by flash chromatography (hexanes/EtOAc/MeOH) followed by preparative HPLC (MeOH or CH_3CN/H_2O with 0.0425% TFA) to afford the target products to afford desired product (6mg, 18%).



XL_9872_159 ^1H NMR (500 MHz, DMSO) δ 9.14 (t, *J* = 5.9 Hz, 1H), 8.23 (s, 1H), 7.98 (d, *J* = 8.4 Hz, 2H), 7.78 (d, *J* = 8.4 Hz, 2H), 7.34 (dd, *J* = 7.7, 5.6 Hz, 4H), 7.29 – 7.22 (m, 1H), 4.50 (d, *J* = 6.0 Hz, 2H). LC/MS (ESI) *m/z* 372.67; $[M+H]^+$ calcd for $C_{17}H_{14}BrN_2O_4S^+$: 373.00

Synthesis of XL_9872_163:



Step 1: 2-amino-5-bromothiazole hydrobromide (1.3g, 5.0mmol), di-tert-butyl decarbonate (1.3g, 6.0mmol), Et_3N (1.4mL, 10.0mmol), and DMAP (0.06g, 0.5mmol) were added into 5mL THF sequentially. The mixture was stirred at room temperature overnight, and diluted with EtOAc (30mL). The solution was washed with saturated sodium bicarbonate (30mL×2). Combined aqueous layers was extracted with EtOAc (50mL). Combined organic layer was washed with brine, dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure to afford crude material. The crude was purified by flash chromatography (hexanes/EtOAc) to afford desired product. LC/MS (ESI) *m/z* 222.88 (M+H−*t*-butyl); $[M+H]^+$ calcd for $C_8H_{12}BrN_2O_2S^+$: 278.98

Step 2: tert-Butyl (5-bromothiazol-2-yl)carbamate (0.57g, 2.1mmol), PPh_3 (1.18g, 4.5mmol), and p-methoxybenzyl alcohol (0.57g, 4.1mmol) were mixed up in 10mL anhydrous THF. The mixture was stirred at 0°C while DIAD (0.91g, 4.5mmol) was added dropwisely. Upon completion, the mixture was warmed to room temperature, and stirred for 2 hours. The resulting reaction mixture was diluted with EtOAc (30mL). The solution was washed with saturated sodium bicarbonate (30mL×2). Combined aqueous layers was extracted with EtOAc (50mL). Combined organic layer was washed with brine, dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure to afford crude material. The crude was purified by flash chromatography (hexanes/EtOAc) to afford desired product (0.60g, 74%). LC/MS (ESI) *m/z* 342.24 (M+H−*t*-butyl); $[M+H]^+$ calcd for $C_{16}H_{20}BrN_2O_3S^+$: 399.04

Step 3: tert-butyl (5-bromothiazol-2-yl)(4-methoxybenzyl)carbamate (0.2g, 0.5mmol), (3,5-dimethylisoxazol-4-yl)boronic acid (0.14g, 1.0mmol), potassium carbonate (0.14g, 1.0mmol) and $Pd(dppf)Cl_2$ (0.04g, 0.05mmol) were added into 1,4-dioxane/H_2O (3mL/0.5mL). The mixture was purged with N_2 for 10min before stirring at 95°C overnight under N_2. Then the mixture was concentrated under reduced pressure, diluted with EtOAc (30mL), and washed with saturated NH_4Cl (30mL×2). Combined aqueous layer was extracted with EtOAc (50mL). Combined organic

layer was washed once with brine, dried over anhydrous Na_2SO_4, filtered, and evaporated under reduced pressure to afford crude material, which was then purified by flash chromatography (hexanes/EtOAc/MeOH) to afford desired product (0.19g, 91%). LC/MS (ESI) *m/z* 359.77 (M+H−*t*-butyl); $[M+H]^+$ calcd for $C_{21}H_{26}N_3O_4S^+$: 416.16

Step 4: tert-butyl (5-(3,5-dimethylisoxazol-4-yl)thiazol-2-yl)(4-methoxybenzyl)carbamate (0.19g, 0.45mmol) was dissolved in 3mL DCM. Into the solution was added 3mL TFA. The mixture was stirred at 80°C for 5 hours, then concentrated under reduced pressure, and re-dissolved in DCM (50mL). The organic solution was basified using saturated $NaHCO_3$ (50mL×2). Combined aqueous layer was extracted with DCM (50mL). Combined organic layer was washed with brine, dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure to afford crude material (0.13g, quant.) which was used directly in the following step without further purification. LC/MS (ESI) *m/z* 195.98; $[M+H]^+$ calcd for $C_8H_{10}N_3OS^+$: 196.05

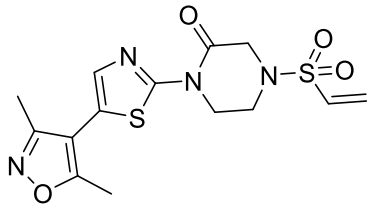
Step 5: *Step 6*: 4-(2-aminothiazol-5-yl)-N-benzylbenzamide (0.078g, 0.4mmol), $CuBr_2$ (0.115g, 0.8mmol), and t-butyl nitrite (0.082g, 0.8mmol) were added into 3mL anhydrous MeCN at 0°C. The mixture was then warmed up to room temperature then 65°C, and stirred for 4 hours. The reaction was cooled to room temperature, and diluted with water (30mL). The mixture was acidified with HBr (48%wt in H_2O) to pH=2 and extracted with EtOAc (30mL×2). The combined organic layer was washed with brine, dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure to afford the crude material. The material was purified by flash chromatography (hexanes/EtOAc) to afford the target product (0.076g, 75%). LC/MS (ESI) *m/z* 258.77; $[M+H]^+$ calcd for $C_8H_8BrN_2OS^+$: 258.95

Step 6: 4-(2-bromothiazol-5-yl)-3,5-dimethylisoxazole (0.076g, 0.3mmol), 1-Boc-3-oxopiperazine (0.12g, 0.6mmol), cesium carbonate (0.39g, 1.2mmol), $Pd_2(dba)_3$ (0.028g, 0.03mmol), and Xantphos (0.035g, 0.06mmol) were added into 4mL 1,4-dioxane. The mixture was degassed by bubbling in N_2 for 10-15min before heated at 95°C overnight. Then the mixture was cooled to room temperature before diluted with EtOAc (30mL). Organic layer was washed with 20% citric acid (20mL×2). Combined aqueous layer was extracted with EtOAc (30mL). Combined organic layer was washed with brine, dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure to afford crude material. The crude material was purified by flash

chromatography (hexanes/EtOAc/MeOH) to afford desired product (0.035g, 31%). LC/MS (ESI) *m/z* 378.87; $[M+H]^+$ calcd for $C_{17}H_{23}N_4O_4S^+$: 379.14

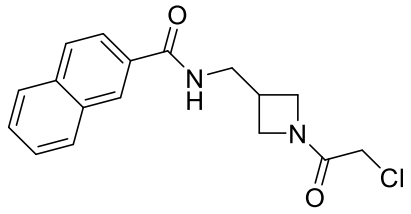
Step 7: tert-butyl 4-(5-(3,5-dimethylisoxazol-4-yl)thiazol-2-yl)-3-oxopiperazine-1-carboxylate (0.035g, 0.09mmol) were dissolved in DCM (1mL) and treated with 4M HCl in 1,4-dioxane (1mL). The mixtures were stirred at room temperature until the tert-butyloxycarbonyl protecting group was cleaved tracking by UPLC-MS. The mixture was concentrated and flushed by flash column chromatography (EtOAc/MeOH/0.5%Et_3N) to afford desired product (0.025g, 97%).

Step 8: 1-(5-(3,5-dimethylisoxazol-4-yl)thiazol-2-yl)piperazin-2-one (0.025g, 0.087mmol), Et_3N (0.1mL, 0.7mmol) were added in to 2mL DCM at 0°C. Into the solution was added 2-chloroethane sulfonyl chloride (0.023g, 0.14mmol). The mixture was stirred at 0°C for 1 hour before purified by flash chromatography (hexanes/EtOAc/MeOH) followed by preparative HPLC (MeOH or CH_3CN/H_2O with 0.0425% TFA) to afford the target products to afford desired product (7mg, 22%).



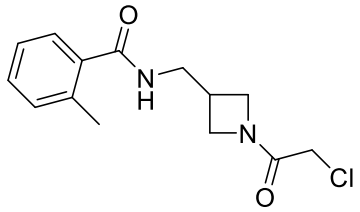
XL_9872_163 ^{1}H NMR (500 MHz, DMSO) δ 7.69 (s, 1H), 6.99 (dd, *J* = 16.5, 10.0 Hz, 1H), 6.41 – 6.12 (m, 2H), 4.28 – 4.18 (m, 2H), 4.12 (s, 2H), 3.68 – 3.59 (m, 2H), 2.48 (s, 3H), 2.29 (s, 3H). LC/MS (ESI) *m/z* 368.87; $[M+H]^+$ calcd for $C_{14}H_{17}N_4O_4S_2^+$: 369.07

Synthesis of XL_11478_092A, 092B, 092C, 092D, 092E, 093A, 093B, 093C, 093D, 093E, 093F:

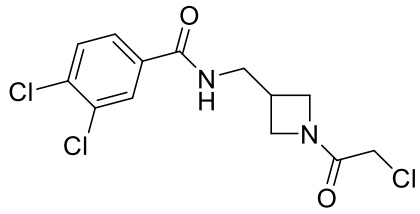


Synthesis of XL_11478_092A followed the General Procedure 1 by using 2-naphthoic acid (0.10g, 0.6mmol) and t-butyl-3-(aminoethyl)acetidine-1-carboxylate (0.09g, 0.5mmol) to afford

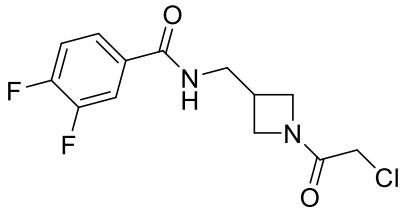
XL_11478_092A (0.027g, 17%) ^{1}H NMR (500 MHz, DMSO) δ 8.82 (t, *J* = 5.6 Hz, 1H), 8.44 (s, 1H), 8.01 (m,3H), 7.93 (dd, *J* = 8.6, 1.6 Hz, 1H), 7.68 – 7.54 (m, 2H), 4.28 (t, *J* = 8.6 Hz, 1H), 4.16 – 4.05 (s, 2H), 4.00 (dd, *J* = 16.6, 7.8 Hz, 2H), 3.74 (dd, *J* = 9.8, 5.4 Hz, 1H), 3.59 – 3.49 (m, 2H), 3.00 – 2.80 (m, 1H). LC/MS (ESI) *m/z* 316.67; $[M+H]^+$ calcd for $C_{17}H_{18}ClN_2O_2^+$: 317.11



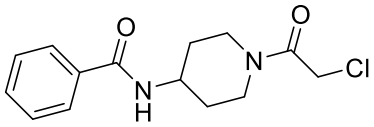
Synthesis of XL_11478_092B followed the General Procedure 1 by using 2-methylbenzoic acid (0.08g, 0.6mmol) and t-butyl-3-(aminoethyl)acetidine-1-carboxylate (0.09g, 0.5mmol) to afford XL_11478_092B (0.042g, 30%) ^{1}H NMR (500 MHz, DMSO) δ 8.44 (m, 1H), 7.41 – 7.28 (m, 2H), 7.28 – 7.18 (m, 2H), 4.32 – 4.21 (m, 1H), 4.15 – 4.05 (m, 2H), 4.02 – 3.91 (m, 2H), 3.76 – 3.63 (m, 1H), 3.44 (m, 2H), 2.91 – 2.77 (m, 1H), 2.33 (s, 3H). LC/MS (ESI) *m/z* 280.77; $[M+H]^+$ calcd for $C_{14}H_{18}ClN_2O_2^+$: 281.11



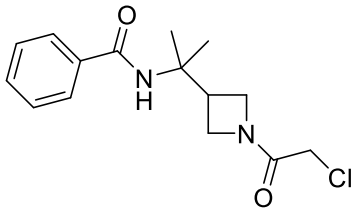
Synthesis of XL_11478_092C followed the General Procedure 1 by using 3,4-dichlorobenzoic acid (0.12g, 0.6mmol) and t-butyl-3-(aminoethyl)acetidine-1-carboxylate (0.09g, 0.5mmol) to afford XL_11478_092C (0.058g, 35%) ^{1}H NMR (500 MHz, DMSO) δ 8.80 (t, *J* = 5.6 Hz, 1H), 8.07 (d, *J* = 2.0 Hz, 1H), 7.82 (dd, *J* = 8.4, 2.0 Hz, 1H), 7.80 – 7.69 (m, 1H), 4.32 – 4.20 (m, 1H), 4.16 – 4.04 (s, 2H), 3.99 – 3.89 (m, 2H), 3.68 (dd, *J* = 9.9, 5.4 Hz, 1H), 3.49 (dd, *J* = 6.8, 5.9 Hz, 2H), 2.91 – 2.78 (m, 1H). LC/MS (ESI) *m/z* 334.87; $[M+H]^+$ calcd for $C_{13}H_{14}Cl_3N_2O_2^+$: 335.01



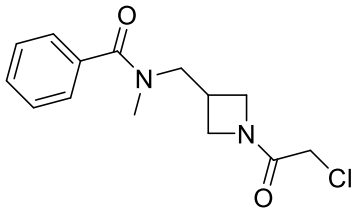
Synthesis of XL_11478_092D followed the General Procedure 1 by using 3,4-difluorobenzoic acid (0.095g, 0.6mmol) and t-butyl-3-(aminoethyl)acetidine-1-carboxylate (0.09g, 0.5mmol) to afford XL_11478_092D (0.037g, 25%) ^{1}H NMR (500 MHz, DMSO) δ 8.77 (m, 1H), 8.00 – 7.85 (m, 1H), 7.80 – 7.68 (m, 1H), 7.61 – 7.49 (m, 1H), 4.29 – 4.22 (m, 1H), 4.14 – 4.04 (m, 2H), 3.99 – 3.91 (m, 2H), 3.73 – 3.64 (m, 1H), 3.48 (dd, *J* = 11.7, 5.4 Hz, 2H), 2.92 – 2.77 (m, 1H). LC/MS (ESI) *m/z* 302.97; $[M+H]^+$ calcd for $C_{13}H_{14}ClF_2N_2O_2^+$: 303.07



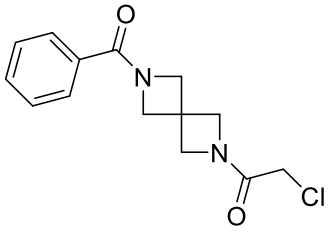
Synthesis of XL_11478_093A followed the General Procedure 1 by using benzoic acid (0.07g, 0.6mmol), t-butyl-4-amino piperidine-1-carboxylate (0.10g, 0.5mmol) to afford XL_11478_093A (0.07g, 52%) ^{1}H NMR (500 MHz, DMSO) δ 8.32 (d, *J* = 7.7 Hz, 1H), 7.85 (d, *J* = 7.4 Hz, 2H), 7.52 (t, *J* = 7.2 Hz, 1H), 7.46 (t, *J* = 7.5 Hz, 2H), 4.44 (d, *J* = 12.8 Hz, 1H), 4.35 (d, *J* = 12.9 Hz, 1H), 4.30 (d, *J* = 13.2 Hz, 1H), 4.14 – 4.00 (m, 1H), 3.87 (d, *J* = 13.6 Hz, 1H), 3.19 (t, *J* = 11.8 Hz, 1H), 2.81 (t, *J* = 11.5 Hz, 1H), 1.96 – 1.74 (m, 2H), 1.57 (qd, *J* = 12.2, 4.0 Hz, 1H), 1.42 (qd, *J* = 12.3, 4.2 Hz, 1H). LC/MS (ESI) *m/z* 280.77; $[M+H]^+$ calcd for $C_{14}H_{18}ClN_2O_2^+$: 281.11



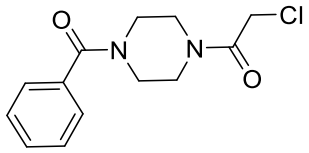
Synthesis of XL_11478_093B followed the General Procedure 1 by using benzoic acid (0.07g, 0.6mmol), t-butyl-3-(2-aminopropan-2-yl)azetidine-1-carboxylate (0.10g, 0.5mmol) to afford XL_11478_093B (5mg, 3.4%) LC/MS (ESI) *m/z* 295.18; $[M+H]^+$ calcd for $C_{15}H_{20}ClN_2O_2^+$: 295.12



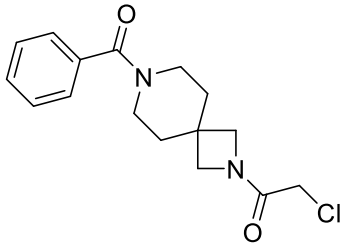
Synthesis of XL_11478_093C followed the General Procedure 1 by using benzoic acid (0.07g, 0.6mmol), t-butyl-ethylaminoazetidine-1-carboxylate (0.10g, 0.5mmol) to afford XL_11478_093C (0.032g, 23%) 1H NMR (500 MHz, CDCl3) δ 7.49 – 7.35 (m, 5H), 4.42 (br, 1H), 4.27 (br, 2H), 4.08 (br, 1H), 3.91 (br, 3H), 3.60 (dd, J = 13.5, 6.6 Hz, 1H), 3.02 (s, 4H). LC/MS (ESI) m/z 281.08; [M+H]+ calcd for C14H18ClN2O2+: 281.11



Synthesis of XL_11478_093D followed the General Procedure 1 by using benzoic acid (0.07g, 0.6mmol), t-butyl-2,6-diazaspiro[3,3]heptane-2-carboxylate (0.10g, 0.5mmol) to afford XL_11478_093D (0.054g, 39%) ^{1}H NMR (500 MHz, DMSO) δ 7.67 – 7.56 (m, 2H), 7.50 (m, 3H), 4.46 (br, 2H), 4.39 – 4.29 (m, 2H), 4.18 (m, 2H), 4.15 – 4.00 (m, 4H). LC/MS (ESI) *m/z* 278.67; $[M+H]^+$ calcd for $C_{14}H_{16}ClN_2O_2^+$: 279.09

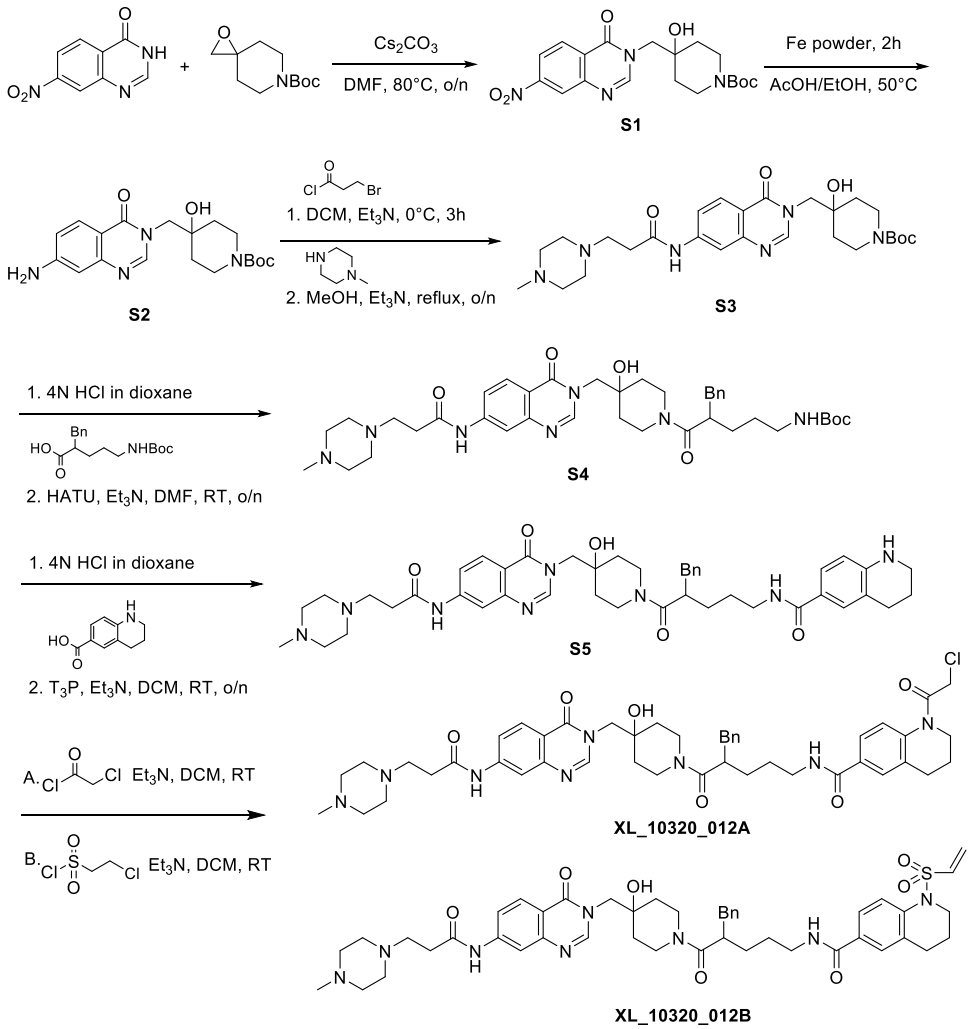


Synthesis of XL_11478_093E followed the General Procedure 1 by using benzoic acid (0.07g, 0.6mmol), t-butylpiperazine-1-carboxylate (0.093g, 0.5mmol) to afford XL_11478_093E (0.11g, 83%) ^{1}H NMR (500 MHz, DMSO) δ 7.51 – 7.45 (m, 3H), 7.45 – 7.40 (m, 2H), 4.37 (s, 2H), 3.54 (br, 8H). LC/MS (ESI) *m/z* 266.77; $[M+H]^+$ calcd for $C_{13}H_{16}ClN_2O_2^+$: 267.09



Synthesis of XL_11478_093F followed the General Procedure 1 by using benzoic acid (0.07g, 0.6mmol), t-butyl-2,7-diazaspiro[3,5]nonane-2-carboxylate (0.11g, 0.5mmol) to afford XL_11478_093F (0.053g, 35%) 1H NMR (500 MHz, CDCl3) δ 7.51 – 7.35 (m, 5H), 4.05 (br, 2H), 3.93 (s, 2H), 3.86 (br, 2H), 3.68 (br, 1H), 3.41 (br, 2H), 1.85 (br, 4H). LC/MS (ESI) m/z 307.08; [M+H]+ calcd for $C_{16}H_{20}ClN_2O_2^+$: 307.12.

<u>Synthesis of XL_10320_012A, 012B:</u>



Step 1 (Synthesis of **S1**): 7-nitroquinazolin-4(*3H*)-one (1.55g, 8.1mmol) and *tert*-butyl-1-oxa-6-azaspiro[2.5]octane-6-carboxylate (1.90g, 8.9mmol) were added into 20mL DMF. Cesium carbonate (7.82g, 24.0mmol) was added in one portion. The mixture was heated at 80°C overnight. The mixture was diluted with EtOAc, then washed with sat. NaCl. Combined organic layer was concentrated under reduced pressure. The crude product was purified by flash chromatography (EtOAc: hexanes: 50%-70%) to afford 2.42g **S1** (75%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.54 (d, *J* = 2.1 Hz, 1H), 8.44 (t, *J* = 8.6 Hz, 1H), 8.26 (dd, *J* = 8.8, 2.2 Hz, 1H), 8.20 (s, 1H), 4.10 (s, 2H), 3.88 (s, 2H), 3.14 (t, *J* = 11.6 Hz, 2H), 1.73 – 1.48 (m, 5H), 1.44 (s, 9H). LCMS (ESI) *m/z* 304.97 (M + H - Boc) [$(M+H)^+$ $C_{19}H_{25}N_4O_6^+$ calcd for 405.18]

Step 2 (Synthesis of **S2**): Compound **S1**(2.4g, 6.0mmol) was suspended in 20mL solvent (EtOH/AcOH=1:1). 4 eq. of Fe powder was added in portions. The mixture was stirred for 1 hour at 55°C. Then the reaction was cooled down to room temperature, and filtered through a pad of Celite. The filtrate was concentrated under reduced pressure to afford the crude product, which was then purified by flash chromatography (10%MeOH in EtOAc) to afford 2.1g product **S2** (93%) ^{1}H NMR (500 MHz, DMSO) δ 8.04 (s, 1H), 7.79 (d, *J* = 8.7 Hz, 1H), 6.72 (dd, *J* = 8.7, 1.9 Hz, 1H), 6.61 (d, *J* = 2.0 Hz, 1H), 6.09 (s, 2H), 4.87 (s, 1H), 3.89 (s, 2H), 3.64 (d, *J* = 12.0 Hz, 2H), 3.05 (s, 2H), 1.54 – 1.24 (m, 13H).LCMS (ESI) *m/z* 374.97 [$(M+H)^+$ $C_{19}H_{27}N_4O_4^+$ calcd for 375.20]

Step 3 (Synthesis of **S3**): Compound **S2** (2.1g, 5.6mmol) was dissolved in anhydrous 10mL dichloromethane under N_2 at 0°C. 3.0 eq. of Et_3N was added. Then 3-bromopropionyl chloride (1.15g, 6.7mmol) was added dropwisely. The mixture was stirred at 0°C for 1 hour, then quenched with MeOH, and concentrated under reduced pressure. The solid residue was directly used for the following step without further purification. The crude product from last step was dissolved in 10mL MeOH, then 3.0eq of Et_3N was added. Into the stirred mixture was added 1-methylpiperazine (0.67g, 6.7mmol) dropwisely. After the addition completed, the mixture was stirred for 1 hour at 50°C. Then the reaction mixture was cooled down to room temperature and concentrated under reduced pressure, then directly subjected to HPLC purification (MeOH/H_2O with 4‰ TFA) to afford 2.1g product **S3** (73% in two steps) ^{1}H NMR (500 MHz, CD_3OD) δ 8.28 (s, 1H), 8.20 (d, *J* = 8.8 Hz, 1H), 8.12 (d, *J* = 1.9 Hz, 1H), 7.69 (dd, *J* = 8.7, 2.0 Hz, 1H), 4.11 (s, 2H), 3.82

(d, *J* = 13.4 Hz, 2H), 3.23 (m, 2H), 2.85 (t, *J* = 7.0 Hz, 2H), 2.79 – 2.50 (m, 10H), 2.37 (s, 3H), 1.72 – 1.62 (m, 2H), 1.50 (d, *J* = 17.4 Hz, 11H). LCMS (ESI) *m/z* 529.08 [$(M+H)^+$ $C_{27}H_{41}N_6O_5^+$ calcd for 529.31].

Step 4 (Synthesis of **S4**): **S3** (0.53g, 1.0mmol) was dissolved in 3mL DCM, then 5mL 4M HCl in 1,4-dioxane was added in portions. The solution was stirred for 1 hour at room temperature. Then the mixture was concentrated under reduced pressure, and left on high vacuum overnight to remove residual acid. Then the product (0.11g, 0.25mmol) was dissolved in 3mL DMF, and basified by adding 10 eq of Et_3N. Into the solution 2-benzyl-5-((tert-butoxycarbonyl)amino)pentanoic acid (0.11g, 0.35mmol) and HATU (0.16g, 0.4mmol) sequentially The resultant solution was stirred overnight. Then the mixture was directly subjected to HPLC purification (MeOH/H_2O with 4‰ TFA) to afford 183mg **S4** (quantitative) ^1H NMR (500 MHz, DMSO) δ 10.51 (s, 1H), 8.16-8.04 (m, 2H, conformer), 8.01 (dd, *J* = 5.0, 1.9 Hz, 1H), 7.61 (ddd, *J* = 9.8, 8.3, 2.0 Hz, 1H), 7.29 – 7.19 (m, 2H), 7.19 – 7.07 (m, 3H), 6.76 (m, 1H), 4.84 (s, 1H), 4.11 (m, 1H), 4.05 – 3.87 (m, 2H), 3.80 (d, *J* = 13.7 Hz, 1H), 3.67 – 3.49 (m, 3H), 3.16 – 3.03 (m, 2H), 2.98 (s, 1H), 2.85 (m, 3H), 2.77 – 2.58 (m, 5H), 2.55 (m, 3H), 2.22 (s, 3H), 1.61 – 1.44 (m, 1H), 1.43 – 1.33 (m, 9H, conformer), 1.33 – 1.22 (m, 3H), 1.22 – 1.02 (m, 4H), 0.39 (m, 1H). LCMS (ESI) *m/z* 718.00 [$(M+H)^+$ 718.43 calcd for $C_{39}H_{56}N_7O_6^+$]

Step 5 (Synthesis of S5): **S4** (0.15g, 0.2mmol) was dissolved in 2mL DCM, then 2mL 4M HCl in 1,4-dioxane was added in portions. The solution was stirred for 1 hour at room temperature. Then the mixture was concentrated under reduced pressure, and left on high vacuum overnight to remove residual acid. Then the product (0.12g, 0.2mmol) was dissolved in 3mL DMF, and basified by adding 10 eq of Et_3N. Into the solution 1,2,3,4-tetrahydroquinoline carboxylic acid (0.064g, 0.3mmol) and HATU (0.12g, 0.3mmol) sequentially. The resultant solution was stirred overnight. Then the mixture was directly subjected to HPLC purification (MeOH/H_2O with 4‰ TFA) to afford 108mg **S5** (56%). LCMS (ESI) *m/z* 776.91 [$(M+H)^+$ 777.44 calcd for $C_{44}H_{57}N_8O_5^+$]

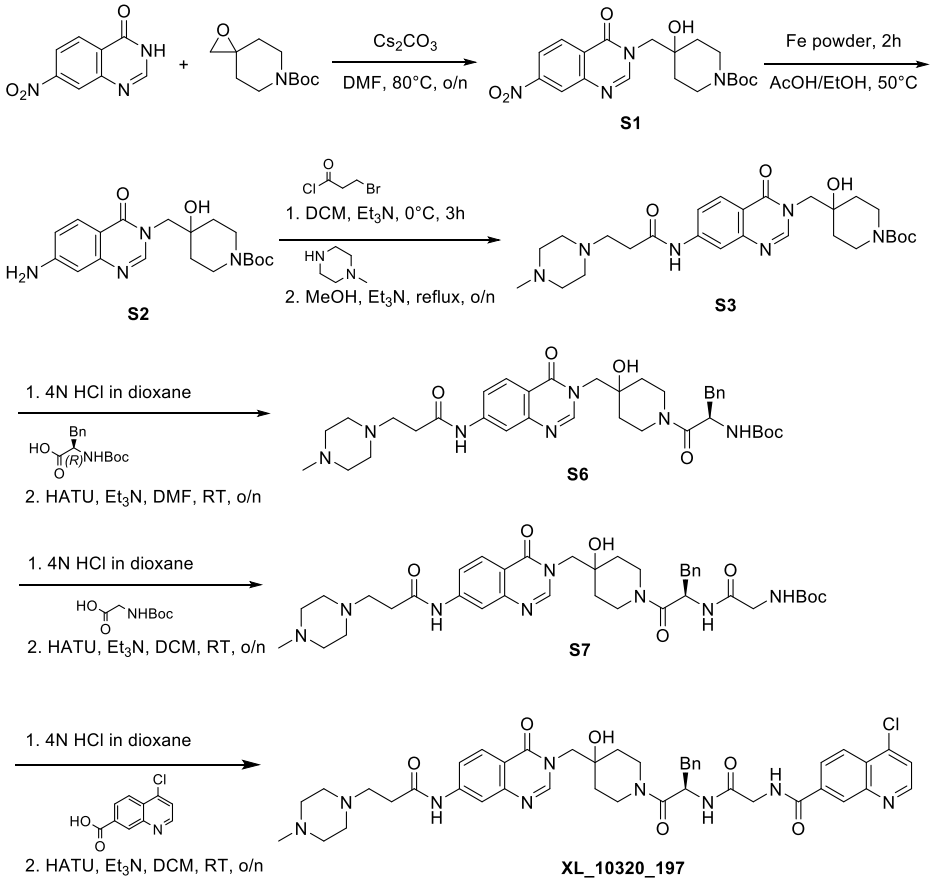
Step 6 (Synthesis of XL_10320_012A, 012B): S5 (0.054g, 0.07mmol) was dissolved in anhydrous DCM/DMF co-solvent (2mL, v/v=1:1). Et_3N (0.05mL, 0.34mmol) was added into the solution, followed by addition of 2-chloroacetyl chloride (9.4mg, 0.08mmol) or 2-chloroethane-1-sulfonyl chloride (22mg, 0.14mmol). The mixture was stirred at 0°C for 1 hour before purified by flash

chromatography (hexanes/EtOAc/MeOH) followed by preparative HPLC (MeOH or CH_3CN/H_2O with 0.0425% TFA) to afford the target products.

XL_10320_012A (31mg, 53%, mixture of rotamers) ^{1}H NMR (500 MHz, DMSO) δ 10.66 (s, 1H), 8.41 (dt, *J* = 10.9, 5.4 Hz, 1H), 8.22 – 8.00 (m, 3H), 7.66 (m, 4H), 7.30 – 7.17 (m, 2H), 7.17 – 7.05 (m, 3H), 4.57 (s, 2H), 4.12 (d, *J* = 12.4 Hz, 0.5H), 3.94 (m, 1H), 3.81 (d, *J* = 13.6 Hz, 0.5H), 3.71 (t, *J* = 6.2 Hz, 2H), 3.66 – 3.54 (m, 2H), 3.36 (m, 3H), 3.32 – 3.15 (m, 5.5H), 3.05 (m, 3H), 2.80 (m, 5.5H), 2.78 – 2.57 (m, 5.5H), 1.98 – 1.81 (m, 2H), 1.59 (m, 1H), 1.51 – 1.25 (m, 5H), 1.13 (m, 3H), 0.39 (m, 1H). LCMS (ESI) *m/z* 853.02 $[(M+H)^+$ 853.42 calcd for $C_{46}H_{58}ClN_8O_6^+]$

XL_10320_012B (34mg, 57%, mixture of rotamers) ^{1}H NMR (500 MHz, DMSO) δ 10.67 (s, 1H), 8.37 (dt, *J* = 11.0, 5.5 Hz, 1H), 8.24 – 8.00 (m, 3H), 7.59 (m, 4H), 7.21 (m, 2H), 7.16 – 7.04 (m, 3H), 6.90 (dd, *J* = 16.3, 9.9 Hz, 1H), 6.14 (d, *J* = 15.0 Hz, 1H), 6.10 (d, *J* = 10.0 Hz, 1H), 4.12 (d, *J* = 12.1 Hz, 0.5H), 4.05 – 3.86 (m, 1H), 3.81 (d, *J* = 13.6 Hz, 0.5H), 3.75 – 3.59 (m, 3.5H), 3.48 (m, 3.5H), 3.28 (m, 4H), 3.19 (m, 2.5H), 3.10 (m, 2.5H), 2.93 – 2.76 (m, 7H), 2.67 (m, 3H), 1.99 – 1.82 (m, 2H), 1.56 (m, 1H), 1.51 – 1.23 (m, 5H), 1.23 – 0.99 (m, 3H), 0.39 (t, *J* = 10.5 Hz, 1H). LCMS (ESI) *m/z* 866.82 $[(M+H)^+$ 867.42 calcd for $C_{46}H_{59}N_8O_7S^+]$

<u>Synthesis of XL_10320_197:</u>



Step 1-3 (Synthesis of S1-S3): the same as the steps in synthetic routes toward XL_10320_012A, B.

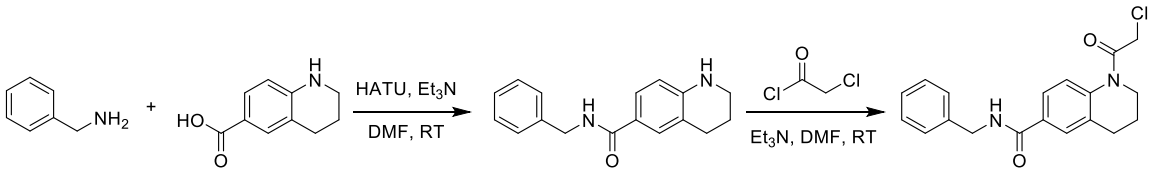
Step 4: (Synthesis of **S6**): **S3** (0.20g, 0.4mmol) was dissolved in 2mL DCM, then 2mL 4M HCl in 1,4-dioxane was added in portions. The solution was stirred for 1 hour at room temperature. Then the mixture was concentrated under reduced pressure, and left on high vacuum overnight to remove residual acid. Then the product (0.16g, 0.4mmol) was dissolved in 5mL DMF, and basified by adding 10 eq of Et_3N. Into the solution (tert-butoxycarbonyl)-D-phenylalanine (0.15g, 0.56mmol) and HATU (0.29g, 0.76mmol) sequentially. The resultant solution was stirred overnight. Then the mixture was directly subjected to HPLC purification (MeOH/H_2O with 4‰ TFA) to afford 110mg **S6** (43%). LCMS (ESI) *m/z* 676.00 [$(M+H)^+$ 676.38 calcd for $C_{36}H_{50}N_7O_6^+$]

Step 5: (Synthesis of **S7**): **S6** (0.11g, 0.16mmol) was dissolved in 2mL DCM, then 2mL 4M HCl in 1,4-dioxane was added in portions. The solution was stirred for 1 hour at room temperature. Then the mixture was concentrated under reduced pressure, and left on high vacuum overnight to remove residual acid. Then the product (0.092g, 0.16mmol) was dissolved in 3mL DMF, and basified by adding 10 eq of Et_3N. Into the solution *N*-Boc glycine (0.053g, 0.3mmol) and HATU (0.15g, 0.4mmol) sequentially. The resultant solution was stirred overnight. Then the mixture was directly subjected to HPLC purification (MeOH/H_2O with 4‰ TFA) to afford 87mg **S7** (63%). LCMS (ESI) *m/z* 732.81 [$(M+H)^+$ 733.40 calcd for $C_{38}H_{53}N_8O_7^+$]

Step 6 (Synthesis of XL_10320_197): **S7** (0.087g, 0.12mmol) was dissolved in 2mL DCM, then 2mL 4M HCl in 1,4-dioxane was added in portions. The solution was stirred for 1 hour at room temperature. Then the mixture was concentrated under reduced pressure, and left on high vacuum overnight to remove residual acid. Then the product (0.076g, 0.12mmol) was dissolved in 3mL anhydrous DCM, and basified by adding 10 eq of Et_3N. Into the solution 4-chloroquinoline-7-carboxylic acid (0.03g, 0.14mmol) and T3P (0.23g, 0.7mmol) sequentially. The resultant solution was stirred overnight. Then the mixture was directly subjected to flash chromatography (EtOAc/MeOH) and HPLC purification (MeOH/H_2O with 4‰ TFA) to afford 25mg **XL_10320_197** (25%). ^{1}H NMR (500 MHz, MeOD) δ 8.87 (dd, *J* = 4.7, 2.5 Hz, 1H), 8.60 (dd, *J* = 4.1, 1.6 Hz, 1H), 8.38 (dd, *J* = 8.7, 6.7 Hz, 1H), 8.27 – 8.14 (m, 3H), 8.14 – 8.03 (m, 1H), 7.78 (dd, *J* = 4.8, 2.4 Hz, 1H), 7.65 (ddd, *J* = 34.7, 8.7, 2.0 Hz, 1H), 7.36 (t, *J* = 7.6 Hz, 1H), 7.32 – 7.16 (m, 4H), 5.25 – 5.08 (m, 1H), 4.27 – 4.06 (m, 3H), 4.00 (dd, *J* = 19.8, 13.9 Hz, 1H),

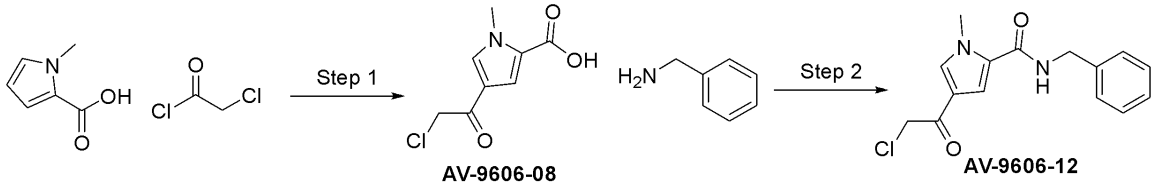
3.74 (m, 2H), 3.15 – 2.91 (m, 4H), 2.86 – 2.77 (m, 2H), 2.74 – 2.40 (m, 7H), 2.32 (s, 3H), 1.83 – 1.55 (m, 2H), 1.54 – 1.21 (m, 3H), 0.71 – 0.49 (m, 1H). LCMS (ESI) *m/z* 821.69 [$(M+H)^+$ 822.35 calcd for $C_{43}H_{49}ClN_9O_6^+$]

Synthesis of AC-10180-19:



Synthesis of AC-10180-19 followed the General Procedure 1 by using benzylamine (0.29mL, 2.7mmol), 1,2,3,4-tetrahydroquinoline-6-carboxylic acid (0.52g, 2.9mmol), HATU (1.12g, 2.9mmol), and Et_3N (1.88mL, 13.4mmol). AC-10180-19 (53mg, 75%) 1H NMR (500 MHz, DMSO) δ 8.99 (t, J = 6.0 Hz, 1H), 7.76 (s, 1H), 7.71 (m, 2H), 7.36 – 7.29 (m, 4H), 7.26 – 7.22 (m, 1H), 4.59 (s, 2H), 4.48 (d, J = 6.0 Hz, 2H), 3.78 – 3.69 (t, J = 10.0 Hz, 2H), 2.78 (t, J = 10.0 Hz, 2H), 1.98 – 1.87 (m, 2H). LC/MS (ESI) m/z 342.79; $[M+H]^+$ calcd for $C_{19}H_{20}ClN_2O_2^+$: 343.12

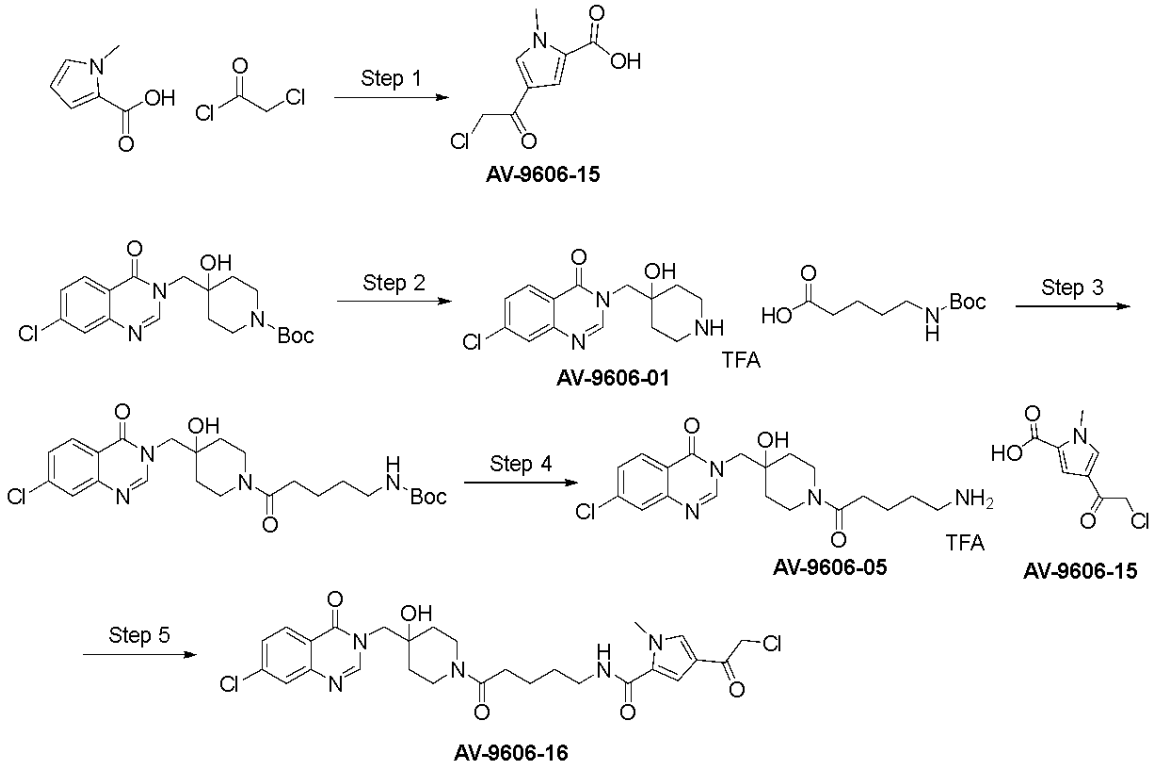
Synthesis of AV-9606-12:



Step 1: 1-methyl-1H-pyrrole-2-carboxylic acid (125.1 mg, 1.0 mmol) was added to a heat dried pressure vial flushed with nitrogen and dissolved in DCM (5 mL). The vial was then placed in an ice bath and $AlCl_3$ (288.0 mg, 2.5 mmol) was then added as a solid to the reaction and the reaction was stirred for 30 minutes on ice. After 30 minutes, chloroacetyl chloride (87.6 µL, 1.1 mmol) was then added and the reaction was heated to 45°C for 18 hours. The reaction was quenched with saturated sodium bicarbonate solution until pH was greater than 7. The reaction was washed with DCM. The aqueous layer was then acidified with concentrated hydrochloric acid until pH 0-2 at which point the desired product as a white precipitate crashed out. The reaction was filtered to isolate the desired product (126.7 mg, 63% yield). ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.90 (d, *J* = 2.0 Hz, 1H), 7.24 (d, *J* = 2.0 Hz, 1H), 4.81 (s, 2H), 3.89 (s, 3H). LC/MS (ESI) *m/z* 202.08 [M+H]+; calcd for $C_8H_9ClNO_3^+$: 202.03.

Step 2: AV-9606-08 (60.4 mg, 0.30 mmol) and HATU (137.2 mg, 0.36 mmol) were combined and suspended in THF (2 mL). Et_3N (83.5 µL, 0.6 mmol) was then added and the reaction was stirred under nitrogen. To the reaction was then added a solution of benzylamine (39.2 µL, 0.36 mmol) in THF (1 mL). The reaction was stirred at room temperature for 2 hours. The reaction was diluted with EtOAc and washed with water, saturated sodium bicarbonate, and brine. The organics were collected, dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The crude material was purified flash chromatography (DCM/MeOH) to afford the desired product (56.6 mg, 65% yield). ^{1}H NMR (500 MHz, DMSO-d6) δ 8.84 (t, J = 6.1 Hz, 1H), 7.86 (d, J = 1.9 Hz, 1H), 7.36 – 7.28 (m, 5H), 7.27 – 7.21 (m, 1H), 4.76 (s, 2H), 4.40 (d, J = 6.0 Hz, 2H), 3.89 (s, 3H). LC/MS (ESI) *m/z* 291.18 [M+H]+; calcd for $C_{15}H_{16}ClN_2O_2^+$: 291.09.

Synthesis of AV-9606-16:



Step 1: 1-methyl-1H-pyrrole-2-carboxylic acid (376.7 mg, 3.0 mmol) was dissolved in DCM (15 mL) and the reaction was placed in an ice bath. To the reaction was then added $AlCl_3$ (822.6 mg, 6.2 mmol) and the reaction was stirred for 30 minutes on ice. After 30 minutes, chloroacetyl chloride (263.0 µL, 3.3 mmol) was then added and the reaction was heated to 45°C for 22 hours. The

reaction was quenched with saturated sodium bicarbonate solution until pH was greater than 7. The reaction was washed with DCM. The aqueous layer was then acidified with concentrated hydrochloric acid until pH 0-2 at which point the desired product as a white precipitate crashed out. Filter to isolate the desired product (371.5 mg, 61% yield). LC/MS (ESI) *m/z* 202.18 [M+H]+; calcd for $C_8H_9ClNO_3^+$: 202.03.

Step 2: tert-butyl 4-((7-chloro-4-oxoquinazolin-3(4H)-yl)methyl)-4-hydroxypiperidine-1-carboxylate (1004.3 mg, 2.55 mmol) was suspended in TFA. The reaction was stirred under nitrogen gas at room temperature for 2 hours. The reaction was concentrated under reduced pressure and used in the subsequent step without further purification. LC/MS (ESI) *m/z* 294.08 [M+H]+; calcd for $C_{14}H_{17}ClN_3O_2^+$: 294.10.

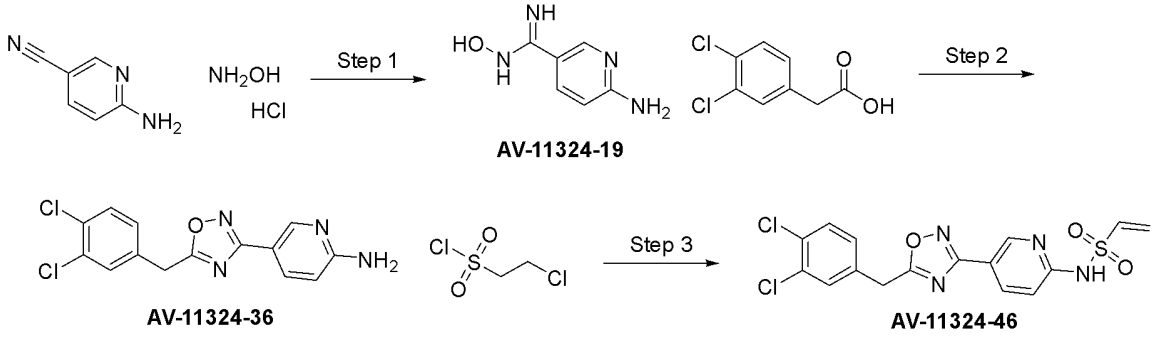
Step 3: 5-(Boc-amino)pentanoic acid (223.7 mg, 1.03 mmol) and HATU (653.2 mg, 2.0 mmol) were combined and suspended in DMF (1 mL). Et_3N (829.2 uL, 5.95 mmol) was added and the reaction was stirred at room temperature for 10 minutes. A solution of AV-9606-01 (249.7 mg, 0.85 mmol) in DMF (1 mL) was then added to the reaction. The reaction was stirred at room temperature under nitrogen for 3 hours. The reaction was diluted with EtOAc and washed with water and brine. The organics were combined, dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude material was purified via flash chromatography (hexanes/EtOAc/MeOH) to afford the desired product (370 mg, 88%). LC/MS (ESI) *m/z* 493.20 [M+H]+; calcd for $C_{24}H_{34}ClN_4O_5^+$: 493.22.

Step 4: AV-9606-02 (370.0 mg, 0.75 mmol) was dissolved in TFA (3 mL). The reaction was stirred at room temperature for 2 hours. The reaction was concentrated under reduced pressure and the crude material was purified via flash chromatography (EtOAc/MeOH/20%NH_4OH) to afford the desired product (261.3 mg, 89%). LC/MS (ESI) *m/z* 393.29 $[M+H]^+$; calcd for $C_{19}H_{26}ClN_4O_3^+$: 393.17.

Step 5: AV-9606-15 (65.2 mg, 0.32 mmol) was suspended in DCM (1 mL). Thionyl chloride (1 mL) was added and the reaction was stirred at room temperature for 1 hour. The reaction was

concentrated under reduced pressure. The crude material was suspended in DCM (1 mL) and the reaction was placed in an ice bath. To the reaction was then added Et_3N (198.7 µL, 1.43 mmol) followed by a solution of AV-9606-05 in DCM (1 mL). The reaction was warmed to room temperature and stirred for 1 hour. The reaction was quenched with water (0.5 mL). The crude material was purified via silica gel chromatography (DCM/MeOH) and fractions containing desired product were collected and concentrated under reduced pressure. The material was washed with hexanes, sonicated, and then filtered and dried under vacuum to afford the desired product (8.4 mg, 4.6%). ^{1}H NMR (500 MHz, Methanol-d_4) δ 8.81 (s, 1H), 8.35 – 8.23 (m, 1H), 7.76 (d, *J* = 2.0 Hz, 1H), 7.73 – 7.62 (m, 2H), 7.21 (d, *J* = 1.9 Hz, 1H), 4.62 (s, 2H), 4.25 – 4.10 (m, 3H), 3.78 (d, *J* = 13.8 Hz, 1H), 3.49 – 3.39 (m, 1H), 3.34 (d, *J* = 6.2 Hz, 2H), 3.21 (q, *J* = 7.3 Hz, 1H), 3.17 – 3.08 (m, 1H), 2.47 (q, *J* = 7.0 Hz, 2H), 2.01 (s, 8H), 1.75 – 1.61 (m, 6H), 1.61 – 1.51 (m, 2H), 1.35 – 1.27 (m, 2H). LC/MS (ESI) *m/z* 576.22 $[M+H]^+$; calcd for $C_{27}H_{32}Cl_2N_5O_5^+$: 576.18.

Synthesis of AV-11324-46:



Step 1: 2-amino-5-cyanopyridine (503.7 mg, 4.20 mmol) and hydroxylamine HCl (354.1 mg, 5.10 mmol) were combined and suspended in EtOH (8 mL). Et_3N (2926 µL, 21.0 mmol) was added and the reaction was stirred at 65°C for 6 hours. The reaction was then removed from heat and stirred at room temperature for 12 hours. The desired product precipitated out of the reaction and was isolated via filtration (316.4 mg, 49.5%). The material was used in the next step without further purification. LC/MS (ESI) *m/z* 153.00 $[M+H]^+$; calcd for $C_6H_9N_4O^+$: 153.08.

Step 2: 3,4-dichlorophenylacetic acid (390.6 mg, 1.90 mmol) and CDI (462.1 mg, 2.85 mmol) were combined and suspended in MeCN (5 mL). The reaction was stirred at room temperature for 2 hours, at which point AV-11324-19 (316.4 mg, 2.08 mmol) was added to the reaction. The reaction was stirred at room temperature for 30 minutes. DBU was then added (568.3 µL, 3.80 mmol) and the reaction was heated to 60°C for 40 minutes. The reaction was diluted with EtOAc ad washed with water, saturated sodium bicarbonate solution, and brine. The organics were collected, dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude material was purified via flash chromatography (DCM/MeOH) followed by preparative HPLC (MeOH or CH_3CN/H_2O with 0.0425% TFA) to afford the desired product (32.5 mg, 5.3%). LC/MS (ESI) *m/z* 320.77 $[M+H]^+$; calcd for $C_{14}H_{11}Cl_2N_4O^+$: 321.03.

Step 3: AV-11324-36 (32.5 mg, 0.10 mmol) was suspended in DCM (2 mL) and placed in an ice bath. To the reaction was then added 2-chloroethanesulfonyl chloride (26.4 µL, 0.25 mmol) followed by Et_3N (69.7 µL, 0.50 mmol). The reaction was stirred at 0°C for 3 hours. The reaction was quenched with water, extracted with DCM, and then washed with brine. The organics were collected, dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude material was purified via flash chromatography (DCM/MeOH) followed by preparative HPLC (MeOH or CH_3CN/H_2O with 0.0425% TFA) to afford the desired product (0.6 mg, 1.5%) LC/MS (ESI) *m/z* 410.87 [M+H]+; calcd for $C_{16}H_{12}Cl_2N_4O_3S^+$: 410.00. ^{1}H NMR (500 MHz, DMSO-d6) δ 8.38 (d, J = 2.2 Hz, 1H), 7.96 (dd, J = 9.3, 2.1 Hz, 1H), 7.72 (d, J = 2.1 Hz, 1H), 7.66 (d, J = 8.3 Hz, 1H), 7.41 (dd, J = 8.3, 2.1 Hz, 1H), 6.72 (d, J = 9.4 Hz, 1H), 4.76 – 4.65 (m, 2H), 4.49 (s, 2H), 3.55 – 3.43 (m, 2H).

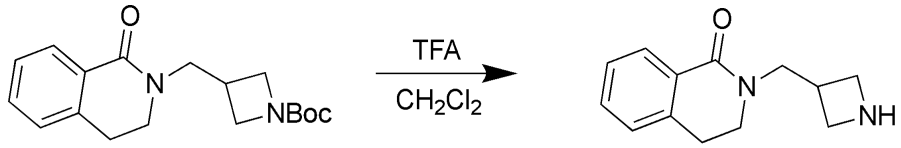
<u>Synthetic Procedure for AXA-01-045:</u>



1 **2** **3**

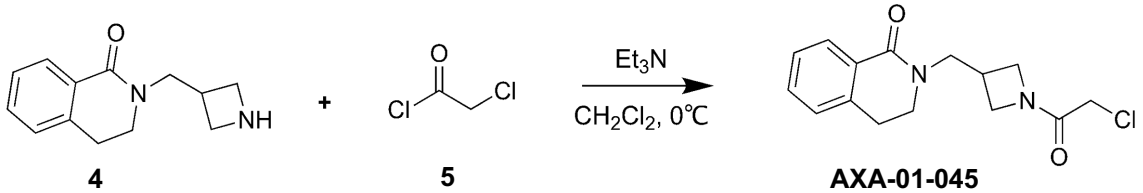
3,4-dihydroisoquinolin-1(2H)-one (1, 73.6 mg, 0.5 mmol) was taken up in DMF (2mL) and cooled to 0°C and NaH (22 mg, 0.55 mmol) was added. The mixture was stirred at 0°C for 30 minutes.

Tert-butyl 3-(bromomethyl)azetidine-1-carboxylate (2,149.5 mg, 0.60 mmol) was added to the mixture, and the reaction was warmed to RT and stirred for 2.5 hrs. The reaction was diluted with water (25mL) and extracted with ethyl acetate (25mL x 2). The combined organics were washed with brine (1 x 50mL), dried over anhydrous Na_2SO_4, filtered, concentrated and purified by column chromatography on silica gel (0% to 100% Hexanes/ EtOAc) to afford 3 as a pale-yellow solid (89.9 mg, yield 56.7%). LCMS (m/z): 317.80 $[M + H]^+$; calcd for $C_{18}H_{25}N_2O_3^+$: 317.19.



3 **4**

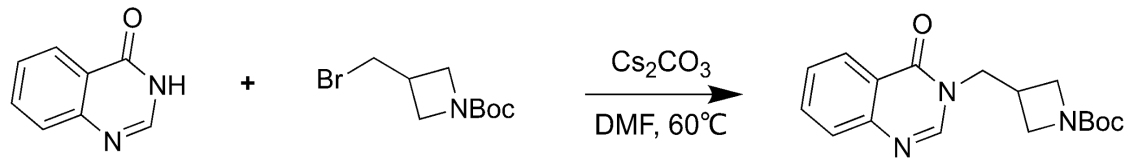
A mixture of 3 (89.9 mg, 0.284 mmol), CH_2Cl_2 (1.8 mL), and TFA (0.2 mL) was stirred at rt for 3 hrs, the reaction was concentrated in vacuum to leave the crude 4 as a white solid (quantitative yield). LCMS (m/z): 217.80 $[M + H]^+$; calcd for $C_{13}H_{17}N_2O^+$: 217.13.



4 **5** **AXA-01-045**

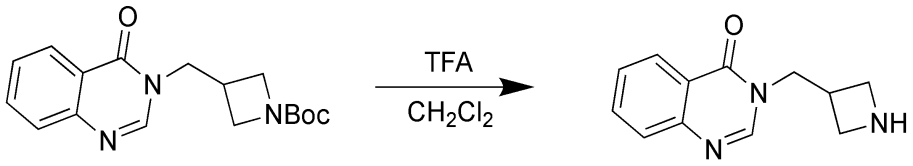
A mixture of 4 (50.0 mg, 0.151 mmol) and Et_3N (46.4 µL, 0.333 mmol) was taken up in CH_2Cl_2 (4 mL) and cooled to 0℃. 2-chloroacetyl chloride (5, 24 µL, 0.303 mmol) was added to the reaction mixture. The reaction was warmed to rt and stirred overnight. The reaction was concentrated in vacuum. The mixture was purified by preparative HPLC (MeCN/H2O with 0.0425% TFA) to afford AXA-01-045 as a white solid (17.3 mg, yield 28.1%). LCMS ESI (m/z): 292.67 $[M + H]^+$; calcd for $C_{15}H_{17}Cl_2N_2O_2^+$: 292.76. ^{1}H NMR (500 MHz, DMSO-d6) δ 7.92 (dd, J = 8.1, 1.2 Hz, 1H), 7.68 (td, J = 7.5, 1.2 Hz, 1H), 7.58 – 7.25 (m, 2H), 4.45 (s, 2H), 4.36 – 4.15 (m, 2H), 3.76 (ddd, J = 13.7, 4.8, 1.9 Hz, 1H), 3.56 (dd, J = 13.7, 8.0 Hz, 1H), 3.40 – 3.29 (m, 2H), 3.07 (t, J = 6.8 Hz, 2H), 2.69 – 2.55 (m, 1H).

<u>Synthetic Procedure for AXA-01-055:</u>



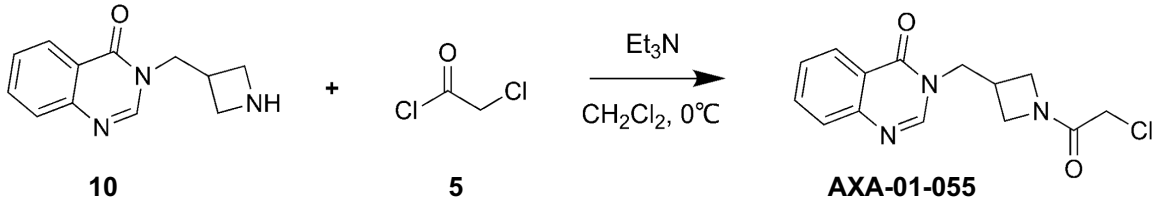
6 **2** **8**

Quinazolin-4(3H)-one (6, 73.1 mg, 0.5 mmol), tert-butyl 3-(bromomethyl)azetidine-1-carboxylate (2, 150.0 mg, 0.6 mmol), and Cs_2CO_3 (325.82 mg, 1.0 mmol) were taken up in DMF (2 mL) and stirred at 60℃ for 3 hours. The reaction was diluted with water (25 mL) and extracted with ethyl acetate (25 mL x 2). The combined organics were washed with brine (50 mL x 1), dried over anhydrous Na_2SO_4, filtered, concentrated and purified by flash chromatography (50% to 100% EtOAc /hexanes) to afford 152.6 mg product 8 (96.8%). LCMS (m/z): 316.37 $[M + H]^+$; calcd for $C_{17}H_{22}N_3O_3^+$: 316.17.



8 **9**

A mixture of tert-butyl-3-((4-oxoquinazolin-3(4H)-yl)methyl)azetidine-1-carboxylate (8, 152.6 mg, 0.484 mmol), CH_2Cl_2 (1.8 mL), and TFA (0.2 mL) was stirred at rt for 3 hours. The reaction was concentrated in vacuum to leave the crude 9 as a white solid (quantitative yield). LCMS (m/z): 216.37 $[M + H]^+$. calcd for $C_{12}H_{14}N_3O^+$: 216.11.

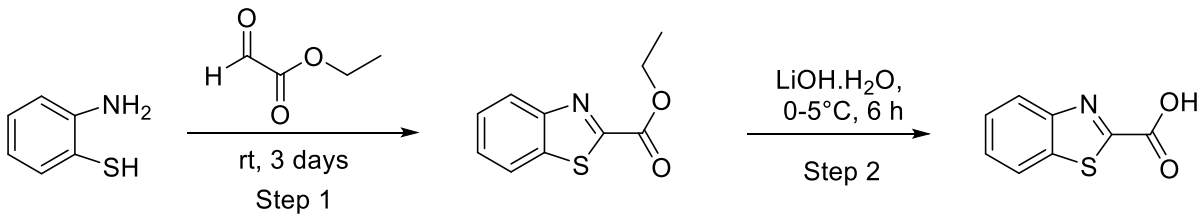


10 **5** **AXA-01-055**

A mixture of 10 (25.0 mg, 0.116 mmol) and Et_3N (81.0 µL, 0.580 mmol) was taken up in CH_2Cl_2 (2 mL) and cooled to 0℃. 2-chloroacetyl chloride (5, 18.5 µL, 0.232 mmol) was added to the reaction

mixture. The reaction was warmed to rt and stirred for 2 hours. The mixture was purified by flash chromatography (5% to 20% MeOH/EtOAc), followed by preparative HPLC (MeCN/H2O with 0.0425% TFA) to afford 12.6 mg product (37.8%). LCMS ESI (m/z): 292.74 $[M + H]^+$. ; calcd for $C_{15}H_{17}Cl_2N_3O_2^+$: 292.76. ^{1}H NMR (500 MHz, DMSO-d6) δ 8.48 (s, 1H), 8.17 (dd, J = 8.1, 1.5 Hz, 1H), 7.85 (ddd, J = 8.4, 7.1, 1.6 Hz, 1H), 7.69 (dd, J = 8.2, 1.1 Hz, 1H), 7.57 (ddd, J = 8.2, 7.1, 1.2 Hz, 1H), 4.32 – 4.21 (m, 3H), 4.07 (dd, J = 8.9, 5.6 Hz, 2H), 3.97 (t, J = 9.2 Hz, 1H), 3.80 (dd, J = 9.9, 5.7 Hz, 1H), 3.17 – 3.07 (m, 1H).

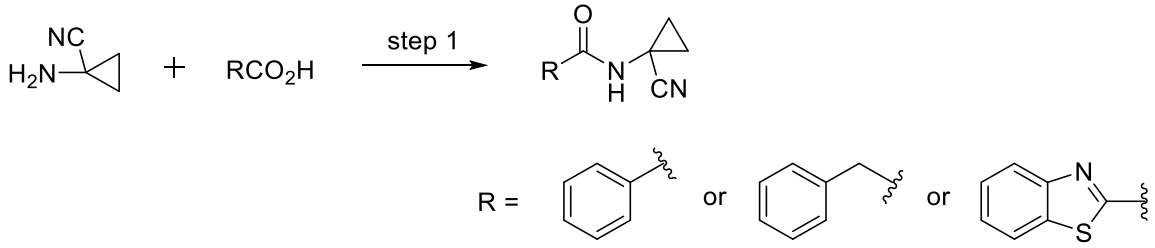
Synthesis of the scaffold: benzo[d]thiazole-2-carboxylic acid



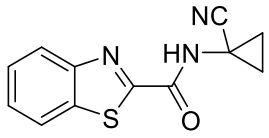
Step 1: A mixture of 2-aminobenzenethiol (2.76mL, 25.6mmol) and ethyl 2-oxoacetate (50% in toluene) (6.28mL, 30.7mmol) was stirred at room temperature for 3 days. The mixture was diluted with EtOAc, and washed with H_2O three times. The organic layer was then washed with brine and dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude material was purified by flash column chromatography (EtOAc in hexanes, 20% to 60%) and preparative HPLC (MeCN/H_2O with 0.0425% TFA) to afford desired product (1.8g, 34%). LCMS ESI (*m/z*): 208.18; $[M+H]^+$ calcd for $C_{10}H_{10}NO_2S^+$: 208.04

Step 2: To a solution of ethyl benzo[d]thiazole-2-carboxylate (0.80g, 3.8mmol) in H_2O (16mL) and THF (12mL) was added a solution of lithium hydroxide monohydrate (0.16g, 3.8mmol) in water at 0~5°C, then stirred at 0~5°C for 6 hours. The mixture was diluted with H2O (~50mL), and adjusted to pH=4~5 with 2N HCl. The precipitate was collected by filtration, washed with water, and dried in vacuo to afford off-white solid product (0.52g). The filtrate was extracted with DCM 3 times. Combined organic layers were washed with brine and dried over Na_2SO_4, filtered and concentrated to afford desired product as yellowish solid (0.13g). LCMS ESI (*m/z*): 180.01; $[M+H]^+$ calcd for $C_8H_6NO_2S^+$: 180.01

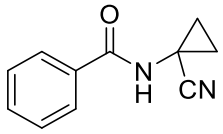
Synthesis of WH-9943-094, 098A, 098B:



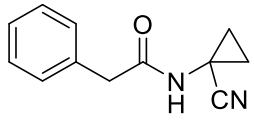
1-aminocyclopropane-1-carbonitrile (1.0 eq.), carboxylic acids (1.2 eq.) Et_3N (5.0 eq.) and HATU (1.5 eq.) were added into DMF (3-5mL). The mixture was stirred at room temperature overnight. If necessary, the mixture was diluted with EtOAc (50mL), and washed with brine (30mL×2) to remove excess DMF. Organic layer was dried over anhydrous sodium sulfate (Na_2SO_4), filtered, and concentrated under reduced pressure. The crude material was then purified by flash column chromatography (hexanes/EtOAc/MeOH) and preparative HPLC (MeCN/H_2O with 0.0425% TFA) to afford desired products.



WH-9943-094 (N-(1-cyanocyclopropyl)benzo[d]thiazole-2-carboxamide (0.03g, 55%)). 1H NMR (500 MHz, DMSO-d6) δ 10.20 (s, 1H), 8.27 (dd, J = 8.2, 1.4 Hz, 1H), 8.16 (dd, J = 8.0, 1.6 Hz, 1H), 7.75 – 7.55 (m, 2H), 1.60 (q, 2H), 1.42 (q, 2H). LCMS ESI (*m/z*): 243.98; $[M+H]^+$ calcd for $C_{12}H_{10}N_3OS^+$: 244.05

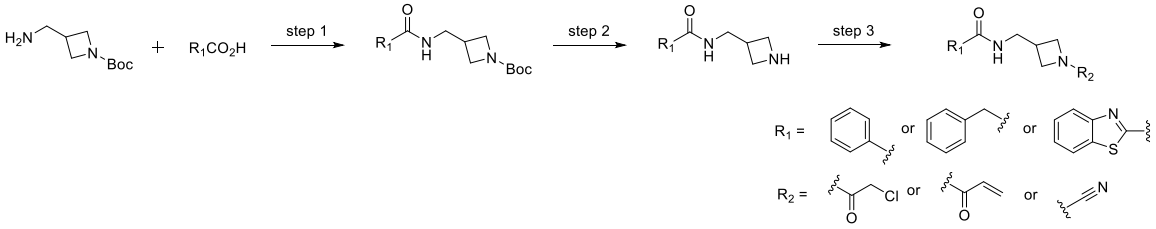


WH9943-098A (N-(1-cyanocyclopropyl)benzamide (0.067g, 85%)). 1H NMR (500 MHz, DMSO-d6) δ 9.32 (s, 1H), 7.89 – 7.80 (m, 2H), 7.63 – 7.54 (m, 1H), 7.54 – 7.45 (m, 2H), 1.57 (q, 2H), 1.29 (q, 2H). LCMS ESI (*m/z*): 187.19; $[M+H]^+$ calcd for $C_{11}H_{11}N_2O^+$: 187.09



WH-9943-098B (N-(1-cyanocyclopropyl)-2-phenylacetamide (0.087g, 100%)). 1H NMR (500 MHz, DMSO-d6) δ 9.00 (s, 1H), 7.37 – 7.28 (m, 2H), 7.28 – 7.20 (m, 3H), 3.42 (s, 2H), 1.47 (q, 2H), 1.12 (q, 2H). LCMS ESI (*m/z*): 201.08; $[M+H]^+$ calcd for $C_{12}H_{13}N_2O^+$: 201.09

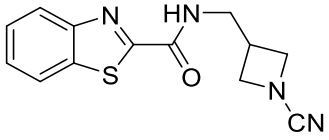
Synthesis of WH-9943-102A, 103B, 103C, 104B:



Step 1: The synthesis was preformed according to General Procedure 1 with tert-butyl 3-(aminomethyl)azetidine-1-carboxylate (1.2eq.), carboxylic acids (1.0eq.), Et_3N (3.0eq.), HATU (1.5eq.). Desired compounds were obtained (when R=phenyl, 0.34g (71%), LCMS ESI (*m/z*): 235.21 (m−*t*-butyl); $[M+H]^+$ calcd for $C_{16}H_{23}N_2O_3^+$: 291.17, when R=benzyl, 0.44g (98%), LCMS ESI (*m/z*): 249.08 (m−*t*-butyl); $[M+H]^+$ calcd for $C_{17}H_{25}N_2O_3^+$: 305.19, when R=benzylthiazol, 0.27g (42%) LCMS ESI (*m/z*): 292.02 (m−*t*-butyl); $[M+H]^+$ calcd for $C_{17}H_{22}N_3O_3S^+$: 348.14

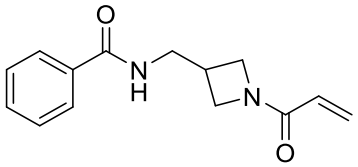
Step 2: The synthesis was performed according to the General Procedure 1 with Boc-protected azetidine derivatives using 4N HCl in dioxane. Reaction mixtures were concentrated under reduced pressure to afford crude material which were used directly without any further purification.

Step 3: The synthesis was performed according to the General Procedure 2 with those azetidines (1.0eq.) and chloroacetyl chloride (1.5eq.), or acryloyl chloride (1.5eq.), or cyanogen bromide (1.5eq.)

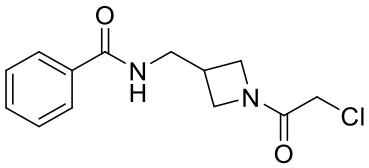


WH-9943-102A (N-((1-cyanoazetidin-3-yl)methyl)benzo[d]thiazole-2-carboxamide (3mg, 8%)). 1H NMR (500 MHz, DMSO-d6) δ 9.44 (t, J = 6.1 Hz, 1H), 8.24 (d, J = 7.9, 1.2 Hz, 1H), 8.15 (d, 1H),

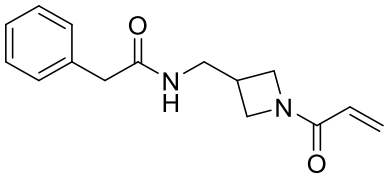
7.71 – 7.54 (m, 2H), 4.22 (t, J = 8.0 Hz, 2H), 3.97 (dd, J = 7.7, 5.7 Hz, 2H), 3.55 (t, J = 6.5 Hz, 2H), 3.02 – 2.89 (m, 1H). LCMS ESI (*m/z*): 272.97; $[M+H]^+$ calcd for $C_{13}H_{13}N_4OS^+$: 273.08



WH-9943-103B (N-((1-acryloylazetidin-3-yl)methyl)benzamide (6mg, 9%)). 1H NMR (500 MHz, DMSO-d6) δ 8.31 (t, J = 5.7 Hz, 1H), 7.93 – 7.80 (m, 2H), 7.52 – 7.46 (m, 1H), 7.46 – 7.40 (m, 2H), 6.28 (dd, J = 17.1, 10.1 Hz, 1H), 6.14 (dd, J = 17.1, 2.2 Hz, 1H), 5.64 (dd, J = 10.1, 2.2 Hz, 1H), 4.49 – 4.34 (m, 1H), 4.11 – 4.04 (m, 1H), 3.66 – 3.57 (m, 1H), 3.30 (dd, J = 16.7, 7.4 Hz, 1H), 3.25 – 3.19 (m, 2H), 2.24 – 2.08 (m, 1H). LCMS ESI (*m/z*): 245.28; $[M+H]^+$ calcd for $C_{14}H_{17}N_2O_2^+$: 245.13

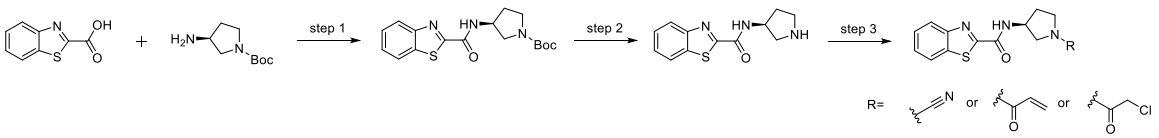


WH-9943-103C (N-((1-(2-chloroacetyl)azetidin-3-yl)methyl)benzamide (18mg, 26%)). 1H NMR (500 MHz, DMSO-d6) δ 8.42 (t, J = 5.9 Hz, 1H), 7.91 – 7.78 (m, 2H), 7.49 – 7.44 (m, 1H), 7.44 – 7.37 (m, 2H), 4.40 – 4.32 (m, 1H), 4.09 (s, 2H), 4.03 (dd, J = 10.6, 8.2 Hz, 1H), 3.62 – 3.53 (m, 1H), 3.26 (dd, J = 16.7, 7.8 Hz, 1H), 3.16 (t, J = 6.7 Hz, 2H), 2.21 – 2.01 (m, 1H). LCMS ESI (*m/z*): 267.27; $[M+H]^+$ calcd for $C_{13}H_{16}ClN_2O_2^+$: 267.09



WH-9943-104B N-((1-acryloylazetidin-3-yl)methyl)-2-phenylacetamide (20mg, 27%). 1H NMR (500 MHz, DMSO-d6) δ 8.25 (t, J = 5.9 Hz, 1H), 7.37 – 7.16 (m, 5H), 6.26 (dd, J = 16.9, 10.3 Hz, 1H), 6.08 (dd, J = 16.9, 2.3 Hz, 1H), 5.65 (dd, J = 10.3, 2.3 Hz, 1H), 4.20 (t, J = 8.5 Hz, 1H), 3.90 (t, J = 10.1, 8.5 Hz, 1H), 3.84 (dd, J = 8.7, 5.4 Hz, 1H), 3.61 (dd, J = 10.1, 5.5 Hz, 1H), 3.42 (s, 2H), 3.32 – 3.22 (m, 2H), 2.80 – 2.64 (m, 1H). LCMS ESI (*m/z*): 259.27; $[M+H]^+$ calcd for $C_{15}H_{19}N_2O_2^+$: 259.14

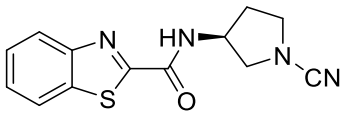
Synthesis of WH-9943-105A, 105B, 105C:



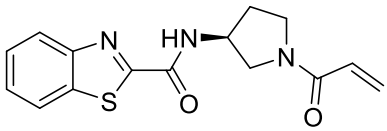
Step 1: The synthesis was preformed according to General Procedure 1 with benzo[d]thiazole-2-carboxylic acid (0.2g, 1.1mmol) and tert-butyl (S)-3-aminopyrrolidine-1-carboxylate (0.25g, 1.3mmol). 0.3g desired compound tert-butyl (S)-3-(benzo[d]thiazole-2-carboxamido)pyrrolidine-1-carboxylate was obtained (77%). LCMS ESI (*m/z*): 292.12 (m−*t*-butyl); $[M+H]^+$ calcd for $C_{17}H_{22}N_3O_3S^+$: 348.14

Step 2: The synthesis was performed according to the General Procedure 1 with tert-butyl (S)-3-(benzo[d]thiazole-2-carboxamido)pyrrolidine-1-carboxylate (0.3g, mmol). 0.29g (S)-N-(pyrrolidin-3-yl)benzo[d]thiazole-2-carboxamide (quant.). LCMS ESI (*m/z*): 247.88; $[M+H]^+$ calcd for $C_{12}H_{14}N_3OS^+$: 248.09

Step 3: The synthesis was performed according to the General Procedure 1 with (S)-N-(pyrrolidin-3-yl)benzo[d]thiazole-2-carboxamide (0.05g, 0.2mmol) and cyanogen bromide (35mg, 0.3mmol) , or acryloyl chloride (30mg, 0.3mmol), or chloroacetyl chloride (37mg, 0.3mmol).

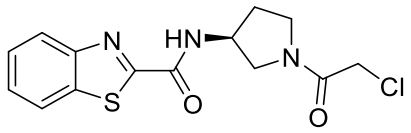


WH-9943-105A ((*S*)-N-(1-cyanopyrrolidin-3-yl)benzo[d]thiazole-2-carboxamide (34mg, 57%)). 1H NMR (500 MHz, DMSO-d6) δ 9.52 (s, 1H), 8.24 (d, J = 8.3, 1.3 Hz, 1H), 8.16 (d, J = 8.1, 1.2 Hz, 1H), 7.73 – 7.51 (m, 2H), 4.60 – 4.49 (m, 1H), 3.65 (dd, J = 9.7, 6.7 Hz, 1H), 3.63 – 3.56 (m, 1H), 3.49 – 3.42 (m, 2H), 2.22 – 2.11 (m, 1H), 2.11 – 2.02 (m, 1H). LCMS ESI (*m/z*): 273.27; $[M+H]^+$ calcd for $C_{13}H_{13}N_4OS^+$: 273.08



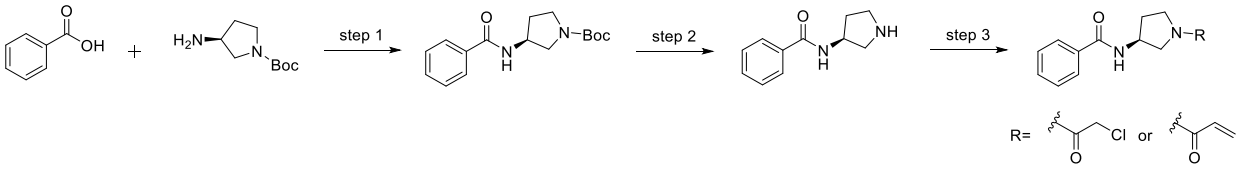
WH-9943-105B ((S)-N-(1-acryloylpyrrolidin-3-yl)benzo[d]thiazole-2-carboxamide (12mg, 18%)). 1H NMR (500 MHz, DMSO-d6) δ 8.42 (t, J = 5.9 Hz, 1H), 7.91 – 7.78 (m, 2H), 7.49 – 7.44 (m, 1H), 7.44 – 7.37 (m, 2H), 4.40 – 4.32 (m, 1H), 4.09 (s, 2H), 4.03 (dd, J = 10.6, 8.2 Hz, 1H), 3.62 – 3.53

(m, 1H), 3.26 (dd, J = 16.7, 7.8 Hz, 1H), 3.16 (t, J = 6.7 Hz, 2H), 2.21 – 2.01 (m, 1H). LCMS ESI (*m/z*): 302.27; $[M+H]^+$ calcd for $C_{15}H_{16}N_3O_2S^+$: 302.10



WH-9943-105C ((S)-N-(1-(2-chloroacetyl)pyrrolidin-3-yl)benzo[d]thiazole-2-carboxamide (26mg, 36%)) 1H NMR (500 MHz, DMSO-d6) δ 9.48 (dd, J = 14.3, 7.1 Hz, 1H), 8.24 (d, 1H), 8.15 (d, J = 8.1 Hz, 1H), 7.63 (t, 1H), 7.59 (t, 1H), 4.63 – 4.48 (m, 1H), 4.39 – 4.28 (m, 2H), 3.80 (dd, J = 10.6, 6.5 Hz, 0.5H), 3.75 – 3.61 (m, 1H), 3.61 – 3.49 (m, 1.5H), 3.50 – 3.37 (m, 1H), 2.33 – 2.03 (m, 2H). LCMS ESI (*m/z*): 324.17; $[M+H]^+$ calcd for $C_{14}H_{15}ClN_3O_2S^+$: 324.06

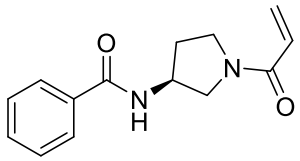
Synthesis of WH-9943-107B, 107C:



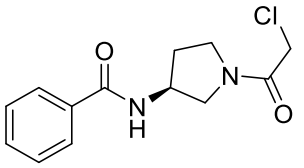
Step 1: The synthesis was preformed according to General Procedure 1 with benzoic acid (0.2g, 1.6mmol) and tert-butyl (S)-3-aminopyrrolidine-1-carboxylate (0.37g, 2.0mmol). 0.45g desired compound tert-butyl tert-butyl (S)-3-benzamidopyrrolidine-1-carboxylate was obtained (95%). LCMS ESI (*m/z*): 235.21 (m–*t*-butyl); $[M+H]^+$ calcd for $C_{16}H_{23}N_2O_3^+$: 291.17

Step 2: The synthesis was performed according to the General Procedure 1 with tert-butyl (S)-3-benzamidopyrrolidine-1-carboxylate (0.45g, 1.5mmol). 0.34g (S)-N-(pyrrolidin-3-yl)benzamide (quant.). LCMS ESI (*m/z*): 191.09; $[M+H]^+$ calcd for $C_{11}H_{15}N_2O^+$: 191.12

Step 3: The synthesis was performed according to the General Procedure 1 with (S)-N-(pyrrolidin-3-yl)benzamide (50mg, 0.2mmol) and chloroacetyl chloride (37mg, 0.3mmol), or acryloyl chloride (30mg, 0.3mmol).

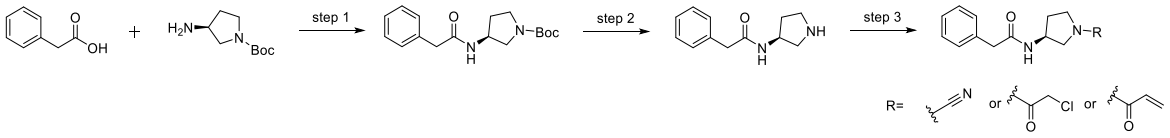


WH-9943-107B ((S)-N-(1-acryloylpyrrolidin-3-yl)benzamide (29mg, 54%). 1H NMR (500 MHz, DMSO-d6) δ 8.66 (dd, J = 18.9, 6.7 Hz, 1H), 8.05 – 7.88 (m, 2H), 7.68 – 7.58 (m, 1H), 7.55 (td, J = 7.5, 2.2 Hz, 2H), 6.78 – 6.54 (m, 1H), 6.32 – 6.13 (m, 1H), 5.75 (td, J = 10.7, 2.4 Hz, 1H), 4.71 – 4.46 (m, 1H), 3.95 (dd, J = 10.6, 6.6 Hz, 0.5H), 3.85 – 3.47 (m, 3.5H), 2.37 – 2.14 (m, 1H), 2.15 – 1.95 (m, 1H). LCMS ESI (*m/z*): 245.28; $[M+H]^+$ calcd for $C_{14}H_{17}N_2O_2^+$: 245.13



WH-9943-107C ((S)-N-(1-(2-chloroacetyl)pyrrolidin-3-yl)benzamide (30mg, 51%)). 1H NMR (500 MHz, DMSO-d6) δ 8.56 (dd, J = 20.4, 6.6 Hz, 1H), 7.91 – 7.80 (m, 2H), 7.58 – 7.51 (m, 1H), 7.50 – 7.42 (m, 2H), 4.60 – 4.40 (m, 1H), 4.38 – 4.24 (m, 2H), 3.81 (dd, J = 10.6, 6.5 Hz, 0.5H), 3.72 – 3.49 (m, 2H), 3.48 – 3.35 (m, 1.5H), 2.26 – 2.06 (m, 1H), 2.07 – 1.88 (m, 1H). LCMS ESI (*m/z*): 267.17; $[M+H]^+$ calcd for $C_{13}H_{16}ClN_2O_2^+$: 267.09

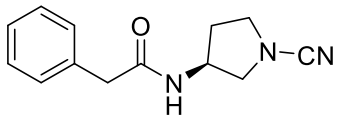
Synthesis of WH-9943-108A, 108B, 108C:



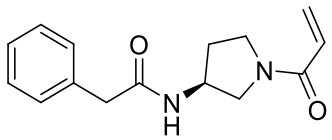
Step 1: The synthesis was preformed according to General Procedure 1 with 2-phenylacetic acid (0.20g, 1.5mmol) and tert-butyl (S)-3-aminopyrrolidine-1-carboxylate (0.33g, 1.8mmol). 0.42g desired compound tert-butyl (S)-3-(2-phenylacetamido)pyrrolidine-1-carboxylate was obtained (94%). LCMS ESI (*m/z*): 249.08 (m−*t*-butyl); $[M+H]^+$ calcd for $C_{17}H_{25}N_2O_3^+$: 305.19

Step 2: The synthesis was performed according to the General Procedure 1 with tert-butyl (S)-3-(2-phenylacetamido)pyrrolidine-1-carboxylate (0.42g, 1.4mmol). 0.39g (S)-2-phenyl-N-(pyrrolidin-3-yl)acetamide (quant.) . LCMS ESI (*m/z*): 204.98; $[M+H]^+$ calcd for $C_{12}H_{17}N_2O^+$: 205.13

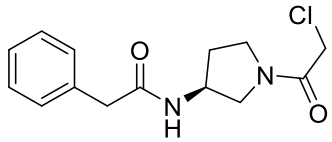
Step 3: The synthesis was performed according to the General Procedure 1 (S)-2-phenyl-N-(pyrrolidin-3-yl)acetamide (0.06g, 0.25mmol) and cyanogen bromide (0.04g, 0.37mmol), or chloroacetyl chloride (0.04g, 0.37mmol), or acryloyl chloride (0.03g,0.37 mmol).



WH-9943-108A ((S)-N-(1-cyanopyrrolidin-3-yl)-2-phenylacetamide (31mg, 54%)). 1H NMR (500 MHz, DMSO-d6) δ 8.41 (d, J = 6.7 Hz, 1H), 7.40 – 7.04 (m, 5H), 4.27 – 4.16 (m, 1H), 3.54 (dd, J = 9.7, 6.1 Hz, 1H), 3.51 – 3.46 (m, 1H), 3.46 – 3.39 (m, 3H), 3.14 (dd, J = 9.7, 3.9 Hz, 1H), 2.11 – 1.97 (m, 1H), 1.84 – 1.70 (m, 1H). LCMS ESI (*m/z*): 230.38; $[M+H]^+$ calcd for $C_{13}H_{16}N_3O^+$: 230.13

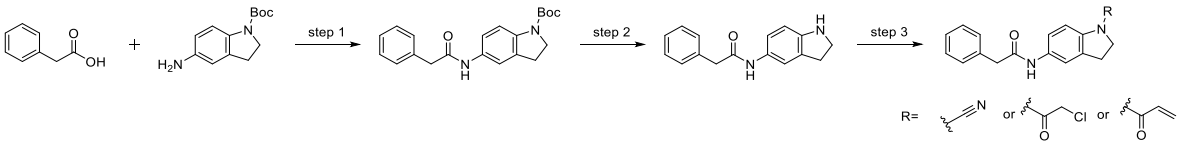


WH-9943-108B ((S)-N-(1-acryloylpyrrolidin-3-yl)-2-phenylacetamide (36mg, 56%)) 1H NMR (500 MHz, DMSO-d6) δ 8.38 (dd, J = 17.3, 6.7 Hz, 1H), 7.41 – 7.02 (m, 5H), 6.73 – 6.41 (m, 1H), 6.14 (ddd, J = 16.7, 10.6, 2.4 Hz, 1H), 5.67 (ddd, J = 15.8, 10.3, 2.4 Hz, 1H), 4.40 – 4.11 (m, 1H), 3.75 (dd, J = 10.6, 6.1 Hz, 0.5H), 3.68 – 3.15 (m, 3.5H), 2.20 – 1.93 (m, 1H), 1.92 – 1.63 (m, 1H). LCMS ESI (*m/z*): 259.27; $[M+H]^+$ calcd for $C_{15}H_{19}N_2O_2^+$: 259.14



WH-9943-108C ((S)-N-(1-(2-chloroacetyl)pyrrolidin-3-yl)-2-phenylacetamide (26mg, 37%)) 1H NMR (500 MHz, DMSO-d6) δ 8.37 (dd, J = 13.4, 6.7 Hz, 1H), 7.42 – 7.11 (m, 5H), 4.36 – 4.14 (m, 3H), 3.70 (dd, J = 10.5, 6.2 Hz, 0.5H), 3.61 – 3.34 (m, 4.5H), 3.31 – 3.18 (m, 1H), 2.19 – 1.92 (m, 1H), 1.92 – 1.65 (m, 1H). LCMS ESI (*m/z*): 281.27; $[M+H]^+$ calcd for $C_{14}H_{18}ClN_2O_2^+$: 281.11

<u>Synthesis of WH-9943-118A, 118B,118C:</u>

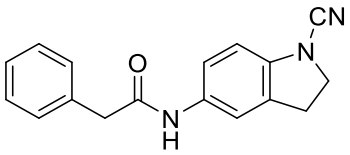


Step 1: The synthesis was preformed according to General Procedure 1 with 2-phenylacetic acid (0.20g, 1.5mmol) and tert-butyl 5-aminoindoline-1-carboxylate (0.42g, 1.7mmol). 0.52g desired

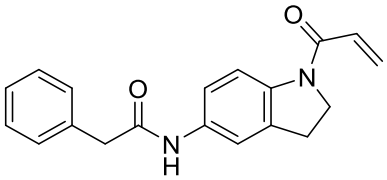
compound tert-butyl 5-(2-phenylacetamido)indoline-1-carboxylate was obtained (quant.). LCMS ESI (*m/z*): 296.87 (m−*t*-butyl); $[M+H]^+$ calcd for $C_{21}H_{25}N_2O_3^+$: 353.19

Step 2: The synthesis was performed according to the General Procedure 1 with tert-butyl 5-(2-phenylacetamido)indoline-1-carboxylate (0.52g, 1.47mmol). 0.37g N-(indolin-5-yl)-2-phenylacetamide (quant.). LCMS ESI (*m/z*): 252.87; $[M+H]^+$ calcd for $C_{16}H_{17}N_2O^+$: 253.13

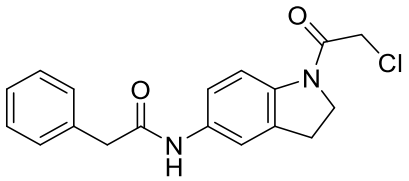
Step 3: The synthesis was performed according to the General Procedure 1 N-(indolin-5-yl)-2-phenylacetamide (0.07g, 0.24mmol) and cyanogen bromide (0.04g, 0.37mmol), or chloroacetyl chloride (0.033g, 0.37mmol), or acryloyl chloride (0.04g, 0.37mmol).



WH-9943-118A (N-(1-cyanoindolin-5-yl)-2-phenylacetamide (36mg, 53%)) 1H NMR (500 MHz, DMSO-d6) δ 10.14 (s, 1H), 7.58 (d, J = 1.9 Hz, 1H), 7.42 (dd, J = 8.5, 2.1 Hz, 1H), 7.38 – 7.30 (m, 4H), 7.29 – 7.23 (m, 1H), 6.87 (d, J = 8.5 Hz, 1H), 4.08 (t, J = 8.5 Hz, 2H), 3.63 (s, 2H), 3.15 (t, J = 8.5 Hz, 2H). LCMS ESI (*m/z*): 277.87; $[M+H]^+$ calcd for $C_{17}H_{16}N_3O^+$: 278.13

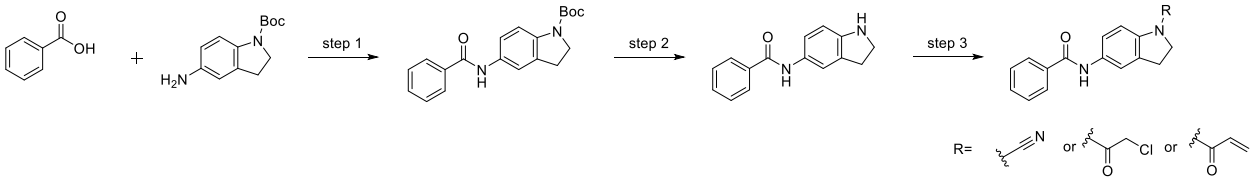


WH-9943-118B (N-(1-acryloylindolin-5-yl)-2-phenylacetamide (49mg, 66%)) 1H NMR (500 MHz, DMSO-d6) δ 10.12 (s, 1H), 8.07 (d, J = 8.8 Hz, 1H), 7.60 (d, J = 2.1 Hz, 1H), 7.38 – 7.28 (m, 5H), 7.28 – 7.21 (m, 1H), 6.72 (dd, J = 16.7, 10.3 Hz, 1H), 6.28 (dd, J = 16.7, 2.3 Hz, 1H), 5.79 (dd, J = 10.4, 2.2 Hz, 1H), 4.19 (t, J = 8.4 Hz, 2H), 3.62 (s, 2H), 3.13 (t, J = 8.4 Hz, 2H). LCMS ESI (*m/z*): 306.87; $[M+H]^+$ calcd for $C_{19}H_{19}N_2O_2^+$: 307.14



WH-9943-118C (N-(1-(2-chloroacetyl)indolin-5-yl)-2-phenylacetamide (33mg, 41%)) 1H NMR (500 MHz, DMSO-d6) δ 10.14 (s, 1H), 7.96 (d, J = 8.7 Hz, 1H), 7.61 (d, J = 2.0 Hz, 1H), 7.38 – 7.29 (m, 5H), 7.28 – 7.20 (m, 1H), 4.50 (s, 2H), 4.12 (t, J = 8.4 Hz, 2H), 3.62 (s, 2H), 3.14 (t, J = 8.4 Hz, 2H). LCMS ESI (*m/z*): 328.87; $[M+H]^+$ calcd for $C_{18}H_{18}ClN_2O_2^+$: 329.11

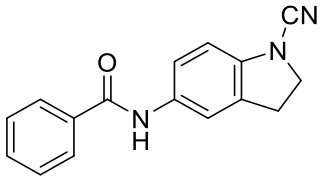
Synthesis of WH-9943-119A, 119B, 119C:



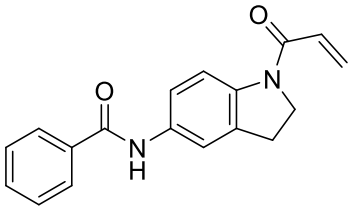
Step 1: The synthesis was preformed according to General Procedure 1 with benzoic acid (0.20g, 1.6mmol) and tert-butyl 5-aminoindoline-1-carboxylate (0.38g, 1.6mmol). 0.36g desired compound tert-butyl 5-benzamidoindoline-1-carboxylate was obtained (67%). LCMS ESI (*m/z*): **241.11** (m−*t*-butyl); $[M+H]^+$ calcd for $C_{20}H_{23}N_2O^+$: 339.17

Step 2: The synthesis was performed according to the General Procedure 1 with tert-butyl 5-benzamidoindoline-1-carboxylate (0.34g, 1.0mmol). 0.24g N-(indolin-5-yl)benzamide (quant.). LCMS ESI (*m/z*): 238.98; $[M+H]^+$ calcd for $C_{15}H_{15}N_2O^+$: 239.12

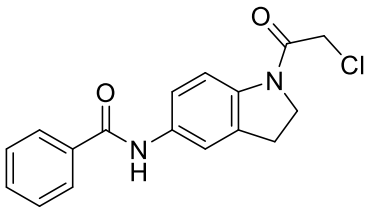
Step 3: The synthesis was performed according to the General Procedure 1 N-(indolin-5-yl)benzamide (0.05g, 0.18mmol) and cyanogen bromide (0.058g, 0.5mmol), or chloroacetyl chloride (0.031g, 0.27mmol), or acryloyl chloride (0.025g, 0.27mmol).



WH-9943-119A (N-(1-cyanoindolin-5-yl)benzamide (34mg, 71%) 1H NMR (500 MHz, DMSO-d6) δ 10.22 (s, 1H), 8.01 – 7.87 (m, 2H), 7.73 (d, J = 1.9 Hz, 1H), 7.65 – 7.56 (m, 2H), 7.57 – 7.48 (m, 2H), 6.92 (d, J = 8.5 Hz, 1H), 4.11 (t, J = 8.5 Hz, 2H), 3.19 (t, J = 8.5 Hz, 2H). LCMS ESI (*m/z*): 263.87; $[M+H]^+$ calcd for $C_{16}H_{14}N_3O^+$: 264.11

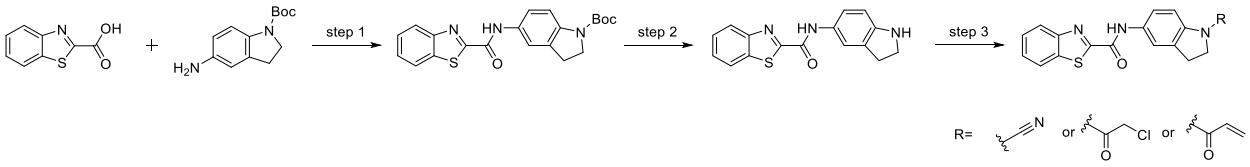


WH-9943-119B (N-(1-acryloylindolin-5-yl)benzamide (39mg, 73%)) 1H NMR (500 MHz, DMSO-d6) δ 10.22 (s, 1H), 8.13 (d, J = 8.7 Hz, 1H), 8.03 – 7.87 (m, 2H), 7.76 (d, J = 2.1 Hz, 1H), 7.63 – 7.56 (m, 1H), 7.56 – 7.48 (m, 3H), 6.75 (dd, J = 16.6, 10.3 Hz, 1H), 6.31 (dd, J = 16.6, 2.4 Hz, 1H), 5.81 (dd, J = 10.2, 2.3 Hz, 1H), 4.23 (t, J = 8.4 Hz, 2H), 3.19 (t, J = 9.4, 7.0 Hz, 2H). LCMS ESI (*m/z*): 292.97; $[M+H]^+$ calcd for $C_{18}H_{17}N_2O_2^+$: 293.13



WH-9943-119C (N-(1-(2-chloroacetyl)indolin-5-yl)benzamide (43mg, 75%)) 1H NMR (500 MHz, DMSO-d6) δ 10.23 (s, 1H), 8.01 (d, J = 8.8 Hz, 1H), 7.98 – 7.91 (m, 2H), 7.78 (d, J = 2.1 Hz, 1H), 7.62 – 7.57 (m, 1H), 7.53 (dd, J = 8.2, 6.6 Hz, 3H), 4.53 (s, 2H), 4.16 (t, J = 8.4 Hz, 2H), 3.20 (t, J = 8.5 Hz, 2H). LCMS ESI (*m/z*): 314.87; $[M+H]^+$ calcd for $C_{17}H_{16}ClN_2O_2^+$: 315.09

Synthesis of WH-9943-120A, 120B, 120C:

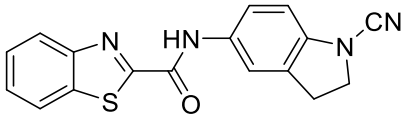


Step 1: The synthesis was preformed according to General Procedure 1 with benzo[d]thiazole-2-carboxylic acid (0.20g, 1.1mmol) and tert-butyl 5-aminoindoline-1-carboxylate (0.31g, 1.3mmol). 0.32g desired compound tert-butyl 5-(benzo[d]thiazole-2-carboxamido)indoline-1-carboxylate was obtained (73%). LCMS ESI (*m/z*): 340.20 (m−*t*-butyl); $[M+H]^+$ calcd for $C_{21}H_{22}N_3O_3S^+$: 396.14

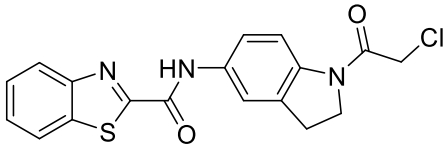
Step 2: The synthesis was performed according to the General Procedure 2 with tert-butyl 5-(benzo[d]thiazole-2-carboxamido)indoline-1-carboxylate (0.32g, mmol). 0.24g N-(indolin-5-

yl)benzo[d]thiazole-2-carboxamide (quant.). LCMS ESI (*m/z*): 295.87; $[M+H]^+$ calcd for $C_{16}H_{14}N_3OS^+$: 296.09

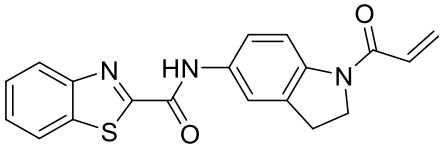
Step 3: The synthesis was performed according to the General Procedure N-(indolin-5-yl)benzo[d]thiazole-2-carboxamide (0.05g, 0.15mmol) and cyanogen bromide (0.048g, 0.45mmol), or chloroacetyl chloride (0.026g, 0.23mmol), or acryloyl chloride (0.021g, 0.23mmol).



WH-9943-120A (N-(1-cyanoindolin-5-yl)benzo[d]thiazole-2-carboxamide (24mg, 50%))1H NMR (500 MHz, DMSO-d6) δ 11.01 (s, 1H), 8.17 (dd, 2H), 7.76 (d, J = 2.1 Hz, 1H), 7.70 (dd, J = 8.5, 2.1 Hz, 1H), 7.64 – 7.57 (m, 1H), 7.57 – 7.50 (m, 1H), 6.86 (d, J = 8.3 Hz, 1H), 4.05 (t, J = 8.5 Hz, 2H), 3.13 (t, J = 8.5 Hz, 2H). LCMS ESI (*m/z*): 320.87; $[M+H]^+$ calcd for $C_{17}H_{13}N_4OS^+$: 321.08



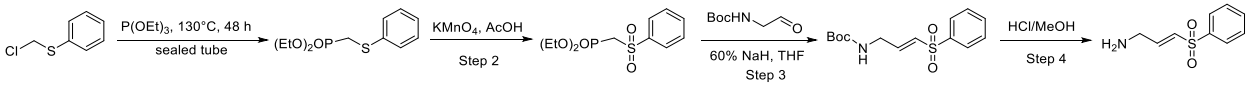
WH-9943-120C (N-(1-(2-chloroacetyl)indolin-5-yl)benzo[d]thiazole-2-carboxamide (35mg, 67%)). 1H NMR (500 MHz, DMSO-d6) δ 11.05 (s, 1H), 8.25 (dd, J = 8.0, 1.3 Hz, 1H), 8.21 (d, 1H), 8.14 (d, J = 8.7 Hz, 1H), 7.85 (d, J = 2.1 Hz, 1H), 7.71 (dd, J = 8.6, 2.2 Hz, 1H), 7.69 – 7.63 (m, 1H), 7.63 – 7.57 (m, 1H), 6.74 (dd, J = 16.6, 10.2 Hz, 1H), 6.31 (d, J = 16.6, 2.2 Hz, 1H), 5.81 (dd, J = 10.3, 2.2 Hz, 1H), 4.23 (t, J = 8.5 Hz, 2H), 3.19 (t, 2H). LCMS ESI (*m/z*): 371.77; $[M+H]^+$ calcd for $C_{18}H_{15}ClN_3O_2S^+$: 372.06



WH-9943-120B (N-(1-acryloylindolin-5-yl)benzo[d]thiazole-2-carboxamide (34mg, 61%)). 1H NMR (500 MHz, DMSO-d6) δ 11.07 (s, 1H), 8.26 (d, J = 8.0, 1.3 Hz, 1H), 8.22 (d, 1H), 8.03 (d, J = 8.7 Hz, 1H), 7.87 (d, J = 2.4 Hz, 1H), 7.72 (dd, J = 8.7, 2.2 Hz, 1H), 7.70 – 7.64 (m, 1H), 7.64 – 7.57

(m, 1H), 4.54 (s, 2H), 4.16 (t, J = 8.4 Hz, 2H), 3.20 (t, J = 8.2 Hz, 2H). LCMS ESI (*m/z*): 349.87; $[M+H]^+$ calcd for $C_{19}H_{16}N_3O_2S^+$: 350.10

Synthesis of the scaffold: (E)-3-(phenylsulfonyl)prop-2-en-1-amine

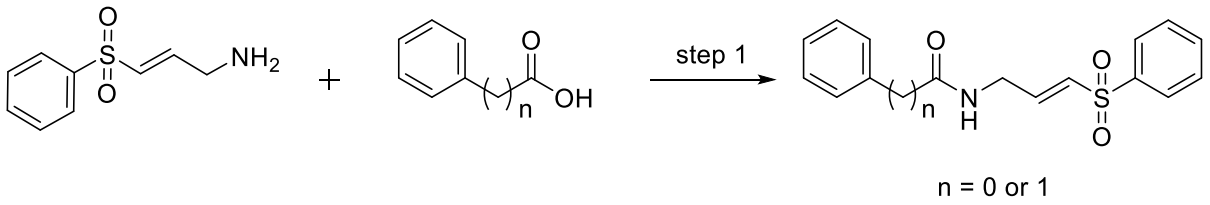


Step 1 and 2: A mixture of (chloromethyl)(phenyl)sulfane (1.26mL, 12.5mmol) and (2.14mL 12.3mmol) in a sealed tube was stirred at 130°C for 48 hours. The mixture was cooled to room temperature and dissolved in AcOH (10mL), $KMnO_4$ (3.74g, 23.6mmol) in H_2O (20mL) was added dropwisely to the solution slowly at 5 to 15°C. Then stirred at room temperature for 1 hour. Saturated $NaHCO_3$ aqueous solution was added until the mixture became colorless (below 15°C). The mixture was extracted with EtOAc twice. Combined organic layers were washed with water and brine, dried over Na2SO4, filtered and concentrated to afford crude material as colorless oil, which was purified by flash chromatography (EtOAc in hexanes, 0 to 100%) to afford product as colorless oil (XX) LCMS ESI (*m/z*): 292.77; $[M+H]^+$ calcd for $C_{11}H_{18}O_5PS^+$: 293.06

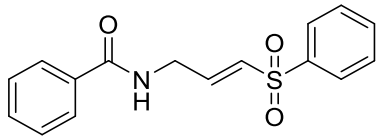
Step 3: Under N_2 atomsphere, to the solution of diethyl ((phenylsulfonyl)methyl)phosphonate (0.60g, 2.0mmol) in THF was added 60% NaH (0.099g, 2.5mmol) at at 0 to 5°C and the solution was stirred for 40min. N-Boc-2-aminoacetaldehyde (0.33g, 2.0mmol) was then added. The reaction was stirred for 15min at 5°C. The mixture was then quenched with saturated NH_4Cl and extracted with EtOAc. The combined organic layers were washed with brine and dried over Na_2SO_4, then filtered and concentrated to afford crude material which was purified by flash chromatography (EtOAc in hexanes, 0 to 100%) to afford product (0.27g, 45%). LCMS ESI (*m/z*): 197.87 (m−Boc); $[M+H]^+$ calcd for $C_{14}H_{20}NO_4S^+$: 298.11

Step 4: A solution of tert-butyl (E)-(3-(phenylsulfonyl)allyl)carbamate (0.27g, 0.9mmol) was dissolved in 2N HCl in MeOH. The solution was stirred at room temperature for 2 hours. The mixture was then concentrated, and the residue was triturated with acetonitrile. Precipitate was collected by filtration, washed with more acetonitrile and dried in vacuo to afford off-white solid as product (0.21g, quant.), LCMS ESI (*m/z*): 197.98; $[M+H]^+$ calcd for $C_9H_{12}NO_2S^+$: 198.06

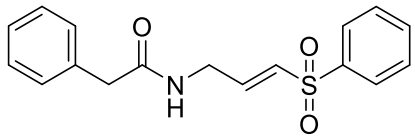
Synthesis of WH-9943-127A, 127B:



Step 1: (E)-3-(phenylsulfonyl)prop-2-en-1-amine (0.07g, 0.3mmol, 1.0 eq.), carboxylic acids (1.5 eq.) Et_3N (5.0 eq.) and HATU (1.5 eq.) were added into DMF (3-5mL). The mixture was stirred at room temperature overnight. If necessary, the mixture was diluted with EtOAc (50mL), and washed with brine (30mL×2) to remove excess DMF. Organic layer was dried over anhydrous sodium sulfate (Na_2SO_4), filtered, and concentrated under reduced pressure. The crude material was then purified by flash column chromatography (hexanes/EtOAc) and preparative HPLC (preparative HPLC (MeCN/H_2O with 0.0425% TFA)

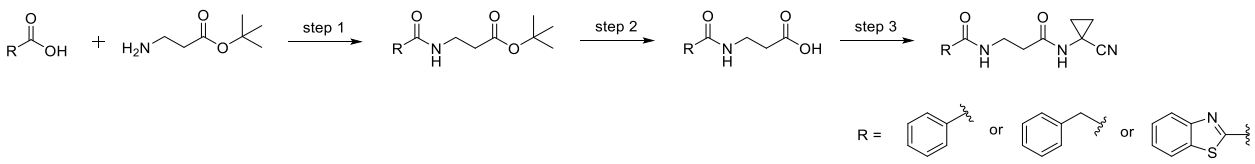


WH-9943-127A ((E)-N-(3-(phenylsulfonyl)allyl)benzamide.(10mg, 11%)) ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.72 (d, *J* = 10.7 Hz, 1H), 7.96 – 7.90 (m, 2H), 7.77 – 7.72 (m, 2H), 7.71 – 7.66 (m, 1H), 7.66 – 7.57 (m, 3H), 7.52 (t, *J* = 8.2, 6.9 Hz, 2H), 7.01 (t, *J* = 10.7, 9.1 Hz, 1H), 4.66 (q, *J* = 8.5 Hz, 1H), 4.49 (d, *J* = 8.2 Hz, 2H). LCMS ESI (*m/z*): 301.87; $[M+H]^+$ calcd for $C_{16}H_{16}NO_3S^+$: 302.08



WH-9943-127B ((E)-2-phenyl-N-(3-(phenylsulfonyl)allyl)acetamide (12mg, 13%)) ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.69 (d, *J* = 11.1 Hz, 1H), 7.93 – 7.84 (m, 2H), 7.77 – 7.68 (m, 1H), 7.65 – 7.59 (m, 2H), 7.36 – 7.28 (m, 2H), 7.27 – 7.18 (m, 3H), 6.81 – 6.68 (m, 1H), 4.51 (q, *J* = 8.5 Hz, 1H), 4.33 (dd, *J* = 8.1, 1.0 Hz, 2H), 3.48 (s, 2H). LCMS ESI (*m/z*): 315.97; $[M+H]^+$ calcd for $C_{17}H_{18}NO_3S^+$: 316.10

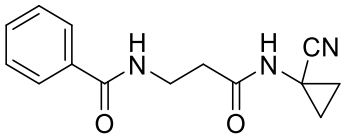
Synthesis of WH-9943-157A, 157B and WH-10417-038A:



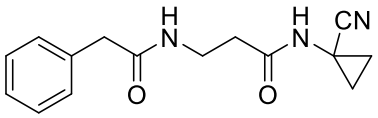
Step 1: The synthesis was preformed according to General Procedure 1 with tert-butyl 3-aminopropanoate (0.98g, 5.4mmol) and acids (1.0 eq.) and desired compounds were obtained (when R=phenyl, 1.3g (96%), LCMS ESI (*m/z*): 193.99 (m−*t*-butyl); $[M+H]^+$ calcd for $C_{14}H_{20}NO_3^+$: 250.14, when R=benzyl, 0.64g (45%), LCMS ESI (*m/z*): 208.08 (m−*t*-butyl); $[M+H]^+$ calcd for $C_{15}H_{22}NO_3^+$: 264.16, when R=benzylthiazol, 0.25g (42%), LCMS ESI (*m/z*): 250.88 (m−*t*-butyl); $[M+H]^+$ calcd for $C_{15}H_{19}N_2O_3S^+$: 307.11)

Step 2: The synthesis was performed according to the General Procedure 1 with t-Butyl ester intermediates synthesized in Step 1. Free acid products were obtained by flash column chromatography (EtOAc in hexanes 0 to 100%) (when R=phenyl, 0.70g (70%), LCMS ESI (*m/z*): 193.99; $[M+H]^+$ calcd for $C_{10}H_{12}NO_3^+$: 194.08, when R=benzyl, 0.44g (82%), LCMS ESI (*m/z*): 208.08; $[M+H]^+$ calcd for $C_{11}H_{14}NO_3^+$: 208.10, when R=benzylthiazol, 0.2g (quant.), LCMS ESI (*m/z*): 250.98; $[M+H]^+$ calcd for $C_{11}H_{11}N_2O_3S^+$: 251.05)

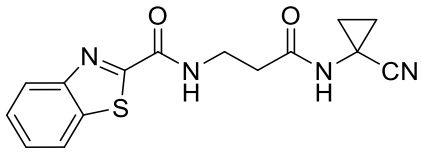
Step 3: The synthesis was performed according to the General Procedure 1 with 1-aminocyclopropane-1-carbonitrile (0.04g, 0.34mmol) and acids (1.0eq.) synthesized in step 2, Et_3N (5.0eq.) and HATU (1.5eq.) The crude materials were purified by preparative HPLC ((MeCN/H_2O with 0.0425% TFA) to afford the products after concentration under reduced pressure, which were then triturated with DCM and collected by filtration to obtain the products as white solid.



WH-9943-157A N-(3-((1-cyanocyclopropyl)amino)-3-oxopropyl)benzamide (42mg, 48%). ^1H NMR (500 MHz, DMSO-d_6) δ 8.82 (s, 1H), 8.53 (t, *J* = 5.7 Hz, 1H), 7.90 – 7.73 (m, 2H), 7.55 – 7.49 (m, 1H), 7.46 (dd, *J* = 8.2, 6.7 Hz, 2H), 3.46 (td, *J* = 7.0, 5.5 Hz, 2H), 2.39 (t, *J* = 7.0 Hz, 2H), 1.44 (q, 2H), 1.12 (q, 2H). LCMS ESI (*m/z*): 257.97; $[M+H]^+$ calcd for $C_{14}H_{16}N_3O_2^+$: 258.12

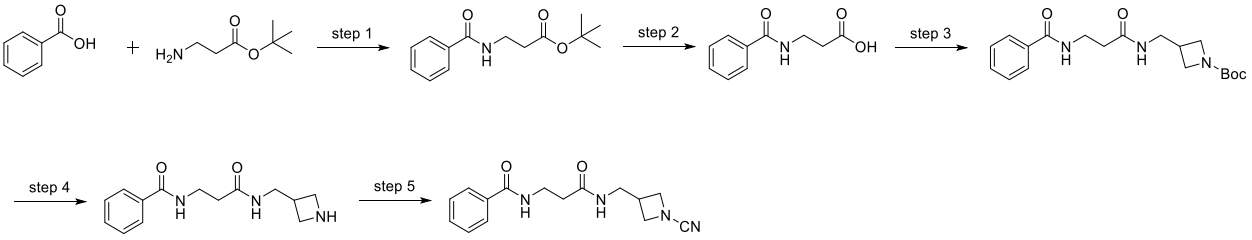


WH-9943-157B N-(1-cyanocyclopropyl)-3-(2-phenylacetamido)propenamide (22mg, 24%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.77 (s, 1H), 8.12 (t, *J* = 5.8 Hz, 1H), 7.33 – 7.26 (m, 2H), 7.25 – 7.17 (m, 3H), 3.38 (s, 2H), 3.35 (s, 3H), 3.24 (q, *J* = 6.7 Hz, 2H), 2.25 (t, *J* = 6.9 Hz, 2H), 1.43 (q, 2H), 1.08 (q, 2H). LCMS ESI (*m/z*): 271.97; $[M+H]^+$ calcd for $C_{15}H_{18}N_3O_2^+$: 272.14



WH-10417-038A N-(3-((1-cyanocyclopropyl)amino)-3-oxopropyl)benzo[d]thiazole-2-carboxamide (31mg, 49%). 1H NMR (500 MHz, DMSO-d6) δ 9.14 (t, J = 5.9 Hz, 1H), 8.86 (s, 1H), 8.23 (d, J = 7.8 Hz, 1H), 8.14 (d, J = 8.0 Hz, 1H), 7.64 (t, 1H), 7.59 (t, J = 7.5 Hz, 1H), 3.53 (q, J = 6.9 Hz, 2H), 2.45 (t, J = 7.1 Hz, 2H), 1.45 (q, 2H), 1.13 (q, 2H). LCMS ESI (*m/z*): 315.17; $[M+H]^+$ calcd for $C_{15}H_{15}N_4O_2S^+$: 315.09

<u>Synthesis of WH-9943-186:</u>



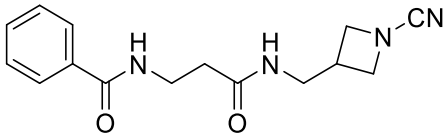
Step 1: The synthesis was preformed according to General Procedure 1 with tert-butyl 3-aminopropanoate (0.98g, 5.4mmol) and benzyl (1.0 eq.) and desired compounds were obtained (1.3g, 96%). LCMS ESI (*m/z*): 193.99 (m−*t*-butyl) ; $[M+H]^+$ calcd for $C_{14}H_{20}NO_3^+$: 250.14

Step 2: The synthesis was performed according to the General Procedure 1 with t-Butyl ester intermediates synthesized in Step 1. Free acid product was obtained by flash column chromatography (EtOAc in hexanes 0 to 100%): 0.70g (70%). LCMS ESI (*m/z*): 193.99 (m−*t*-butyl); $[M+H]^+$ calcd for $C_{10}H_{12}NO_3^+$: 194.08

Step 3: The synthesis was preformed according to General Procedure 1 with 3-benzamidopropanoic acid (0.07g, 0.36mmol) and tert-butyl 3-(aminomethyl)azetidine-1-carboxylate (0.08g, 0.44mmol), Et3N (0.15mL, 1.1mmol) and HATU (0.21g, 0.54mmol) 0.14g desired compound tert-butyl 3-((3-benzamidopropanamido)methyl)azetidine-1-carboxylate was obtained (quant.). LCMS ESI (*m/z*): 361.97; $[M+H]^+$ calcd for $C_{19}H_{28}N_3O^+$: 362.21

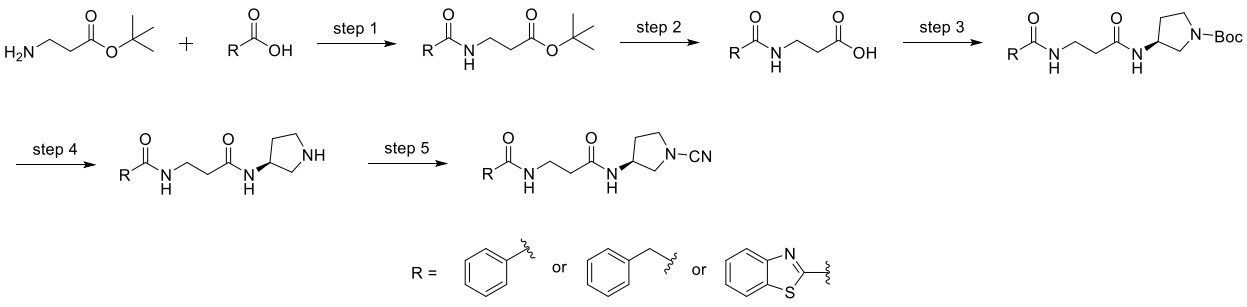
Step 4: The synthesis was performed according to the General Procedure 1 with tert-butyl 3-((3-benzamidopropanamido)methyl)azetidine-1-carboxylate (0.14g, 0.38mmol). 0.1g N-(3-((azetidin-3-ylmethyl)amino)-3-oxopropyl)benzamide (quant.). LCMS ESI (*m/z*): 261.97; $[M+H]^+$ calcd for $C_{14}H_{20}N_3O_2^+$: 262.16

Step 5: The synthesis was performed according to the General Procedure 1 N-(3-((azetidin-3-ylmethyl)amino)-3-oxopropyl)benzamide (0.23g, 0.6mmol) and cyanogen bromide (0.06g, 0.6mmol).



WH-9943-186 N-(3-(((1-cyanoazetidin-3-yl)methyl)amino)-3-oxopropyl)benzamide (57mg, 33%). 1H NMR (500 MHz, DMSO-d6) δ 8.44 (t, J = 5.6 Hz, 1H), 8.01 (t, J = 5.8 Hz, 1H), 7.81 – 7.70 (m, 2H), 7.48 – 7.41 (m, 1H), 7.41 – 7.33 (m, 2H), 4.04 (t, J = 7.9 Hz, 2H), 3.74 (dd, J = 7.6, 5.8 Hz, 2H), 3.39 (q, J = 7.1, 5.6 Hz, 2H), 3.19 (t, J = 6.3 Hz, 2H), 2.77 – 2.61 (m, 1H), 2.32 (t, J = 7.2 Hz, 2H). LCMS ESI (*m/z*): 286.97; $[M+H]^+$ calcd for $C_{15}H_{19}N_4O_2^+$: 287.15

<u>Synthesis of WH-9943-188, 189 and WH-10417-046A:</u>



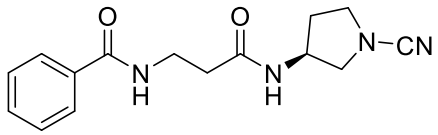
Step 1: The synthesis was preformed according to General Procedure 1 with tert-butyl 3-aminopropanoate (0.98g, 5.4mmol) and acids (1.0 eq.) and desired compounds were obtained (when R=phenyl, 1.3g (96%), LCMS ESI (*m/z*): 193.99 (m−*t*-butyl); $[M+H]^+$ calcd for $C_{14}H_{20}NO_3^+$: 250.14, when R=benzyl, 0.64g (45%), LCMS ESI (*m/z*): 208.08 (m−*t*-butyl); $[M+H]^+$ calcd for $C_{15}H_{22}NO_3^+$: 264.16, when R=benzylthiazol, 0.25g (42%), LCMS ESI (*m/z*): 250.88 (m-t-butyl); $[M+H]^+$ calcd for $C_{15}H_{19}N_2O_3S^+$: 307.11)

Step 2: The synthesis was performed according to the General Procedure 1 with t-Butyl ester intermediates synthesized in Step 1. Free acid products were obtained by flash column chromatography (EtOAc in hexanes 0 to 100%) (when R=phenyl, 0.70g (70%), LCMS ESI (*m/z*): 193.99; $[M+H]^+$ calcd for $C_{10}H_{12}NO_3^+$: 194.08, when R=benzyl, 0.44g (82%), LCMS ESI (*m/z*): 208.08; $[M+H]^+$ calcd for $C_{11}H_{14}NO_3^+$: 208.10, when R=benzylthiazol, 0.2g (quant.), LCMS ESI (*m/z*): 250.98; $[M+H]^+$ calcd for $C_{11}H_{11}N_2O_3S^+$: 251.05)

Step 3: The synthesis was preformed according to General Procedure 1 with tert-butyl (S)-3-aminopyrrolidine-1-carboxylate (0.08g, 0.44mmol) and acids (0.8eq.). Desired compounds were obtained (when R=phenyl, 0.14g (quant.), LCMS ESI (*m/z*): 261.87 (m−*Boc*); $[M+H]^+$ calcd for $C_{19}H_{28}N_3O_4^+$: 362.21, when R=benzyl, 0.15g (quant.), LCMS ESI (*m/z*): 275.97 (m−*Boc*); $[M+H]^+$ calcd for $C_{20}H_{30}N_3O_4^+$: 376.22, when R=benzylthiazol, 0.09g (quant.) LCMS ESI (*m/z*): 319.17 (m−*Boc*); $[M+H]^+$ calcd for $C_{20}H_{27}N_4O_4S^+$: 419.17).

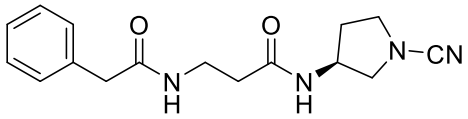
Step 4: The synthesis was performed according to the General Procedure 1 with Boc protected intermediate (0.4mmol). The amines (TFA salt) were obtained (quant.)

Step 5: The synthesis was performed according to the General Procedure 1 with amines (1 eq.) and cyanogen bromide (1 eq.), Et_3N (10.0 eq.) in DMSO.

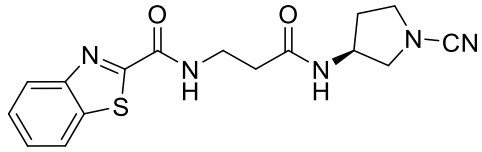


WH-9943-188 (S)-N-(3-((1-cyanopyrrolidin-3-yl)amino)-3-oxopropyl)benzamide (31mg, 40%). 1H NMR (500 MHz, DMSO-d6) δ 8.52 (t, J = 5.6 Hz, 1H), 8.24 (d, J = 6.7 Hz, 1H), 7.91 – 7.76 (m, 2H), 7.55 – 7.49 (m, 1H), 7.49 – 7.42 (m, 2H), 4.35 – 4.19 (m, 1H), 3.53 (dd, J = 9.7, 6.1 Hz, 1H), 3.51 –

3.36 (m, 4H), 3.20 – 3.12 (m, 1H), 2.38 (t, J = 7.1 Hz, 2H), 2.11 – 1.95 (m, 1H), 1.84 – 1.68 (m, 1H). LCMS ESI (*m/z*): 286.87; $[M+H]^+$ calcd for $C_{15}H_{19}N_4O_2^+$: 287.14

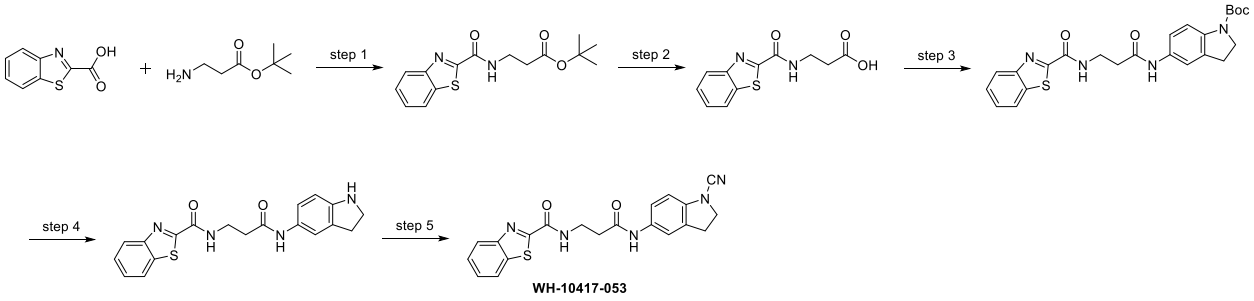


WH-9943-189 (S)-N-(1-cyanopyrrolidin-3-yl)-3-(2-phenylacetamido)propenamide (19mg, 25%). 1H NMR (500 MHz, DMSO-d6) δ 8.18 (d, J = 6.6 Hz, 1H), 8.09 (t, J = 5.8 Hz, 1H), 7.44 – 7.10 (m, 5H), 4.30 – 4.18 (m, 1H), 3.52 (dd, J = 9.7, 6.1 Hz, 1H), 3.48 – 3.36 (m, 4H), 3.25 (q, J = 6.7 Hz, 2H), 3.11 (dd, J = 9.7, 3.9 Hz, 1H), 2.25 (t, J = 6.9 Hz, 2H), 2.14 – 1.90 (m, 1H), 1.82 – 1.63 (m, 1H). LCMS ESI (*m/z*): 300.97; $[M+H]^+$ calcd for $C_{16}H_{21}N_4O_2^+$: 301.17



WH-10417-046A (S)-N-(3-((1-cyanopyrrolidin-3-yl)amino)-3-oxopropyl)benzo[d]thiazole-2-carboxamide (43mg, 48%). 1H NMR (500 MHz, DMSO-d6) δ 9.11 (t, J = 6.0 Hz, 1H), 8.27 (d, J = 6.7 Hz, 1H), 8.23 (d, J = 7.9 Hz, 1H), 8.14 (d, J = 8.0 Hz, 1H), 7.64 (t, 1H), 7.59 (t, 1H), 4.33 – 4.19 (m, 1H), 3.59 – 3.49 (m, 3H), 3.49 – 3.36 (m, 2H), 3.15 (dd, J = 9.7, 3.9 Hz, 1H), 2.44 (t, J = 7.2 Hz, 2H), 2.10 – 1.92 (m, 1H), 1.86 – 1.67 (m, 1H). LCMS ESI (*m/z*): 344.17; $[M+H]^+$ calcd for $C_{16}H_{18}N_5O_2S^+$: 344.12

<u>Synthesis of WH-10417-053:</u>



Step 1: The synthesis was preformed according to General Procedure 1 with tert-butyl 3-aminopropanoate (0.46g, 2.5mmol) and benzo[d]thiazole-2-carboxylic acid (0.35g, 1.9mmol.)

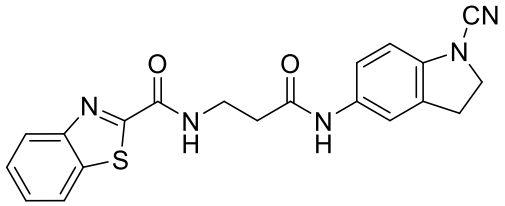
and desired compounds were obtained (0.25g, 42%). LCMS ESI (*m/z*): 250.88 (m−*t*-butyl); $[M+H]^+$ calcd for $C_{15}H_{19}N_2O_3S^+$: 307.11

Step 2: The synthesis was performed according to the General Procedure 1 with t-Butyl ester intermediates synthesized in Step 1. Free acid products were obtained by flash column chromatography (EtOAc in hexanes 0 to 100%) (0.2g, quant.) LCMS ESI (*m/z*): 250.98; $[M+H]^+$ calcd for $C_{11}H_{11}N_2O_3S^+$: 251.05

Step 3: The synthesis was preformed according to General Procedure 1 with 3-(benzo[d]thiazole-2-carboxamido)propanoic acid (0.05g, 0.2mmol) and tert-butyl 5-aminoindoline-1-carboxylate (0.05g, 0.2mmol). 0.06g desired compound tert-butyl 5-(3-(benzo[d]thiazole-2-carboxamido)propanamido)indoline-1-carboxylate was obtained (61%). LCMS ESI (*m/z*): 467.28; $[M+H]^+$ calcd for $C_{24}H_{27}N_4O_4S^+$: 467.17

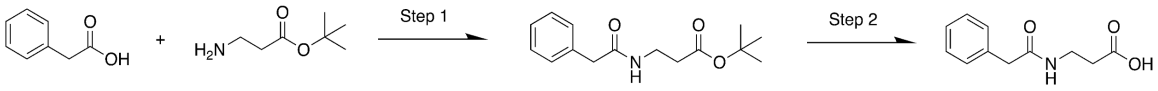
Step 4: The synthesis was performed according to the General Procedure 1 with tert-butyl 5-(3-(benzo[d]thiazole-2-carboxamido)propanamido)indoline-1-carboxylate (0.06g, 0.13mmol) in 2mL 4N HCl in dioxane. 0.04g N-(3-(indolin-5-ylamino)-3-oxopropyl)benzo[d]thiazole-2-carboxamide (85%). LCMS ESI (*m/z*): 367.17; $[M+H]^+$ calcd for $C_{19}H_{19}N_4O_2S^+$: 367.12

Step 5: The synthesis was performed according to the General Procedure 3 N-(3-(indolin-5-ylamino)-3-oxopropyl)benzo[d]thiazole-2-carboxamide (0.04g, 0.1mmol) and cyanogen bromide (0.023g, 0.2mmol).



WH-10417-053 N-(3-((1-cyanoindolin-5-yl)amino)-3-oxopropyl)benzo[d]thiazole-2-carboxamide (15mg, 35%). 1H NMR (500 MHz, DMSO-d6) δ 9.73 (s, 1H), 8.24 (t, J = 5.7 Hz, 1H), 7.82 (d, J = 8.6 Hz, 1H), 7.47 (s, 1H), 7.38 – 7.24 (m, 3H), 7.22 (d, J = 8.7, 2.1 Hz, 1H), 3.87 (t, J = 8.8 Hz, 2H), 3.54 – 3.45 (m, 2H), 2.99 (t, J = 8.8 Hz, 2H), 2.51 (t, J = 6.8 Hz, 2H). LCMS ESI (*m/z*): 392.17; $[M+H]^+$ calcd for $C_{20}H_{18}N_5O_2S^+$: 392.12

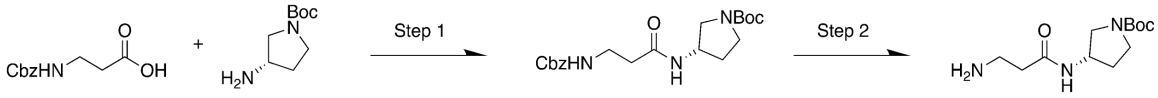
Synthesis of 3-(2-phenylacetamido)propanoic acid



Step 1: The synthesis was preformed according to General Procedure 1 with *tert*-butyl 3-aminopropanoate (0.75 g, 4.13mmol) and 2-phenylacetic acid (1.12 g, 8.26mmol). 1.02 g desired compound (*tert*-butyl 3-(2-phenylacetamido)propanoate was obtained (94%).

Step 2: The synthesis was performed according to the General Procedure 1 with (*tert*-butyl 3-(2-phenylacetamido)propanoate (1.02 g, 3.87mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.802 g 3-(2-phenylacetamido)propanoic acid was obtained (quant.)

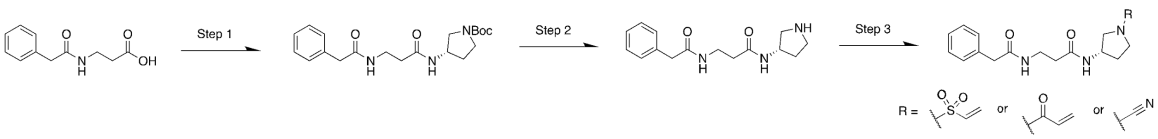
Synthesis of *tert*-butyl (*S*)-3-(3-aminopropanamido)pyrrolidine-1-carboxylate



Step 1: The synthesis was preformed according to General Procedure 1 with *tert*-butyl (*S*)-3-aminopyrrolidine-1-carboxylate (0.625g, 3.36mmol) and 3-(((benzyloxy)carbonyl)amino)propanoic acid, (0.75g, 3.36mmol). 1.3g desired compound *tert*-butyl (*S*)-3-(3-(((benzyloxy)carbonyl)amino)propanamido)pyrrolidine-1-carboxylate was obtained (99%).

Step 2: The synthesis was performed according to the General Procedure 1 with *tert*-butyl (*S*)-3-(3-(((benzyloxy)carbonyl)amino)propanamido)pyrrolidine-1-carboxylate (1.3g, 0.357mmol) except for using activated palladium on carbon and methanol instead of TFA/DCM. 0.918g *tert*-butyl (*S*)-3-(3-aminopropanamido)pyrrolidine-1-carboxylate was obtained (quant.)

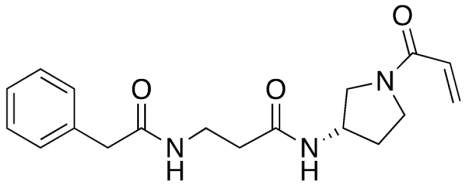
Synthesis of AF_11010_64, 112:



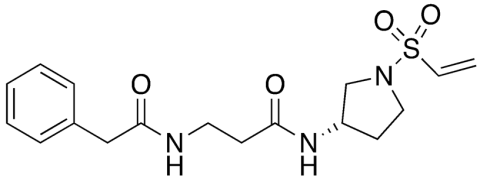
Step 1: The synthesis was performed according to the General Procedure 1 with 3-(2-phenylacetamido)propanoic acid (0.075g, 0.36mmol) and *tert*-butyl (*S*)-3-aminopyrrolidine-1-carboxylate (0.066mL, 0.36 mmol). 0.096g desired compound *tert*-butyl (*S*)-3-(3-(2-phenylacetamido)propanamido)pyrrolidine-1-carboxylate was obtained (71%).

Step 2: The synthesis was performed according to the General Procedure 1 with *tert*-butyl (*S*)-3-(3-(2-phenylacetamido)propanamido)pyrrolidine-1-carboxylate (0.096 g, 0.255 mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.070 g (*S*)-3-(2-phenylacetamido)-*N*-(pyrrolidin-3-yl)propanamide was obtained (quant.)

Step 3: The synthesis was performed according to the General Procedure 1 with *S*)-3-(2-phenylacetamido)-*N*-(pyrrolidin-3-yl)propanamide (0.05 g, 0.18 mmol) and ethenesulfonyl chloride (0.046mL, 0.44mmol), or acryloyl chloride (0.036mL, 0.44mmol).



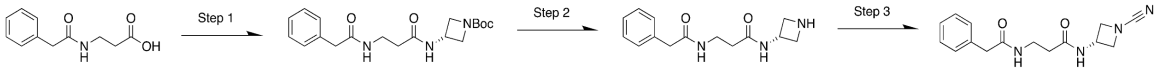
AF_11010_64 ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.33 – 7.17 (m, 5H), 6.56 (ddd, *J* = 39.0, 16.8, 10.4 Hz, 1H), 6.27 (ddd, *J* = 16.8, 6.2, 2.0 Hz, 1H), 5.74 (ddd, *J* = 10.4, 7.5, 1.9 Hz, 1H), 4.35 (dp, *J* = 15.9, 5.6 Hz, 1H), 3.73 – 3.65 (m, 2H), 3.62 – 3.52 (m, 1H), 3.49 – 3.46 (m, 2H), 3.43 (td, *J* = 6.7, 1.8 Hz, 2H), 3.41 – 3.34 (m, 1H), 2.38 (td, *J* = 6.7, 1.0 Hz, 2H), 2.26 – 2.06 (m, 1H), 1.95 – 1.79 (m, 1H). LC/MS (ESI) *m/z* 329.77; $[M+H]^+$ calcd for $C_{18}H_{23}N_3O_3^+$: 330.17



AF_11010_112 ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.32 – 7.21 (m, 5H), 6.69 (dd, *J* = 16.6, 10.0 Hz, 1H), 6.19 (d, *J* = 16.6 Hz, 1H), 6.09 (d, *J* = 10.1 Hz, 1H), 4.29 – 4.24 (m, 1H), 3.48 (d, *J* = 3.3 Hz,

2H), 3.47 – 3.36 (m, 4H), 3.09 (ddd, *J* = 10.5, 4.7, 0.8 Hz, 1H), 2.38 (t, *J* = 6.7 Hz, 2H), 2.13 (ddt, *J* = 13.1, 8.1, 6.6 Hz, 1H), 1.86 – 1.77 (m, 1H). LC/MS (ESI) *m/z* 365.97; $[M+H]^+$ calcd for $C_{17}H_{23}N_3O_4S^+$: 366.14

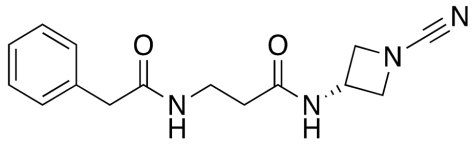
Synthesis of AF_11010_66:



Step 1: The synthesis was performed according to the General Procedure 1 with 3-(2-phenylacetamido)propanoic acid (0.075g, 0.36mmol) and *tert*-butyl 3-aminoazetidine-1-carboxylate (0.057mL, 0.36mmol). 0.129g desired compound *tert*-butyl 3-(3-(2-phenylacetamido)propanamido)azetidine-1-carboxylate was obtained (99%).

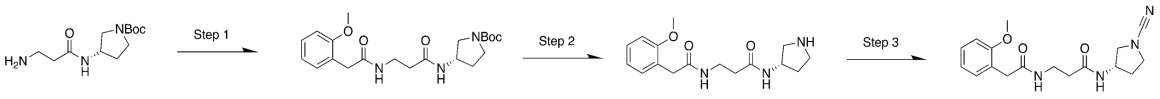
Step 2: The synthesis was performed according to the General Procedure 1 with *tert*-butyl 3-(3-(2-phenylacetamido)propanamido)azetidine-1-carboxylate (0.129g, 0.357mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.093g *N*-(azetidin-3-yl)-3-(2-phenylacetamido)propanamide was obtained (quant.)

Step 3: The synthesis was performed according to the General Procedure 1 with *N*-(azetidin-3-yl)-3-(2-phenylacetamido)propanamide (0.093g, 0.357mmol) and cyanogen bromide (0.424mL, 1.27mmol).



AF_11010_66 ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.35 – 7.24 (m, 5H), 4.59 (tt, *J* = 7.8, 5.9 Hz, 1H), 4.37 (t, *J* = 7.9 Hz, 2H), 4.03 – 3.98 (m, 2H), 3.50 (s, 2H), 3.45 (t, *J* = 6.6 Hz, 2H), 2.41 (t, *J* = 6.6 Hz, 2H). LC/MS (ESI) *m/z* 286.97; $[M+H]^+$ calcd for $C_{15}H_{18}N_4O_2^+$: 287.14

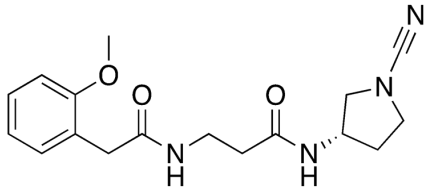
Synthesis of AF_11010_82:



Step 1: The synthesis was performed according to the General Procedure 2-(2-methoxyphenyl)acetic acid (0.0485g, 0.291mmol) and *tert*-butyl (*S*)-3-(3-aminopropanamido)pyrrolidine-1-carboxylate (0.05g, 0.194mmol). 0.079g desired compound *tert*-butyl (*S*)-3-(3-(2-(2-methoxyphenyl)acetamido)propanamido)pyrrolidine-1-carboxylate was obtained (99%).

Step 2: The synthesis was performed according to the General Procedure 1 with *tert*-butyl (*S*)-3-(3-(2-(2-methoxyphenyl)acetamido)propanamido)pyrrolidine-1-carboxylate (0.079g, 0.194mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.059g (*S*)-3-(2-(2-methoxyphenyl)acetamido)-*N*-(pyrrolidin-3-yl)propanamide was obtained (quant.)

Step 3: The synthesis was performed according to the General Procedure 1 (*S*)-3-(2-(2-methoxyphenyl)acetamido)-*N*-(pyrrolidin-3-yl)propanamide (0.059g, 0.194mmol) and cyanogen bromide (0.129mL, 0.388mmol).



AF_11010_82 ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.26 (td, *J* = 7.8, 1.7 Hz, 1H), 7.19 (dd, *J* = 7.4, 1.7 Hz, 1H), 6.96 (dd, *J* = 8.2, 1.1 Hz, 1H), 6.91 (td, *J* = 7.5, 1.1 Hz, 1H), 4.30 (ddt, *J* = 8.6, 6.1, 3.0 Hz, 1H), 3.83 (s, 3H), 3.58 (dd, *J* = 9.9, 6.0 Hz, 1H), 3.54 – 3.50 (m, 1H), 3.49 (s, 2H), 3.48 – 3.44 (m, 1H), 3.42 (t, *J* = 6.7 Hz, 2H), 3.19 (ddd, *J* = 9.9, 3.9, 0.9 Hz, 1H), 2.38 (t, *J* = 6.6 Hz, 2H), 2.13 (dtd, *J* = 13.1, 8.0, 6.2 Hz, 1H), 1.89 – 1.80 (m, 1H). LC/MS (ESI) *m/z* 330.97; $[M+H]^+$ calcd for $C_{17}H_{22}N_4O_3^+$: 331.17

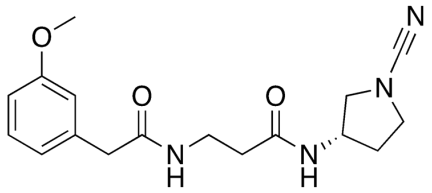
Synthesis of AF_11010_83:



Step 1: The synthesis was performed according to the General Procedure 2-(3-methoxyphenyl)acetic acid (0.0485g, 0.291mmol) and *tert*-butyl (*S*)-3-(3-aminopropanamido)pyrrolidine-1-carboxylate (0.05g, 0.194mmol). 0.079g desired compound *tert*-butyl (*S*)-3-(3-(2-(3-methoxyphenyl)acetamido)propanamido)pyrrolidine-1-carboxylate was obtained (99%).

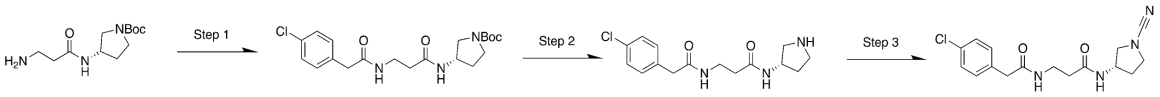
Step 2: The synthesis was performed according to the General Procedure 1 with *tert*-butyl (*S*)-3-(3-(2-(3-methoxyphenyl)acetamido)propanamido)pyrrolidine-1-carboxylate (0.079g, 0.194mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.059g (*S*)-3-(2-(3-methoxyphenyl)acetamido)-*N*-(pyrrolidin-3-yl)propanamide was obtained (quant.)

Step 3: The synthesis was performed according to the General Procedure 1 (*S*)-3-(2-(3-methoxyphenyl)acetamido)-*N*-(pyrrolidin-3-yl)propanamide (0.059g, 0.194mmol) and cyanogen bromide (0.129mL, 0.388mmol).



AF_11010_83 ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.29 – 7.21 (m, 1H), 7.18 (dt, *J* = 7.4, 2.1 Hz, 1H), 6.96 (dd, *J* = 8.2, 1.2 Hz, 1H), 6.90 (tt, *J* = 7.5, 1.4 Hz, 1H), 4.30 (dd, *J* = 8.8, 4.3 Hz, 1H), 3.83 (d, *J* = 1.5 Hz, 3H), 3.58 (dd, *J* = 10.0, 6.0 Hz, 1H), 3.56 – 3.50 (m, 1H), 3.49 (d, *J* = 2.2 Hz, 2H), 3.48 – 3.44 (m, 1H), 3.42 (td, *J* = 6.7, 2.6 Hz, 2H), 3.22 – 3.14 (m, 1H), 2.38 (td, *J* = 6.6, 3.2 Hz, 2H), 2.19 – 2.06 (m, 1H), 1.89 – 1.79 (m, 1H). LC/MS (ESI) *m/z* 331.07; $[M+H]^+$ calcd for $C_{17}H_{22}N_4O_3^+$: 331.17

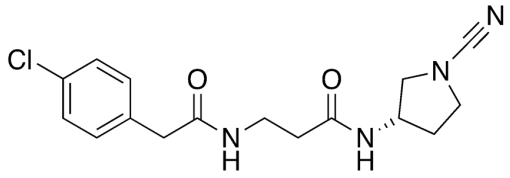
Synthesis of AF_11010_84:



Step 1: The synthesis was performed according to the General Procedure 2-(4-chlorophenyl)acetic acid (0.0485g, 0.291mmol) and *tert*-butyl (*S*)-3-(3-aminopropanamido)pyrrolidine-1-carboxylate (0.05g, 0.194mmol). 0.079g desired compound *tert*-butyl (*S*)-3-(3-(2-(4-chlorophenyl)acetamido)propanamido)pyrrolidine-1-carboxylate was obtained (99%).

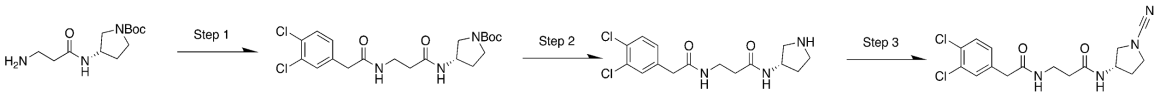
Step 2: The synthesis was performed according to the General Procedure 1 with *tert*-butyl (*S*)-3-(3-(2-(4-chlorophenyl)acetamido)propanamido)pyrrolidine-1-carboxylate (0.079g, 0.194mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.059g (*S*)-3-(2-(4-chlorophenyl)acetamido)-*N*-(pyrrolidin-3-yl)propanamide was obtained (quant.)

Step 3: The synthesis was performed according to the General Procedure 1 ((*S*)-3-(2-(4-chlorophenyl)acetamido)-*N*-(1-cyanopyrrolidin-3-yl)propanamide (0.059g, 0.194mmol) and cyanogen bromide (0.129mL, 0.388mmol).



AF_11010_84 ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.33 – 7.25 (m, 3H), 4.33 – 4.28 (m, 1H), 3.59 (dd, *J* = 9.9, 6.0 Hz, 1H), 3.54 – 3.48 (m, 1H), 3.48 – 3.45 (m, 2H), 3.44 (t, *J* = 6.6 Hz, 2H), 3.22 – 3.17 (m, 1H), 2.38 (t, *J* = 6.6 Hz, 2H), 2.18 – 2.09 (m, 1H), 1.84 (ddt, *J* = 12.6, 7.4, 4.9 Hz, 1H), 1.30 (s, 1H). LC/MS (ESI) *m/z* 334.97; $[M+H]^+$ calcd for $C_{16}H_{19}ClN_4O_2^+$: 335.12

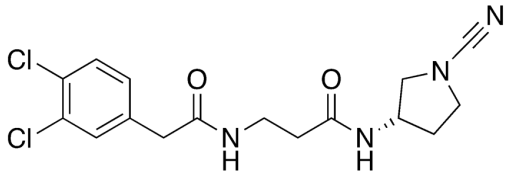
Synthesis of AF_11010_124:



Step 1: The synthesis was performed according to the General Procedure 2-(3,4-dichlorophenyl)acetic acid (0.06g, 0.291mmol) and *tert*-butyl (*S*)-3-(3-aminopropanamido)pyrrolidine-1-carboxylate (0.05g, 1.0mmol). 0.059g desired compound *tert*-butyl (*S*)-3-(3-(2-(3,4-dichlorophenyl)acetamido)propanamido)pyrrolidine-1-carboxylate was obtained (68%).

Step 2: The synthesis was performed according to the General Procedure 1 with *tert*-butyl (*S*)-3-(3-(2-(3,4-dichlorophenyl)acetamido)propanamido)pyrrolidine-1-carboxylate (0.059g, 0.133mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.046g (*S*)-3-(2-(3,4-dichlorophenyl)acetamido)-*N*-(pyrrolidin-3-yl)propanamide was obtained (quant.)

Step 3: The synthesis was performed according to the General Procedure 1 (*S*)-3-(2-(3,4-dichlorophenyl)acetamido)-*N*-(pyrrolidin-3-yl)propanamide (0.045g, 0.133mmol) and cyanogen bromide (0.112mL, 0.33mmol).



AF_11010_124 ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.47 (d, *J* = 2.1 Hz, 1H), 7.46 (d, *J* = 8.2 Hz, 1H), 7.21 (dd, *J* = 8.3, 2.1 Hz, 1H), 4.31 (ddd, *J* = 10.3, 6.0, 4.3 Hz, 1H), 3.59 (dd, *J* = 9.9, 6.0 Hz, 1H), 3.54 – 3.49 (m, 1H), 3.48 (s, 2H), 3.44 (t, *J* = 6.6 Hz, 2H), 3.21 (dd, *J* = 9.9, 3.8 Hz, 1H), 2.39 (t, *J* = 6.6 Hz, 2H), 2.14 (dtd, *J* = 15.9, 7.9, 6.2 Hz, 1H), 1.84 (ddt, *J* = 12.6, 7.5, 4.9 Hz, 1H).LC/MS (ESI) *m/z* 368.97; $[M+H]^+$ calcd for $C_{16}H_{18}Cl_2N_4O_2^+$: 369.08

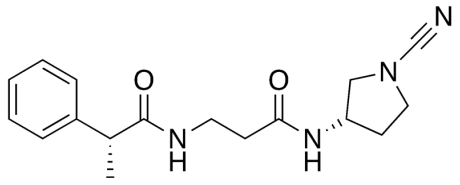
<u>Synthesis of AF_11010_125:</u>



Step 1: The synthesis was performed according to the General Procedure (*R*)-2-phenylpropanoic acid (0.044g, 0.29mmol) and *tert*-butyl (*S*)-3-(3-aminopropanamido)pyrrolidine-1-carboxylate (0.050g, 0.194mmol). 0.052g desired compound *tert*-butyl (*S*)-3-(3-((*R*)-2-phenylpropanamido)propanamido)pyrrolidine-1-carboxylate was obtained (69%).

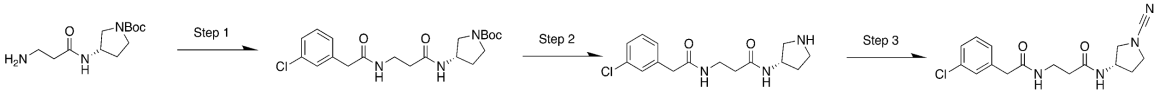
Step 2: The synthesis was performed according to the General Procedure 1 with *tert*-butyl (*S*)-3-(3-((*R*)-2-phenylpropanamido)propanamido)pyrrolidine-1-carboxylate (0.052g, 0.135mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.040g ((*R*)-*N*-(3-oxo-3-(((*S*)-pyrrolidin-3-yl)amino)propyl)-2-phenylpropanamide was obtained (quant.)

Step 3: The synthesis was performed according to the General Procedure 1 (*R*)-*N*-(3-oxo-3-(((*S*)-pyrrolidin-3-yl)amino)propyl)-2-phenylpropanamide (0.04g, 0.135mmol) and cyanogen bromide (0.112mL, 0.34mmol).



AF_11010_125 ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.36 – 7.30 (m, 4H), 7.27 – 7.23 (m, 1H), 4.29 (ddd, *J* = 6.2, 4.5, 1.7 Hz, 1H), 3.63 (q, *J* = 7.1 Hz, 1H), 3.56 (dd, *J* = 9.9, 6.1 Hz, 1H), 3.53 – 3.44 (m, 3H), 3.40 (q, *J* = 6.8 Hz, 1H), 3.09 (ddd, *J* = 9.9, 4.1, 0.8 Hz, 1H), 2.38 (td, *J* = 6.7, 3.9 Hz, 2H), 2.18 – 2.09 (m, 1H), 1.88 – 1.80 (m, 1H), 1.45 (d, *J* = 7.0 Hz, 3H). LC/MS (ESI) *m/z* 314.87; $[M+H]^+$ calcd for $C_{17}H_{22}N_4O_2^+$: 315.17

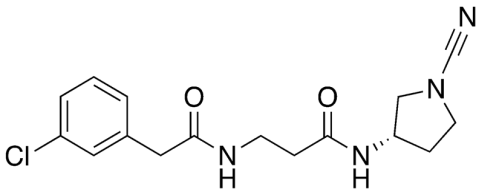
Synthesis of AF_11010_131:



Step 1: The synthesis was performed according to the General Procedure 2-(3-chlorophenyl)acetic acid (0.044g, 0.291mmol) and *tert*-butyl (*S*)-3-(3-aminopropanamido)pyrrolidine-1-carboxylate (0.05g, 0.194mmol). 0.079g desired compound *tert*-butyl (*S*)-3-(3-(2-(3-chlorophenyl)acetamido)propanamido)pyrrolidine-1-carboxylate was obtained (99%).

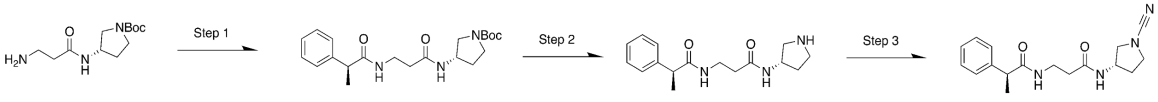
Step 2: The synthesis was performed according to the General Procedure 1 with *tert*-butyl (*S*)-3-(3-(2-(3-chlorophenyl)acetamido)propanamido)pyrrolidine-1-carboxylate (0.079g, 0.194mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.06g (*S*)-3-(2-(3-chlorophenyl)acetamido)-*N*-(pyrrolidin-3-yl)propanamide was obtained (quant.)

Step 3: The synthesis was performed according to the General Procedure 1 ((*S*)-3-(2-(3-chlorophenyl)acetamido)-*N*-(1-cyanopyrrolidin-3-yl)propanamide (0.06g, 0.194mmol) and cyanogen bromide (0.162mL, 0.485mmol).



AF_11010_131 ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.34 – 7.20 (m, 4H), 4.32 (tt, *J* = 6.1, 4.2 Hz, 1H), 3.59 (dd, *J* = 9.9, 6.0 Hz, 1H), 3.54 – 3.50 (m, 1H), 3.49 (s, 2H), 3.48 – 3.46 (m, 1H), 3.45 (t, *J* = 6.7 Hz, 2H), 3.20 (dd, *J* = 9.9, 3.9 Hz, 1H), 2.40 (t, *J* = 6.7 Hz, 2H), 2.19 – 2.09 (m, 1H), 1.90 – 1.81 (m, 1H). LC/MS (ESI) *m/z* 334.97; $[M+H]^+$ calcd for $C_{16}H_{19}ClN_4O_2^+$: 334.12

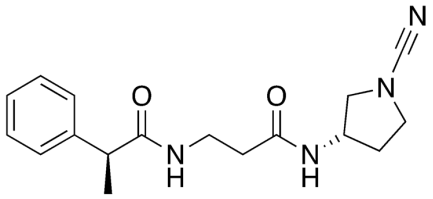
Synthesis of AF_11010_132:



Step 1: The synthesis was performed according to the General Procedure (*S*)-2-phenylpropanoic acid (0.044g, 0.291mmol) and *tert*-butyl (*S*)-3-(3-aminopropanamido)pyrrolidine-1-carboxylate (0.05g, 0.194mmol). 0.075g desired compound *tert*-butyl (*S*)-3-(3-((*S*)-2-phenylpropanamido)propanamido)pyrrolidine-1-carboxylate was obtained (99%).

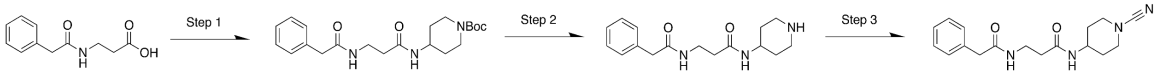
Step 2: The synthesis was performed according to the General Procedure 1 with *tert*-butyl (*S*)-3-(3-((*S*)-2-phenylpropanamido)propanamido)pyrrolidine-1-carboxylate (0.075g, 0.194mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.056g ((*S*)-*N*-(3-oxo-3-(((*S*)-pyrrolidin-3-yl)amino)propyl)-2-phenylpropanamide was obtained (quant.)

Step 3: The synthesis was performed according to the General Procedure 1 (*R*)-*N*-(3-oxo-3-(((*S*)-pyrrolidin-3-yl)amino)propyl)-2-phenylpropanamide (0.056g, 0.194mmol) and cyanogen bromide (0.162mL, 0.485mmol).



AF_11010_132 ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.35 – 7.26 (m, 4H), 7.22 (dddd, *J* = 6.9, 5.7, 3.3, 1.4 Hz, 1H), 4.26 (ddd, *J* = 10.4, 6.1, 4.3 Hz, 1H), 3.63 – 3.53 (m, 2H), 3.49 – 3.34 (m, 4H), 3.17 (ddd, *J* = 9.9, 3.8, 0.9 Hz, 1H), 2.39 – 2.33 (m, 2H), 2.14 – 2.02 (m, 1H), 1.83 – 1.73 (m, 1H), 1.42 (dd, *J* = 7.2, 1.5 Hz, 3H). LC/MS (ESI) *m/z* 314.97; $[M+H]^+$ calcd for $C_{17}H_{22}N_4O_2^+$: 315.17

Synthesis of AF_11010_104:

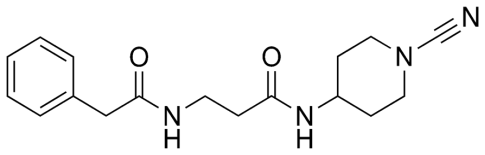


Step 1: The synthesis was performed according to the General Procedure 1 with 3-(2-phenylacetamido)propanoic acid (0.075g, 0.362mmol) and *tert*-butyl 4-aminopiperidine-1-carboxylate (0.06g, 0.3mmol). 0.116g desired compound *tert*-butyl 4-(3-(2-phenylacetamido)propanamido)piperidine-1-carboxylate was obtained (99%).

Step 2: The synthesis was performed according to the General Procedure 1 with *tert*-butyl 4-(3-(2-phenylacetamido)propanamido)piperidine-1-carboxylate (0.116g, 0.3mmol) except for using

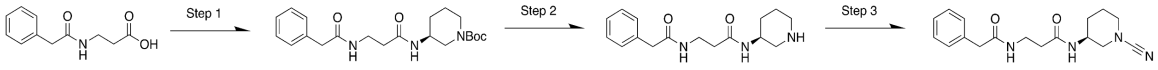
4N HCl in 1,4-dioxane instead of TFA/DCM. 0.087g 3-(2-phenylacetamido)-*N*-(piperidin-4-yl)propanamide was obtained (quant.)

Step 3: The synthesis was performed according to the General Procedure 1 with 3-(2-phenylacetamido)-*N*-(piperidin-4-yl)propanamide (0.087g, 0.3mmol) and cyanogen bromide (0.251mL, 0.75mmol).



AF_11010_104 ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.33 – 7.22 (m, 5H), 3.74 (tt, *J* = 10.8, 4.0 Hz, 1H), 3.48 (s, 2H), 3.43 (t, *J* = 6.7 Hz, 2H), 3.39 (td, *J* = 5.3, 4.4, 3.1 Hz, 2H), 3.13 (ddd, *J* = 13.2, 11.6, 2.8 Hz, 2H), 2.36 (t, *J* = 6.7 Hz, 2H), 1.86 – 1.79 (m, 2H), 1.55 – 1.44 (m, 2H). LC/MS (ESI) *m/z* 314.77; $[M+H]^+$ calcd for $C_{17}H_{22}N_4O_2^+$: 315.17

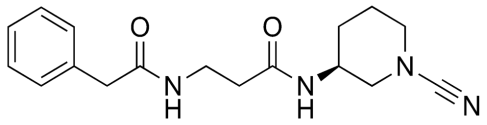
Synthesis of AF_11010_105:



Step 1: The synthesis was performed according to the General Procedure 1 with 3-(2-phenylacetamido)propanoic acid (0.15g, 0.724mmol) and *tert*-butyl (*S*)-3-aminopiperidine-1-carboxylate (0.125g, 0.626mmol). 0.243g desired compound *tert*-butyl (*S*)-3-(3-(2-phenylacetamido)propanamido)piperidine-1-carboxylate was obtained (99%).

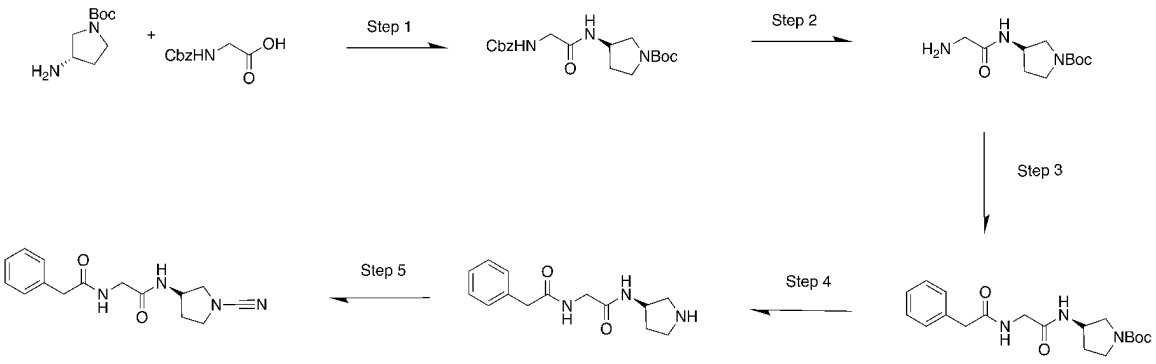
Step 2: The synthesis was performed according to the General Procedure 1 with *tert*-butyl (*S*)-3-(3-(2-phenylacetamido)propanamido)piperidine-1-carboxylate (0.233, 0.6mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.173g (*S*)-3-(2-phenylacetamido)-*N*-(piperidin-3-yl)propanamide was obtained (quant.)

Step 3: The synthesis was performed according to the General Procedure 1 with (*S*)-3-(2-phenylacetamido)-*N*-(piperidin-3-yl)propanamide (0.087g, 0.3mmol) and cyanogen bromide (0.25mL, 0.75mmol).



AF_11010_105 ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.35 – 7.21 (m, 5H), 3.83 (ddd, *J* = 9.0, 5.5, 4.0 Hz, 1H), 3.48 (s, 2H), 3.43 (td, *J* = 6.6, 3.2 Hz, 2H), 3.35 (dd, *J* = 12.4, 4.2 Hz, 1H), 3.30 – 3.24 (m, 1H), 3.09 – 3.03 (m, 1H), 2.79 (dd, *J* = 12.4, 8.6 Hz, 1H), 2.39 (td, *J* = 6.7, 1.1 Hz, 2H), 1.86 – 1.78 (m, 2H), 1.71 – 1.61 (m, 1H), 1.44 – 1.35 (m, 1H). LC/MS (ESI) *m/z* 314.67; $[M+H]^+$ calcd for $C_{17}H_{22}N_4O_2^+$: 315.17

Synthesis of AF_11010_136



Step 1: The synthesis was preformed according to General Procedure 1 with ((benzyloxy)carbonyl)glycine (0.075g, .36mmol) and *tert*-butyl (*S*)-3-aminopyrrolidine-1-carboxylate (0.055mL, 0.3mmol). 0.103g desired compound *tert*-butyl (*S*)-3-(2-(((benzyloxy)carbonyl)amino)acetamido)pyrrolidine-1-carboxylate was obtained (91%).

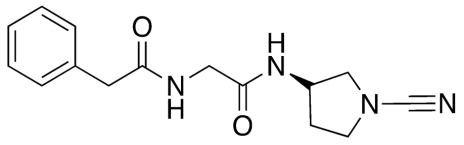
Step 2: The synthesis was performed according to the General Procedure 1 with *tert*-butyl (*S*)-3-(2-(((benzyloxy)carbonyl)amino)acetamido)pyrrolidine-1-carboxylate (0.103g, 0.27mmol) except

for using palladium on carbon and MeOH instead of TFA/DCM. 0.065g *tert*-butyl (*S*)-3-(2-aminoacetamido)pyrrolidine-1-carboxylate was obtained (quant.)

Step 3: The synthesis was performed according to the General Procedure 1 *tert*-butyl (*S*)-3-(2-aminoacetamido)pyrrolidine-1-carboxylate (0.065g, 0.27mmol) and *2-phenyl acetic acid* (0.043g, 0.32mmol). 0.90g desired compound *tert*-butyl (*S*)-3-(2-(2-phenylacetamido)acetamido)pyrrolidine-1-carboxylate was obtained (92%).

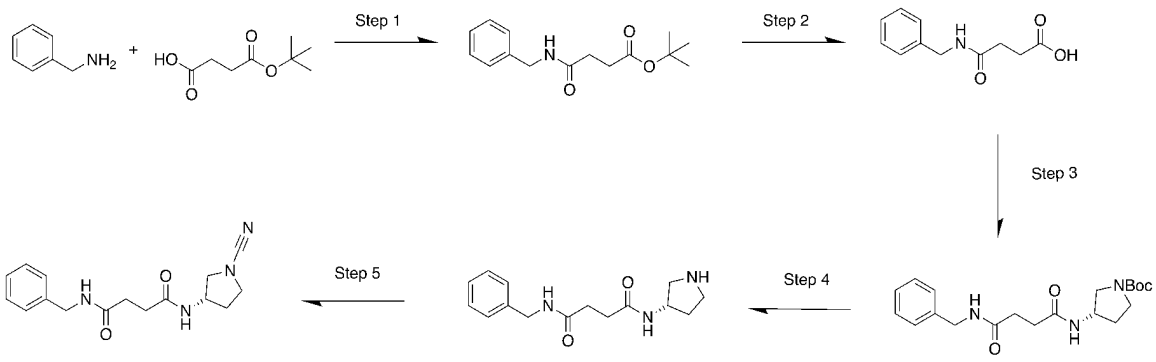
Step 4: The synthesis was performed according to the General Procedure 1 with *S*)-3-(2-(2-phenylacetamido)acetamido)pyrrolidine-1-carboxylate (0.09g, 0.25mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.065g (*S*)-*N*-(2-oxo-2-(pyrrolidin-3-ylamino)ethyl)-2-phenylacetamide was obtained (quant.)

Step 5: The synthesis was performed according to the General Procedure 1 with (*S*)-*N*-(2-oxo-2-(pyrrolidin-3-ylamino)ethyl)-2-phenylacetamide (0.065g, 0.25mmol) and cyanogen bromide (0.250mL, 0.75mmol).



AF_11010_136 ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.35 – 7.32 (m, 3H), 7.29 – 7.24 (m, 1H), 4.38 (tt, *J* = 6.1, 4.2 Hz, 1H), 3.84 (d, *J* = 2.5 Hz, 2H), 3.63 (dd, *J* = 9.9, 6.0 Hz, 1H), 3.60 (s, 2H), 3.55 – 3.46 (m, 2H), 3.26 (ddd, *J* = 10.1, 3.9, 0.9 Hz, 1H), 2.17 (dtd, *J* = 13.0, 8.0, 6.2 Hz, 1H), 1.94 – 1.87 (m, 1H). LC/MS (ESI) *m/z* 286.97; $[M+H]^+$ calcd for $C_{15}H_{18}N_4O_2^+$: 287.14

Synthesis of AF_11010_137



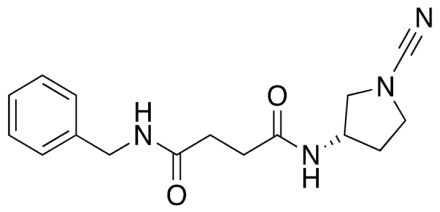
Step 1: The synthesis was preformed according to General Procedure 1 with phenylmethanamine (0.031mL, 0.287mmol) and 4-(*tert*-butoxy)-4-oxobutanoic acid (0.075g, 0.43mmol). 0.105g desired compound *tert*-butyl 4-(benzylamino)-4-oxobutanoate was obtained (93%).

Step 2: The synthesis was performed according to the General Procedure 1 with (*tert*-butyl 4-(benzylamino)-4-oxobutanoate (0.105g, 0.4mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.83g 4-(benzylamino)-4-oxobutanoic acid was obtained (quant.)

Step 3: The synthesis was performed according to the General Procedure 1 4-(benzylamino)-4-oxobutanoic acid (0.083g, 0.4mmol) and *tert*-butyl (*S*)-3-(3-aminopropanamido)pyrrolidine-1-carboxylate (0.088mL, 0.48mmol). 0.144g desired compound *tert*-butyl (*S*)-3-(4-(benzylamino)-4-oxobutanamido)pyrrolidine-1-carboxylate was obtained (99%).

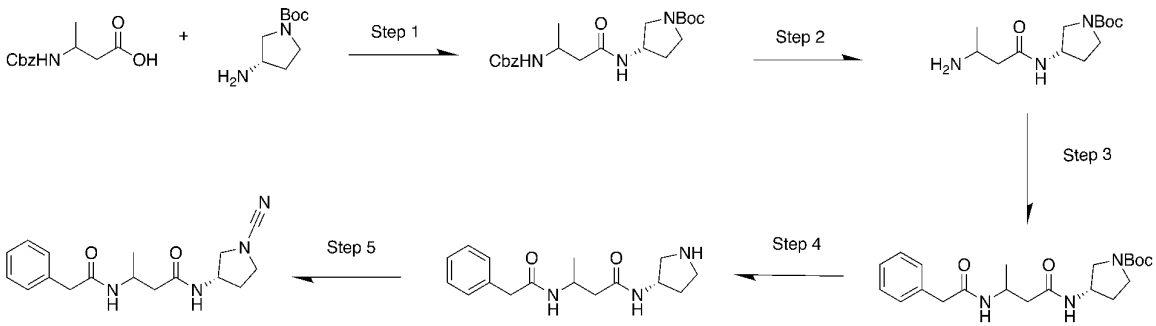
Step 4: The synthesis was performed according to the General Procedure 1 with *tert*-butyl (*S*)-3-(4-(benzylamino)-4-oxobutanamido)pyrrolidine-1-carboxylate (0.144g, 0.4mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.11g (*S*)-N^1-benzyl-N^4-(pyrrolidin-3-yl)succinamide was obtained (quant.)

Step 5: The synthesis was performed according to the General Procedure 1 with *S*)-N^1-benzyl-N^4-(pyrrolidin-3-yl)succinamide (0.11g, 0.4mmol) and cyanogen bromide (0.33mL, 1.0mmol).



AF_11010_137 ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.35 – 7.23 (m, 5H), 4.37 (s, 2H), 4.34 (dq, *J* = 6.0, 4.2, 3.1 Hz, 1H), 3.61 (dd, *J* = 9.9, 6.0 Hz, 1H), 3.58 – 3.46 (m, 2H), 3.24 (dd, *J* = 9.9, 3.9 Hz, 1H), 2.58 – 2.49 (m, 4H), 2.15 (dtd, *J* = 13.1, 8.0, 6.2 Hz, 1H), 1.90 (ddt, *J* = 12.6, 7.3, 4.8 Hz, 1H). LC/MS (ESI) *m/z* 300.97; $[M+H]^+$ calcd for $C_{16}H_{20}N_4O_2^+$: 301.16

Synthesis of AF_11010_139



Step 1: The synthesis was preformed according to General Procedure 1 with 3-(((benzyloxy)carbonyl)amino)butanoic acid (0.072g, 0.3mmol) and *tert*-butyl (*S*)-3-aminopyrrolidine-1-carboxylate (0.046mL, 0.25mmol). 0.094g desired compound *tert*-butyl (3*S*)-3-(3-(((benzyloxy)carbonyl)amino)butanamido)pyrrolidine-1-carboxylate was obtained (93%).

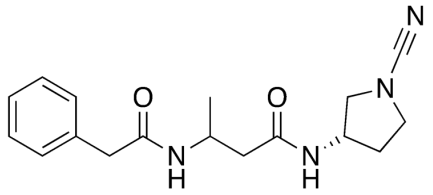
Step 2: The synthesis was performed according to the General Procedure 1 with *tert*-butyl (3*S*)-3-(3-(((benzyloxy)carbonyl)amino)butanamido)pyrrolidine-1-carboxylate (0.094g, 0.233mmol) except for using palladium on carbon and MeOH instead of TFA/DCM. 0.63g *tert*-butyl (3*S*)-3-(3-aminobutanamido)pyrrolidine-1-carboxylate was obtained (quant.)

Step 3: The synthesis was performed according to the General Procedure 1 3-(2-phenylacetamido)butanoic acid (0.038g, 0.28mmol) and *tert*-butyl (3*S*)-3-(3-

aminobutanamido)pyrrolidine-1-carboxylate (0.063g, 0.233mmol). 0.09 desired compound *tert*-butyl (3*S*)-3-(3-(2-phenylacetamido)butanamido)pyrrolidine-1-carboxylate was obtained (98%).

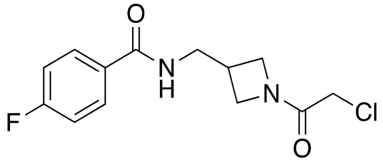
Step 4: The synthesis was performed according to the General Procedure 1 with *tert*-butyl (3*S*)-3-(3-(2-phenylacetamido)butanamido)pyrrolidine-1-carboxylate (0.09g, 0.23mmol) except for using 4N HCl in 1,4-dioxane instead of TFA/DCM. 0.067g 3-(2-phenylacetamido)-*N*-((*S*)-pyrrolidin-3-yl)butanamide was obtained (quant.)

Step 5: The synthesis was performed according to the General Procedure 1 3-(2-phenylacetamido)-*N*-((*S*)-pyrrolidin-3-yl)butanamide (0.067g, 0.23mmol) and cyanogen bromide (0.192mL, 0.575mmol).

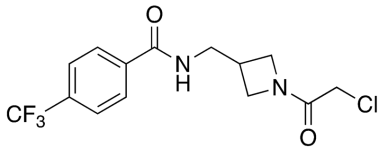


AF_11010_139 ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.34 – 7.23 (m, 5H), 4.27 (qd, *J* = 6.4, 2.5 Hz, 2H), 3.57 (dd, *J* = 9.9, 6.0 Hz, 1H), 3.55 – 3.49 (m, 1H), 3.48 (d, *J* = 2.4 Hz, 2H), 3.45 (tdd, *J* = 9.3, 4.9, 3.3 Hz, 1H), 3.19 – 3.11 (m, 1H), 1.20 (d, *J* = 6.7 Hz, 3H). LC/MS (ESI) *m/z* 314.97; $[M+H]^+$ calcd for $C_{17}H_{22}N_4O_2^+$: 315.17

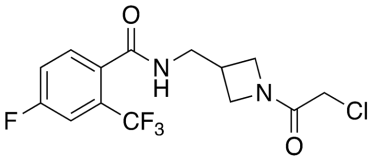
<u>Synthesis of CAS 11487 188, 193, 195, 199, 200, 205:</u>



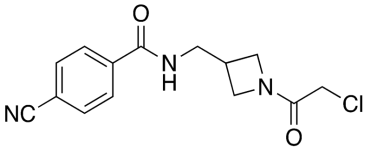
Synthesis of **CAS-11487-188** followed the General Procedure 1 by using 4-fluorobenzoic acid (0.075 g, 0.535 mmol), and *tert*-butyl 3-(aminomethyl)azetidine-1-carboxylate (0.098 mL, 0.535 mmol) to afford CAS-11487-188 (0.0353g, 51.7%) ^{1}H NMR (500 MHz, DMSO-*d*6) δ 8.67 (t, *J* = 5.7 Hz, 1H), 7.99 – 7.81 (m, 2H), 7.42 – 7.18 (m, 2H), 4.31 – 4.17 (m, 1H), 4.09 (d, *J* = 1.4 Hz, 2H), 4.01 – 3.89 (m, 2H), 3.69 (dd, *J* = 9.9, 5.5 Hz, 1H), 3.51 – 3.37 (m, 2H), 2.91 – 2.71 (m, 1H). LC/MS (ESI) *m/z* 285.67; $[M+H]^+$ calcd for $C_{13}H_{15}ClFN_2O_2^+$ 285.08



Synthesis of **CAS-11487-193** followed the General Procedure 1 by using 4-(trifluoromethyl)benzoic acid (0.200 g, 1.05 mmol), and *tert*-butyl 3-(aminomethyl)azetidine-1-carboxylate (0.192 mL, 1.05 mmol) to afford CAS-11487-193 (0.0128g, 11.7%) ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.94 (d, *J* = 8.1 Hz, 2H), 7.72 (d, *J* = 8.2 Hz, 2H), 7.07 (t, *J* = 6.0 Hz, 1H), 4.42 (t, *J* = 8.7 Hz, 1H), 4.20 – 4.08 (m, 2H), 3.96 – 3.91 (m, 1H), 3.89 (d, *J* = 2.5 Hz, 2H), 3.80 (dt, *J* = 14.0, 6.3 Hz, 1H), 3.68 (dt, *J* = 13.7, 6.5 Hz, 1H), 3.04 (dtdd, *J* = 8.3, 6.6, 4.9, 1.5 Hz, 1H). LC/MS (ESI) *m/z* 335.07; $[M+H]^+$ calcd for $C_{14}H_{15}ClF_3N_2O_2^+$: 335.08

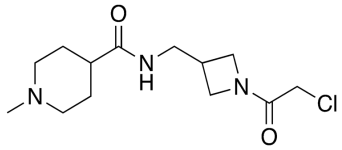


Synthesis of **CAS-11487-195** followed the General Procedure 1 by using 4-fluoro-2-(trifluoromethyl)benzoic acid (0.250 g, 1.20 mmol), and *tert*-butyl 3-(aminomethyl)azetidine-1-carboxylate (0.440 mL, 2.40 mmol) to afford CAS-11487-195 (0.0266 g, 12.6%) ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.57 (dd, *J* = 8.5, 5.2 Hz, 1H), 7.43 (dd, *J* = 8.8, 2.6 Hz, 1H), 7.32 (td, *J* = 8.1, 2.6 Hz, 1H), 6.37 (s, 1H), 4.41 (t, *J* = 8.8 Hz, 1H), 4.17 (t, *J* = 9.5 Hz, 1H), 4.10 (dd, *J* = 9.2, 5.4 Hz, 1H), 3.88 (s, 2H), 3.84 (dd, *J* = 10.6, 5.6 Hz, 1H), 3.79 – 3.60 (m, 2H), 3.08 – 2.98 (m, 1H). LC/MS (ESI) *m/z* 353.06; $[M+H]^+$ calcd for $C_{14}H_{14}ClF_4N_2O_2^+$: 353.07

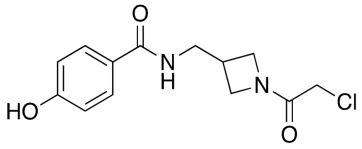


Synthesis of **CAS-11487-199** followed the General Procedure 1 by using 4-cyanobenzoic acid (1.0 g, 6.80 mmol), and *tert*-butyl 3-(aminomethyl)azetidine-1-carboxylate (1.4 mL, 7.48 mmol) to afford CAS-11487-199 (0.011 g, 11.5%) ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.97 – 7.88 (m, 2H), 7.81 – 7.73 (m, 2H), 6.82 (s, 0H), 4.43 (t, *J* = 8.7 Hz, 1H), 4.23 – 4.15 (m, 1H), 4.12 (dd, *J* = 9.2, 5.2 Hz, 1H), 3.96 – 3.90 (m, 1H), 3.90 (d, *J* = 2.0 Hz, 2H), 3.81 (dt, *J* = 13.1, 6.4 Hz, 1H), 3.69

(dt, *J* = 13.6, 6.4 Hz, 1H), 3.10 – 2.98 (m, 1H). LC/MS (ESI) *m/z* 292.38; $[M+H]^+$ calcd for $C_{14}H_{15}ClN_3O_2^+$: 292.08

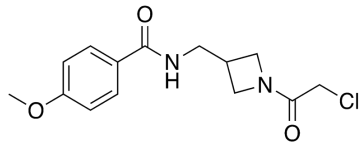


Synthesis of **CAS-11487-200** followed the General Procedure 1 by using 1-methylpiperidine-4-carboxylic acid (0.300 g, 2.10 mmol), and *tert*-butyl 3-(aminomethyl)azetidine-1-carboxylate (0.421 mL, 2.30 mmol) to afford CAS-11487-200 (0.0266 g, 12.6%). ^{1}H NMR (500 MHz, Methanol-d_4) δ 4.37 (t, *J* = 8.8 Hz, 1H), 4.13 – 4.07 (m, 1H), 4.07 – 3.98 (m, 4H), 3.76 (dd, *J* = 10.4, 5.5 Hz, 1H), 3.62 – 3.54 (m, 2H), 3.51 – 3.37 (m, 3H), 3.03 (td, *J* = 13.0, 3.1 Hz, 2H), 2.88 (s, 5H), 2.53 (tt, *J* = 12.1, 3.8 Hz, 1H), 2.15 (d, *J* = 16.0 Hz, 1H), 2.10 – 2.01 (m, 2H), 2.01 – 1.88 (m, 2H). LC/MS (ESI) *m/z* 288.07; $[M+H]^+$ calcd for $C_{13}H_{23}ClN_3O_2^+$: 288.15



Synthesis of **CAS-11487-205** followed the General Procedure 1 by using 4-hydroxybenzoic acid (0.200 g, 1.45 mmol), and *tert*-butyl 3-(aminomethyl)azetidine-1-carboxylate (0.291 mL, 1.59 mmol) to afford CAS-11487-205 (0.001 g, 1.4%) ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.76 – 7.65 (m, 2H), 7.66 – 7.51 (m, 0H), 7.05 – 6.93 (m, 0H), 6.90 – 6.78 (m, 2H), 4.46 – 4.34 (m, 1H), 4.19 – 4.06 (m, 2H), 4.02 (d, *J* = 1.1 Hz, 2H), 3.91 – 3.81 (m, 1H), 3.67 – 3.55 (m, 2H), 3.00 (dddd, *J* = 12.3, 8.4, 5.0, 2.2 Hz, 1H). LC/MS (ESI) *m/z* 283.15; $[M+H]^+$ calcd for $C_{13}H_{16}ClN_2O_3^+$: 283.08

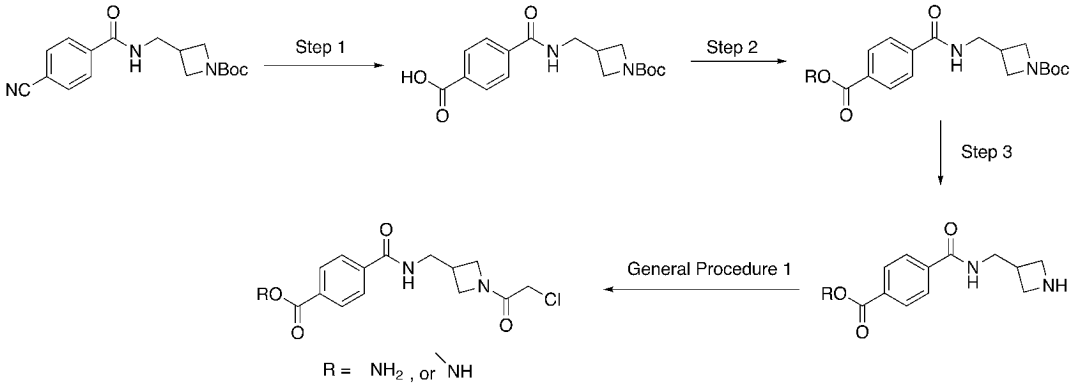
Synthesis of CAS 12290 001



Synthesis of **CAS-12290-001** followed the General Procedure 1 by using 4-methoxybenzoic acid (0.250 g, 1.64 mmol), and *tert*-butyl 3-(aminomethyl)azetidine-1-carboxylate (0.330 mL, 1.81 mmol) to afford CAS-12290-001 (0.060 g, 24.8%) ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.83 –

7.63 (m, 2H), 7.18 (t, *J* = 6.1 Hz, 1H), 7.00 – 6.80 (m, 2H), 4.36 (t, *J* = 8.7 Hz, 1H), 4.18 – 4.04 (m, 2H), 3.89 (q, *J* = 5.4, 5.0 Hz, 1H), 3.86 (d, *J* = 7.5 Hz, 5H), 3.72 (dt, *J* = 13.1, 6.4 Hz, 1H), 3.62 (dt, *J* = 13.6, 6.4 Hz, 1H), 3.05 – 2.94 (m, 1H). LC/MS (ESI) *m/z* 297.18; $[M+H]^+$ calcd for $C_{14}H_{18}ClN_2O_3^+$: 297.10

Syntheses of CAS 12290 010, and 011



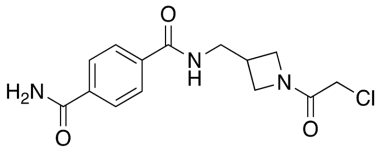
Step 1: Nitrile *tert*-butyl 3-((4-cyanobenzamido)methyl)azetidine-1-carboxylate (0.350 g, 1.10 mmol) was dissolved in 80% EtOH (8 mL), KOH (0.311 g, 5.55 mmol) was added and the reaction was heated to reflux for 16h. Crude was purified directly by flash chromatography using eluent gradient 0-40% MeOH/EtOAc. 140 mg of desired product 4-(((1-(*tert*-butoxycarbonyl)azetidin-3-yl)methyl)carbamoyl)benzoic acid was obtained (38%).

Step 2: 4-(((1-(*tert*-butoxycarbonyl)azetidin-3-yl)methyl)carbamoyl)benzoic acid (0.0703 g, .210 mmol) was dissolved in DCM (1-2 mL) with diisopropylethylamine (0.074 mL, 0.4205 mmol), and HATU (0.094 g, 0.2522 mmol) at room temperature. Ammonium chloride (0.045 g, 0.8409 mmol), or 2M Methylamine (0.030 mL, 0.6307 mmol) were added. The mixture then stirred at room temperature overnight, and was directly purified by flash chromatography 10% MeOH/ EtOAc. 34 mg of desired product, *tert*-butyl 3-((4-carbamoylbenzamido)methyl)azetidine-1-carboxylate, and 39 mg of desired product *tert*-butyl 3-((4-(methylcarbamoyl)benzamido)methyl)azetidine-1-carboxylate were obtained (49%, 53%, respectively).

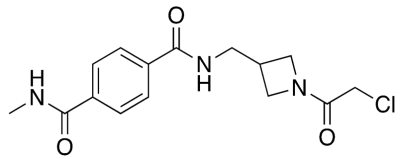
Step 3: Products from the last step were dissolved in DCM (1-2 mL) at room temperature and treated with TFA (1 mL). The mixtures stirred at room temperature until the tert-

butyloxycarbonyl protecting group was cleaved tracking by UPLC-MS. The mixtures were concentrated and placed *in vacuo* for 12 hours.

Step 4: Acylations using chloroacetyl chloride were performed according to Step 3 of General Procedure 1 to yield the target compounds **CAS-12290-010** (5.65%), and **CAS-12290-011** (5.7%).

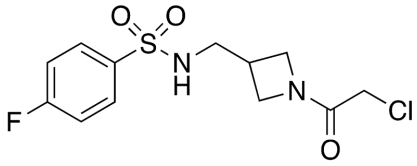


CAS-12290-010 ^{1}H NMR (500 MHz, Methanol-d_4) δ 8.82 (s, 1H), 7.99 – 7.95 (m, 1H), 7.93 – 7.89 (m, 1H), 4.42 (t, *J* = 8.8 Hz, 1H), 4.16 (d, *J* = 9.5 Hz, 1H), 4.13 (dd, *J* = 8.3, 4.6 Hz, 1H), 4.03 (d, *J* = 2.5 Hz, 2H), 3.87 (dd, *J* = 10.3, 5.5 Hz, 1H), 3.71 – 3.63 (m, 2H), 3.05 – 3.00 (m, 1H). LC/MS (ESI) *m/z* 310.07; $[M+H]^+$ calcd for $C_{14}H_{17}ClN_3O_3^+$: 310.10



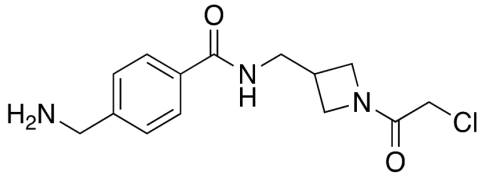
Synthesis of **CAS-12290-011** ^{1}H NMR (500 MHz, Methanol-d_4) δ 8.81 (s, 1H), 7.91 (s, 4H), 4.43 (d, *J* = 8.8 Hz, 1H), 4.17 – 4.11 (m, 2H), 4.03 (d, *J* = 2.6 Hz, 2H), 3.87 (dd, *J* = 10.3, 5.5 Hz, 2H), 3.66 (dd, *J* = 7.0, 3.2 Hz, 2H), 3.03 (d, *J* = 7.1 Hz, 1H), 2.95 (s, 3H), 2.68 (s, 2H). LC/MS (ESI) *m/z* 324.17; $[M+H]^+$ calcd for $C_{15}H_{19}ClN_3O_3^+$: 324.11

<u>Syntheses of CAS-12290 024, 039, 073, 076, and 077</u>

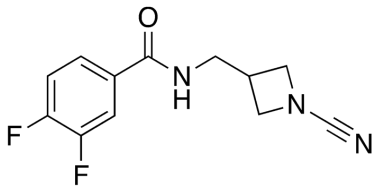


Synthesis of **CAS-12290-024** was performed by combining 4-fluorobenzenesulfonyl chloride (.200 g, 1.03 mmol), *tert*-butyl 3-(aminomethyl)azetidine-1-carboxylate (0.181 mL, 1.03 mmol), and K_2CO_3 (0.285 g, 2.06 mmol) in DCM at rt and was allowed to react overnight. Crude was then directly purified by flash chromatography 80% EtOAc / Hexanes then conc. *in vacuo* to afford *tert*-butyl 3-(((4-fluorophenyl)sulfonamido)methyl)azetidine-1-carboxylate (92.2 %).

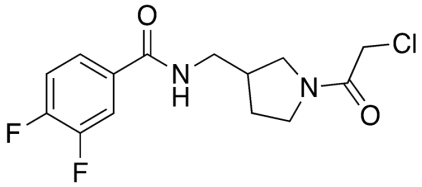
Subsequent removal of the Boc protecting group and acylation using chloroacetyl chloride were executed following Step(s) 2, and 3 of General Procedure 1 to afford **CAS-12290-024** (2.5 %). ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.96 – 7.78 (m, 1H), 7.52 – 7.40 (m, 1H), 6.57 (s, 1H), 4.16 (t, *J* = 8.6 Hz, 1H), 4.07 (s, 1H), 3.90 – 3.76 (m, 1H), 3.55 (dd, *J* = 9.9, 5.4 Hz, 1H), 3.03 – 2.90 (m, 1H), 2.67 (dddd, *J* = 13.8, 6.7, 4.2, 1.5 Hz, 1H). LC/MS (ESI) *m/z* 321.04; $[M+23]^+$ calcd for $C_{12}H_{15}ClFN_2O_3S^+$: 321.05



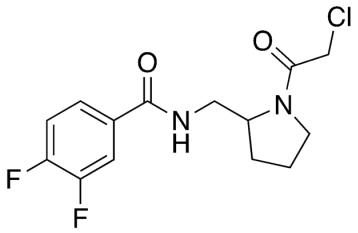
Synthesis of **CAS-12290-039** was executed by combining CAS-11487-199 (0.0263 g, 0.0902 mmol), and $NiCl_2 \cdot 6H_2O$ (0.02145 g, 0.0902 mmol) in 15:1 EtOH/DCM (1 mL) and placed on ice. $NaBH_4$ (0.0102 g, 0.2708 mmol) was added slowly, and then the mixture was warmed to room temperature and allowed to react for 2h. The crude was purified by flash chromatography (hexanes/EtOAc/MeOH) and preparative HPLC (MeOH/H_2O w/ 0.0425% TFA) to afford the target molecule (5.6 %). ^{1}H NMR (500 MHz, Methanol-d_4) δ 10.09 (s, 1H), 8.05 – 7.98 (m, 2H), 7.86 – 7.81 (m, 1H), 7.57 (dd, *J* = 18.7, 8.1 Hz, 2H), 4.45 – 4.38 (m, 2H), 4.15 (ddd, *J* = 14.2, 9.3, 5.2 Hz, 3H), 4.03 (t, *J* = 2.5 Hz, 2H), 3.87 (dt, *J* = 9.9, 4.6 Hz, 2H), 3.69 – 3.62 (m, 3H), 3.03 (q, *J* = 6.5, 5.9 Hz, 2H). LC/MS (ESI) *m/z* 296.32; $[M+H]^+$ calcd for $C_{14}H_{19}ClN_3O_2^+$: 296.12



Synthesis of **CAS-12290-073** followed the General Procedure 1 by using 3,4-difluorobenzoic acid (0.500 g, 3.16 mmol), and *tert*-butyl 3-(aminomethyl)azetidine-1-carboxylate (0.636 mL, 3.48 mmol) to afford CAS-12290-073 (0.0062 g, 11.6%) ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.67 (ddd, *J* = 10.6, 7.5, 2.2 Hz, 1H), 7.54 (dddd, *J* = 8.5, 3.9, 2.2, 1.4 Hz, 1H), 7.28 – 7.22 (m, 1H), 6.46 (s, 1H), 4.30 (t, *J* = 8.0 Hz, 2H), 3.98 (dd, *J* = 7.8, 5.6 Hz, 2H), 3.69 (t, *J* = 6.4 Hz, 2H), 3.05 (ttt, *J* = 8.2, 6.7, 5.6 Hz, 1H). LC/MS (ESI) *m/z* 252.19; $[M+H]^+$ calcd for $C_{12}H_{12}F_2N_3O^+$: 252.09

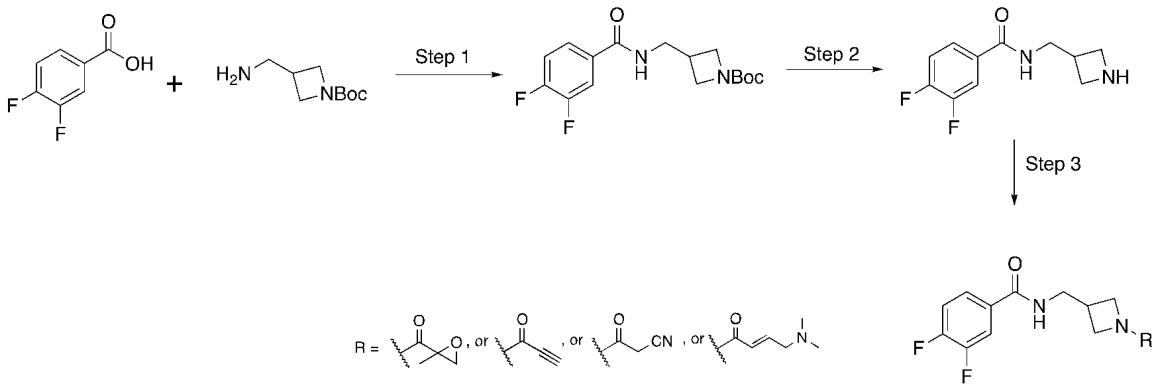


Synthesis of **CAS-12290-076** followed the General Procedure 1 by using 3,4-difluorobenzoic acid (0.150 g, 0.9487 mmol), and *tert*-butyl 3-(aminomethyl)pyrrolidine-1-carboxylate (0.200 mL, 1.04 mmol) to afford CAS-12290-076 (0.0151 g, 14.8%) ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.71 (dddd, *J* = 16.4, 10.7, 7.5, 2.2 Hz, 1H), 7.64 – 7.51 (m, 1H), 7.28 – 7.20 (m, 1H), 6.83 (d, *J* = 141.9 Hz, 1H), 4.08 – 4.03 (m, 3H), 3.71 (pd, *J* = 6.0, 4.5, 1.7 Hz, 3H), 3.67 – 3.54 (m, 1H), 3.51 (ddd, *J* = 8.9, 5.5, 3.3 Hz, 1H), 3.47 – 3.28 (m, 2H), 2.77 – 2.60 (m, 1H), 2.26 – 2.08 (m, 1H), 1.83 (ddq, *J* = 55.3, 12.8, 7.7 Hz, 1H). LC/MS (ESI) *m/z* 317.38; $[M+H]^+$ calcd for $C_{14}H_{16}ClF_2N_2O_2^+$: 317.09



Synthesis of **CAS-12290-077** followed the General Procedure 1 by using 3,4-difluorobenzoic acid (0.150 g, 0.9487 mmol), and *tert*-butyl 2-(aminomethyl)pyrrolidine-1-carboxylate (0.200 mL, 1.04 mmol) to afford CAS-12290-077 (0.0291 g, 27.3%) ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.40 (s, 1H), 7.75 (ddd, *J* = 11.0, 7.6, 2.2 Hz, 1H), 7.60 (dddd, *J* = 8.6, 3.9, 2.2, 1.4 Hz, 1H), 7.23 (ddd, *J* = 9.9, 8.6, 7.7 Hz, 1H), 5.90 (s, 1H), 4.48 (ddt, *J* = 10.7, 7.4, 3.3 Hz, 1H), 4.10 (d, *J* = 1.9 Hz, 2H), 3.76 – 3.61 (m, 3H), 3.44 (ddd, *J* = 14.0, 10.4, 3.7 Hz, 1H), 2.21 – 1.97 (m, 3H), 1.93 – 1.70 (m, 1H). LC/MS (ESI) *m/z* 317.08; $[M+H]^+$ calcd for $C_{14}H_{16}ClF_2N_2O_2^+$: 317.09

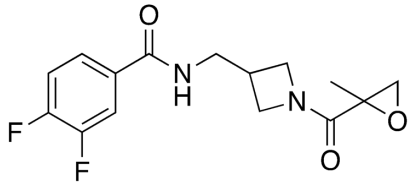
<u>Synthesis of CAS 12290 092, 087, 089, 091</u>



Step 1: Following methods described in Step 1 of General Procedure 1, 3,4-difluorobenzoic acid (0.500 g, 3.16 mmol), and *tert*-butyl 3-(aminomethyl)azetidine-1-carboxylate (0.636 mL, 3.48 mmol) were combined to form *tert*-butyl 3-((3,4-difluorobenzamido)methyl)azetidine-1-carboxylate (57.3%).

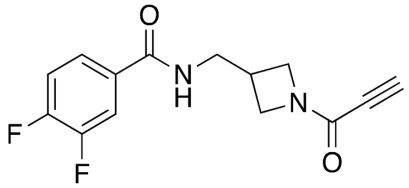
Step 2: Synthesis was performed adhering to methods describes in Step 2 of General Procedure 1 to afford *N*-(azetidin-3-ylmethyl)-3,4-difluorobenzamide (quant.)

Step 3: The synthesis was performed according to Step 1 of General Procedure 3 with *N*-(azetidin-3-ylmethyl)-3,4-difluorobenzamide (0.100 g, 0.3064 mmol), and 2-methyloxirane-2-carboxylic acid (0.045 mL, 0.4864 mmol), or propiolic acid (0.022 mL, 0.337 mmol), or 2-cyanoacetic acid (0.050 g, 0.574 mmol), or (*E*)-4-(dimethylamino)but-2-enoic acid (0.062 g, 0.4862 mmol). Combined organic layer was dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude material was then purified by flash column chromatography (hexanes/EtOAc/MeOH) followed by preparative HPLC (MeOH or CH_3CN/H_2O with 0.0425% TFA) to afford the target products **CAS-12290-092** (7%), **CAS-12290-087**(15.3%), **CAS-12290-089** (17.8%), and **CAS-12290-091**(13.8%).

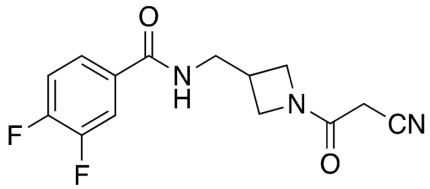


CAS-12290-092 ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.67 (dddt, *J* = 11.4, 7.5, 4.1, 2.0 Hz, 1H), 7.56 (ddh, *J* = 8.7, 6.6, 2.2 Hz, 1H), 7.24 (dddt, *J* = 11.7, 10.2, 8.0, 2.1 Hz, 1H), 7.04 – 6.55 (m, 1H), 4.65 – 4.46 (m, 1H), 4.46 – 4.24 (m, 1H), 4.16 (td, *J* = 9.3, 8.6, 4.7 Hz, 1H), 3.97 – 3.81 (m,

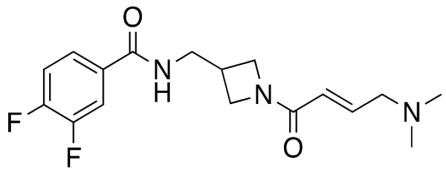
1H), 3.80 – 3.61 (m, 3H), 3.05 – 2.89 (m, 1H), 2.89 – 2.74 (m, 1H), 1.54 (d, *J* = 8.1 Hz, 1H), 1.51 – 1.37 (m, 1H), 1.31 (s, 1H). LC/MS (ESI) *m/z* 311.11; $[M+H]^+$ calcd for $C_{15}H_{17}F_2N_2O_3^+$: 311.12



CAS-12290-087 ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.73 (ddd, *J* = 10.8, 7.5, 2.2 Hz, 1H), 7.61 (dddd, *J* = 8.7, 3.9, 2.2, 1.4 Hz, 1H), 7.24 (ddd, *J* = 9.9, 8.6, 7.6 Hz, 1H), 7.13 (d, *J* = 6.2 Hz, 1H), 4.35 (dd, *J* = 9.7, 8.2 Hz, 1H), 4.14 (dd, *J* = 10.7, 8.4 Hz, 1H), 4.02 (ddd, *J* = 9.8, 5.2, 1.2 Hz, 1H), 3.91 (ddd, *J* = 10.8, 5.3, 1.2 Hz, 1H), 3.78 (dt, *J* = 13.9, 6.3 Hz, 1H), 3.61 (ddd, *J* = 13.6, 7.5, 5.8 Hz, 1H), 3.02 (ddd, *J* = 14.0, 7.4, 3.3 Hz, 1H), 2.71 (s, 1H). LC/MS (ESI) *m/z* 279.09; $[M+H]^+$ calcd for $C_{14}H_{13}F_2N_2O_2^+$: 279.09

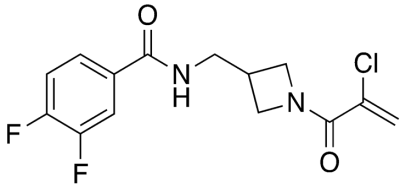


CAS-12290-089 ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.67 (ddd, *J* = 10.1, 7.4, 2.1 Hz, 1H), 7.53 (d, *J* = 8.8 Hz, 1H), 7.30 – 7.23 (m, 4H), 4.42 (t, *J* = 8.6 Hz, 1H), 4.21 (t, *J* = 9.4 Hz, 1H), 4.14 (dd, *J* = 8.8, 5.2 Hz, 1H), 3.88 (dd, *J* = 10.3, 5.4 Hz, 1H), 3.79 (dt, *J* = 13.6, 6.7 Hz, 1H), 3.68 (dt, *J* = 13.6, 6.2 Hz, 1H), 3.29 (s, 2H), 3.05 (s, 1H).). LC/MS (ESI) *m/z* 294.18; $[M+H]^+$ calcd for $C_{14}H_{14}F_2N_3O_2^+$: 294.10



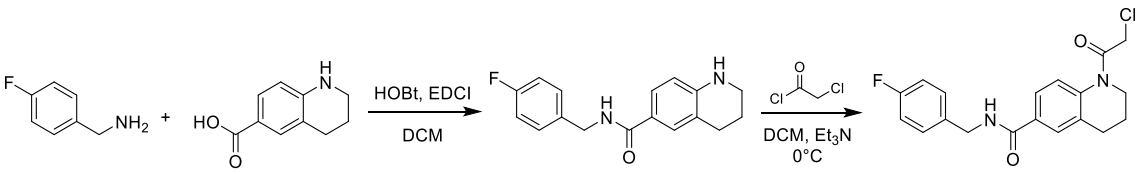
CAS-12290-091 ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.05 (s, 1H), 8.78 (t, *J* = 5.7 Hz, 1H), 7.88 (ddd, *J* = 11.5, 7.8, 2.2 Hz, 1H), 7.77 – 7.65 (m, 1H), 7.57 (dt, *J* = 10.5, 8.3 Hz, 1H), 6.57 (dt, *J* = 15.2, 7.1 Hz, 2H), 6.42 – 6.34 (m, 1H), 4.28 (t, *J* = 8.5 Hz, 1H), 4.02 – 3.93 (m, 2H), 3.87 (dd, *J* = 7.1, 1.2 Hz, 2H), 3.72 (dd, *J* = 10.3, 5.4 Hz, 1H), 3.56 – 3.43 (m, 2H), 2.91 – 2.81 (m, 1H), 2.76 (s, 6H). LC/MS (ESI) *m/z* 338.16; $[M+H]^+$ calcd for $C_{17}H_{22}F_2N_3O_2^+$: 338.17

Synthesis of CAS-12290-094



Synthesis of **CAS-12290-094** followed Step 3 of General Procedure 1 and used previously synthesized intermediate *N*-(azetidin-3-ylmethyl)-3,4-difluorobenzamide (0.050 g, 0.221 mmol), and 2-chloroacryloyl chloride (0.0226 mL, 0.243 mmol) to afford CAS-12290-094 (0.0124 g,17.9 %) ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.69 (ddd, *J* = 10.7, 7.5, 2.3 Hz, 1H), 7.57 (ddt, *J* = 8.0, 4.1, 1.8 Hz, 1H), 7.28 – 7.19 (m, 1H), 6.86 (s, 1H), 6.11 (d, *J* = 1.8 Hz, 1H), 5.80 (d, *J* = 1.7 Hz, 1H), 4.55 (t, *J* = 9.1 Hz, 1H), 4.21 (d, *J* = 9.8 Hz, 2H), 3.96 (dd, *J* = 11.1, 5.5 Hz, 1H), 3.85 – 3.54 (m, 2H), 3.00 (dddd, *J* = 13.5, 6.8, 3.0, 1.5 Hz, 1H). LC/MS (ESI) *m/z* 315.07; $[M+H]^+$ calcd for $C_{14}H_{14}ClF_2N_2O_2^+$: 315.06

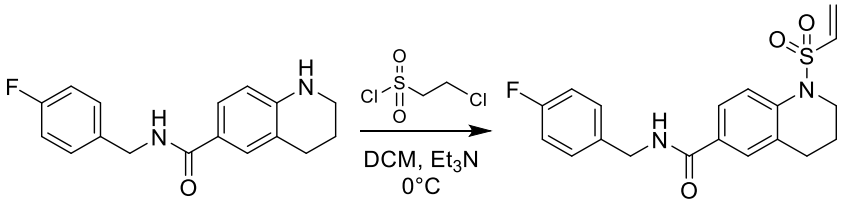
Synthesis of ED-11370-017, 018:



Step 1: 1,2,3,4-tetrahydroquinoline-6-carboxylic acid (0.088g, 0.5mmol), HOBt (80%, 0.12g, 0.6mmol), and EDCl (0.14g, 0.75mmol) were added sequentially into 3mL anhydrous DCM. Into the solution was added 4-fluorobenzylamine (0.063g, 0.5mmol). The mixture was stirred at room temperature overnight, and purified directly by flash column chromatography (hexanes/$EtOA_C$/MeOH) to afford 0.15g product 4-fluorobenzyl 1,2,3,4-tetrahydroquinoline-6-carboxylate (quant.) LC/MS (ESI) *m/z* 284.67; $[M+H]^+$ calcd for $C_{17}H_{18}FN_2O^+$: 285.14

Step 2: 4-fluorobenzyl 1,2,3,4-tetrahydroquinoline-6-carboxylate (0.075g, 0.25mmol) was dissolved in 2.5mL anhydrous DCM. Into the solution was added Et_3N (0.18mL, 1.3mmol), and 2-chloroacetyl chloride (0.036g, 0.32mmol) at 0°C. The mixture was stirred for 10min before direct purification by flash column chromatography (hexanes/$EtOA_C$/MeOH) and followed by preparative HPLC (CH_3CN/H_2O with 0.0425% TFA) to afford desired product ED_11370_017

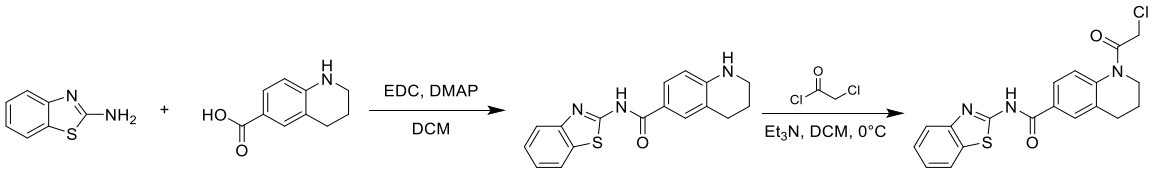
(28mg, 30%). 1H NMR (500 MHz, DMSO) δ 9.00 (t, J = 5.8 Hz, 1H), 7.79 – 7.62 (m, 3H), 7.35 (dd, J = 8.5, 5.7 Hz, 2H), 7.21 – 7.10 (m, 2H), 4.59 (s, 2H), 4.45 (d, J = 5.9 Hz, 2H), 3.78 – 3.68 (t, J = 6.3 Hz, 2H), 2.78 (t, J = 6.6 Hz, 2H), 1.96 – 1.88 (m, 2H). LC/MS (ESI) *m/z* 360.77; $[M+H]^+$ calcd for $C_{19}H_{19}ClFN_2O_2^+$: 361.11



ED_11370_018

Step 1: 4-fluorobenzyl 1,2,3,4-tetrahydroquinoline-6-carboxylate (0.075g, 0.25mmol) was dissolved in 2.5mL anhydrous DCM. Into the solution was added Et_3N (0.18mL, 1.3mmol), and 2-chloroethanesulfonyl chloride (0.052g, 0.32mmol) at 0°C. The mixture was stirred for 10min before direct purification by flash column chromatography (hexanes/$EtOA_C$/MeOH) and followed by preparative HPLC (CH_3CN/H_2O with 0.0425% TFA) to afford desired product ED_11370_018 (10mg, 10%). 1H NMR (500 MHz, DMSO) δ 8.95 (t, J = 5.9 Hz, 1H), 7.73 – 7.63 (m, 2H), 7.56 (d, J = 8.7 Hz, 1H), 7.34 (dd, J = 8.8, 5.6 Hz, 2H), 7.15 (t, J = 8.9 Hz, 2H), 6.92 (dd, J = 16.4, 9.9 Hz, 1H), 6.14 (d, J = 16.4 Hz, 9.9 Hz, 2H), 4.44 (d, J = 6.1 Hz, 2H), 3.80 – 3.69 (m, 2H), 2.82 (t, J = 6.6 Hz, 2H), 1.93 (dd, J = 12.1, 6.1 Hz, 2H). LC/MS (ESI) *m/z* 374.97; $[M+H]^+$ calcd for $C_{19}H_{20}FN_2O_3S^+$: 375.12

<u>Synthesis of ED_11370_027:</u>

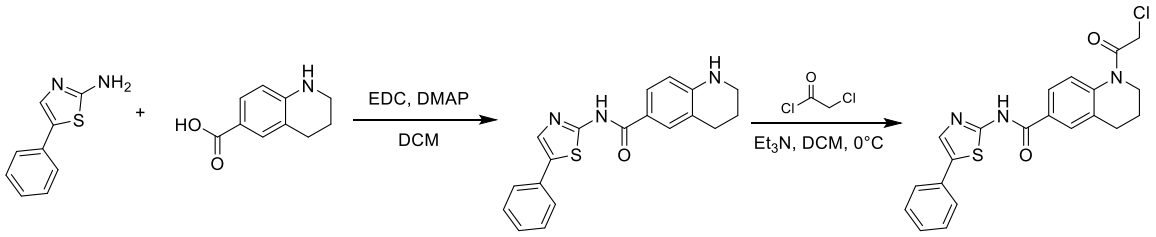


Step 1: 1,2,3,4-tetrahydroquinoline-6-carboxylic acid (0.26g, 1.5mmol), benzothiazole-2-amine (0.12g, 0.75mmol) were added into 3mL anhydrous DCM. Into the solution was added EDC (0.29g, 1.5mmol), and DMAP(0.37g, 3.0mmol) sequentially. The mixture was stirred at room temperature overnight, and purified directly by flash column chromatography (hexanes/EtOAc/MeOH) to afford 0.026g product N-(benzo[d]thiazol-2-yl)-1,2,3,4-

tetrahydroquinoline-6-carboxamide (11%). LC/MS (ESI) *m/z* 309.87; $[M+H]^+$ calcd for $C_{17}H_{16}N_3OS^+$: 310.10

Step 2: N-(benzo[d]thiazol-2-yl)-1,2,3,4-tetrahydroquinoline-6-carboxamide (0.026g, 0.08mmol) was dissolved in 5mL anhydrous DCM. Into the solution was added Et_3N (0.085mL, 0.6mmol), and 2-chloroacetyl chloride (0.01uL, 0.14mmol) at 0°C. The mixture was stirred for 10min before direct purification by flash column chromatography (hexanes/$EtOA_C$/MeOH) and followed by preparative HPLC (CH_3CN/H_2O with 0.0425% TFA) to afford desired product ED_11370_027 (27mg, 88%). 1H NMR (500 MHz, DMSO) δ 12.80 (s, 1H), 8.06 – 8.00 (m, 2H), 7.98 (dd, J = 8.6, 2.1 Hz, 1H), 7.81 (dd, J = 16.2, 8.1 Hz, 2H), 7.48 (t, J = 7.7 Hz, 1H), 7.35 (t, J = 7.6 Hz, 1H), 4.65 (s, 2H), 3.81 – 3.72 (t, J = 5.8 Hz, 2H), 2.84 (t, J = 6.5 Hz, 2H), 2.03 – 1.87 (m, 2H). LC/MS (ESI) *m/z* 385.97; $[M+H]^+$ calcd for $C_{19}H_{17}ClN_3O_2S^+$: 386.07

Synthesis of ED_11370_030:



Step 1: 1,2,3,4-tetrahydroquinoline-6-carboxylic acid (0.26g, 1.5mmol), 5-phenylthiazol-2-amine (0.13g, 0.75mmol) were added into 3mL anhydrous DCM. Into the solution was added EDC (0.29g, 1.5mmol), and DMAP(0.28g, 2.3mmol) sequentially. The mixture was stirred at room temperature overnight, and purified directly by flash column chromatography (hexanes/EtOAc/MeOH) to afford 0.064g N-(5-phenylthiazol-2-yl)-1,2,3,4-tetrahydroquinoline-6-carboxamide (25%). LC/MS (ESI) *m/z* 335.87; $[M+H]^+$ calcd for $C_{19}H_{18}ClN_3OS^+$: 336.12

Step 2: N-(5-phenylthiazol-2-yl)-1,2,3,4-tetrahydroquinoline-6-carboxamide (0.064g, 0.2mmol) was dissolved in 5mL anhydrous DCM. Into the solution was added Et_3N (0.13mL, 0.9mmol), and 2-chloroacetyl chloride (0.017uL, 0.23mmol) at 0°C. The mixture was stirred for 10min before direct purification by flash column chromatography (hexanes/$EtOA_C$/MeOH) and followed by preparative HPLC (CH_3CN/H_2O with 0.0425% TFA) to afford desired product ED_11370_030

(12mg, 15%). 1H NMR (500 MHz, DMSO) δ 12.63 (s, 1H), 8.01 (s, 1H), 7.99 – 7.92 (m, 2H), 7.80 (d, J = 8.0 Hz, 1H), 7.66 (d, J = 7.6 Hz, 2H), 7.44 (t, J = 7.7 Hz, 2H), 7.33 (t, J = 7.4 Hz, 1H), 4.64 (s, 2H), 3.82 – 3.70 (t, J = 6.1 Hz, 2H), 2.84 (t, J = 6.4 Hz, 2H), 2.04 – 1.88 (m, 2H). LC/MS (ESI) *m/z* 411.97; $[M+H]^+$ calcd for $C_{21}H_{19}ClN_3O_2S^+$: 412.09

6.1.2 Assay methods

Constructs

UCHL1 (residues 1-223, full length) was cloned into a pGEX6P1 expression vector with an N-terminal GST tag.

UCHL3 (residues 1-230, full length) was cloned into a pET28PP expression vector with an N-terminal 6xHis tag.

USP7 (residues 208–560, catalytic domain) was cloned as described.[30]

USP28 (residues 149-704, catalytic domain) was cloned into a SUMO-pETDUET expression vector with a N-terminal 6xHis-SUMO tag was purchased from Genewiz.

USP30 (residues 65–517, catalytic domain) was cloned into a pET28PP expression vector with an N-terminal 6xHis tag.

OTUD7A (residues1-462, catalytic domain+UBA) in a pOPINK vector with an N-terminal GST tag was purchased from Addgene (#61582).

VCPIP1 (residues 25-561, catalytic domain) in a pOPINK vector with an N-terminal GST tag was purchased from Addgene (#61583).

Recombinant protein

USP20 (UBI-64-0039-050) and USP27x (UBI-46-0046-050) were ordered from Ubiquigent.

Recombinant USP9x (E-552-052), USP22 (E-608-050), USP15 (E-594-050), and USP48 (E-614-050) were all purchased from R&D Systems, Inc.

Reagents

Ub-AMC (U-550) and HA-Ub-VS (U-212) were obtained from Boston Biochem.

Bio-Ub-PA (UbiQ-076) and Bio-Ub-VME (UbiQ-054) were obtained from UbiQ Bio.

Antibodies

USP25 (ab187156) antibody was obtained from abcam. GAPDH (2118s), UCHL1(13179S), UCHL3 (3525S), USP28 (4217S), USP7 (4833s) antibody was obtained from Cell Signaling Technology. VCPIP1 (A302-933) and USP48 (A301-190A-M) antibody was obtained from Bethyl Laboratories.

Protein Expression

All constructs were overexpressed in E. coli BL21 (DE3). Cells were grown at 37 °C to an OD of 0.9, cooled to 16 °C, induced with 500µM isopropyl -1-thio-D-galactopyranoside (IPTG), incubated overnight at 16 °C, collected by centrifugation, and stored at −80 °C. Cell pellets were sonicated in lysis bufer (25 mM Tris pH 8, 1 M NaCl, and 10 mM BME) supplemented with 10 µg/ml phenylmethanesulfonylfuoride (PMSF) and the resulting lysate was centrifuged at 30,000

g for 40 min. Lysate from His-tagged proteins were mixed with Ni-NTA beads (Qiagen) 2 hours, and washed with lysis buffer supplemented with 25mM imidazole. The bound protein was eluted with lysis buffer supplemented with 300mM imidazole.

Lysate from GST-tagged proteins were mixed with glutathione beads (company) for 2 hours, washed with lysis buffer, and eluted overnight with 3C protease. The samples were then concentrated to 1 ml (30 kDa concentrator; Amicon Ultra, Millipore), and run on a Superdex 200 (GE healthcare) in buffer containing 25 mM HEPES pH 7.5, 200 mM NaCl, and 1mM DTT. Fractions were pooled, concentrated and frozen at −80 °C.

Biochemical Assays

Enzymes were tested for activity in Ubiquitin-Rhodamine assay in presence or absence of inhibitors. Enzyme (UCHL1: 2nM; UCHL3: 200pm; USP7: 10nM; USP28: 5nM; USP48: 10nM; VCPIP1: 100nM, JOSD1: 25nM, OTUD7A: 50nM, USP15: 0.1nM, USP9X:0.1nM, USP27X: 125nM, USP20: 1nM, USP21: 2nM) was pre-incubated for 6 hours at room temperature with different concentrations of inhibitors or DMSO as a control in 50mM TRIS pH 8, 0.5 mM EDTA, 10 µM ovalbumin, and 5mM TCEP. Ubiquitin-Rhodamine (Boston Biochem) was then added to a final concentration of 250 nM. The initial rate of the reaction was measured by collecting fluorescence data at one minute intervals over 30-minute period using a Clariostar fluorescence plate reader at excitation and emission wavelength of 345 and 445nm respectively. The calculated initial rate values were plotted against inhibitor concentrations to determine IC50s. All the experimental data were plotted using Prism GraphPad. All assays for each compound were performed at least twice for each compound.

Biochemical Assays

Selectivity profiling (DUBProfiler) was performed by Ubiquigent according to manufacturer protocol.

Cell culture

HEK293T cells were cultured in DMEM supplemented with 10% FBS, 1% penicillin-streptomycin, and 4 mM L-glutamine. Cell were maintained in 10 cm tissue-culture treated dishes 37°C in a 5% CO_2 incubator. Cells were treated with indicated compounds for the time and amount indicated.

Western blot target engagement

Flag-Ub-PA experiments were performed as previously described in Lamberto et al.[30] Briefly, target engagement lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$, 0.5 mM EDTA, 0.5% NP-40, 10% glycerol, 1mM TCEP, protease and phosphatase inhibitors) was added to cell pellets on ice. Lysate was cleared by centrifugation and diluted to 2 mg/mL. Where indicated, 30 µL lysate was then incubated with inhibitors or DMSO for the indicated time points. 2 µM Flag-Ub-PA was then added to the lysate and incubated at RT for the indicated time points. Labeling reactions were quenched with 4x LDS sample buffer (Termo Fisher B0007) supplemented with 10% BME, vortexed vigorously, and heated to 95°C for 5 minutes. Samples were resolved by SDS-PAGE and analyzed by Western blot with the indicated antibodies.

DUB Activity Based Protein Profiling Primary Screening Assay

DUB Activity based protein profiling was performed using conditions modified from those in Schaeur et al., based on work by Lawson *et al.*.[33,46] HEK 293T cells were lysed (50 mM Tris pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$, 0.5 mM EDTA, 0.5% NP-40, 10% glycerol, 1 mM TCEP, protease and phosphatase inhibitors) and the lysate was clarified by centrifugation, then diluted to 10 mg/mL. 200 μL aliquots were incubated at the indicated compound concentrations or DMSO for 5 hours at RT, final DMSO concentration 0.5%. Afterwards, the treated lysates were incubated with 1 μM each of Biotin-Ub-PA and Biotin-Ub-VME for 90 minutes at RT. 25 μL magnetic streptavidin sepharose slurry was added to each sample, followed by incubation at RT for 30 minutes with end-to-end rotation. After immobilizing the beads using a magnetic rack, the supernatant was subjected to an additional streptavidin pulldown as described above, and the pooled beads were washed (3x 0.2% SDS, 3x PBS, 2x ddH_2O). After the final wash, supernatant was removed, and the resin was flash frozen and stored at -80° C.

Sample Prep for Mass Spectrometry Analysis

Streptavidin beads were resuspended in 95 μL 100 mM Tris pH 8.0. Each sample was denatured with 0.1% rapigest, reduced (10 mM dithiothreitol), alkylated (22.5 mM iodoacetamide), and digested with trypsin at 37 °C overnight. The next day, beads were captured using a magnetic rack, and supernatants were acidified with 10% TFA, incubated at 37° C for 30 minutes, and centrifuged at 14,000 rpm for 15 minutes at 4° C to remove rapigest. Peptides were then desalted by C18 and dried by vacuum centrifugation.

Dried peptides were reconstituted in 40μL 50mM pH 8.0 TEAB, and 1/4 unit of TMT reagent was added and reactions incubated at RT for 1 hour. TMT reactions were pooled and treated with hydroxylamine according to the manufacturer's instructions. Peptide mixtures were then dried, reconstituted in 100 mM ammonium bicarbonate and desalted by SP3.[93] Eluted peptides were then analyzed by nanoLC-MS as described in Ficarro *et al.* with a NanoAcquity UPLC system (Waters, Milford, MA) interfaced to a QExactive HF mass spectrometer (Thermofisher Scientific, San Jose, CA).[94] TMT labeled peptides were injected onto a precolumn (4 cm POROS 10R2, Applied Biosystems, Framingham, MA), resolved on an analytical column (30 μm I.D. x 50 cm packed with 5 μm Monitor C18) and introduced to the mass spectrometer by ESI (spray voltage = 3.5 kV, flow rate ~30 nL/min). The mass spectrometer was operated in data dependent mode such that the 15 most abundant ions in each MS scan (*m/z* 300-2000, 120K resolution, target=3E6, lock mass for 445.120025 enabled) were subjected to MS/MS (m/z 100-2000, 30K resolution, target=1E5, max fill time=100 ms). Dynamic exclusion was selected with a repeat count of 1 and an exclusion time of 30 seconds. MS/MS data was extracted to .mgf using mulitplierz scripts and searched against a forward-reverse human NCBI refseq database using Mascot version 2.6.2.[95,96] Search parameters specified fixed cysteine carbamidomethylation, fixed N-terminal and lysine TMT labelling, and variable methionine oxidation. Additional multiplierz scripts were used to filter results to 1% FDR and derive protein-level aggregate reporter ion intensities using peptides mapping uniquely into the genome. Proteins with fewer than two unique peptides were disregarded for quantification due to low signal-to-noise ratio.

% ABP labelling blockage" is calculated by:

$$\left(1 - \frac{\text{aggregate TMT reporter ion intensity for protein in condition}}{\text{average aggregate TMT reporter ion intensity for protein in DMSO controls}}\right) \times 100\%$$

Intact MS Analysis

5μg of indicated DUBs were treated with DMSO or a 10-fold molar excess of compound for 1 hour. Reactions were then injected onto a self-packed reversed phase column (1/32" O.D. x 500 um I.D., 5 cm of POROS 10R2 resin), desalted, and eluted with an HPLC gradient (0-100% B in 4 minutes, A=0.2M acetic acid in water, B=0.2 M acetic acid in acetonitrile, flow rate ~30 μL/min) into an LTQ ion trap mass spectrometer (ThermoFisher Scientific, San Jose, CA). Profile mass spectra (*m/z* 300-2000) were deconvoluted using MagTran1.03b2 software[97].

CE-MS Analysis

To identify sites of covalent modification, treated protein was reduced (10 mM TCEP), alkylated (22.5 mM MMTS), and digested with trypsin overnight at 37 °C. Peptides were desalted using SP3, dried by vacuum centrifugation, and reconstituted in 1% formic acid/50% acetonitrile with 100 mM ammonium acetate[93]. Peptides were then analyzed by CE-MS using a ZipChip CE system and autosampler (908 Devices, Boston, MA) interfaced to a QExactive HF mass spectrometer (ThermoFisher Scientific, San Jose, CA). Peptide solution was loaded for 30 seconds, and the mass spectrometer was operated in data dependent mode and subjected the 5 most abundant ions in each MS scan (60k resolution, 3E6 target, lock mass enabled) to MS/MS (15k resolution, 1E5 target, 100 ms max inject time). Dynamic exclusion was enabled with a repeat count of 1 and an exclusion time of 6 seconds. MS/MS data was extracted to .mgf using mulitplierz scripts and searched against a forward-reverse human NCBI refseq database

using Mascot version 2.6[95,96]. Search parameters specified fixed carbamidomethylation of cysteine, and variable oxidation (methionine) and compound modification. Precursor mass tolerance was set to 10 ppm and product ion tolerance was 25 mmu. Spectral validation was performed using mzStudio[98].

Competition with biotinylated inhibitor analog for global off-target profiling

HEK 293T cells were lysed as described above, and the lysate was cleared by centrifugation. Samples were diluted to 10 mg/mL, and 200 µL lysate (2 mg protein total) was incubated with the indicated concentrations of F70 for 4 hours at RT, then 2 µM of DTB-F-70 for 4 additional hours. SDS was added to a final concentration of 1.2% and the sample was boiled for 5 minutes. After cooling to RT, DPBS was added to dilute SDS concentration to a final of 0.2%. 50 µL streptavidin agarose slurry was added to each sample, followed by incubation at RT for 90 minutes. After streptavidin enrichment, samples were washed (3×0.2% SDS, 3x PBS, 2x ddH2O). After the final wash, all supernatant was removed and the resin was flash frozen and stored at −80 °C until workup for TMT labeling. See “Sample Prep for Mass Spectrometry Analysis” section in the Methods section of the main text for further steps.

7. References

1. Barnash, K. D., James, L. I. & Frye, S. V. Target class drug discovery. *Nat. Chem. Biol.* **13**, 1053–1056 (2017).
2. Singh, J., Petter, R. C., Baillie, T. A. & Whitty, A. The resurgence of covalent drugs. *Nat. Rev. Drug Discov.* **10**, 307–317 (2011).
3. Mitchell, J. R. *et al.* Acetaminophen-Induced Hepatic Necrosis. I. Role of Drug Metabolism. *J. Pharmacol. Exp. Ther.* **187**, 185–194 (1973).
4. Jollow, D. J. *et al.* Acetaminophen-Induced Hepatic Necrosis. Ii. Role of Covalent Binding in Vivo. *J. Pharmacol. Exp. Ther.* **187**, 195–202 (1973).
5. Potter, W. Z. *et al.* Acetaminophen-Induced Hepatic Necrosis. Iii. Cytochrome P-450-Mediated Covalent Binding in Vitro. *J. Pharmacol. Exp. Ther.* **187**, 203–210 (1973).
6. Bondeson, D. P. *et al.* Catalytic in vivo protein knockdown by small-molecule PROTACs. *Nat. Chem. Biol.* **11**, 611–617 (2015).
7. Winter, G. E. *et al.* Phthalimide conjugation as a strategy for in vivo target protein degradation. *Science* **348**, 1376–1381 (2015).
8. Mullard, A. Targeted protein degraders crowd into the clinic. *Nat. Rev. Drug Discov.* **20**, 247–250 (2021).
9. Komander, D. & Rape, M. The ubiquitin code. *Annu. Rev. Biochem.* **81**, 203–229 (2012).
10. Goldstein, G. *et al.* Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proc. Natl. Acad. Sci. U. S. A.* **72**, 11–15 (1975).

11. Komander, D., Clague, M. J. & Urbé, S. Breaking the chains: structure and function of the deubiquitinases. *Nat. Rev. Mol. Cell Biol.* **10**, 550–563 (2009).
12. Pickart, C. M. & Rose, I. A. Ubiquitin carboxyl-terminal hydrolase acts on ubiquitin carboxyl-terminal amides. *J. Biol. Chem.* **260**, 7903–7910 (1985).
13. Hewings, D. S. *et al.* Reactive-site-centric chemoproteomics identifies a distinct class of deubiquitinase enzymes. *Nat. Commun.* **9**, 1162 (2018).
14. Kwasna, D. *et al.* Discovery and Characterization of ZUFSP/ZUP1, a Distinct Deubiquitinase Class Important for Genome Stability. *Mol. Cell* **70**, 150-164.e6 (2018).
15. Abdul Rehman, S. A. *et al.* MINDY-1 Is a Member of an Evolutionarily Conserved and Structurally Distinct New Family of Deubiquitinating Enzymes. *Mol. Cell* **63**, 146–155 (2016).
16. Bailey-Elkin, B. A., Kasteren, P. B. van, Snijder, E. J., Kikkert, M. & Mark, B. L. Viral OTU Deubiquitinases: A Structural and Functional Comparison. *PLOS Pathog.* **10**, e1003894 (2014).
17. Shin, D. *et al.* Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature* **587**, 657–662 (2020).
18. Ndubaku, C. & Tsui, V. Inhibiting the Deubiquitinating Enzymes (DUBs). *J. Med. Chem.* **58**, 1581–1595 (2015).
19. Schauer, N. J., Magin, R. S., Liu, X., Doherty, L. M. & Buhrlage, S. J. Advances in Discovering Deubiquitinating Enzyme (DUB) Inhibitors. *J. Med. Chem.* **63**, 2731–2750 (2020).
20. Yang, J. *et al.* Small molecule inhibition of deubiquitinating enzyme JOSD1 as a novel targeted therapy for leukemias with mutant JAK2. *Leukemia* 1–11 (2021) doi:10.1038/s41375-021-01336-9.

21. Weisberg, E. L. *et al.* Inhibition of USP10 induces degradation of oncogenic FLT3. *Nat. Chem. Biol.* **13**, 1207–1215 (2017).
22. Coyne, E. S. *et al.* Knockout of USP19 Deubiquitinating Enzyme Prevents Muscle Wasting by Modulating Insulin and Glucocorticoid Signaling. *Endocrinology* **159**, 2966–2977 (2018).
23. Kluge, A. F. *et al.* Novel highly selective inhibitors of ubiquitin specific protease 30 (USP30) accelerate mitophagy. *Bioorg. Med. Chem. Lett.* **28**, 2655–2659 (2018).
24. Shimada, Y., Okano, T., Ohira, T. & Ikeda, N. UCHL1 has prognostic relevance and is a therapeutic target in high-grade neuroendocrine lung cancers. *Ann. Oncol.* **29**, (2018).
25. Lane, D. P. Cancer. p53, guardian of the genome. *Nature* **358**, 15–16 (1992).
26. Li, M. *et al.* Deubiquitination of p53 by HAUSP is an important pathway for p53 stabilization. *Nature* **416**, 648–653 (2002).
27. Bhattacharya, S., Chakraborty, D., Basu, M. & Ghosh, M. K. Emerging insights into HAUSP (USP7) in physiology, cancer and other diseases. *Signal Transduct. Target. Ther.* **3**, 1–12 (2018).
28. Altun, M. *et al.* Activity-Based Chemical Proteomics Accelerates Inhibitor Development for Deubiquitylating Enzymes. *Chem. Biol.* **18**, 1401–1412 (2011).
29. Reverdy, C. *et al.* Discovery of Specific Inhibitors of Human USP7/HAUSP Deubiquitinating Enzyme. *Chem. Biol.* **19**, 467–477 (2012).
30. Lamberto, I. *et al.* Structure-Guided Development of a Potent and Selective Non-covalent Active-Site Inhibitor of USP7. *Cell Chem. Biol.* **24**, 1490-1500.e11 (2017).
31. Turnbull, A. P. *et al.* Molecular basis of USP7 inhibition by selective small molecule inhibitors. *Nature* **550**, 481–486 (2017).

32. Di Lello, P. *et al.* Discovery of Small-Molecule Inhibitors of Ubiquitin Specific Protease 7 (USP7) Using Integrated NMR and in Silico Techniques. *J. Med. Chem.* **60**, 10056–10070 (2017).
33. Schauer, N. J. *et al.* Selective USP7 inhibition elicits cancer cell killing through a p53-dependent mechanism. *Sci. Rep.* **10**, 1–15 (2020).
34. Mistry, H. *et al.* Small-molecule inhibitors of USP1 target ID1 degradation in leukemic cells. *Mol. Cancer Ther.* **12**, 2651–2662 (2013).
35. Popov, N. *et al.* The ubiquitin-specific protease USP28 is required for MYC stability. *Nat. Cell Biol.* **9**, 765–774 (2007).
36. Berlin, I., Schwartz, H. & Nash, P. D. Regulation of Epidermal Growth Factor Receptor Ubiquitination and Trafficking by the USP8·STAM Complex *. *J. Biol. Chem.* **285**, 34909–34921 (2010).
37. Bédard, N. *et al.* Inactivation of the ubiquitin-specific protease 19 deubiquitinating enzyme protects against muscle wasting. *FASEB J.* **29**, 3889–3898 (2015).
38. Rountree, J. S. S. *et al.* Usp19 Inhibitors for Use in Therapy. (2020).
39. Rountree, J. S. S. *et al.* 4-hydroxypiperidine derivatives and their use as inhibitors of ubiquitin specific protease 19 (usp19). (2019).
40. Fu, Z. *et al.* The complex structure of GRL0617 and SARS-CoV-2 PLpro reveals a hot spot for antiviral drug discovery. *Nat. Commun.* **12**, 488 (2021).
41. Ritorto, M. S. *et al.* Screening of DUB activity and specificity by MALDI-TOF mass spectrometry. *Nat. Commun.* **5**, 4763 (2014).

42. Kooij, R. *et al.* Small-Molecule Activity-Based Probe for Monitoring Ubiquitin C-Terminal Hydrolase L1 (UCHL1) Activity in Live Cells and Zebrafish Embryos. *J. Am. Chem. Soc.* **142**, 16825–16841 (2020).
43. Panyain, N. *et al.* Discovery of a Potent and Selective Covalent Inhibitor and Activity-Based Probe for the Deubiquitylating Enzyme UCHL1, with Antifibrotic Activity. *J. Am. Chem. Soc.* **142**, 12020–12026 (2020).
44. Clancy, A. *et al.* The deubiquitylase USP9X controls ribosomal stalling. *J. Cell Biol.* **220**, (2021).
45. Borodovsky, A. *et al.* Small-Molecule Inhibitors and Probes for Ubiquitin- and Ubiquitin-Like-Specific Proteases. *ChemBioChem* **6**, 287–291 (2005).
46. Lawson, A. P. *et al.* Identification of deubiquitinase targets of isothiocyanates using SILAC-assisted quantitative mass spectrometry. *Oncotarget* **8**, 51296–51316 (2017).
47. 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).
48. Kwok, B. H. B., Koh, B., Ndubuisi, M. I., Elofsson, M. & Crews, C. M. The anti-inflammatory natural product parthenolide from the medicinal herb Feverfew directly binds to and inhibits IκB kinase. *Chem. Biol.* **8**, 759–766 (2001).
49. Berdan, C. A. *et al.* Parthenolide Covalently Targets and Inhibits Focal Adhesion Kinase in Breast Cancer Cells. *Cell Chem. Biol.* **26**, 1027-1035.e22 (2019).
50. Greenbaum, D., Medzihradszky, K. F., Burlingame, A. & Bogyo, M. Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. *Chem Biol* **7**, 569–81 (2000).

51. Liu, Y., Patricelli, M. P. & Cravatt, B. F. Activity-based protein profiling: the serine hydrolases. *Proc Natl Acad Sci U A* **96**, 14694–9 (1999).
52. Jessani, N. & Cravatt, B. F. The development and application of methods for activity-based protein profiling. *Curr Opin Chem Biol* **8**, 54–9 (2004).
53. Bachovchin, D. A. *et al.* Superfamily-wide portrait of serine hydrolase inhibition achieved by library-versus-library screening. *Proc. Natl. Acad. Sci.* **107**, 20941–20946 (2010).
54. Weerapana, E. *et al.* Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **468**, 790–795 (2010).
55. Bak, D. W. & Weerapana, E. Interrogation of Functional Mitochondrial Cysteine Residues by Quantitative Mass Spectrometry. *Methods Mol. Biol. Clifton NJ* **1967**, 211–227 (2019).
56. Akter, S. *et al.* Chemical proteomics reveals new targets of cysteine sulfinic acid reductase. *Nat. Chem. Biol.* **14**, 995–1004 (2018).
57. Yang, J., Gupta, V., Carroll, K. S. & Liebler, D. C. Site-specific mapping and quantification of protein S -sulphenylation in cells. *Nat. Commun.* **5**, 1–12 (2014).
58. Yang, J. *et al.* Global, in situ , site-specific analysis of protein S -sulfenylation. *Nat. Protoc.* **10**, 1022–1037 (2015).
59. 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).
60. Marino, S. M. & Gladyshev, V. N. Cysteine Function Governs Its Conservation and Degeneration and Restricts Its Utilization on Protein Surfaces. *J. Mol. Biol.* **404**, 902–916 (2010).

61. Resnick, E. *et al.* Rapid Covalent-Probe Discovery by Electrophile-Fragment Screening. *J. Am. Chem. Soc.* **141**, 8951–8968 (2019).

62. Kuljanin, M. *et al.* Reimagining high-throughput profiling of reactive cysteines for cell-based screening of large electrophile libraries. *Nat. Biotechnol.* 1–12 (2021) doi:10.1038/s41587-020-00778-3.

63. Wright, A. T. & Cravatt, B. F. Chemical Proteomic Probes for Profiling Cytochrome P450 Activities and Drug Interactions In Vivo. *Chem. Biol.* **14**, 1043–1051 (2007).

64. Backus, K. M. *et al.* Proteome-wide covalent ligand discovery in native biological systems. *Nature* **534**, 570–574 (2016).

65. Hacker, S. M. *et al.* Global profiling of lysine reactivity and ligandability in the human proteome. *Nat. Chem.* **9**, 1181–1190 (2017).

66. Browne, C. M. *et al.* A Chemoproteomic Strategy for Direct and Proteome-Wide Covalent Inhibitor Target-Site Identification. *J. Am. Chem. Soc.* (2018) doi:10.1021/jacs.8b07911.

67. van Rooden, E. J. *et al.* Mapping *in vivo* target interaction profiles of covalent inhibitors using chemical proteomics with label-free quantification. *Nat. Protoc.* **13**, 752–767 (2018).

68. Wijeratne, A. *et al.* Chemical Proteomic Characterization of a Covalent KRASG12C Inhibitor. *ACS Med. Chem. Lett.* **9**, 557–562 (2018).

69. Hassiepen, U. *et al.* A sensitive fluorescence intensity assay for deubiquitinating proteases using ubiquitin-rhodamine110-glycine as substrate. *Anal. Biochem.* **371**, 201–207 (2007).

70. Zeng, M. *et al.* Potent and Selective Covalent Quinazoline Inhibitors of KRAS G12C. *Cell Chem. Biol.* **24**, 1005-1016.e3 (2017).

71. Wispelaere, M. de *et al.* Inhibition of Flaviviruses by Targeting a Conserved Pocket on the Viral Envelope Protein. *Cell Chem. Biol.* **25**, 1006-1016.e8 (2018).

72. Xu, G. *et al.* Deconvolution in mass spectrometry based proteomics. *Rapid Commun. Mass Spectrom.* **32**, 763–774 (2018).

73. Amico, V., Foti, S., Saletti, R., Cambria, A. & Petrone, G. Identification of iodination sites in cytochrome c by high-performance liquid chromatography and fast atom bombardment mass spectrometry. *Biomed. Environ. Mass Spectrom.* **16**, 431–437 (1988).

74. Ficarro, S. B. *et al.* Leveraging Gas-Phase Fragmentation Pathways for Improved Identification and Selective Detection of Targets Modified by Covalent Probes. *Anal. Chem.* **88**, 12248–12254 (2016).

75. Ward, J. A. *et al.* Quantitative Chemical Proteomic Profiling of Ubiquitin Specific Proteases in Intact Cancer Cells. *ACS Chem. Biol.* **11**, 3268–3272 (2016).

76. Gehringer, M. & Laufer, S. A. Emerging and Re-Emerging Warheads for Targeted Covalent Inhibitors: Applications in Medicinal Chemistry and Chemical Biology. *J. Med. Chem.* **62**, 5673–5724 (2019).

77. Weikart, N. D., Sommer, S. & Mootz, H. D. Click synthesis of ubiquitin dimer analogs to interrogate linkage-specific UBA domain binding. *Chem. Commun.* **48**, 296–298 (2011).

78. Flierman, D. *et al.* Non-hydrolyzable Diubiquitin Probes Reveal Linkage-Specific Reactivity of Deubiquitylating Enzymes Mediated by S2 Pockets. *Cell Chem. Biol.* **23**, 472–482 (2016).

79. Hu, M. *et al.* Crystal Structure of a UBP-Family Deubiquitinating Enzyme in Isolation and in Complex with Ubiquitin Aldehyde. *Cell* **111**, 1041–1054 (2002).

80. Báez-Santos, Y. M. *et al.* X-ray Structural and Biological Evaluation of a Series of Potent and Highly Selective Inhibitors of Human Coronavirus Papain-like Proteases. *J. Med. Chem.* **57**, 2393–2412 (2014).
81. Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science* **347**, (2015).
82. Agrotis, A. *et al.* Human ATG4 autophagy proteases counteract attachment of ubiquitin-like LC3/GABARAP proteins to other cellular proteins. *J. Biol. Chem.* **294**, 12610–12621 (2019).
83. Xie, X. *et al.* PPPDE1 is a novel deubiquitinase belonging to a cysteine isopeptidase family. *Biochem. Biophys. Res. Commun.* **488**, 291–296 (2017).
84. Liu, J. *et al.* Beclin1 Controls the Levels of p53 by Regulating the Deubiquitination Activity of USP10 and USP13. *Cell* **147**, 223–234 (2011).
85. Boutouja, F., Brinkmeier, R., Mastalski, T., El Magraoui, F. & Platta, H. W. Regulation of the Tumor-Suppressor BECLIN 1 by Distinct Ubiquitination Cascades. *Int. J. Mol. Sci.* **18**, (2017).
86. Tsai, Y. C. *et al.* Deubiquitinating enzyme VCIP135 dictates the duration of botulinum neurotoxin type A intoxication. *Proc. Natl. Acad. Sci.* **114**, E5158–E5166 (2017).
87. Totsukawa, G. *et al.* VCIP135 deubiquitinase and its binding protein, WAC, in p97ATPase-mediated membrane fusion. *EMBO J.* **30**, 3581–3593 (2011).
88. Zhang, X.-Y. *et al.* The putative cancer stem cell marker USP22 is a subunit of the human SAGA complex required for activator-driven transcription and cell cycle progression. *Mol. Cell* **29**, 102–111 (2008).
89. Lee, B.-H. *et al.* Enhancement of Proteasome Activity by a Small-Molecule Inhibitor of Usp14. *Nature* **467**, 179–184 (2010).

90. Varca, A. C. *et al.* Identification and validation of selective deubiquitinase inhibitors. *Cell Chem. Biol.* (2021) doi:10.1016/j.chembiol.2021.05.012.

91. Konradi, A. W. & Lin, T. T.-L. T. Oxadiazole compounds. (2019).

92. Chen, Y.-N. P. *et al.* Allosteric inhibition of SHP2 phosphatase inhibits cancers driven by receptor tyrosine kinases. *Nature* **535**, 148–152 (2016).

93. Hughes, C. S. *et al.* Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol. Syst. Biol.* **10**, 757 (2014).

94. Ficarro, S. B. *et al.* Improved Electrospray Ionization Efficiency Compensates for Diminished Chromatographic Resolution and Enables Proteomics Analysis of Tyrosine Signaling in Embryonic Stem Cells. *Anal. Chem.* **81**, 3440–3447 (2009).

95. Alexander, W. M., Ficarro, S. B., Adelmant, G. & Marto, J. A. multiplierz v2.0: A Python-based ecosystem for shared access and analysis of native mass spectrometry data. *PROTEOMICS* **17**, 1700091 (2017).

96. Askenazi, M., Parikh, J. R. & Marto, J. A. mzAPI: a new strategy for efficiently sharing mass spectrometry data. *Nat. Methods* **6**, 240–241 (2009).

97. Zhang, Z. & Marshall, A. G. A universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra. *J. Am. Soc. Mass Spectrom.* **9**, 225–233 (1998).

98. Ficarro, S. B., Alexander, W. M. & Marto, J. A. mzStudio: A Dynamic Digital Canvas for User-Driven Interrogation of Mass Spectrometry Data. *Proteomes* **5**, 20 (2017).