

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

### Title

Chemoproteomics-enabled Discovery of Covalent Inhibitors and Novel Induced Proximity Platforms

### Permalink

<https://escholarship.org/uc/item/0f28z56z>

### Author

Boike, Lydia

### Publication Date

2022

### Supplemental Material

<https://escholarship.org/uc/item/0f28z56z#supplemental>

Peer reviewed|Thesis/dissertation

Chemoproteomics-enabled Discovery of Covalent Inhibitors and Novel Induced  
Proximity Platforms

By

Lydia Boike

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

In

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Daniel K. Nomura, Chair

Professor Evan Miller

Professor James A. Olzmann

Spring 2022



## ABSTRACT

### Chemoproteomics-enabled Discovery of Covalent Inhibitors and Novel Induced Proximity Platforms

By

Lydia Boike

Doctor of Philosophy in Chemistry  
University of California, Berkeley

Under Supervision of Professor Daniel K. Nomura, Chair

Covalent drugs incorporate a mildly electrophilic functional group that reacts with protein targets to confer additional affinity beyond the non-covalent interactions involved in drug binding. In the past, concerns about these reactive molecules' interference with biological assays and lack of selectivity often discouraged further investigation. However, the emergence of intentionally designed covalent drugs against major disease targets over the last ten years showcases the strengths of this field. The first part of this thesis will review the historical and modern milestones in covalent drug discovery with emphasis on chemoproteomics techniques that enable success, and the remaining chapters will address my contributions to the field through the development of tool covalent compounds as novel inhibitors and new induced proximity platforms.

Chemoproteomics-enabled covalent ligand screening platforms can be used to identify novel ligands against undruggable protein targets like MYC. MYC is a major oncogenic transcriptional driver of most human cancers that has remained intractable to direct targeting because much of MYC is intrinsically disordered. I performed a cysteine-reactive covalent ligand screen to identify compounds that could disrupt the binding of MYC to its DNA consensus sequence *in vitro* and also impair MYC transcriptional activity *in situ* in cells. I identified a covalent ligand EN4 that targets cysteine 171 (C171) of MYC within a predicted intrinsically disordered region of the protein. I show that EN4 directly targets MYC in cells, reduces MYC and MAX thermal stability, inhibits MYC transcriptional activity, downregulates multiple MYC transcriptional targets, and impairs tumorigenesis. I also show initial structure-activity relationships of EN4 and identify compounds that show improved potency. Overall, I identified a novel ligandable site within an intrinsically disordered region of MYC that leads to inhibition of MYC transcriptional activity.

In addition to facilitating the discovery of covalent small-molecule allosteric modulators of disease relevant proteins and pathways, chemoproteomics-enabled covalent drug discovery also proves useful in identifying inhibitors of critical disease proteins that contain catalytic cysteines. Among the various genes and proteins encoded by all coronaviruses, one particularly "druggable" or relatively easy-to-drug target is the coronavirus Main Protease (3CLpro or Mpro), a cysteine-protease that is

involved in cleaving a long peptide translated by the viral genome into its individual protein components that are then assembled into the virus to enable viral replication in the cell. Inhibiting Mpro's catalytic cysteine with a small-molecule antiviral would effectively stop the ability of the virus to replicate, providing therapeutic benefit across coronaviruses. As part of a collaborative effort, I utilized activity-based protein profiling (ABPP)- chemoproteomic approaches to discover and further optimize cysteine-reactive pyrazoline-based covalent inhibitors for the SARS-CoV-2 Mpro. Structure-guided medicinal chemistry and modular synthesis of di- and tri-substituted pyrazolines bearing either chloroacetamide or vinyl sulfonamide cysteine-reactive warheads enabled the expedient exploration of structure-activity relationships (SAR), yielding nanomolar potency inhibitors against Mpro from not only SARS-CoV-2, but across many previous coronaviruses. Our studies highlight promising chemical scaffolds that may contribute to future pan-coronavirus inhibitors.

Finally, chemoproteomics-enabled drug discovery platforms facilitate the expansion of targeted protein degradation and related fields. Many diseases are driven by proteins that are aberrantly ubiquitinated and degraded. These diseases would be therapeutically benefited by targeted protein stabilization (TPS). We designed deubiquitinase-targeting chimeras (DUBTACs), heterobifunctional small molecules consisting of a deubiquitinase recruiter linked to a protein-targeting ligand, to stabilize the levels of specific proteins degraded in a ubiquitin-dependent manner. Using chemoproteomic approaches, we discovered the covalent ligand EN523 that targets a non-catalytic allosteric cysteine C23 in the K48-ubiquitin-specific deubiquitinase OTUB1. We showed that a DUBTAC consisting of our EN523 OTUB1 covalent recruiter linked to lumacaftor, a drug used to treat cystic fibrosis that binds  $\Delta F508$ -cystic fibrosis transmembrane conductance regulator (CFTR), robustly stabilized  $\Delta F508$ -CFTR protein levels, leading to improved chloride channel conductance in human cystic fibrosis bronchial epithelial cells. We also demonstrated stabilization of the tumor suppressor kinase WEE1 in hepatoma cells. Our study showcases covalent chemoproteomic approaches to develop new induced proximity-based therapeutic modalities and introduces the DUBTAC platform for TPS.

## TABLE OF CONTENTS

DEDICATION.....	iii
ACKNOWLEDGEMENTS.....	iv
LIST OF FIGURES.....	v
LIST OF APPENDICES.....	vii
LIST OF ABBREVIATIONS.....	viii

### CHAPTER 1

Advancements in Covalent Drug Discovery.....	1
1.1 History of Covalent Drugs.....	2
1.2 Targeted Covalent Inhibitors.....	3
1.2.1 Covalent EGFR Inhibitors.....	3
1.2.2 Covalent BTK Inhibitors.....	5
1.2.3 Other Covalent Kinase Inhibitors.....	7
1.2.4 Covalent KRAS-G12C Inhibitors.....	7
1.3 Covalent Protease Inhibitors.....	9
1.3.1 SARS-CoV2 Main Protease Inhibitors.....	9
1.3.2 HCV NS3/4a Protease Inhibitors.....	12
1.3.3 Covalent Proteasome Inhibitors.....	12
1.4 The Covalent Drug Discovery Toolbox.....	12
1.4.1 Chemoproteomics-Enabled Discovery.....	12
1.4.1.1 Covalent Ligand Discovery using ABPP.....	13
1.4.1.2 ABPP-Enabled Induced Proximity: Targeted Protein Degradation and Beyond.....	15
1.4.2 Screening Platforms beyond ABPP.....	15
1.4.2.1 Intact Protein-MS for Covalent Ligand Screening.....	16
1.4.2.2 Covalent DNA Encoded Libraries.....	17
1.4.2.3 Covalent Docking.....	17
1.4.3 Lysine-directed Covalent Ligands.....	18
1.5 Conclusion.....	19
1.6 Figures and Tables.....	20

### CHAPTER 2

Discovery of a Functional Covalent Ligand Targeting an Intrinsically Disordered Cysteine within Myc.....	27
2.1 Covalent Ligand Screening to Identify MYC Inhibitors.....	28
2.2 Characterization of EN4 Interactions with c-MYC <i>In Vitro</i> .....	29
2.3 Characterization of EN4 Engagement of MYC in Cells.....	30
2.4 Characterization of EN4 Inhibitory Activity against MYC in Cells.....	31
2.5 Structure-Activity Relationship of EN4 Analogs.....	32
2.6 Discussion.....	32
2.7 Figures and Tables.....	35

### CHAPTER 3

Development of Potent Pyrazoline-Based Covalent Main Protease Inhibitors.....	54
---	----

3.1 Covalent Ligand Screening to Discover Pyrazoline-based SARS-CoV-2 Main Protease Inhibitors.....	55
3.1.1 Covalent Screening and Main Protease Activity Assay Development.....	56
3.1.2 Assessing Proteome-Wide Cysteine-Reactivity of Hit Compound EN82...	56
3.2 Exploration of 3,5-Disubstituted Pyrazoline SAR.....	57
3.3 Exploration of Trisubstituted Pyrazoline Inhibitors.....	57
3.5 Exploration of Cysteine Reactive Warheads.....	58
3.6 Testing Lead Compounds Against Coronavirus Panel.....	58
3.7 Figures and Tables.....	60
<b>CHAPTER 4</b>	
Deubiquitinase-targeting Chimeras for Targeted Protein Stabilization.....	74
4.1 Identifying Allosteric Ligandable Sites within DUBs.....	75
4.2 Discovery of a Covalent Recruiter Against OTUB1.....	77
4.3 Showing Proof of Concept of DUBTACs with Mutant CFTR.....	79
4.4 Using DUBTACs to Stabilize WEE1.....	82
4.5 Discussion.....	82
4.6 Figures and Tables.....	84
<b>CHAPTER 5</b>	
Concluding Remarks.....	108
<b>REFERENCES.....</b>	<b>110</b>

## DEDICATION

For Trisha Morrison and Jack Crane.

Dear Trisha,

The support and love that you have given me throughout our long friendship consistently builds me up and makes me feel like I can accomplish my dreams. At the toughest moments during my PhD you were there for me. You helped me practice my qualifying exam, worked alongside me to keep me company while I put together supporting info for papers, and always invited me out to all the fun things in San Francisco even when I was at my busiest. Thank you for your unconditional understanding, support, and love. You mean the world to me.

Dear Jack,

Experiencing life with you is the best gift I will ever receive. The ways in which you take care of me daily enable me to thrive. You give me your undivided and most thoughtful attention, and how you live your life inspires me to push the limits of what I think is possible to achieve. Thank you for being my partner. I love you and am excited to grow with you forever.



## ACKNOWLEDGEMENTS

My PhD was truly a collaborative effort. I would like to thank everyone who has worked with me, inspired me, and been a friend to me while completing it. I would like to acknowledge a few specific people by name:

Dan Nomura,

Thank you for being an advisor who always strives to solve challenging problems and expands the limits of translational chemical biology. I'm grateful to have trained under you, and thankful for all the guidance you have given me.

Nate Henning,

Working with you has been the best part of my PhD. I am so grateful for the mentorship you have given me in synthetic chemistry; I learned so much from you. Your friendship is invaluable, and I am so excited to continue working with you.

Jen Co, Felix Majewski, Katelyn Randal, and Anoohya Panidapu,

Thank you all so much for the effort you put into all our projects. I've enjoyed mentoring all of you and I'm excited to see all the amazing things you will do in the future.

I also would like to acknowledge all the coauthors on the papers that I am adapting in my thesis:

Dan Nomura, Nate Henning, Jessica Spradlin, Scott Brittain, Carl Ward, Feng Wang, Dustin Dovala, Daniel Fuller, Jeffrey McKenna, John Tallarico, Markus Schirle, Gang Liu, Erika Zhang, Alexander Cioffi, Felix Majewski, Jennifer Co, Michael Jones, Patrick Moon, Mark Knapp, Helene Wolleb, Charlotte Zammit, Gabrielle Blake, Yipin Lu, Dean Toste.

## LIST OF FIGURES AND TABLES

### CHAPTER 1

<b>Figure 1</b> Timeline of the development of major covalent drugs.....	20
<b>Table 1</b> Highlighted covalent kinase inhibitors.....	21
<b>Figure 2</b> Aligned structures (PDB: 6ut0 and 6oim) of KRAS <sup>G12C</sup> co-crystalized with MRTX-849 (adagrasib) and AMG-510 (sotorasib).....	22
<b>Figure 3</b> Nirmatrelvir in complex with SARS-CoV-2 Mpro.....	23
<b>Table 2</b> Highlighted covalent drugs.....	24
<b>Figure 3</b> isotopic tandem orthogonal proteolysis–activity-based protein profiling.....	25
<b>Figure 5</b> Screening methods for covalent drug discovery.....	26

### CHAPTER 2

<b>Figure 1</b> Covalent ligand screen to identify compounds that inhibit MYC activity <i>in vitro</i> and <i>in situ</i> .....	35
<b>Figure 2</b> Characterization of EN4 binding to MYC.....	37
<b>Figure 3</b> Characterizing EN4 engagement of MYC in cells.....	39
<b>Figure 4</b> Characterizing EN4 inhibitory activity against MYC in cells.....	41
<b>Table S1</b> Covalent ligand structures and screening data for data shown in Figure 1a, 1b.....	43
<b>Table S2</b> isoTOP-ABPP analysis of EN4 in 231MFP breast cancer cells shown in Figure S3.....	43
<b>Table S3</b> RNA sequencing analysis of EN4-mediated transcriptional changes in 231MFP cells. 231MFP cells were treated with DMSO vehicle or EN4 (50 mM) for 24 h, n=3 biological replicates/group.....	43
<b>Table S4</b> Structure-Activity Relationships of EN4 Analogs.....	44
<b>Figure S1</b> Characterization of EN4 inhibition of c-MYC/MAX DNA binding and interaction of EN4 with c-MYC.....	46
<b>Figure S2</b> Sequence alignment of human JUN, c-MYC, N-MYC and L-MYC.....	48
<b>Figure S3</b> Characterization of EN4 in cells.....	51
<b>Figure S4</b> Characterization of EN4-18 in DNA binding assays.....	53

### CHAPTER 3

<b>Figure 1</b> Discovery of pyrazoline-based SARS-CoV-2 main protease inhibitors.....	60
<b>Figure 2</b> Proteome-wide cysteine-reactivity of EN82 in HEK293T cell lysate with spiked-in SARS-CoV-2 Mpro.....	62
<b>Figure 3</b> Exploration of 3,5-disubstituted pyrazoline SAR.....	63
<b>Figure 4</b> Exploration of trisubstituted pyrazoline inhibitors.....	64
<b>Figure 5</b> Exploration of Cysteine Reactive Warheads.....	66
<b>Figure 6</b> Exploration of trisubstituted pyrazoline inhibitors with a vinyl sulfonamide cysteine-reactive warhead.....	67
<b>Figure 7</b> Most promising hit compound and corresponding IC <sub>50</sub> values.....	68
<b>Figure S1</b> Gel-based ABPP screen of cysteine-reactive covalent ligands against Mpro enzyme.....	69
<b>Figure S2</b> Hit confirmation of initial gel-based ABPP screen.....	71
<b>Figure S3</b> Testing reproducible hit compounds in dose-response gel-based ABPP and Mpro substrate activity assays.....	72

<b>Figure S4</b> Potency (IC <sub>50</sub> , mM) of inhibitors against MPro enzymes from other coronaviruses.....	73
---	----

CHAPTER 4

<b>Figure 1</b> DUBTAC platform.....	84
<b>Figure 2</b> Discovery of covalent ligands that target OTUB1.....	86
<b>Figure 3</b> DUBTAC against mutant CFTR.....	88
<b>Figure 4</b> Characterizing the mechanism of the CFTR DUBTAC NJH-2-057.....	90
<b>Figure 5</b> WEE1 DUBTAC.....	92
<b>Extended Data Figure 1</b> Primary covalent ligand screen against OTUB1.....	93
<b>Extended Data Figure 2</b> NMR analysis of OTUB1, EN523, and UBE2D2.....	95
<b>Extended Data Figure 3</b> Structure-activity relationships of EN523 analogs with.....	97
<b>Extended Data Figure 4</b> EN523 does not alter OTUB1 levels and NJH-2-075 engages OTUB1 in CFBE41o-4.7 cells expressing $\Delta$ F508-CFTR.....	99
<b>Extended Data Figure 5</b> Effect of DUBTACs on mutant CFTR levels.....	100
<b>Extended Data Figure 6</b> Effect of DUBTACs on mutant CFTR levels.....	101
<b>Extended Data Figure 7</b> Effect of DUBTACs on mutant CFTR levels in siControl and siCFTR cells.....	102
<b>Extended Data Figure 8</b> Linker dependence of CFTR DUBTACs.....	103
<b>Extended Data Figure 9</b> CFTR DUBTACs with rigid linkers.....	104
<b>Extended Data Figure 10</b> Negative control DUBTACs.....	106
<b>Table S1</b> Chemoproteomic analysis of DUBs.....	107
<b>Table S2</b> Structures of covalent ligands screened against OTUB1.....	107
<b>Table S3</b> IsoTOP-ABPP analysis of NJH-2-057.....	107
<b>Table S4</b> TMT-based quantitative proteomic profiling of NJH-2-057 treatment.....	107

## LIST OF APPENDICES

APPENDIX 1	
CHAPTER 2 methods.....	131
APPENDIX 2	
CHAPTER 3 methods.....	155
APPENDIX 3	
CHAPTER 4 methods.....	203

## LIST OF ABBREVIATIONS

(As they appear, in order, in main text)

### CHAPTER 1

EGFR	Epidermal growth factor receptor
KRAS <sup>G12C</sup>	KRAS glycine to cysteine mutation at position 12
SARS-CoV-2 Mpro	Severe acute respiratory syndrome coronavirus 2 main protease
BTK	Bruton tyrosine kinase
NSAID	Non-steroidal anti-inflammatory drug
COX-1	Cyclooxygenase 1
PBPs	Penicillin-binding proteins
MurA	UDP-N-acetylglucosamine enolpyruvyl transferase
P450	Cytochrome P450 (enzyme superfamily)
P2Y12	Purinergic receptor P2Y12
FDA	Food and drug administration
NSCLC	Non-small-cell lung carcinoma
L858R	Lysine to arginine mutant at position 858
T790M	Threonine to methionine mutant at position 790
C797	Cysteine at position 797
Her2	Human epidermal growth factor receptor 2
TKI	Targeted kinase inhibitor
C797X	Cysteine point mutation at 797
C797S	Cysteine to serine mutant at position 797
Lyn	Lyn proto-oncogene
Syk	Spleen tyrosine kinase
XLA	Bruton type agammaglobulinemia or X-linked agammaglobulinemia
C481	Cysteine at position 481
ATP	Adenosine triphosphate
CLL	Chronic lymphocytic leukemia
MCL	Mantle cell lymphoma
GSH	Glutathione
LCMS	Liquid chromatography mass spectrometry
PK/PD	Pharmacokinetics/Pharmacodynamics
JAK3	Janus kinase 3
Cys909	Cysteine at position 909
FGFR4	Fibroblast growth factor receptor 4
HCC	Hepatocellular cancer
Cys552	Cysteine at position 552
FGFR	Fibroblast growth factor receptor
CDK	Cyclin dependent kinase
CDK7	Cyclin dependent kinase 7
KRAS	Kirstin rat sarcoma virus oncogene
GTP	Guanosine triphosphate
GDP	Guanosine diphosphate
UCSF	University of California San Francisco

S-IIP	Switch II binding pocket
H95	Histidine at position 95
Y96	Tyrosine at position 96
Q99	Glutamine at position 99
COVID-19	Coronavirus disease 2019
EUA	Emergency use authorization
RNA	Ribonucleic acid
3CLpro	3-chymotrypsin like protease
Cys145	Cysteine at position 145
His41	Histidine at position 41
TGEV	Transmissible gastroenteritis virus
CMK	Chloromethyl ketone
HRV	Human rhinovirus
FRET	Fluorescence resonance energy transfer
hACE2	Human angiotensin-converting enzyme 2
CYP3A	Cytochrome P450, family 3, subfamily A
NS3/4a	Nonstructural protein 3/4a serine protease
HCV	Hepatitis C virus
NS5a	Nonstructural protein 5a
NS5b	Nonstructural protein 5b
ABPP	Activity-based protein profiling
E3	Ubiquitin ligase
isoTOP-ABPP	Isotopic tandem orthogonal proteolysis-activity-based protein profiling
TEV	Tobacco etch virus
PTM	Post-translational modification
IA-alkyne	Iodoacetamide-alkyne
CASP8	Caspase 8
CASP10	Caspase 10
KEAP1	Kelch-like ECH-associated protein 1
NRF2	Nuclear factor-erythroid factor 2-related factor 2
NR0B1	Nuclear receptor subfamily 0 group B member 1
BIRC2/3	Baculoviral IAP repeat containing 2/3
NuRD	Nucleosome remodeling and deacetylase complex
ITK	Tyrosine-protein kinase
CYTIP	Cytohesin 1 interacting protein
DMF	Dimethyl fumarate
IRAK4	Interleukin 1 receptor associated kinase 4
Myd88	MYD88 innate immune signal transduction adaptor
ATP6V1A	ATPase H <sup>+</sup> transporting V1 subunit A
TDP-43	TAR DNA binding protein 43
PDAC	Pancreatic ductal adenocarcinoma
ABHD6	Monoacylglycerol lipase ABHD6
MYC	Myc proto-oncogene
SP3	Sample preparation steps
FAIMS	High field asymmetric waveform ion mobility spectrometry

TMT	Tandem mass tag
SLC-ABPP	Streamlined cysteine-activity based protein profiling
DUB	Deubiquitinase
PROTAC	Proteolysis targeting chimera
CRBN	Cereblon
VHL	Von Hippel-Lindau Tumor Suppressor
RNF114	Ring finger protein 114
RNF4	Ring finger protein 4
DCAF16	DDB1 and CUL4 associated factor 16
DCAF11	DDB1 and CUL4 associated factor 11
FKBP12	12-kDa FK506-binding protein
BRD4	Bromodomain-containing protein 4
DUBTAC	Deubiquitinase targeting chimera
OTUB1	OTU domain-containing ubiquitin aldehyde-binding protein 1
CFTR	Cystic fibrosis transmembrane conductance regulator
WEE1	Mitosis inhibitor protein kinase Wee1
TPD	Targeted protein stabilization
HECT	Homologous to the E6-AP carboxyl terminus
Nedd4-1	Neural precursor cell expressed developmentally down-regulated protein 4
RBR	RING-in-between-RING
HOIP	HOIL-1L interacting protein
NUDT7	Nudix hydrolase 7
OTUB2	OTU domain-containing ubiquitin aldehyde-binding protein 2
FBDD	Fragment based drug discovery
Pin1	Peptidyl-prolyl cis/trans isomerase, NIMA-interacting 1
CITe-Id	Covalent Inhibitor Target-site Identification
DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
DEL	DNA encoded library
JNK1	c-Jun N-terminal kinase 1
MEK2	Mitogen-activated protein kinase kinase 2
MAP2K6	Mitogen-activated protein kinase kinase 6
DOCKoValent	Covalent virtual screening software
MKK7	Mitogen-activated protein kinase kinase 7
NHS ester	N-hydroxysuccinimide ester
HbS	Sickle hemoglobin
a-Hb	Alpha hemoglobin
pKa	Acid dissociation constant
ADME	Absorption, Distribution, Metabolism and Excretion
PI3Kdelta	Phosphoinositide 3-kinase
CDK2	Cyclin dependent kinase 2
Hsp90	Heat shock protein 90
mRNA	Messenger RNA
DcpS	Scavenger mRNA-decapping enzyme

## CHAPTER 2

MYC or c-MYC	Myc is a family of regulator genes and proto-oncogenes that code for transcription factors
MAX	Myc associated protein X
E-box	Enhancer box
BET	Bromodomain and extraterminal protein family
CDK7/9	Cyclin dependent kinase 7/9
mTORC1	Mechanistic target of rapamycin complex 1
AKT/PI3K	Protein kinase B/Phosphoinositide 3-kinase
BRD4	Bromodomain-containing 4
JQ1	Inhibitor of BET family of bromodomain proteins
GSK525762	BET inhibitor
4EBP1	Translation initiation factor 4E binding protein 1
ABPP	Activity-based protein profiling
MS	Mass spectrometry
HEK293T	A derivative human cell line that expresses a mutant version of the SV40 large T antigen
TATA box	Sequence of DNA found in the core promoter region of genes in archaea and eukaryotes
DNA	Deoxyribonucleic acid
LC-MS/MS	Liquid chromatography-tandem MS
C171	Cysteine at position 171
NTD	N-terminal domain
CuAAC	Copper-catalyzed azide-alkyne cycloaddition
C171Y	Cysteine to tyrosine mutant at position 171
N-MYC	N-myc proto-oncogene protein
L-MYC	L-myc proto-oncogene protein
WT	Wild type
231MFP	Triple-negative mammalian breast cancer cell line
isoTOP-ABPP	Isotopic tandem orthogonal proteolysis-enabled activity-based protein profiling
JUN	Jun proto-oncogene
C99	Cysteine at position 99
FLAG	DYKDDDDK-tag
MDA-MD-231	Triple-negative mammalian breast cancer cell line
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
siRNA	Small interfering RNA
ChIP-seq	Chromatin immunoprecipitation sequencing
E2F	E2 Factor
G2M	G2-M DNA damage checkpoint
CDK2	Cyclin dependent kinase 2
CDC25A	Cell division cycle 25A (M-phase inducer phosphatase 1)
MIZ1 or ZBTB17	Zinc finger and BTB domain containing 17
MCF10A	Non-tumorigenic epithelial mammalian cell line
RNF114	Ring-finger protein 114



### CHAPTER 3

SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
Mpro or 3CL <sup>pro</sup>	Main protease or 3C-like protease
ABPP	Activity-based protein profiling
SAR	Structure-activity relationships
COVID-19	Coronavirus disease 2019
pp1a and pp1ab	Replicase polyproteins 1a and 1ab
C145	Cysteine at position 145
IA	Iodoacetamide
FRET	Fluorescence resonance energy transfer
MS-ABPP	Mass spectrometry-based activity-based protein profiling
HEK293T	A derivative human cell line that expresses a mutant version of the SV40 large T antigen
HMOX2	Heme oxygenase 2
C282	Cysteine at position 282
IC50	Half-maximal inhibitory concentration
SARS-CoV-1	Severe acute respiratory syndrome coronavirus 1
HCoV-HKU1	Human coronavirus HKU1
HCoV-229E	Human coronavirus 229E
MERS-CoV	Middle East respiratory syndrome-related coronavirus
PorCoV-HKU15	Porcine coronavirus HKU15
HCoV-NL63	Human coronavirus NL63
IBV	Infectious bronchitis virus
HCoV-OC43	Human coronavirus OC43

### CHAPTER 4

TPD	Targeted protein degradation
PROTACs	Proteolysis-targeting chimeras
E3	E3 ubiquitin ligases
TP53	Tumor protein 53
CDKN1A	Cyclin dependent kinase inhibitor 1A
CKN1C	Cyclin dependent kinase inhibitor 1C
BAX	Bcl-2-associated X protein
ΔF508	Deletion mutation caused by frameshift mutation at codon 508 resulting in absent phenylalanine residue
CFTR	Cystic fibrosis transmembrane conductance regulator
TPS	Targeted protein stabilization
DUBTAC	Deubiquitinase-targeting chimera
DUB	Deubiquitinase
OTUB1	OUT domain-containing ubiquitin aldehyde-binding protein 1
K48	Lysine at position 48
ABPP	Activity-based protein profiling
isoTOP-ABPP	Isotopic tandem orthogonal proteolysis-ABPP

IA	Iodoacetamide
OTU	Ovarian tumor proteases
C23	Cysteine at position 23
C91	Cysteine at position 99
C212	Cysteine at position 212
PONDR	Predictor of natural disordered regions
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
UBE2D1	Ubiquitin-conjugating enzyme E2 D1
HMQC	Heteronuclear multiple quantum coherence
SAR	Structure-activity relationships
RNF5	Ring finger protein 5
CFBE41o-4.7	Wt-CFTR human CF bronchial epithelial cell line
K63	Lysine at position 63
VDAC2 C76	Voltage dependent anion channel 2, cysteine at position 76
TUBB1 C201	Tubulin beta 1 class VI, cysteine at position 201
RLF C744	Rearranged L-myc fusion zinc finger protein, cysteine at position 744
VDAC2 C47	Voltage dependent anion channel 2, cysteine at position 47
VCAC3 C66	Voltage dependent anion channel 3, cysteine at position 66
Nedd4	Neural precursor cell expressed developmentally down-regulated protein 4
VX770	Vertex's CFTR potentiator, Ivacaftor
WEE1	Nuclear kinase belonging to the Ser/Thr family of protein kinases
CDK1	Cyclin dependent kinase 1
AZD1775	AstraZeneca's Wee1 kinase inhibitor
HEP3B	Hepatoma cancer cell line
PEG	Polyethylene glycol
AlphaLISA	Bead-based luminescent amplification assay

## CHAPTER 1

### Advancements in Covalent Drug Discovery

(Adapted from a currently in peer review Nature Reviews Drug Discovery article written by Lydia Boike, Nate Henning, and Dan Nomura)

## Advancements in Covalent Drug Discovery

The past decade has witnessed several remarkable shifts in covalent drug discovery. Many historical covalent drugs were discovered serendipitously and bind active sites to inhibit enzymatic activity.<sup>1,2,3,4</sup> These drugs often mimic a substrate transition state that allows covalent modification of a catalytic amino acid residue. Over the past twenty years, covalently targeting non-conserved amino acids to improve selectivity has become commonplace. Covalent drugs' prolonged target engagement also provides distinct pharmacodynamic profiles and exceptional potency.<sup>5</sup>

These benefits of potency and selectivity have inspired medicinal chemists to explore covalent drug space despite concerns about reactivity. In many cases, compromises between these properties have produced safe and effective drugs. Examples we will discuss here include the BTK inhibitor ibrutinib (Abbvie) and EGFR inhibitor 2simertinib (AstraZeneca), with \$8.43 billion and \$4.33 billion in 2020 sales, respectively.<sup>6,7</sup> Potent inhibition through covalent modification has enabled targeting traditionally 'undruggable' proteins, such as the mutant KRAS<sup>G12C</sup>, a GTPase that has resisted decades of drug discovery efforts.<sup>8,9</sup> At the same time, more traditional covalent targeting of protease active sites has continued to yield valuable drugs, such as nirmatrelvir (Pfizer; Paxlovid), which inhibits the SARS-CoV-2 main protease (Mpro) (**Fig. 1**).<sup>10</sup>

While targeted covalent inhibitors are frequently discovered through structure-guided design based initially on a reversible ligand, screening libraries of electrophilic compounds to discover covalent ligands has become more common.<sup>6,11,12</sup> This "electrophile-first" approach is enabled by the development of chemoproteomic platforms that facilitate rapid target identification and selectivity profiling of covalent protein targets.<sup>13-17</sup> Combining the advances in covalent fragment screening and chemoproteomics with structural biology to empower medicinal chemistry has the potential to rapidly generate molecules that selectively bind challenging targets.

In this review we highlight historical examples of covalent drugs and the often-serendipitous discovery of their mechanisms of action. We elaborate on the major milestones in covalent drug discovery over the past decade and describe the role of covalency in these molecules' discovery and activity. Finally, we summarize the powerful techniques, with emphasis on screening strategies and selectivity profiling, that have been developed to enable covalent drug discovery.

### 5.5 History of Covalent Drugs

Compounds containing protein-reactive functional groups have often been avoided in medicinal chemistry and excluded from compound screening collections due to their potential for assay interference and off-target promiscuity. Many historical examples of covalent drugs were discovered to act through covalent mechanisms after their use was already widespread. The most prominent among these is the non-

steroidal 3simertinib3atory drug (NSAID) aspirin which has been marketed since 1899 (**Fig. 1**).<sup>18</sup> Aspirin's mechanism of action was unknown until 1971, when it was discovered to exert its 3simertinib3atory effects by acetylating Ser529 in the substrate binding channel of cyclooxygenase 1 (COX-1), preventing conversion of the substrate arachidonic acid into subsequent prostaglandins.<sup>19</sup>

Early covalent drugs also tend to be derived from or inspired by natural sources. B-lactam antibiotics, like Penicillin, bind to Penicillin-binding proteins (PBPs) which are involved in bacterial cell wall synthesis (**Fig. 1**).<sup>20</sup> All PBPs contain active-site serine residues that can be acylated by penicillin, inhibiting PBP activity and leading to cell membrane rupture.<sup>20</sup> Another covalent antibiotic is the epoxide-containing 3simertini, which acts by reacting with the catalytic cysteine of MurA, also disrupting peptidoglycan synthesis and inducing membrane rupture (**Fig. 1**).<sup>21-23</sup> In a similar manner to beta-lactam antibiotics, the lactone-containing Orlistat reacts with the catalytic serine of gastric and pancreatic lipases to reduce hydrolysis of dietary fat.<sup>24</sup>

Proton pump inhibitors, which address gastrointestinal reflux disease, are prodrugs whose thiol-containing metabolites form disulfide bonds to inactivate their targets.<sup>25</sup> Omeprazole, approved by the US FDA in 1988, is an example of a proton pump inhibitor brought to market before its mechanism of action was understood to be covalent (**Fig. 1**). Clopidogrel, a medication used to prevent stroke and heart attacks, is also activated by cytochrome P450 enzymes in the liver (**Fig. 1**).<sup>26</sup> The resulting thiol-containing metabolite covalently inhibits the purinergic receptor P2Y<sub>12</sub> on the surface of platelets to prevent clotting.

Covalent drugs have also been historically significant in cancer therapy. The pyrimidine nucleoside analog prodrugs 5-fluorouracil<sup>27,28</sup> and gemcitabine<sup>29</sup> are used to inhibit thymidylate synthase and ribonucleotide reductase I, respectively, to treat a wide range of cancers (**Fig. 1**). Bortezomib, a dipeptide boronic acid which covalently binds to and inhibits the catalytic threonine residue of the 26S proteasome, was the first proteasome inhibitor approved by the FDA in 2003 to treat multiple myeloma patients (**Fig. 1**).<sup>30</sup>

## 1.2 Targeted Covalent Inhibitors

As covalent drug discovery has improved, and issues of on-target (and off-target) toxicity have been systematically addressed, major milestones of covalent drug development have been reached. These milestones include the development of covalent EGFR inhibitors like Afatinib (approved in 2013), BTK inhibitors like Ibrutinib (approved in 2013), selective KRAS<sup>G12C</sup> inhibitors like sotorasib (approved in 2021), and the SARS-CoV-2 Mpro inhibitor nirmatrelvir (approved in 2021) (**Fig. 1**). These oncology and antiviral drug programs highlight the progress that covalent drug discovery has made towards developing potent and selective covalent small molecules designed to avoid toxicity issues and balance reactivity with potency and selectivity.

### 1.2.1 Covalent EGFR Inhibitors

Inhibitors of the receptor tyrosine kinase EGFR were originally developed to address EGFR's role in driving of non-small cell lung cancer (NSCLC) through its overexpression.<sup>31</sup> During clinical development in the early 2000s, the first approved EGFR reversible (non-covalent) inhibitors gefitinib and erlotinib were discovered to be effective against EGFR harboring somatic activating mutations, either deletions in exon 19 or the L858R point mutation which occur in 10-30% of patients with NSCLC.<sup>31-33</sup> Unfortunately, patients eventually still progressed in disease; in 60% of cases this was due to acquisition of the T790M "gatekeeper mutation".<sup>34,35</sup> The gate keeper mutation in the ATP binding site of EGFR not only reduces inhibitor binding affinity to EGFR but also increases EGFR's binding affinity to ATP.<sup>36</sup>

To overcome this problem, second generation covalent inhibitors were designed with acrylamide Michael acceptors to covalently engage EGFR's C797, which is located adjacent to the ATP binding site.<sup>33</sup> In addition to modest activity against T790M, the more durable suppression of EGFR signaling through covalent inhibition and activity against Her2 (an EGFR dimerization partner) suggested that second generation covalent EGFR inhibitors might be more efficacious than first generation reversible inhibitors like erlotinib.<sup>33</sup> Afatinib (Boehringer Ingelheim) was approved in 2013 as a first-line treatment for patients with metastatic NSCLC with activating mutations in EGFR (**Fig. 1**).<sup>37,38</sup> Afatinib showed improvement over first-generation erlotinib and gefitinib, showing increased potency and prolonged target engagement due to its covalent mechanism. Despite the increased potency that covalent engagement brought against the disease target, dose-limiting toxicity likely due to wild type (wt) EGFR inhibition prevented afatinib from being as effective as platinum-based chemotherapy in treating cancers bearing the T790M gatekeeper mutation. Other second generation inhibitors include neratinib, which was approved for HER2 positive breast cancer in 2017, as it potently inhibits HER2 through covalent binding to the homologous C805 and dacomitinib (Pfizer), which was approved to treat NSCLC in 2018.<sup>39-41</sup>

A new third generation of covalent EGFR inhibitors followed afatinib, with the aim to selectively target for the T790M mutation over wt-EGFR (WZ-4001 (Dana-Farber),<sup>42</sup> Osimertinib (AstraZeneca),<sup>43,44</sup> and rociletinib (Clovis, CO-1686) (**Table 1**).<sup>45</sup> These compounds maintain the acrylamide to covalently bind C797 but exchange the quinazoline of first generation compounds for a pyrimidine to better accommodate the T790M mutation.<sup>46</sup> Higher affinity for T790M over wt-EGFR not only enabled efficacy in cancers with the EGFR gatekeeper mutation but also contributed to an improved safety profile and higher recommended dose than afatinib.<sup>47</sup> Osimertinib was granted accelerated approval by the FDA in 2015 as a second-line treatment for NSCLC and was superior to platinum-based chemotherapy in patients who had progressed during first-line EGFR TKI therapy.<sup>48</sup> It was then approved as a first-line treatment for metastatic NSCLC in 2018. Osimertinib does depend on C797 for covalent binding, and C797X mutations account for 15% of cases of resistance to second-line osimertinib.<sup>49-51</sup> Therapeutic options are limited for the resulting triple mutant L858R/T790M/C797S, though non-covalent fourth generation EGFR inhibitors (such as EAI045) have been discovered that bind reversibly to an allosteric pocket created by the displacement of

the C-helix in the inactive conformation of EGFR.<sup>52</sup> Allosteric inhibitors could also potentially be used in combination with covalent orthosteric inhibitors such as 5simertinib, though these allosteric compounds have yet to enter clinical trials.<sup>53</sup> In the meantime, 5simertinib remains the standard of care for many types of NSCLC and demonstrates the power of covalent drugs to maintain potency despite mutations around the binding site.

The success of covalent EGFR drugs validated the approach of engaging non-catalytic, non-conserved cysteines adjacent to kinase active sites in order to boost potency and modulate pharmacodynamics. In particular, development of these drugs showed that the acrylamide electrophile was reactive enough to engage a cysteine adjacent to a compound binding site but not so reactive as to induce hapteneization and adverse immune response. Incorporating covalent binding in EGFR inhibitors also enabled selectivity among kinases by designing interactions with a non-conserved cysteine rather than active-site residues that typically interact with ATP.

### 1.2.2 Covalent BTK Inhibitors

Bruton's tyrosine kinase became a target of interest in chronic lymphocytic leukemia due to its critical role downstream of the B-cell receptor.<sup>54</sup> Activation of the B-cell receptor induces phosphorylation of BTK through Lyn and Syk kinases, and eventually activates transcription factors related to B-cell proliferation and differentiation, as well as cell migration and adhesion.<sup>55</sup> Additionally, BTK is mutated in X-linked agammaglobulinemia (XLA) where loss-of-function mutations result in B-cell deficiency.<sup>56</sup> This critical role in B-cell development indicated it was a relevant target for B-cell malignancies.

In the early 2000s, scientists at Celera Genomics used a structure-based approach to discover an acrylamide-containing inhibitor of the BTK kinase domain that could be used as a tool compound to fluorescently label BTK, and were interested in using BTK inhibitors to treat rheumatoid arthritis.<sup>57</sup> After acquisition by Pharmacyclics, it was discovered that the tool compound itself, later named ibrutinib, had sufficient affinity and selectivity for BTK to advance into clinical studies.<sup>58,59</sup> Ibrutinib (Abbvie) was approved by the FDA for the treatment of chronic lymphocytic leukemia (CLL) in 2013 following trials showing remarkable overall response rate of 70%, and progression free survival of 75% after 26 months (**Table 1**).<sup>60,61</sup> This was followed by a number of approvals as a second line therapy for B-cell cancers including for relapsed/refractory mantle cell lymphoma, Waldenstrom's macroglobulinemia, and chronic graft versus host disease.<sup>62-64</sup>

Similar to EGFR inhibitors, ibrutinib binds BTK's C481 adjacent to the ATP binding site and only a handful of kinases have a homologous cysteine.<sup>65</sup> Ibrutinib's rapid clearance ( $t_{1/2} = 2-3$  hours) also enables remarkable kinase selectivity: ibrutinib maintains activity against BTK due to prolonged covalent engagement while reversible off-target inhibition is minimized.<sup>66</sup> This combination of fast covalent engagement of BTK with rapid clearance allows for selectivity *in vivo* even despite off-targets observed in biochemical assays. Ibrutinib has a remarkable safety profile as a result, with

maximum tolerated dose not reached in phase 1 studies; doses of 420 mg and 560 mg daily were recommended for CLL and MCL, respectively.<sup>66</sup> Despite historical concerns over potential toxicity of covalent drugs, ibrutinib demonstrates that rationally designed covalent drugs are able to achieve acceptable safety profiles and blockbuster status. In 2020 alone, Abbvie's Imbruvica revenues totaled \$8.53 billion, making it the 5<sup>th</sup> highest grossing drug that year.<sup>67</sup> Ibrutinib's success builds on that of covalent EGFR inhibitors such as 6simertinib by showing that kinase inhibitors targeting non-conserved cysteines adjacent to the ATP binding site are capable of being developed into successful drugs.

A number of other BTK kinase inhibitors have been discovered that improve on various aspects of ibrutinib. Several second-generation BTK inhibitors highlight the variety of Michael acceptors that can be used as alternatives to acrylamides for covalently binding cysteine residues. Most prominent among these is acalabrutinib (AstraZeneca), approved by the US FDA in 2019 to treat CLL, which contains a butyramide electrophile that is less reactive towards BTK C481 and GSH.<sup>68</sup> One crucial tool that is commonly used to assess cysteine reactivity of covalent inhibitors is a glutathione (GSH) reactivity assay, where the compound is incubated with excess GSH and consumption of the compound is observed via LCMS.<sup>68-70</sup> The lower reactivity of the butyramide, in addition to other substitutions, enables acalabrutinib's superior selectivity over ibrutinib for BTK. This improved selectivity could be responsible for acalabrutinib's improved safety profile, in particular reducing cardiovascular adverse events.<sup>71,72</sup> Other covalent BTK inhibitors include tirabrutinib (Ono/Gilead, approved in Japan 2020) and branebrutinib (Bristol-Myers Squibb) which also contain a butyramide electrophile, while spebrutinib (Bristol-Myers Squibb) and zanubrutinib (Beigene, approved 2019 in US) and evobrutinib (Merck KgaA) contain acrylamides.<sup>73-75</sup> In general, these second-generation BTK inhibitors have been developed through structure-based design and each bear strong resemblance to ibrutinib.

One interesting development in designing new electrophiles for covalent drugs is the discovery that cyanoacrylamides are able to form reversible-covalent bonds with cysteine residues, and are effective in inhibiting BTK.<sup>76-78</sup> This covalent yet reversible interaction would ideally be able to increase potency while reducing off-target reactivity, though the cyanoacrylamide BTK inhibitor rilzabrutinib (Principia/Sanofi) was recently shown to be ineffective in treating the autoimmune disease pemphigus.<sup>79</sup> While BTK has been a target in inflammatory and autoimmune disease since the early 2000s, no BTK inhibitor has yet been approved for any indication in these fields.<sup>80</sup> This suggests that while rilzabrutinib was not successful, the tunable covalency of cyanoacrylamides may still hold value for other targets.

The discovery of ibrutinib and other covalent BTK inhibitors reemphasizes the lessons learned from covalent EGFR inhibitors. Employing structure-based design to target a non-conserved cysteine residue adjacent to the ATP binding site allowed for selective targeting of BTK. Taking advantage of the PK/PD characteristics of ibrutinib enabled by its covalent mechanism of action allowed prolonged BTK blockade while reducing off-target kinase inhibition through rapid clearance *in vivo*. Newer BTK inhibitors based on ibrutinib more finely tune the cysteine-reactivity of the electrophile



and improve selectivity. Most significantly, the performance of ibrutinib in treating B-cell malignancies emphasizes that molecules once considered chemical biology tool compounds can be highly effective drugs in their own right.

### 1.2.3 Other Covalent Kinase Inhibitors

Success with covalent EGFR and BTK inhibitors has driven interest in covalent inhibitors of other kinases with non-conserved cysteines adjacent to ATP-binding sites to gain selectivity, especially within kinase families.<sup>81,82</sup> One example is the non-receptor tyrosine kinase JAK3, which is primarily expressed in leukocytes and whose function in cytokine signaling has led to its investigation as a target in autoimmune disease.<sup>83</sup> Covalent targeting of the non-conserved Cys909 of JAK3 has been used to gain selectivity over the other JAK family members by a number of covalent JAK3 inhibitors.<sup>83–87</sup> One of these, ritlecitinib (Pfizer, PF-06651600), has shown promising results for rheumatoid arthritis patients in a phase II clinical trial.<sup>88</sup>

Another target of interest is the receptor tyrosine kinase FGFR4, which is implicated as a target in hepatocellular carcinoma (HCC) due to overexpression of its ligand FGF19.<sup>89</sup> In a very similar manner to discovery of JAK3 inhibitors, compounds have been discovered that covalently target the non-conserved Cys552 to confer selectivity for FGFR4 over FGFR1-3, and can overcome mutations that confer resistance to first-generation FGFR inhibitors.<sup>90</sup> The acrylamide-containing FGFR inhibitor fisogatinib (Blueprint, BLU-554) is currently the subject of a Ph2 clinical trial (NCT04194801), while the reversible-covalent aldehyde roblitinib (Novartis, FGF401) which also reacts with Cys552 is also currently under clinical investigation (NCT02325739) (**Table 1**).<sup>89,91</sup>

In addition to JAK3 and FGFR4, a number of compounds have been developed to selectively target various cyclin-dependent kinases (CDKs), including CDK7 and CDK12/13 and explore these kinases' roles in transcriptional regulation.<sup>92,93</sup> Overall, rational design of covalent kinase inhibitors targeting non-conserved cysteines adjacent to the ATP-binding site has become a common approach to enhance kinase inhibitor potency and selectivity.

### 1.2.4 Covalent KRAS<sup>G12C</sup> Inhibitors

One of the most exciting discovery-to-clinic stories for covalent drugs in the last ten years is the design of KRAS<sup>G12C</sup> mutant selective covalent inhibitors. KRAS is an oncogenic GTPase that is mutated in about 25% of all cancers, most notably in pancreatic, colorectal, and lung cancers.<sup>94</sup> Wild-type KRAS is carefully regulated between active GTP-bound states and inactive GDP-bound states, but mutated KRAS has a much higher affinity for the GTP-bound active conformation leading to low rates of ATP hydrolysis and elevated Ras signaling, driving tumorigenesis and clinical resistance.<sup>95</sup> Since the discovery of the role of KRAS in cancer nearly 30 years ago, attempts to drug it directly using traditional drug discovery methods have been unsuccessful.<sup>8,9</sup> KRAS does not have accessible pockets for reversible inhibitors to

bind, competitive inhibitors would need to overcome picomolar binding affinities of GTP/GDP, and the development of pan-KRAS inhibitors could lead to on-target toxicity.<sup>94,96</sup>

In 2013, Kevan Shokat at UCSF published the first example of a selective covalent inhibitor of the mutant KRAS<sup>G12C</sup>.<sup>96</sup> Developing covalent KRAS inhibitors against the G12C mutation is appealing for several reasons: 1) targeting mutant KRAS could allow for selective cytotoxicity to cancer cells, 2) the affinity enabled by covalent binding would be advantageous as KRAS lacks easily ligandable pockets, 3) 45% to 50% of KRAS mutations in NSCLC are KRAS<sup>G12C</sup>, presenting a promising patient group that would directly benefit from KRAS<sup>G12C</sup> inhibition, and 4) position 12 in KRAS sits closely beneath the GTP/GDP effector binding region as well as the nucleotide binding pocket, suggesting covalent KRAS<sup>G12C</sup> ligands might affect KRAS function.<sup>95</sup>

In 2013, Ostrem et al. at UCSF used a disulfide-fragment-based screening approach called tethering, and screened a library of 480 disulfides against KRAS<sup>G12C</sup> in the GDP-bound state using intact protein mass spectrometry.<sup>96</sup> Co-crystal structures of KRAS<sup>G12C</sup> showed the tethering hits bound to the Switch-II region, and further medicinal chemistry yielded acrylamide and vinyl sulfonamide KRAS<sup>G12C</sup> inhibitors active *in vitro*. Compound 12, their lead compound, induced apoptosis in KRAS<sup>G12C</sup> mutant cell lines at 10  $\mu$ M treatment conditions over 48 hours.<sup>96</sup> Compound binding to the S-IIP (switch II pocket), which was not apparent in previous structures of RAS, impairs KRAS by shifting the nucleotide affinity from favoring GTP to GDP and leads to the accumulation of KRAS in its inactive state.<sup>97</sup> Beyond identifying inhibitors of KRAS, the Shokat lab's breakthrough discovery of the allosteric S-IIP proved critical in understanding the mechanism of covalent inhibition of KRAS<sup>G12C</sup>.

This novel mechanism for KRAS<sup>G12C</sup> selective inhibition paved the way for the development of clinical covalent KRAS<sup>G12C</sup> inhibitors. In 2016, Wellspring Biosciences disclosed ARS-853, a selective, covalent inhibitor of KRAS<sup>G12C</sup> with *in cellulo* efficacy in the low micromolar range.<sup>98</sup> Patricelli et al. used a cellular LC/MS-MS-based assay to determine the degree of KRAS<sup>G12C</sup> engagement in H358 cells by quantifying decreasing amounts of the Cys12-containing peptide from tryptic digests of these mutant cells upon treatment with candidate compounds. Structure-guided optimization of Ostrem et al.'s Compound 12 yielded ARS-853, which when crystalized with KRAS<sup>G12C</sup>, showed significant improvements in binding of the S-IIP. ARS-853 treatment in mutant cell lines reduces active KRAS<sup>G12C</sup>, inhibits downstream Ras signaling, and induces apoptosis with 1 mM treatment conditions in KRAS<sup>G12C</sup> cell lines. While KRAS<sup>G12C</sup> was previously thought to be constitutively active, with no GTPase activity, the selective binding of ARS-853 to GDP-bound, inactive KRAS<sup>G12C</sup> provided evidence that KRAS mutants indeed cycle between GTP- and GDP-bound states.

The development of improved KRAS<sup>G12C</sup> inhibitors continued with the publication of ARS-1620 by Wellspring. ARS-1620 was the result of an effort to overcome metabolic stability and bioavailability limitations of ARS-853 to facilitate *in vivo* studies of KRAS<sup>G12C</sup> inhibition.<sup>95</sup> ARS-1620 is based on a novel quinazoline core scaffold,

designed to better occupy the S-IIP and thus rigidify a more favorable conformation for covalent reaction with the acrylamide warhead, and additionally act as a more diversifiable molecule with better drug-like properties. ARS-1620 showed efficacy in patient derived tumor xenograft studies with G12C mutations, treating at 200mg/kg. Ultimately, ARS-1620 was identified as the first example of a direct KRAS<sup>G12C</sup> small-molecule inhibitor that was orally bioavailable and suited for *in vivo* studies. Improved potency of this series of KRAS<sup>G12C</sup> inhibitors and success in *in vivo* models indicated that it might be possible to design clinically efficacious drugs.

Sotorasib (Amgen, AMG-510, Lumakras/Lumykras) was the first selective KRAS<sup>G12C</sup> inhibitor to enter clinical trials in 2018 (NCT03600883) and was developed by Amgen in partnership with Carmot Therapeutics (**Figure 2, Table 2**).<sup>99</sup> They screened a custom library of covalent cysteine reactive small-molecules, utilizing Carmot's Chemotype Evolution platform and structure-based design to identify sotorasib. Sotorasib was designed to occupy and exploit a previously unknown sub-pocket in the S-IIP of KRAS<sup>G12C</sup>, interacting with H95, Y96, and Q99 (**Figure 2**).<sup>100</sup> A phase II clinical trial investigating sotorasib was successfully completed in 2020, followed by FDA approval to treat adults with KRAS<sup>G12C</sup> -mutated locally advanced or metastatic non-small cell lung cancer (NSCLC) in May 2021.<sup>101</sup> Other covalent KRAS<sup>G12C</sup> inhibitors are quickly following sotorasib through clinical trials, including JNJ-74699157 (ARS-3248; J&J/Wellspring), which is being investigated in patients with several types of advanced solid tumors expressing KRAS<sup>G12C</sup> including NSCLC and colorectal cancer (NCT04006301). Mirati has also developed adagrasib (MRTX-849) with Array BioPharma by identifying tetrahydropyridopyrimidines as irreversible covalent inhibitors of KRAS<sup>G12C</sup> and using structure-based design approaches to optimize adagrasib, which entered clinical trials in January 2019 (NCT03785249) (**Figure 2**).<sup>102,103</sup> Adagrasib was granted a breakthrough therapy designation by the FDA in June 2021 for use as a potential therapeutic option for patients with KRAS<sup>G12C</sup> NSCLC following previous systemic therapy.<sup>104</sup> Designing small-molecule covalent KRAS<sup>G12C</sup> selective inhibitors provides an elegant solution to drugging an undruggable cancer target. This story represents a major step forward in oncology, and the process by which KRAS<sup>G12C</sup> inhibitors were designed can provide a map for targeting other common KRAS mutations like G12D, G13D and G13C. In the case of other diseases where a point mutation leads to the substitution of a nucleophilic amino acid, covalent inhibitors present an optimal opportunity to provide precision therapy for patients.

## 1.3 Covalent Protease Inhibitors

### 1.3.1 SARS-CoV2 Main Protease Inhibitors

Since the beginning of the COVID-19 pandemic, over 420 million cases of COVID-19 have been reported worldwide and over 5.86 million people have died from the disease.<sup>105</sup> Vaccines against COVID-19 were developed at unprecedented speeds and similar research momentum is leading to the development of therapeutics that will benefit infected patients. On December 22, 2021, the FDA issued an Emergency Use

Authorization (EUA) for Pfizer's Paxlovid (nirmatrelvir and ritonavir, co-packaged for oral use) to treat mild-to-moderate COVID-19 (caused by the SARS-CoV-2 coronavirus) in adults and some pediatric patients, marking the first approved oral treatment for the disease.<sup>106</sup> Nirmatrelvir covalently inhibits the main protease (M<sup>pro</sup>) of SARS-CoV-2, and ritonavir is a CYP3A inhibitor used to enhance the pharmacokinetic properties of protease inhibitors in general.<sup>2</sup>

SARS-CoV-2 is a single-stranded RNA virus that encodes for two polyprotein precursors (pp1a and pp1ab) as well as structural and accessory proteins.<sup>107</sup> Viral replication depends on successful cleavage of pp1a and pp1ab by the main protease (M<sup>pro</sup>, also referred to as 3Clpro) into functional viral proteins.<sup>107</sup> Additionally, viral M<sup>pro</sup> cleaves its substrates after Gln, which is a unique cleavage site compared with known human proteases. Therefore, M<sup>pro</sup> is a promising protein target for drug development against SARS-CoV-2. Because M<sup>pro</sup> is a cysteine protease (with a catalytic diad composed of Cys145 and His41), covalently inhibiting the catalytic cysteine should prevent viral replication.<sup>107</sup>

The discovery of covalent inhibitors against SARS-CoV-2 M<sup>pro</sup> emerged from extensive previous work towards developing protease inhibitors for SARS-CoV-1, which is the causative virus for the severe acute respiratory syndrome coronavirus 1 outbreak in 2002-2003.<sup>107</sup> During the 2002-2003 SARS outbreak, caused by the SARS-CoV-1 coronavirus, Anand et. al. used a crystal structure of the homologous porcine TGEV coronavirus M<sup>pro</sup> bound to a hexapeptidyl chloromethylketone (CMK) inhibitor as a model to provide a base for the design of potential SARS-CoV-1 M<sup>pro</sup> inhibitors.<sup>108</sup> While peptidomimetics can pose challenges due to their inherent metabolic instability, Anand et. al. argued that the significantly shared homology across all M<sup>pro</sup> active sites should enable the use of other viral M<sup>pro</sup> inhibitors as the starting point for SARS-CoV-1 specific M<sup>pro</sup> drugs. Rupintrivir, developed originally to target human rhinovirus (HRV) M<sup>pro</sup>, could provide an initial scaffold for further investigation.<sup>108</sup>

Because the SARS-CoV-1 outbreak subsided, work into developing coronavirus M<sup>pro</sup> inhibitors slowed until the emergence of SARS-CoV-2 in 2019. SARS-CoV-2 M<sup>pro</sup> shares 96% sequence identity with SARS-CoV-1 M<sup>pro</sup>, and there is 100% sequence overlap of the catalytic sites.<sup>109</sup> Renewed interest in improving on chloromethyl ketone inhibitors lead to Hoffman et. al. systematically exploring the medicinal chemistry around rupintrivir to develop potent *in vitro* SARS-CoV-2 M<sup>pro</sup> inhibitors (both reversible and irreversible covalent inhibitors). Rupintrivir initially showed little inhibitory activity against SARS-CoV-1 M<sup>pro</sup>, but Hoffman et. al. were able to improve affinity enough to obtain a co-crystal structure with SARS-CoV-1 M<sup>pro</sup>. This enabled identification of the hydroxymethylketone PF-00835231 which demonstrated potent SARS CoV-1 M<sup>pro</sup> inhibition both in a FRET-based activity assay, activity in antiviral cell-based assays, good solubility, stability in plasma, and low clearance *in vivo*. The identification of PF-00835231 as a potent inhibitor of SARS-CoV-2 M<sup>pro</sup> spurred further pre-clinical investigation to determine if it could be used as a COVID-19 therapeutic.

Concurrently with Hoffman et. al., Qiao et. al. designed M<sup>pro</sup> inhibitors based on the approved hepatitis C antivirals 11simertini and boceprevir. The active site of M<sup>pro</sup> is composed of four sites (S1', S1, S2, and S4), which accommodate four components of the various established inhibitors (P1', P1, P2, and P3, respectively). Keeping P1 fixed, Qiao et. al. synthesized compounds that varied at the P2 position only based on HCV NS3/4a inhibitors boceprevir and 11simertini and varied P3 significantly across the series.<sup>107</sup> Activities of the compound series were measured by the standard FRET-based activity assay against recombinant SARS-CoV-2 protein, and direct binding was confirmed by differential scanning fluorimetry. MI-09 and MI-30 (both containing aldehyde warheads) were identified as the most potent hits (with IC<sub>50</sub> values of <20 nM in the *in vitro* FRET assay). MI-09 and MI-30 inhibited SARS-CoV-2 infection in a human angiotensin-converting enzyme 2 (hACE2) transgenic mouse model, showing these compounds could reduce lung lesions *in vivo*.

With a set of potent SARS-CoV-2 M<sup>pro</sup> inhibitors in hand and established *in vitro* and *in vivo* SARS-CoV-2 assays, Owen et. al. set out to improve oral bioavailability of PF-00835231. PF-00835231 showed potent inhibitory activity in the SARS-CoV-2 FRET-based assay as well as antiviral activity (EC<sub>50</sub> = 231nM) in Vero E6 cells. However, improving oral bioavailability would significantly increase the potential for PF-00835231 to impact most COVID-19 patients. Owen et. al. replaced the  $\alpha$ -hydroxymethyl ketone with a nitrile, which is capable of acting as an electrophilic warhead.<sup>110</sup> Nitriles are able to covalently bind particularly reactive nucleophiles, though the ease of thiol elimination from the thioimide adduct makes nitriles more reversible than some other electrophiles such as acrylamides.<sup>111</sup> This compound demonstrated significant improvement in rat oral absorption, while maintaining metabolic stability, but *in vitro* and *in cellulo* antiviral activity was reduced. Hydrogen bond donors were removed by introducing a bicyclic proline derivative at P2, and the trifluoroacetamide capping group to restore potency through occupying the S3 pocket. Combining these changes yielded PF-07321332, a highly potent SARS-CoV-2 reversible covalent inhibitor that displayed potent inhibition in the FRET assay across all human coronaviruses, while no inhibitory effects were seen against human cysteine or serine proteases. *In vivo* efficacy of PF-07321332 was demonstrated in a mouse-adapted SARS-CoV-2 (SARS-CoV-2 MA10) model, and PF-07321332 additionally showed good oral bioavailability, favorable off-target selectivity, and metabolic stability in rats and monkeys. PF-07321332, named nirmatrelvir, was investigated as a monotherapy and in combination with ritonavir (a CYP3A inhibitor that enhances the PK profile of nirmatrelvir) in a randomized, double-blind, placebo-controlled, single ascending dose study in healthy adults (NCT04756531) (**Figure 3, Table 2**).<sup>2</sup>

Paxlovid (combination nirmatrelvir/ritonavir) was shown to be highly effective at preventing symptomatic patients' progression to severe COVID-19 in a Phase II/III clinical trial data released in November 2021.<sup>112</sup> The emergence of the orally bioavailable drug for COVID-19 will help to ameliorate illness for non-hospitalized patients in high-risk groups.<sup>113</sup> As a whole, SARS-CoV-2 M<sup>pro</sup> covalent inhibitors provide a promising avenue to treat coronavirus infections either as monotherapies or in combination with other antiviral drugs. The quick adaptation of previous protease

inhibitors to selectively target SARS-CoV-2 M<sup>pro</sup> is an example of elegant structure-based design paired with the power of covalent drugs.

### 1.3.2 HCV NS3/4a Protease Inhibitors

The NS3/4a serine protease cleaves the HCV polyprotein into multiple non-structural proteins, which are required for replication.<sup>114</sup> While HCV had been treated with a combination of PEGylated interferon alpha and ribavirin, modest response rates and significant adverse events prompted the discovery of NS3/4a protease inhibitors to treat HCV.<sup>115</sup> Based on initial observations that hexapeptide cleavage products could inhibit NS3/4a, the linear peptidomimetic inhibitors boceprevir<sup>116,117</sup> and telaprevir<sup>118,119</sup> were designed. These compounds, along with narlaprevir,<sup>120</sup> use a ketoamide to covalently engage the catalytic serine of NS3/4a. This covalent interaction is relatively reversible due to elimination of the serine alcohol from the protein-inhibitor adduct.<sup>121</sup> Boceprevir and telaprevir were effective in treating HCV and approved in 2011 after successful trials,<sup>122,123</sup> though telaprevir was withdrawn in 2014 due to adverse events and boceprevir was discontinued by Merck in 2015 due to the superiority of newer direct-acting antivirals, in particular ledipasvir/sofosbuvir (Gilead) which target the HCV polymerases NS5a and NS5b.<sup>124–126</sup> However, the success of ketoamide NS3/4a inhibitors in improving efficacy of interferon/ribavirin therapy emphasizes the utility of the ketoamide as a serine-reactive electrophile in designing covalent antivirals.

### 1.3.3 Covalent Proteasome Inhibitors

Bortezomib, a boronic acid proteasome inhibitor, was approved in 2003 to treat multiple myeloma. To identify proteasome inhibitors with improved safety profile, a series of medicinal chemistry efforts transformed the natural product 12simertini to carfilzomib (Proteolix/Onyx), which was approved in 2012.<sup>127</sup> Carfilzomib's epoxyketone is able to form a morpholino ring with the catalytic N-terminal threonine of the 20S proteasome, and this mechanism confers high selectivity because most proteases do not have N-terminal nucleophiles.<sup>128</sup> The mechanism of carfilzomib highlights how covalency can help drive selectivity, rather than encourage promiscuity and off-target inhibition.

## 1.4 The Covalent Drug Discovery Toolbox

Covalent drug discovery is enabled by a myriad of techniques that provide the means for discovery and characterization of electrophilic molecules. In particular, activity-based protein profiling (ABPP) approaches have transformed characterization of electrophilic compounds. Combining ABPP-style techniques with other advances in covalent ligand screening has enabled adoption of “electrophile-first” drug discovery efforts. This type of approach proved successful in drugging KRAS<sup>G12C</sup> and has facilitated rapid discovery of E3-ligase ligands. Looking forward, the immense knowledge of reactive sites across the proteome generated through chemoproteomic experiments over the past decade will provide powerful insights into potential targets of covalent drugs.

## 1.4.1 Chemoproteomics-Enabled Discovery

One of the most important tools in covalent drug discovery is the use of chemoproteomics platforms to identify covalent compounds and their corresponding ligandable sites on target proteins directly in complex biological systems. Advances in chemoproteomics have facilitated the discovery of covalent ligands against undruggable disease targets, selectivity profiling of covalent ligands across the proteome to identify targets and off-targets of these ligands, and discovery of novel covalent E3 ligase recruiters and deubiquitinase recruiters for targeted protein degradation and stabilization platforms.

### 1.4.1.1 Covalent Ligand Discovery using ABPP

Activity-based protein profiling (ABPP) enables the discovery of covalently ligandable sites and corresponding ligands in complex biological samples. ABPP was initially pioneered by Cravatt and Bogoyevitch using active-site directed chemical probes that covalently targeted catalytic residues of various enzyme classes – including hydrolases, proteases, and kinases.<sup>134</sup> This technique, often using gel-based assays, was employed to gain functional readouts of these active enzymes in biological contexts.<sup>13,134,135</sup> ABPP probes contain a warhead that covalently reacts with nucleophilic amino acids (such as cysteine) and a reporter handle to monitor probe binding such as a fluorophore, biotin, or alkyne moiety for subsequent click chemistry-enabled applications (**Figure 4a**).<sup>134</sup>

Instead of focusing on active sites, more recent ABPP approaches use mass spectrometry and broadly reactive chemical probes to map allosteric reactive sites as well.<sup>136</sup> Weerapana et al first described the isoTOP-ABPP (isotopic tandem orthogonal proteolysis–activity-based protein profiling) approach which used an iodoacetamide (IA) probe to identify hyper-reactive cysteines across the proteome.<sup>136,137</sup> In addition to the iodoacetamide electrophile, the probe contains an alkyne which links to an azide- and biotin-containing, TEV protease-cleavable tag with either an isotopically light or heavy valine. These functionalities enable enrichment of probe modified peptides, and tandem analysis of two light/heavy samples with mass-spectrometry, controlling for run-to-run variability and allowing for quantitative comparisons between samples. Weerapana et al. discovered that cysteine hyper-reactivity predicts functionality in catalysis and at sites of PTMs (**Figure 4b**).<sup>137</sup>

Building off this work, Backus et al used competitive isoTOP-ABPP to identify proteome-wide targets of a small covalent fragment library by competing individual acrylamides and chloroacetamides with the IA-alkyne probe.<sup>138</sup> This study identified >700 ligandable cysteines and provided information about the proteome-wide selectivity of each covalent fragment. Backus et al. used the covalent ligands discovered with this approach and their corresponding ligandable sites to help elucidate the role of CASP8 and CASP10 in extrinsic apoptosis in T cells, showing that this approach can rapidly identify compounds that target proteins of biological interest. Bar-Peled et al used isoTOP-ABPP to map cysteine reactivity and ligandable sites in KEAP1 mutant vs wild

type KEAP1 NSCLC lines, and discovered the nuclear receptor NR0B1 has a ligandable cysteine regulated by NRF2, the substrate of KEAP1.<sup>139</sup> Several chloroacetamides were identified that bound NR0B1 selectively and altered the transcriptional profile and growth of NRF2-activated cells. In an impressive study, Vinogradova et al explored cysteine ligandability in activated T cells, using promiscuous acrylamide/chloroacetamide “scout-fragments” to map ligandability, and functional assays to identify elaborated electrophilic compounds that suppress T cell activity.<sup>140</sup> This approach identified several proteins that could be targeted covalently to impair T cell activity, including BIRC2/3, the NuRD complex, and the kinases ITK and CYTIP.

IsoTOP-ABPP can also be used to identify the protein targets of electrophilic drugs, such as the dimethyl fumarate (DMF) used to treat autoimmune disease. Though DMF has been used for decades to treat psoriasis and was approved the FDA in 2013 for multiple sclerosis, the direct covalent targets of DMF remained unclear. Separate studies have used chemoproteomic approaches to identify both protein kinase C- $\alpha$  (PKC $\alpha$ ) and IRAK4 as targets of DMF.<sup>141,142</sup> In both cases, covalent engagement of a cysteine disrupted a protein-protein interaction to modulate immune cell function. Disrupting the PKC $\alpha$ -CD28 interaction reduced T-cell activation, and disrupting the IRAK4-Myd88 interaction suppressed interferon-alpha production in plasmacytoid dendritic cells.<sup>141,142</sup>

IsoTOP-ABPP can also be used for target identification and selectivity assessment of covalently-acting small-molecules that are discovered initially from cell-based phenotypic screens. Chung et al. used this approach to discover an autophagy activator which covalently targets the vacuolar proton pump ATP6V1A.<sup>143</sup> Using isoTOP-ABPP, hit compounds from an autophagic flux assay were shown to target ATP6V1A to induce clearance of Tar-binding protein 43 (TDP-43) aggregates which drive amyotrophic lateral sclerosis (ALS). Novel screening platforms are also often easily paired with isoTOP-ABPP target ID experiments. For example, Gruner et al developed a multiplexed *in vivo* screening platform where barcoded PDAC lines were pretreated with electrophilic compounds and injected into mice to observe compound-dependent decrease in metastatic potential.<sup>144</sup> IsoTOP-ABPP experiments were able to identify the lipase ABHD6 as the target of hit compounds from this screen, even though ABHD6 was not known previously to play a role in metastasis or cancer progression. Beyond identifying the lipase ABHD6 as critical for metastatic fitness, this approach allowed screening in a biological context more relevant to the disease state through adaptation of covalent ligand screening to a multiplexed *in vivo* phenotypic assay. In an example of targeting a traditionally undruggable protein with a covalent molecule, Boike and Cioffi et. al. used ABPP paired with phenotypic covalent ligand screening to identify a covalent MYC ligand EN4.<sup>145</sup> EN4 targets cys171 within a predicted intrinsically disordered region of MYC, and showed selectivity on a proteome-wide scale profiling >1,500 cysteines using competitive isoTOP-ABPP.

Recent adaptations of isoTOP-ABPP have been developed to increase the coverage of cysteines across the proteome and increase throughput. Yan et al. used sample preparation steps (SP3) including off-line fractionation and a FAIMS source,



which allows for additional separation before MS detection, to identify more than 30,000 reactive cysteines across a panel of tumor cell lines.<sup>146</sup> Kuljanin et. al. designed a tandem mass tag (TMT)-based streamlined cysteine activity-based protein profiling (SLC-ABPP) methodology to dramatically increase sample throughput, profiling an electrophilic fragment library at a depth of >8,000 reactive cysteine sites at 18 min per compound.<sup>147</sup>

#### **1.4.1.2 ABPP-Enabled Induced Proximity: Targeted Protein Degradation and Beyond**

ABPP-based chemoproteomics facilitates the expansion of targeted protein degradation approaches and the development of new therapeutic modalities by enabling the discovery of covalent recruiters against classically undruggable protein classes such as E3 ubiquitin ligases and deubiquitinases (DUBs).<sup>13,148–155</sup> Although most bifunctional degrader molecules (also known as proteolysis-targeting chimeras, or PROTACs) recruit the E3 ligases CRBN or VHL to degrade target proteins, there are over 600 E3 ligases with varying substrate scopes. Since 2019, covalent recruiters have been used to validate a large portion of the E3 ligases that have been harnessed for targeted protein degradation, including the E3s RNF114, RNF4, DCAF16, DCAF11, KEAP1, and most recently FEM1B.<sup>148–154</sup> In 2019, Spradlin et al used isoTOP-ABPP to identify RNF114 as the target of the enone-containing natural product nimbolide, which was used to make bifunctional degraders of BRD4 and BCR-ABL.<sup>148</sup> Simultaneously, Zhang et al used “scout fragments” to construct bifunctional molecules that degraded FKBP12 and BRD4, and identified the DCAF16 as the covalent target responsible for degradation.<sup>151</sup> These discoveries led to the variety of covalent E3 recruiters now available, which have been reviewed elsewhere.<sup>156</sup> Based on analyses of chemoproteomic datasets assessing cysteine reactivity, 97% of E3 ligases possess reactive cysteines, suggesting that covalent approaches to harness more E3 ligases could continue to be successful.<sup>156</sup>

ABPP has also enabled the development of new induced proximity-based therapeutic modalities beyond targeted protein degradation. Henning and Boike et. al. recently developed a targeted protein stabilization platform, termed Deubiquitinase Targeting Chimeras (DUBTACs).<sup>155</sup> By assessing aggregate chemoproteomic data describing proteome-wide cysteine reactivity, the authors identified a reactive allosteric cysteine C23 in the DUB OTUB1. An ABPP-based screen against pure OTUB1 yielded a covalent, allosteric OTUB1 ligand that binds C23, EN523, which was used to design heterobifunctional DUBTACs linking OTUB1 to protein-targeting ligands. The authors demonstrated proof-of-concept of the DUBTAC platform with stabilization of mutant CFTR, whose ubiquitination and degradation drives cystic fibrosis, and the tumor suppressor WEE1 kinase for potential therapeutic in cystic fibrosis and cancer, respectively.<sup>155</sup> Overall, these studies have showcased how ABPP and covalent ligand discovery approaches have been used to expand the scope of TPD and develop new therapeutic modalities for drug discovery.

#### **1.4.2 Screening Platforms beyond ABPP**

While ABPP-based screening approaches are valuable in discovering covalent ligands, limited throughput has encouraged alternative screening methods. These often involve MS-based detection, uniquely possible in covalent ligand screening, which offers additional information about binding stoichiometry. Other phenotypic or computational approaches have also been used, which are then paired with MS-based validation, ABPP-based experiments to inform selectivity, and structural biology to enable medicinal chemistry. The growth of commercial libraries of electrophilic fragment-like compounds contribute to the rise of these “electrophile-first” discovery strategies (**Figure 5**).

#### 1.4.2.1 Intact Protein-MS for Covalent Ligand Screening

Originally, MS-based compound screening grew out of “tethering,” a technique employed since 2000 that uses libraries of compounds linked to disulfides to identify fragments that bind cysteine-adjacent pockets.<sup>157,158</sup> Molecules that bind undergo disulfide exchange with the cysteine (which could be endogenous or engineered) to form an adduct with the protein, and pooled screening with MS detection can identify the bound compound. Binding fragments can be combined or grown in a fragment-based approach to identify high-affinity ligands. This strategy was originally designed to identify reversible ligands for challenging targets, but covalent binding can be designed back in by replacing the disulfide with an electrophile such as an acrylamide, as in the case of the covalent KRAS<sup>G12C</sup> inhibitors first discovered at UCSF.<sup>96</sup> This approach has often been employed to discover compounds that modulate protein-protein interactions,<sup>159,160</sup> and one recent study employed a similar tethering strategy using aldehydes to form imines with lysine residues.<sup>161</sup>

Over the past decade covalent ligand discovery has shifted towards screening more drug-like electrophilic fragments, in particular enabled by the commercial availability of acrylamide-focused libraries.<sup>12</sup> In 2014, Kathman et al. appended an acrylate functionality to 100 fragments to identify non-peptidic inhibitors of papain protease.<sup>162</sup> Electrophilic fragments were pooled, and screening through ESI-MS was able to identify hit compounds out of pooled experiments. Use of a similar library of acrylates enabled discovery of covalent inhibitors for the HECT E3 ligase Nedd4-1.<sup>163</sup> Co-crystal structures were critical to understanding the mechanism of these compounds, which prevent association of ubiquitin with the E3 ligase and thus induce a switch from a processive to distributive mechanism. In another example, acrylate-based inhibitors of the RBR E3 HOIP were also discovered using MS-based screening, highlighting how covalent fragment screening approaches can be useful for protein classes which have been challenging to develop ligands against, such as E3 ligases.<sup>164</sup>

More recently, Resnick et al. reported screening a commercial library of 993 acrylamides and chloroacetamides using intact-MS to identify ligands of the deubiquitinase OTUB2 and the pyrophosphatase NUDT7.<sup>165</sup> As is often the case, the authors used co-crystal structures with OTUB2 and NUDT7 in complex with the hit compounds to inform fragment growing to increase potency. Because generally weak

interactions of hit compounds in covalent fragment screens necessitates medicinal chemistry optimization, the pairing of MS-based screening and structure-guided FBDD is particularly helpful.

The same compound collection used by Resnick et al. was also used to screen against the peptidyl-proline *cis-trans* isomerase Pin1, which is overexpressed or activated in a number of tumor types but has been challenging to target selectively.<sup>166</sup> The resulting chloroacetamide Sulfopin was shown to be selective for Pin1 using a CITE-Id chemoproteomics experiment, phenocopied Pin1 knockdown effects in cell lines, and was effective in regressing neuroblastoma growth in mice. This suggests that though chloroacetamides have liabilities, including rapid metabolism, they can be valuable tool compounds with which to assess target relevance in a variety of disease models.

One particularly powerful example of covalent ligand screening is the discovery of initial compounds in the series that led to the first approved KRAS<sup>G12C</sup> inhibitor, sotorasib (Amgen). A library of 3300 acrylamides was screened in three assays, a thiol reactivity assay, a RAF-coupled nucleotide exchange assay, and an intact MS assay.<sup>167</sup> Combined with crystallographic data that revealed un-occupied sub-pockets, this effort provided the basis for rapid discovery of novel KRAS<sup>G12C</sup> inhibitors, discussed above.

#### 1.4.2.2 Covalent DNA Encoded Libraries

DNA-encoded libraries present an alternative screening approach that takes advantage of massive library size. Unlike MS or ABPP approaches there is no specialized advantage of covalency in enabling DEL screening, but the throughput of DELs allows screening of much larger covalent libraries through a similar workflow of immobilization, enrichment, amplification, and sequencing. First reports of electrophilic protein-nucleic acid encoded libraries described targeting protease active sites,<sup>168</sup> but in over the past 5 years cysteine-targeted DNA or PNA-encoded libraries have identified covalent ligands for JNK1, MEK2 and Her2.<sup>169,170</sup> A covalent ligand of MAP2K6 was also identified serendipitously through screening a DEL against a DNA-encoded protein library.<sup>171</sup> More recently, covalent DELs of larger size (~100,000,000 members) have been developed and used to identify acrylamide- and epoxide-based BTK inhibitors with novel scaffolds.<sup>172</sup> This expansion of library size and commercial availability represents an excited development, and DELs containing electrophilic molecules may become more widespread in covalent ligand discovery.

#### 1.4.2.3 Covalent Docking

A small but growing set of studies have used computational methods to screen electrophilic compounds. Development of the DOCKoValent method by London et al. in 2014 allowed for the discovery of boronic acid Amp-C beta-lactamase and cyanoacrylamide inhibitors of JAK3.<sup>173</sup> Since then, the same method has been used to identify novel covalent inhibitors of the kinase MKK7,<sup>174</sup> as well as compounds that bind KRAS<sup>G12C</sup> to destabilize the protein and accelerate nucleic acid exchange.<sup>175</sup> We see

these developments as exciting and expect to see continual improvement and adoption of computational screening methods for covalent compounds.

### 1.4.3 Lysine-directed Covalent Ligands

Cysteine's low abundance enables selectivity but limits opportunities for covalently targeting specific proteins of interest. This has driven scientists to look towards discovering lysine-targeted covalent molecules. Especially due to the lower nucleophilicity relative to cysteine under physiological conditions, efficient lysine-targeting covalent ligand discovery requires identification of unusually reactive lysines. isoTOP-ABPP experiments using several lysine-directed probes have proven especially powerful to profile lysine reactivity across the proteome.<sup>176,177</sup> Through isoTOP-ABPP experiments, Hacker et al. identified >9,000 ligandable lysines and showed alongside Ward et al. that more elaborated pentafluorophenol- or NHS ester-containing compounds could selectively label specific proteins of interest.<sup>176,177</sup> Building on these experiments, Abbasov et al. assembled a library of ~180 electrophiles and used isoTOP-ABPP experiments to assess the selectivity of different chemotypes and identify lysines ligandable with small molecules.<sup>178</sup> Their study yielded more broadly reactive electrophiles, such as dicarboxaldehydes, that could be used for further lysine profiling experiments, but also identified less reactive electrophiles including N-acyl-N-alkyl sulfonamides which had been previously used as tools for bioconjugation in cells.<sup>178,179</sup>

Voxelotor is a lysine-targeted covalent drug, used to treat sickle-cell anemia, whose discovery was dependent upon knowledge of heightened side-chain reactivity. Sickle cell anemia is caused by a single mutation on the beta-hemoglobin chain that induces polymerization of mutant hemoglobin (HbS) under hypoxic conditions.<sup>180</sup> Previously, an aldehyde containing natural product and several synthetic aldehyde analogs were found to prevent polymerization through increasing HbS affinity for oxygen.<sup>181</sup> These aldehydes bind in a reversible covalent manner, forming a Schiff-base with the  $\alpha$ -Hb chain's N-terminal valine.<sup>181</sup> Almost 50 years ago, this N-terminal amine was discovered to have a particularly low pKa of 6.9, indicating that it is primarily unprotonated under physiological conditions, and as a result more nucleophilic.<sup>182</sup> Based off earlier aldehydes, Voxelotor (Global Blood Therapeutics) was discovered through a structure-guided effort to discover compounds that increase HbS oxygen affinity, and was designed to bind the HbS tetramer in a 1:1 stoichiometry unlike the 2:1 ratio of initial structures.<sup>181,183</sup> With a remarkable RBC/plasma ratio of ~150 that likely reduces off-target effects, Voxelotor was approved in the US in 2019 with a recommended dose of 1.5 grams daily, an unusually high dose for a covalent drug.<sup>180,184</sup> Voxelotor's success compounds with ibuprofen's to show that covalent drugs can be dosed at high concentrations given favorable ADME properties. Furthermore, Voxelotor's discovery shows the power of identifying unusually reactive amino acid residues such as the  $\alpha$ -Hb N-terminus.

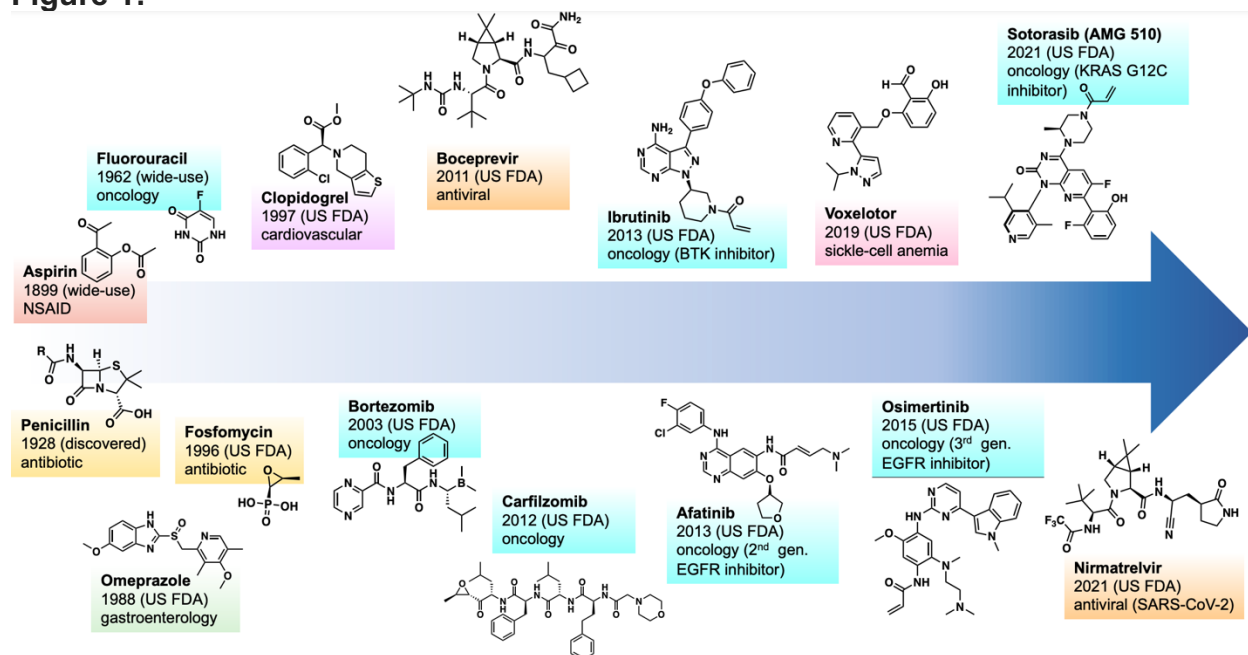
Most other lysine-targeted ligands have been designed using structure-based methods from an existing ligand, often through rationally placing an electrophilic sulfonyl fluoride, fluorosulfate, or vinyl sulfone in an appropriate orientation to react with an

epsilon-amino group of a lysine adjacent to an established binding site. These have included a kinetic transthyretin stabilizer,<sup>185</sup> an isoform selective PI3Kdelta inhibitor,<sup>186</sup> and inhibitors of CDK2<sup>187</sup> and Hsp90.<sup>188</sup> A sulfonyl fluoride-bearing promiscuous kinase inhibitor targeting a conserved lysine in the ATP-binding site was also used as a probe to profile kinase inhibitor selectivity in live cells.<sup>189</sup> Sulfonyl fluoride-based probes are not completely selective for lysine, however. In the case of the mRNA decapping scavenger enzyme DcpS, a sulfonyl fluoride-based tool compound was used to competitively determine intracellular target engagement through covalent targeting of tyrosine.<sup>176</sup> A recent preprint profiled the amino-acid reactivity preference of 54 different electrophiles using chemical proteomics, which will prove to be a great resource for covalent ligand discovery.<sup>191</sup> Combining comprehensive electrophile profiling, lysine-directed chemoproteomics, and structure-guided approaches will allow scientists to leverage the abundance of lysine adjacent to ligand binding sites to enhance covalent drug discovery.

## 1.5 Conclusion

Over the last decade, impressive advancements in covalent drug discovery have led to promising drugs including EGFR inhibitors (afatinib and 19simertinib), BTK inhibitors (ibrutinib), KRAS<sup>G12C</sup> mutant specific inhibitors (sotorasib), and SARS-CoV-2 Mpro inhibitors (nirmatrelvir). These major milestones in drug development showcase the evolution of covalent drug discovery from a serendipitous effort to a field with a well-established roadmap for success. The technologies that enable this success include covalent screening platforms, chemoproteomics discovery platforms including selectivity profiling, and structure-guided drug design. Covalent drug discovery overcomes obstacles in traditional drug discovery, making possible the design of small-molecule ligands against many traditionally “undruggable” disease targets. In particular we were interested in harnessing the toolbox of covalent drug discovery to enable the design of covalent MYC inhibitors.

**Figure 1:**



**Figure 1. Timeline of the development of major covalent drugs.** Each covalent drug is classified according to the type of disease it treats (indicated by color and label).

**Table 1. Highlighted Covalent Kinase Inhibitors**

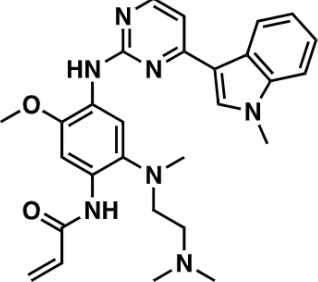
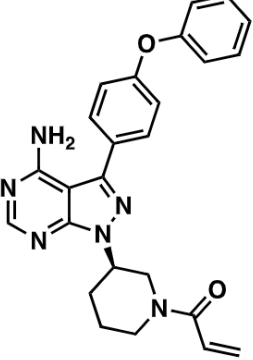
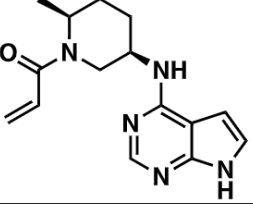
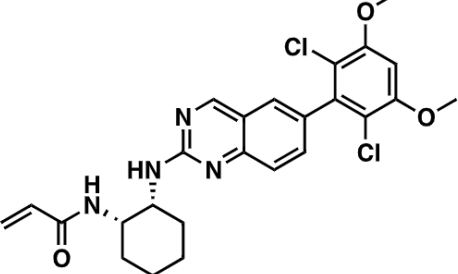
Drug (Company; Former compound name)	Structure	Target	Approval/Trial Status	Ref(s)
Osimertinib (Astrazeneca ; AZD9291)	 <p>The structure of Osimertinib features a central pyridine ring substituted with a 4-methoxyphenyl group, a 1-methyl-2-(methylamino)ethyl group, and a 2-(2-methyl-1H-indol-3-yl) group. It is linked via an amide bond to a 4-(2-methyl-2-propenyl)phenyl group.</p>	Mutant-selective EGFR inhibitor	FDA approved 2015 for NSCLC	Finlay 2014; Cross 2014; Janne 2015; Mok 2016
Ibrutinib (Abbvie; PCI-32765)	 <p>The structure of Ibrutinib consists of a central pyrazolo[1,5-a]pyrimidine core. It is substituted with a 4-phenoxyphenyl group, a 2-amino group, and a 1-(2-methyl-2-propenyl)pyrrolidine group.</p>	BTK inhibitor	First FDA approval in 2013 for mantle cell lymphoma – approved for many other B cell malignancies since 2013	Pan 2007; Honigberg 2010; Byrd 2013
Ritlecitinib (Pfizer; PF-06651600)	 <p>The structure of Ritlecitinib features a central pyrazolo[1,5-a]pyrimidine core. It is substituted with a 1-(2-methyl-2-propenyl)pyrrolidine group and a 2-(2-methyl-2-propenyl)amino group.</p>	JAK3 inhibitor	Ph III for Alopecia Areata, Ph II for many other inflammatory indications	Telliez 2016; Thorarensen 2017; Robinson 2020
Fisogatinib (Blueprint; BLU-554)	 <p>The structure of Fisogatinib features a central pyridine ring substituted with a 4-(2-methyl-2-propenyl)amino group and a 4-(2,4-dichloro-3,5-dimethoxyphenyl) group.</p>	FGFR4 inhibitor	Ph II for hepatocellular carcinoma	Kim 2019

Figure 2:

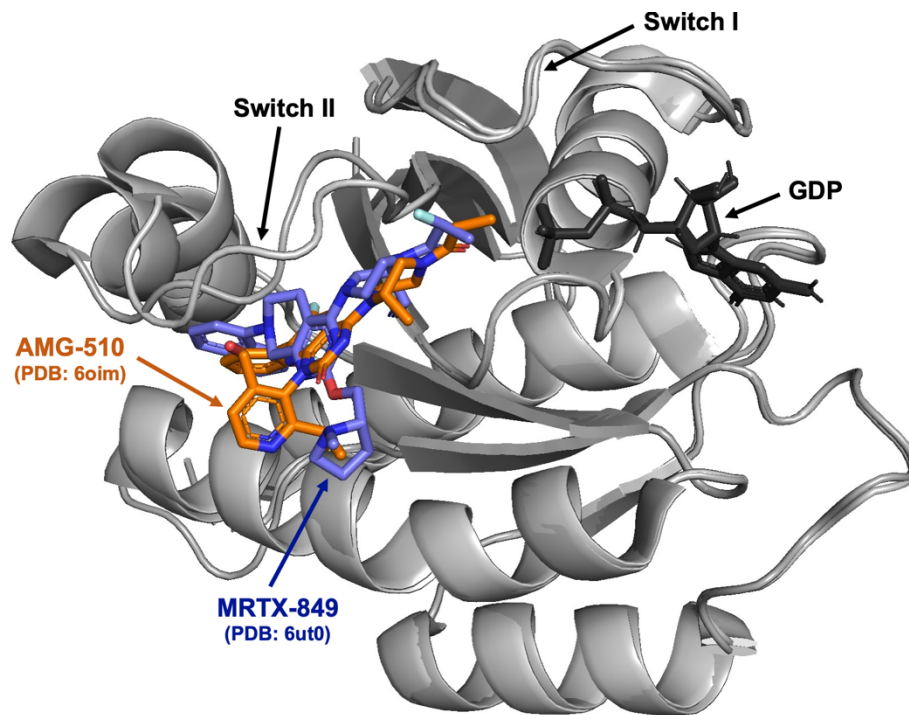
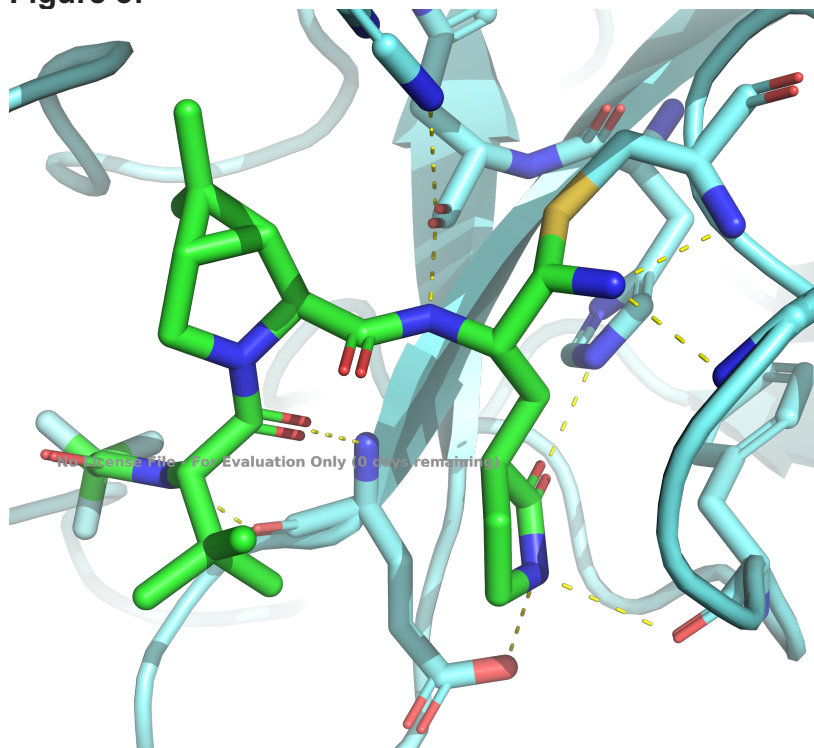


Figure 2: Aligned structures (PDB: 6ut0 and 6oim) of KRAS-G12C co-crystallized with MRTX-849 (adagrasib) and AMG-510 (sotorasib). Inhibitors are bound to the S-IIP, adjacent to the GDP binding pocket. Switch I and II are indicated.



**Figure 3:**



**Figure 3. Nirmatrelvir in Complex with SARS-CoV-2 Mpro.** The nitrile of nirmatrelvir has reacted with Mpro Cys145 to form a covalent thioimide adduct, and extensive hydrogen-bond interactions are shown throughout the pocket (PDB code 7RFW) (Owen 2021).

**Table 2. Highlighted Covalent Drugs**

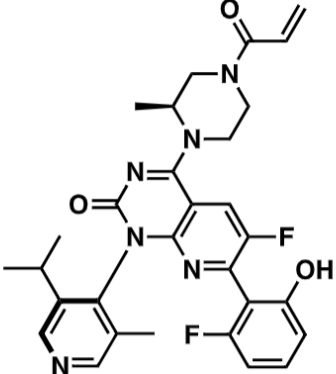
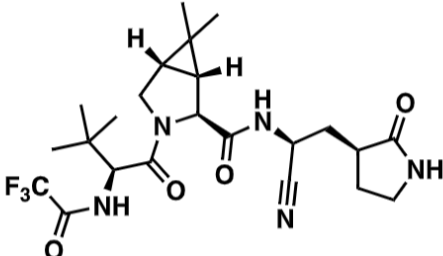
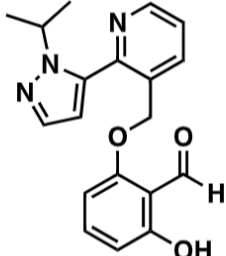
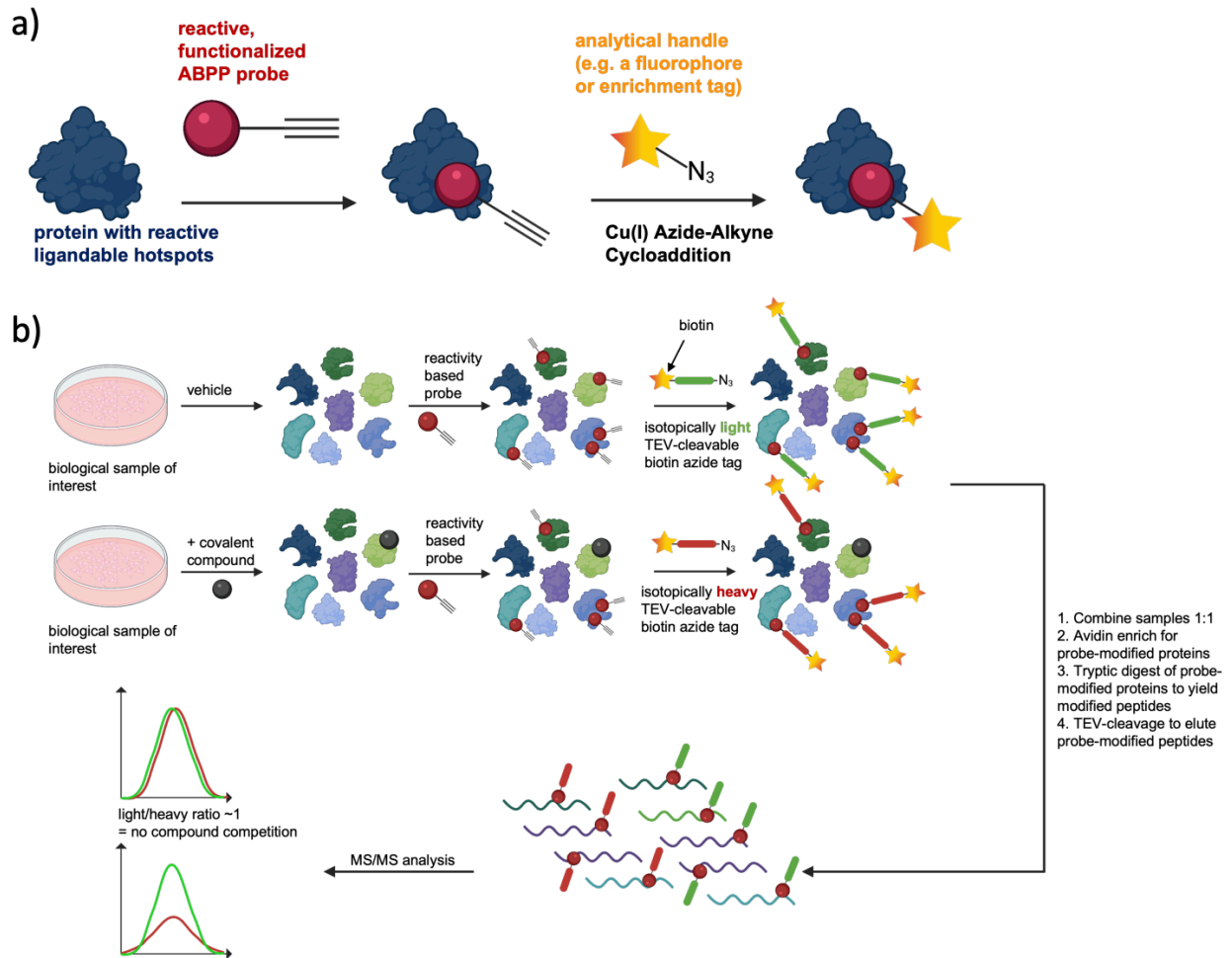
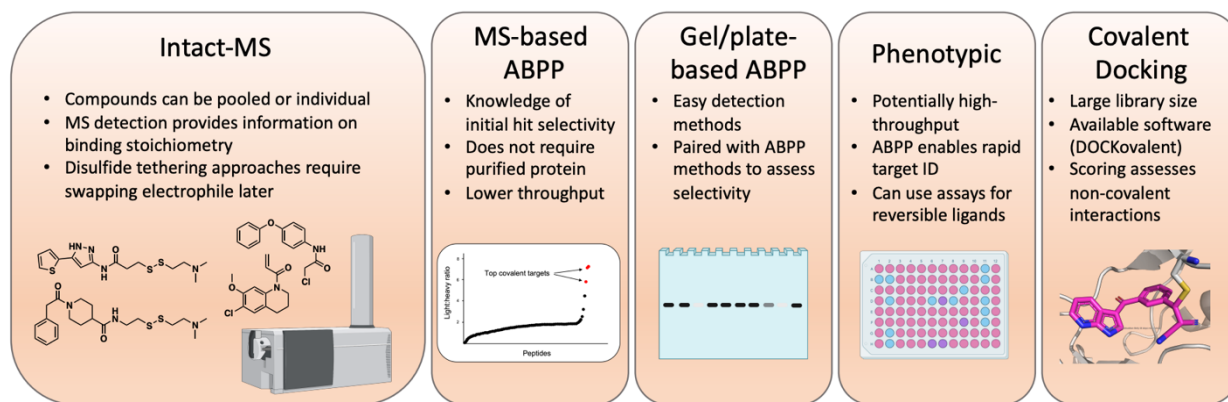
Drug (Company; Former compound name)	Structure	Target	Approval and Indication	Ref(s)
Sotorasib (Amgen; AMG510)	 <p>The chemical structure of Sotorasib (AMG510) is a complex heterocyclic molecule. It features a central pyridine ring substituted with a fluorine atom and a hydroxyl group. This central ring is connected to a piperazine ring system, which is further substituted with a vinyl ketone group and another piperazine ring. The molecule also includes a pyridine ring with a methyl group and a hydroxyl group, and a benzene ring with two fluorine substituents.</p>	KRAS <sup>G12C</sup> inhibitor	FDA Approved in 2021 for NSCLC with <i>KRAS<sup>G12C</sup></i> mutation	Lanman 2020; Canon 2019; Skoulidis 2021
Nirmatrelvir (Pfizer; PF- 07321332)	 <p>The chemical structure of Nirmatrelvir (PF-07321332) is a complex molecule with multiple stereocenters. It features a central piperidine ring system with a methyl group and a hydrogen atom explicitly shown. The structure includes a trifluoromethyl group (F<sub>3</sub>C), a cyano group (C≡N), and a pyrrolidine ring. There are several amide and ester-like linkages connecting different parts of the molecule.</p>	SARS-CoV-2 Mpro inhibitor	FDA approved in 2021 for COVID-19	Owen 2021; Hammond 2022
Voxelotor (Global Blood Therapeutics ; GBT-440)	 <p>The chemical structure of Voxelotor (GBT-440) consists of a central pyridine ring substituted with a methyl group and a hydroxyl group. This central ring is connected to a benzene ring, which is further substituted with a hydroxyl group and a formyl group (CHO). The structure also includes a pyridine ring with a methyl group and a hydroxyl group, and a benzene ring with a hydroxyl group and a formyl group.</p>	Mutant Hemoblobi n modulator	FDA approved in 2019 for sickle cell anemia	Metcalf 2017; Oksenber g 2016; Vichinsky 2019

Figure 4 (a-b):



**Figure 4: isotopic tandem orthogonal proteolysis–activity-based protein profiling. A)** Example of a reactive probe for ABPP designed with a broadly reactive electrophilic warhead linked to an analytical handle. **b)** Schematic of competitive isoTOP-ABPP methodology.

**Figure 5:**



**Figure 5. Screening methods for covalent drug discovery.** For each strategy, unique features and advantages are highlighted.

## CHAPTER 2

### Discovery of a Functional Covalent Ligand Targeting an Intrinsically Disordered Cysteine within Myc

(Adapted from published article by Lydia Boike, Alexander G Cioffi, Felix C Majewski, Jennifer Co, Nathaniel J Henning, Michael D Jones, Gang Liu, Jeffrey M McKenna, John A Tallarico, Markus Schirle, Daniel K Nomura DOI: 10.1016/j.chembiol.2020.09.001)

## Discovery of a Functional Covalent Ligand Targeting an Intrinsically Disordered Cysteine within Myc

MYC is a nuclear transcription factor that regulates the expression of genes involved in cell proliferation, metabolism, and survival and is the most frequently amplified oncogene in human cancers.<sup>205,207</sup> MYC activates its transcriptional targets through obligate heterodimerization with MAX and binding to the E-box sequences. Due to its fundamental importance to cancer pathogenesis, significant effort has been invested in trying to develop inhibitors against MYC or pathways that regulate MYC.<sup>199,207,208,211</sup> Strategies for MYC pathway inhibition have included targeting MYC transcription with bromodomain and extraterminal (BET) family inhibitors or CDK7/9 inhibitors, targeting MYC translation with mTORC1 or AKT/PI3K inhibitors, or targeting the heterodimeric complex between MYC and MAX. Examples of these small-molecules include bromodomain-containing 4 (BRD4) inhibitors such as JQ1 or GSK525762 that displace BRD4 binding to sites of hyperacetylated chromatin that regulate MYC transcription, and mTORC1 inhibitors such as rapamycin and everolimus that inhibit mTORC1-dependent phosphorylation of the translation initiation factor 4E binding protein 1 (4EBP1) and inhibit MYC translation.<sup>195,207</sup> Studies using small-molecule microarrays have yielded MAX/MAX homostabilizers that sequester MAX away from MYC leading to inhibition of MYC transcriptional activity.<sup>210</sup> However, direct targeting of MYC itself has remained challenging and it has often been considered “undruggable” because much of MYC is intrinsically disordered and there are no obvious binding pockets or ligandable sites for pharmacological interrogation.<sup>195,199,207</sup>

Covalent ligand screening using chemoproteomic approaches such as activity-based protein profiling (ABPP) has arisen as a powerful strategy for rapid discovery of covalent ligands, ligandable hotspots, and therapeutic modalities against proteins that are often considered intractable to standard drug discovery efforts.<sup>138,193,143,200,148,137</sup> Because of the irreversible binding of covalent ligands to the target of interest, mass-spectrometry-based proteomic approaches can be used to rapidly identify the site of modification of hit ligands without structural elucidation of the protein-ligand complex—an effort that would be hindered for largely intrinsically disordered proteins like MYC.<sup>148,150</sup> In this study, we sought to identify a covalent compound directly targeting a ligandable site on MYC that would functionally impair MYC transcriptional activity.

### 2.1 Covalent Ligand Screening to Identify MYC Inhibitors

To identify direct and covalently-acting inhibitors against c-MYC, we screened a library of 98 primarily cysteine-reactive covalent ligands consisting of acrylamides and chloroacetamides *in vitro* to identify compounds that would inhibit the binding of purified recombinant human MYC/MAX complex to the E-box DNA consensus sequence (**Fig. 1a, Table S1**). These compounds were from a collection of compounds purchased from Enamine. While human MAX does not contain any cysteine residues, we wanted to avoid potential off-chance reactivity of the covalent ligands with another nucleophilic amino acid on MAX. To maximize our chances of identifying covalent ligands that directly targeted a cysteine on MYC, we pre-incubated covalent ligands with MYC prior

to the addition of MAX and then added this MYC/MAX complex to plates bearing the E-box consensus sequence. From this screen EN4 was identified as the top hit (**Fig. 1a**).

To further identify covalent ligands that could inhibit MYC transcriptional activity in cells *in situ*, we also screened these cysteine-reactive ligands in a MYC luciferase reporter assay in HEK293T cells (**Fig. 1b, Table S1**). This reporter assay employed a MYC-responsive Firefly luciferase reporter, encoding a luciferase reporter gene under the control of a basal promoter element, the TATA box, joined to tandem repeats of the MYC transcriptional E-box response element, alongside a constitutively expressing Renilla luciferase for normalization purposes. We were particularly interested in covalent ligands that inhibited both MYC/MAX DNA binding *in vitro* and MYC transcriptional activity *in situ*, as these compounds may be functionally inhibiting MYC activity in cells through directly binding to MYC itself. From these two screens, we identified one primary hit EN4 that bears a cysteine-reactive acrylamide warhead (**Fig. 1c**). EN4 showed the strongest inhibition of both MYC/MAX binding to its DNA consensus sequence *in vitro* as well as MYC transcriptional activity in cells (**Fig. 1a-1c**). EN4 inhibited MYC/MAX binding to the E-box response element DNA consensus sequence in a dose-responsive manner with a 50 % inhibitory concentration (IC<sub>50</sub>) value of 6.7  $\mu$ M (**Fig. 1d**). We further showed that EN4 inhibited MYC/MAX DNA binding regardless of whether EN4 was pre-incubated with MYC before addition of MYC or incubated with the pre-formed MYC/MAX complex (**Fig. S1a-S1b**). EN4 also inhibited MYC luciferase reporter activity in a dose-responsive manner with an IC<sub>50</sub> value of 2.8  $\mu$ M (**Fig. 1e**). We note that EN4 is fully soluble in aqueous media at 50  $\mu$ M, the maximum concentration used in this study, but this compound is not soluble at concentrations of 100  $\mu$ M or higher.

## 2.2 Characterization of EN4 Interactions with c-MYC *In Vitro*

We expected that EN4 would undergo Michael addition to a cysteine on MYC (**Fig. 2a**). As such, we next used liquid-chromatography-tandem mass spectrometry (LC-MS/MS) to identify the single site of EN4 labeling on pure human MYC as C171 (**Fig. 2b, Fig. S1c**). Intriguingly, C171 belongs to a predicted intrinsically disordered region of MYC for which there is no structural information within the transactivation and transrepression N-terminal domain (NTD) far from the E-Box or MAX binding C-terminal domain (CTD) (**Fig. 2c**). To further confirm that EN4 was covalently binding to MYC, we synthesized 4 different alkyne-functionalized EN4 probes (EN4-alkyne-1, -2, -3, -4) (**Fig. 2d, Supplemental Methods**). All four of these probes showed labeling of pure human MYC protein on a denaturing gel after appending a rhodamine-azide onto probe-labeled proteins by copper-catalyzed azide-alkyne cycloaddition (CuAAC) (**Fig. 2d**). We also showed that EN4 displaced the labeling of EN4-alkyne-2 labeling of pure human MYC protein (**Fig. 2e**). Further confirming interaction of these compounds with C171, EN4-alkyne-2 labeling was not observed with pure human MYC cysteine 171-to-tyrosine mutant (C171Y) mutant protein compared to wild-type MYC (**Fig. 2f**). Further demonstrating that C171 on MYC is the functional target of EN4, EN4-mediated inhibition of MYC/MAX binding to its DNA consensus sequence was significantly attenuated with C171Y mutant MYC protein compared to wild-type MYC protein (**Fig.**

**2g).** Thus, our data show that EN4 directly and covalently modifies pure full-length c-MYC protein at C171 to inhibit MYC/MAX DNA binding *in vitro*. Interestingly, this C171 on c-MYC is not conserved in N-MYC or L-MYC and may suggest that EN4 may be selective for c-MYC over N-MYC and L-MYC (**Fig. S2**).

### 2.3 Characterization of EN4 Engagement of MYC in Cells

We next sought to further characterize EN4 engagement of MYC in cells. We first showed that EN4-alkyne-4 probe treatment *in situ* in 231MFP cells significantly enriched endogenous c-MYC compared to vehicle-treated controls after appending an azide-functionalized biotin handle by CuAAC, avidin-enrichment, and blotting for MYC (**Fig. 3a**). Enrichment was not observed with an unrelated protein vinculin (**Fig. 3a**). We note that these c-MYC enrichment studies with the EN4-alkyne probe required substantially more protein input and high-capacity avidin beads than pulldown studies that we have previously performed. This may be due to low abundance or overall instability of MYC. In this experiment, we were also not able to compete the observed c-MYC enrichment by excess EN4, since EN4 was not soluble above the 50  $\mu$ M concentration of EN4-alkyne-4 needed to observe this c-MYC enrichment. Next, further reinforcing EN4 engagement of endogenous c-MYC in cells, we performed a cellular thermal shift assay (CETSA) and showed that EN4 treatment significantly reduced MYC thermal stability in 231MFP breast cancer cells (**Fig. 3b**).<sup>206,196</sup> EN4 also reduced the thermal stability of the MYC heterodimer partner MAX as well (**Fig. 3b**). The reduced thermal stability of MYC with EN4 treatment is further confirmation that EN4 engages MYC in cells. Furthermore, our data suggest that the mechanism of EN4-mediated MYC inhibition may be through covalent targeting of MYC leading to destabilization of MYC folding. This may in-turn impair the formation of either MYC/MYC homodimer or MYC/MAX heterodimer complexes. The reduction in MAX thermal destabilization would seem to reinforce this hypothesis and would be consistent with MAX instability due to either loss of its heterodimer complex partner MYC leading to instability of free MAX and/or potentially higher levels of a less thermally stable MAX homodimer complex. These data collectively indicate that EN4 or its close analog EN4-alkyne-4 engage endogenous MYC in 231MFP cells.

To understand the proteome-wide cysteine reactivity of EN4 in cells, we next performed isotopic tandem orthogonal proteolysis-enabled ABPP (isoTOP-ABPP) to investigate the proteome-wide cysteine-reactivity and selectivity of EN4 in 231MFP breast cancer cells using the cysteine-reactive alkyne-functionalized iodoacetamide probe (IA-alkyne). While we were not able to observe MYC itself in this isoTOP-ABPP experiment, likely due to its low overall abundance or potential instability, we did show that EN4 was selective across >1400 probe-modified peptides, in which we only observed 8 other off-targets of EN4 showing an isotopically light versus heavy or control versus EN4-treated probe-modified peptide ratios of greater than 3 (**Fig. S3a; Table S2**). Intriguingly, one of the off-targets of EN4 from isoTOP-ABPP studies was another important transcription factor JUN (**Fig. S3a**), which was also an off-target of previous MYC-directed probes.<sup>194</sup> In the isoTOP-ABPP data, EN4 targets C99 of JUN. While sequence alignment analysis does not indicate that C171 is conserved between c-MYC



and JUN, C99 of JUN aligns with C117 of c-MYC, suggesting potential structural features that may be in common between these proteins that give rise to common pharmacology (**Fig. S2**). We confirmed that JUN was an off-target of EN4 by gel-based ABPP showing competition of EN4 against labeling of pure JUN protein with the cysteine-reactive rhodamine-functionalized iodoacetamide probe (IA-rhodamine) (**Fig. S3b**). However, EN4 still exhibited comparable inhibitory effects upon MYC luciferase reporter activity upon JUN knockdown compared to control cells, suggesting that JUN was likely not a major driver of the EN4 inhibitory effects upon MYC transcriptional activity (**Fig. S3c-S3d**).

We also attempted to perform chemoproteomic pulldown studies with the EN4-alkyne-4 probes, but unfortunately, we were not able to observe c-MYC in these studies either and we were not able to perform competition studies with excess EN4 because of solubility issues encountered with beyond 50  $\mu$ M. This again may be due to the relative low abundance and instability of c-MYC in cells. We performed gel-based ABPP analysis of EN4-alkyne-4 labeling in 231MFP cells and showed 5 protein bands that were competed by EN4, which was comparable to the 8 off-targets identified by isoTOP-ABPP (**Fig. S3e**).

## 2.4 Characterization of EN4 Inhibitory Activity against MYC in Cells

Because EN4 was not completely selective for C171 of MYC, we wanted to confirm that this *in situ* MYC inhibitory activity was due to EN4 engagement of C171. Since knocking down the expression of endogenous MYC was not successful, likely due to its fundamental importance in cell viability, we overexpressed either wild-type (WT) or a C171Y mutant FLAG-MYC in HEK293T cells (**Fig. 4a**). The MYC inhibitory effect of EN4 was significantly attenuated in cells expressing C171Y FLAG-MYC compared to WT FLAG-MYC, indicating that the EN4 inhibition of MYC activity in cells was at least in-part driven through targeting C171 on MYC (**Fig. 4b**).

We next investigated whether EN4 affects MYC-regulated transcriptional targets in 231MFP breast cancer cells. These highly aggressive and tumorigenic triple-negative breast cancer cells are derived from MDA-MB-231 cells that have been previously shown to be driven by MYC.<sup>202,213,214</sup> We performed RNA-sequencing on EN4-treated 231MFP cells to identify gene expression changes (**Fig. 4c-4d**). EN4 treatment significantly downregulated the expression of both direct MYC target genes previously confirmed by CHIP-Seq as well as MYC-regulated gene targets found in the Hallmark gene sets from the MSigDB Team (**Fig. 4c-4d**).<sup>215</sup> Consistent with EN4 inhibition of the MYC pathway, Signature Commons search of EN4-mediated transcriptional changes using the MSigDB Collection H hallmark gene set showed MYC target genes as one of the most significantly enriched pathways, alongside E2F and G2M checkpoint target genes that are also related to MYC and cell proliferation pathways (**Fig. 4e**).<sup>197,209</sup> We further demonstrated that protein levels of representative MYC-regulated target genes including CDK2 and CDC25A were significantly lowered with EN4 treatment (**Fig. 4f, 4g**).<sup>192,203,216</sup> Consistent with previous studies showing that MYC inhibitors downregulate MYC levels in cells either through auto-regulation or destabilization and

degradation, we also showed that MYC levels were substantially reduced with EN4 treatment (**Fig. 4f, 4g**).<sup>201,210</sup> In addition, we observed the upregulation of a MYC binding partner MIZ1 or ZBTB17 with EN4 treatment (**Fig. S3f**).<sup>212</sup>

We next investigated whether EN4 affects cell proliferation, cell survival, or tumorigenesis of 231MFP breast cancer cells. EN4 treatment significantly impaired 231MFP breast cancer cell proliferation in a dose- and time-dependent manner with >90 % inhibition of proliferation at 50  $\mu\text{M}$  (**Fig. 4h; Fig. S3g**). Further demonstrating the dependence of observed EN4 effects on MYC, EN4 impaired the cell survival of MYC-transformed mammary epithelial MCF10A cells, but not parental MCF10A cells that are not dependent on MYC (**Fig. 4i**). Daily EN4 treatment initiated after establishment of a 231MFP breast tumor xenograft in immune-deficient mice also significantly attenuated tumor growth *in vivo* (**Fig. 4j**). The striking anti-tumorigenic effects of EN4 *in vivo* despite the relatively poor potency of EN4, may be due to the covalent nature of EN4, in which MYC would remain inhibited by EN4 until protein turnover even if EN4 compound levels may have rapidly cleared from the body. Nonetheless, future studies will be required to further ascertain pharmacokinetic and pharmacodynamic parameters of EN4.

## 2.5 Structure-Activity Relationship of EN4 Analogs

Finally, we sought to preliminarily define the structure-activity relationship between EN4 and MYC in the absence of any structure for the region of MYC encompassing C171. We synthesized 18 analogs and tested these analogs and the previously described EN4-derived alkyne probes in dose response studies in MYC luciferase reporter assays in HEK293T cells (**Supplementary Methods**). Among these analogs, we showed that EN4-2 and EN4-alkyne-4 showed the best activity with EC50s of 1.7 and 1.4  $\mu\text{M}$ , respectively, compared to EN4 EC50 of 12.5  $\mu\text{M}$  (**Table S4**). In contrast, adding the electron-withdrawing trifluoromethyl group on the acrylamide warhead (EN4-6) or changing the acrylamide warhead to a chloroacetamide group (EN4-3) led to loss of activity. Furthermore, the attachment position of the phenoxybenzene moiety is critical for activity. Five analogs (EN4-13, EN4-14, EN4-15, EN4-16, and EN4-17) with either *meta* or *para* positioning relative to the amide all showed loss of activity (**Table S4**). Interestingly, we also demonstrate that the non-reactive EN4 analogs EN4-18 also inhibited MYC luciferase reporter activity, suggesting that EN4 may possess inherent affinity for the apparent site within MYC that contains C171 (**Table S4**). Consistent with this premise, we show that EN4-18 significantly inhibits MYC/MAX DNA binding *in vitro* (**Fig. S4**). Despite EN4 targeting an apparently intrinsically disordered region of MYC with no structural insights, we still observe some structure-activity relationships that has enabled us to improve potency of EN4 by 7-8-fold.

## 2.6 Discussion

Overall, we have screened a cysteine-reactive covalent ligand library against MYC and have identified a hit compound EN4 that targets a new ligandable site C171

within MYC to reduce MYC and MAX thermal stability, inhibit MYC/MAX DNA binding *in vitro*, impair MYC transcriptional activity in cells, downregulate many known MYC target genes as well as MYC itself, and inhibit proliferation and tumorigenesis in breast cancer cells. Despite targeting a predicted intrinsically disordered region within MYC, preliminary medicinal chemistry efforts gave rise to structure-activity relationships and yielded improved compounds. This study adds onto the pharmacopeia of small-molecules that target the MYC pathway, which include various modalities such as BRD4 and mTORC1 inhibitors and MAX/MAX stabilizers.<sup>195,207,210,211</sup> Sequence alignment indicates that C171 of c-MYC is not conserved in N-MYC or L-MYC suggesting that EN4 may selectively target c-MYC over other MYC proteins. Additionally, other forms of MYC have been previously identified such as a cytoplasmic form of MYC, known as MYC-nick, which retains the conserved MYC box regions but lacks nuclear localization signals and the domains necessary for MAX heterodimerization and DNA binding. MYC-nick has unique cytoplasmic functions in regulating actin acetylation and cell morphology.<sup>198</sup> Since this MYC-nick retains the NTD, EN4 which targets C171 within the NTD, may be capable of also targeting this proteolytically cleaved form of MYC. Future studies are necessary to determine whether EN4 and analogs selectively target c-MYC and MYC-nick over other MYC proteins, and whether unique covalent targeting of the NTD by EN4 may open avenues to explore MYC function compared to other indirect or direct MYC inhibitors.

While EN4 showed overall selectivity on a proteome-wide scale in profiling >1500 cysteines using competitive ABPP platforms, EN4 was not fully selective for c-MYC as indicated by 8 potential off-targets identified in this study. One of the off-targets is another important transcription factor involved in cell proliferation JUN.<sup>204</sup> While we show here that the inhibitory action of EN4 towards MYC/MAX DNA binding and MYC transcriptional reporter activity is significantly attenuated upon mutation of C171 to a tyrosine, some of the activity observed here may be due to a combination of inhibiting MYC and some of the other EN4 off-targets. Now that we have shown the feasibility of direct and functional covalent targeting of c-MYC, future medicinal chemistry and expanded covalent ligand screening efforts are necessary to further optimize the potency and selectivity of EN4 as well as to discover additional scaffolds that may target c-MYC.

Our study also highlights how intrinsically disordered regions of classically “undruggable” targets may be functionally targeted with small-molecules and underscores the utility of covalent ligand screening coupled with proteomic approaches to uncover new ligands against undruggable targets like MYC. While more work needs to be done, these data along with our recent case where we discovered that nimbolide functionally targeted a predicted intrinsically disordered region within the E3 ligase RNF114, suggest that covalent ligand discovery might be a viable approach towards drugging intrinsically disordered proteins since it allows facile mass-spectrometry-based identification of modification events and ligandable pockets that might be unique to one of several conformations that these proteins can adopt.<sup>148</sup> Most importantly, covalent ligand binders and site of covalent modification can be identified in the absence of protein structures, unlike with reversible ligands or hits arising from DNA-encoded

libraries where structural biology efforts are often needed to understand mode of ligand-protein binding. While not having structural enablement for a protein such as MYC may hamper medicinal chemistry efforts, we show here that one can still generate preliminary SAR against a predicted intrinsically disordered region within MYC.

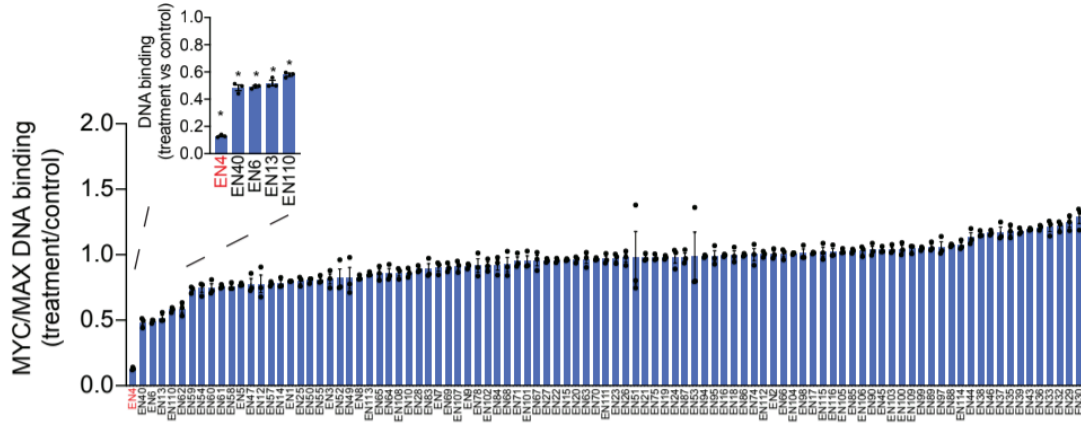
In this study, we coupled *in vitro* biochemical screens of covalent ligands with pure MYC/MAX DNA binding assays and cellular MYC luciferase reporter assays to identify our initial hits. Future hit finding and medicinal chemistry efforts would also be benefited by intact mass adduct screening for covalent ligand/c-MYC protein adducts by mass spectrometry-based screening platforms,<sup>165</sup> thermal shift assays searching for MYC/MAX protein destabilizers as employed in this study, as well as more detailed kinetic studies to investigate the reversible and irreversible binding components of EN4 and optimized inhibitors to MYC. A future optimized covalent ligand targeting C171 of c-MYC would presumably show potent MYC/MAX DNA binding inhibition, possess inherent affinity for the site that encompasses C171 of c-MYC, confer a single mass adduct on c-MYC by intact mass screening, show better proteome-wide selectivity using chemoproteomic methods, inhibit MYC transcriptional activity in cells, confer reduced thermal stability of MYC and MAX in cells, downregulate MYC target genes, and impair cancer pathogenicity in a MYC-dependent manner. Understanding these various parameters will not only aid in optimizing covalent c-MYC inhibitors, but also in future pharmacological targeting of intrinsically disordered regions.

Overall, we demonstrate the utility of covalent ligand screening and chemoproteomic platforms in discovering covalent and functional ligands that target an intrinsically disordered C171 within c-MYC. Our results also further underscore the general utility of covalent ligand screening approaches for rapidly discovering and pharmacologically interrogating unique ligandable sites within intractable disease targets.

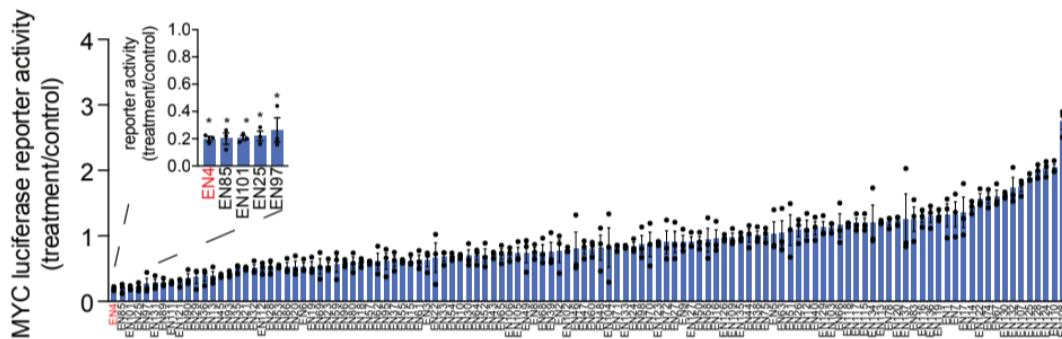
## 2.7 Figures and Tables

Figure 1 (a-e):

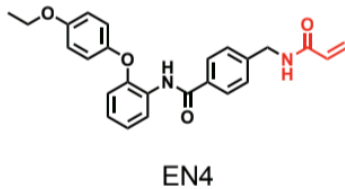
a



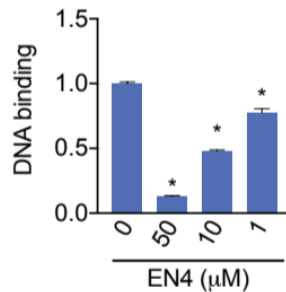
b



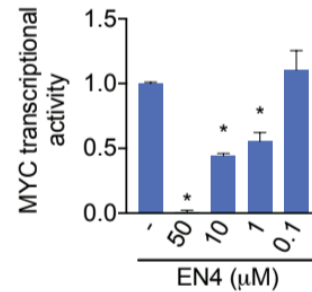
c



d



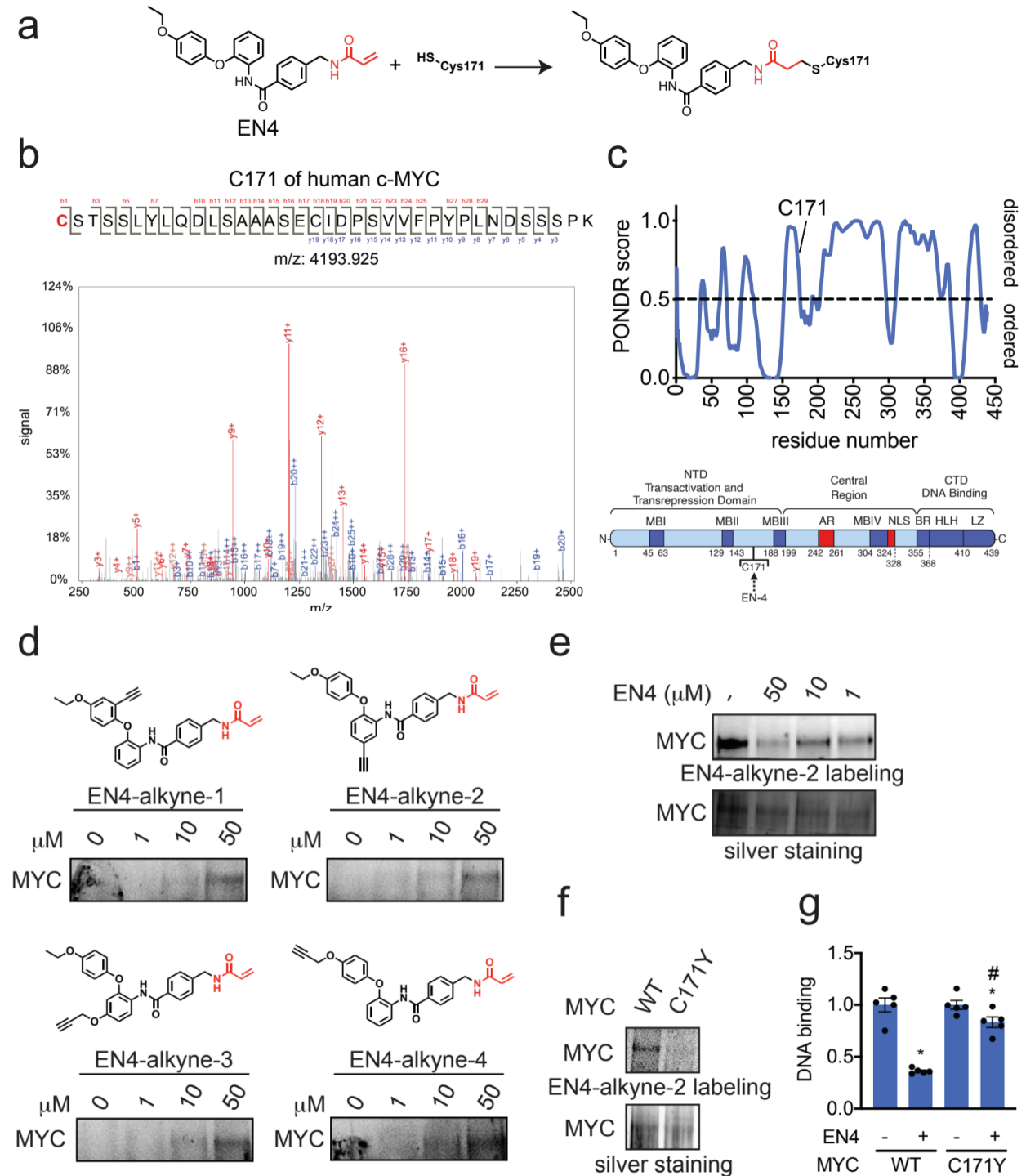
e



**Figure 1. Covalent ligand screen to identify compounds that inhibit MYC activity *in vitro* and *in situ*.** (a) Screening a cysteine-reactive covalent ligand library *in vitro* for compounds that would inhibit MYC/MAX binding to its E-box DNA consensus sequence. DMSO vehicle or covalent ligands (50  $\mu$ M) were pre-incubated with pure human full-

length MYC protein for 30 min before direct addition of MAX and then addition to DNA binding plates for 1 h. Data are shown as ratio relative to DMSO vehicle treated controls set to 1. Shown in red was the top hit. **(b)** MYC luciferase reporter activity in HEK293T cells transfected with a MYC luciferase reporter construct encoding a Firefly luciferase reporter gene under the control of a basal promoter element, the TATA box, joined to tandem repeats of the MYC transcriptional E-box response element, alongside a constitutively expressing Renilla luciferase for normalization purposes. DMSO vehicle or cysteine-reactive covalent ligands (50  $\mu$ M) were treated in HEK293T cells transfected with a MYC luciferase reporter gene for 24 h and Firefly and Renilla luciferase activity was read-out. Firefly luciferase levels were normalized to Renilla luciferase levels and then normalized to vehicle-treated controls. Data are shown as ratios relative to DMSO vehicle treated controls. Shown in red was the top hit. **(c)** Structure of top hit in both *in vitro* and *in situ* assays in **(a)** and **(b)** EN4. Shown in red is the cysteine-reactive acrylamide. **(d)** Dose-response of EN4 inhibition of MYC/MAX binding to its DNA consensus sequence. **(e)** Dose-response of EN4 inhibition of MYC luciferase reporter activity in cells. DMSO vehicle or EN4 was treated in HEK293T cells transfected with a MYC luciferase reporter gene for 24 h and luciferase activity was read-out. Data are shown as ratio relative to DMSO vehicle treated controls set to 1. Data shown in **(a, b, d, e)** are average  $\pm$  sem, n=3 biologically independent samples/group. Statistical significance was calculated with unpaired two-tailed Student's t-tests in **(d, e)** and is expressed as \*p<0.05. This figure is related to **Figure S1**.

**Figure 2 (a-g):**



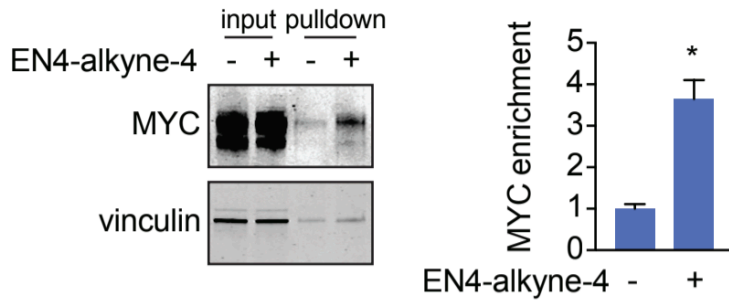
**Figure 2. Characterization of EN4 binding to MYC. (a)** Structure of EN4 cysteine adduct on C171 of MYC. **(b)** MS/MS analysis of EN4 adduct on MYC C171. Pure human full-length MYC was incubated with EN4 (50  $\mu$ M) for 30 min. MYC was then subjected to tryptic digestion and tryptic digests were analyzed by LC-MS/MS. **(c)** PONDR prediction of MYC protein folding and location of C171 showing that it is part of

an intrinsically disordered region of MYC and a domain map of MYC. **(d)** 4 alkyne-functionalized EN4 probes were tested for covalent binding to MYC by gel-based ABPP methods. Pure human full length MYC protein was incubated with DMSO vehicle or each probe for 2 h at 37°C, after which rhodamine-azide was conjugated to probe-labeled proteins by CuAAC, separated by SDS/PAGE, and visualized by in-gel fluorescence. **(e)** Gel-based ABPP analysis of EN4 interaction with MYC. EN4 displaces EN4-alkyne-2 labeling of pure human MYC protein. DMSO vehicle or EN4 was pre-incubated with MYC for 30 min at 37°C, prior to EN4-alkyne-2 labeling (50 µM) for 2 h at 37°C. Rhodamine-azide was appended to probe-modified proteins by CuAAC, subjected to SDS/PAGE and analyzed by in-gel fluorescence. Shown also is silver staining of the gel to show equal protein loading. **(f)** gel-based ABPP analysis of EN4-alkyne-2 labeling of pure human wild-type (WT) and C171Y mutant MYC protein. Probe labeling with pure protein was performed with EN4-alkyne-2 (50 µM) for 2 h at 37°C. **(g)** MYC/MAX DNA binding with EN4 treatment. DMSO vehicle or EN4 (50 µM) were pre-incubated with pure human full-length wild-type or C171Y mutant MYC protein for 30 min before direct addition of MAX and then addition to DNA binding plates for 1 h. Gels for **(d, e, f)** are representative of n=3 biologically independent samples/group. Data shown in **(g)** show individual replicate values and average ± sem, n=5 biologically independent samples/group. Statistical significance was calculated with unpaired two-tailed Student's t-tests in **(g)** and is expressed as \*p<0.05 compared to respective vehicle-treated control groups and #p<0.05 compared to EN4-treated group with WT MYC protein. This figure is related to **Figure S1** and **Figure S2**.

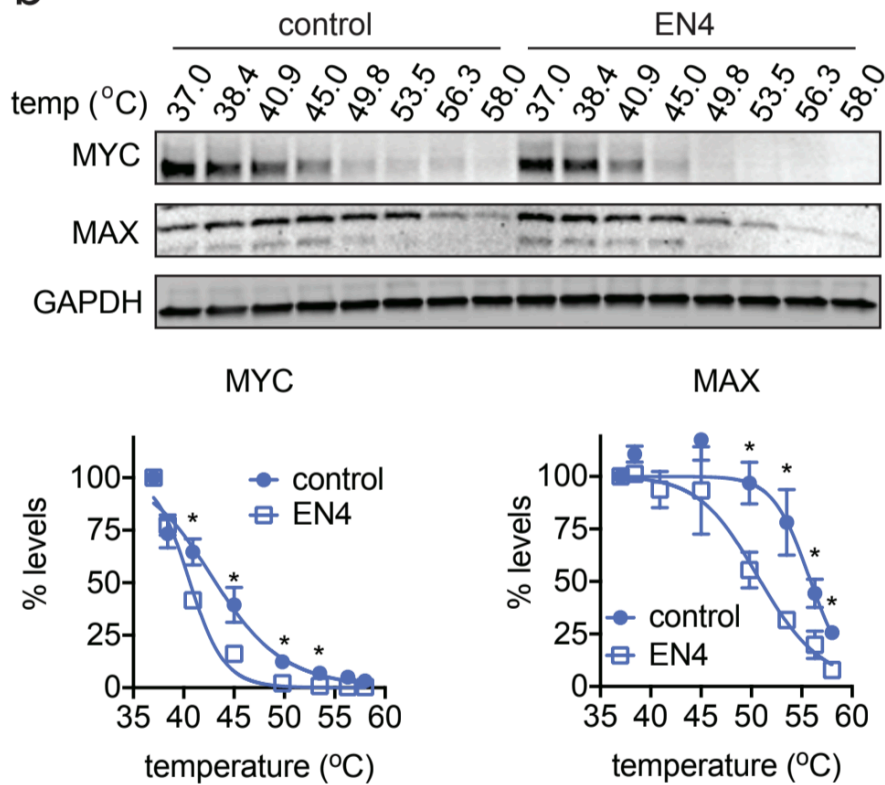


**Figure 3 (a-b):**

**a**



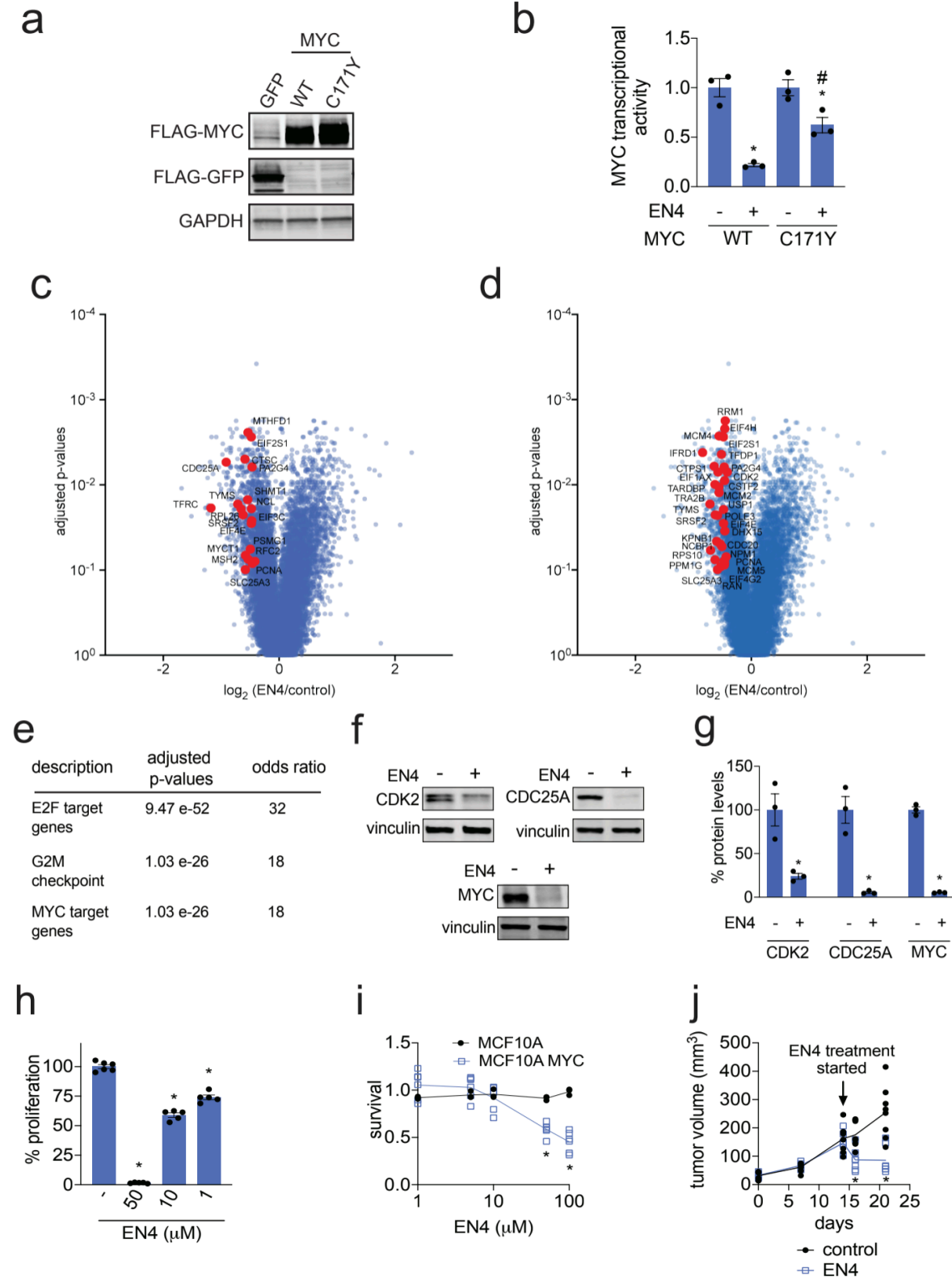
**b**



**Figure 3. Characterizing EN4 engagement of MYC in cells. (a)** EN4-alkyne-4 enrichment of endogenous MYC in 231MFP breast cancer cells. 231MFP cells were treated with DMSO vehicle or EN4-alkyne-4 (50  $\mu$ M) for 4 h. Resulting cell lysates were subjected to CuAAC with an azide-functionalized biotin handle and EN4-alkyne-4-labeled proteins were subsequently avidin-enriched, run on SDS/PAGE, and blotted for c-MYC and loading control vinculin. Also shown is c-MYC and vinculin expression from the input protein that went into the pull-down. Gel shown is a representative image of  $n=3$  biologically independent samples/group and fluorescence is quantified in the bar graph shown on the right. **(b)** Cellular thermal shift assay showing c-MYC, MAX, and loading control GAPDH protein levels in 231MFP cells treated with vehicle DMSO or EN4 (50  $\mu$ M) for 2 h. Shown are representative gels and quantification of MYC and MAX protein levels,  $n=4$  biologically independent samples/group normalized to MYC or MAX

protein levels at 37° C, respectively. Data shown in **(a, b)** are average  $\pm$  sem. Statistical significance was calculated with unpaired two-tailed Student's t-tests and is expressed as \* $p < 0.05$  compared to vehicle-treated controls in **(a)** and vehicle-treated controls at each temperature in **(b)**. This figure is related to **Figure S3**.

**Figure 4 (a-j):**



**Figure 4. Characterizing EN4 inhibitory activity against MYC in cells. (a)** Protein expression of FLAG-tagged GFP, MYC wild-type (WT), or MYC C171Y mutant protein

and GAPDH loading control. FLAG-GFP, FLAG-MYC WT or FLAG-MYC C171Y mutant protein were transiently overexpressed in HEK293T cells. **(b)** MYC luciferase reporter activity in HEK293T cells transiently expressing FLAG-MYC WT or FLAG-MYC C171Y mutant protein. DMSO vehicle or EN4 (50  $\mu$ M) was treated in HEK293T cells for 24 h transiently expressing FLAG-MYC WT or FLAG-MYC C171Y protein transfected with a MYC luciferase reporter gene, and luciferase activity was read-out. Data are shown as ratio relative to DMSO vehicle treated controls set to 1. **(c, d)** RNA sequencing analysis of EN4-mediated transcriptional changes in 231MFP cells. 231MFP cells were treated with DMSO vehicle or EN4 (50  $\mu$ M) for 24 h. Shown are log<sub>2</sub> fold change ratios of EN4 versus DMSO vehicle-treated controls. Shown in red in **(c)** are direct MYC targets confirmed by CHIP-seq and in **(d)** are MYC-regulated gene targets that are downregulated with EN4-treatment with adjusted p-values <0.1 with fold-change <0.75 comparing EN4 to control. All RNAseq data can be found in **Table S3**. **(e)** Signature enrichment analysis of genes with an FDR < 0.1 and Fold change  $\leq$  0.75 (567 genes) using the Signature Commons platform search of the MSigDB Collection H hallmark gene set. Full analysis is shown in **Table S3**. **(f)** Protein levels of known MYC transcriptional targets CDK2 and CDC25A as well as MYC itself and loading control vinculin, assessed by Western blotting. 231MFP breast cancer cells were treated with DMSO vehicle or EN4 (50  $\mu$ M) for 36 h, with a replenishment of media and with vehicle or EN4 treatment at 24 h. Shown are representative blots from n=3 biologically independent replicates/group. **(g)** Quantitation of Western blots from **(f)**. **(h)** Cell proliferation of 231MFP breast cancer cells. Cells were treated once per day daily with DMSO vehicle or EN4 for 72 h and cell proliferation was assessed by Hoechst staining. Data are shown as percent proliferation compared to DMSO vehicle control. **(i)** Serum-free cell survival in MCF10A parental or MYC-expressing cells. MCF10A parental or MYC-overexpressing cells were treated with DMSO vehicle or EN4 for 48 h and cell survival as assessed by Hoechst staining. Data are shown in relation to DMSO vehicle controls for respective lines set to 1. **(j)** Tumor xenograft growth of 231MFP breast cancer cells in immune-deficient SCID mice. 231MFP breast cancer cells were injected subcutaneously into the flank of SCID mice and vehicle or EN4 (50 mg/kg ip) treatment was initiated after 14 days of tumor implantation. Vehicle or EN4 were treated in mice subsequently once per day and tumor volume was measured by calipers. Data shown in **(a, b, f, g, h, j)** are average  $\pm$  sem, n=3-8 biologically independent samples/group. Statistical significance was calculated with unpaired two-tailed Student's t-tests and is expressed as \*p<0.05 compared to vehicle-treated controls and in **(a, b, c, d, f, g, h, and j)** and compared to respective EN4 concentrations in MCF10A cells in **(i)**, #p<0.05 compared to EN4-treated WT group in **(f)**. This figure is related to **Figure S3** and **Table S3**.

**Table S1. Covalent ligand structures and screening data for data shown in Figure 1a, 1b.**

Available in supplementary downloadable files.

**Table S2. IsoTOP-ABPP analysis of EN4 in 231MFP breast cancer cells shown in Figure S3.**

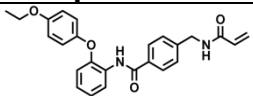
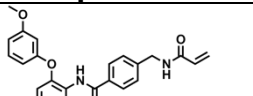
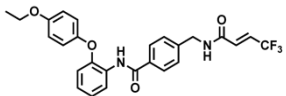
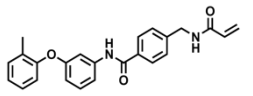
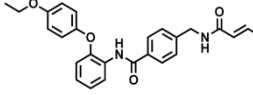
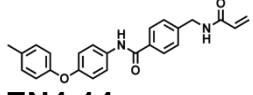
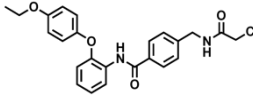
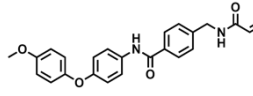
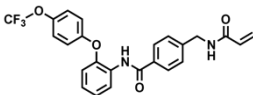
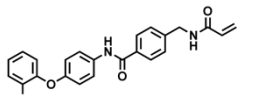
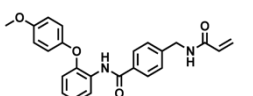
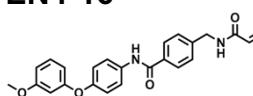
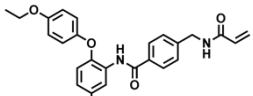
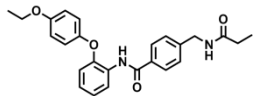
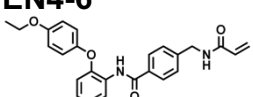
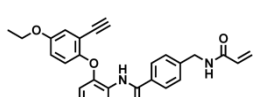
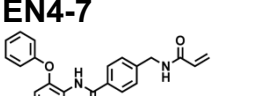
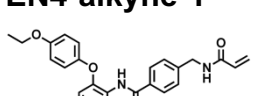
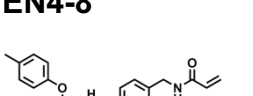
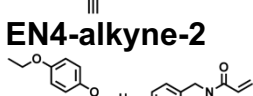

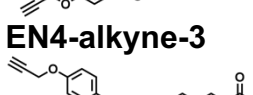
Available in supplementary downloadable files.

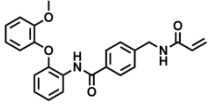
**Table S3. RNA sequencing analysis of EN4-mediated transcriptional changes in 231MFP cells. 231MFP cells were treated with DMSO vehicle or EN4 (50  $\mu$ M) for 24 h, n=3 biological replicates/group.**

Available in supplementary downloadable files.

Tab 1 shows total RNA sequencing data, quantitation, and statistical analysis. Contrast analysis produced t-test p-values, Benjamini-Hochberg false-discovery rates (FDRs) and log 2 fold changes for each gene with treated versus DMSO vehicle-treated controls. Tab 2 shows signature enrichment. For the signature enrichment, genes with an FDR < 0.1 and Fold change  $\leq$  0.75 (567 genes) were selected and signature search was performed using the MSigDB Collection H hallmark gene set: <https://www.gsea-msigdb.org/gsea/msigdb/collections.jsp#H>.

**Table S4. Structure-Activity Relationships of EN4 Analogs.**

Compound	EC50 (μM)	Compound	EC50 (μM)
	12.5		5.0
<b>EN4</b>		<b>EN4-12</b>	
	>50		>50
<b>EN4-1</b>		<b>EN4-13</b>	
	1.7		>50
<b>EN4-2</b>		<b>EN4-14</b>	
	>50		>50
<b>EN4-3</b>		<b>EN4-15</b>	
	13.0		>50
<b>EN4-4</b>		<b>EN4-16</b>	
	2.1		>50
<b>EN4-5</b>		<b>EN4-17</b>	
	>50		2.9
<b>EN4-6</b>		<b>EN4-18</b>	
	2.2		6.3
<b>EN4-7</b>		<b>EN4-alkyne-1</b>	
	11.9		16.4
<b>EN4-8</b>		<b>EN4-alkyne-2</b>	
	10.4		7.8
<b>EN4-9</b>		<b>EN4-alkyne-3</b>	
	>50		1.4
<b>EN4-10</b>		<b>EN4-alkyne-4</b>	



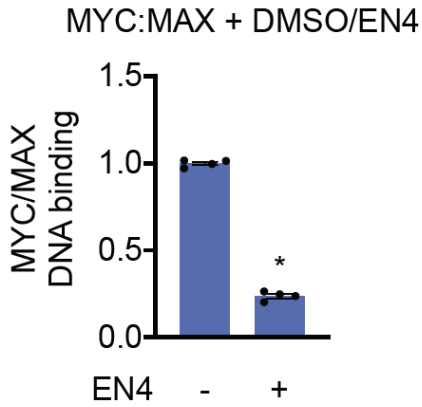
**EN4-11**

**11.3**

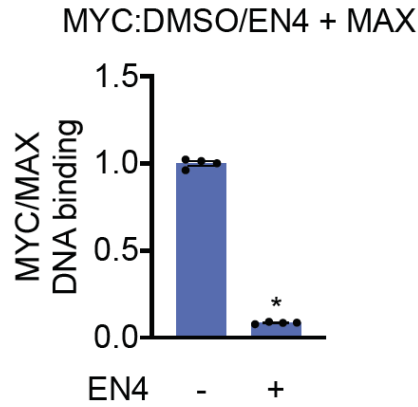


**Figure S1 (a-c):**

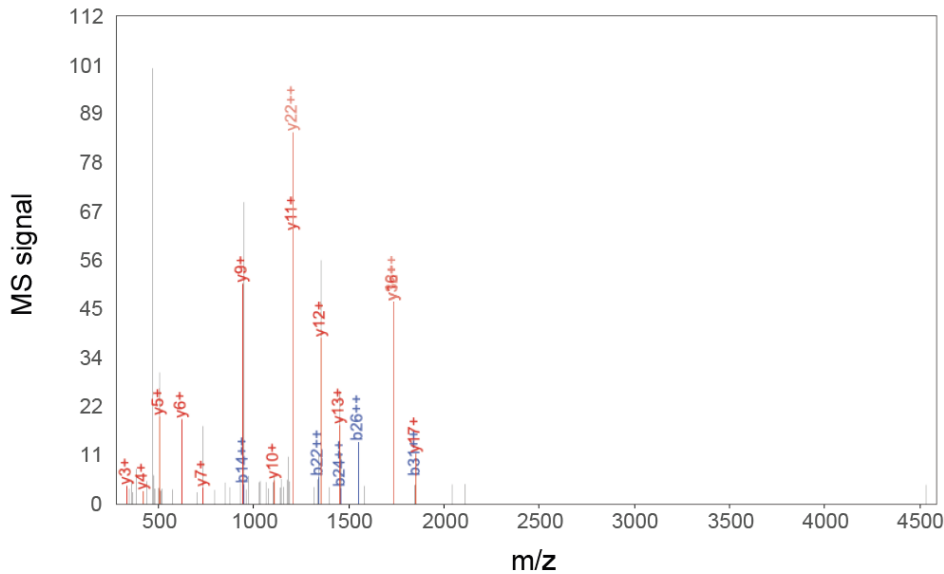
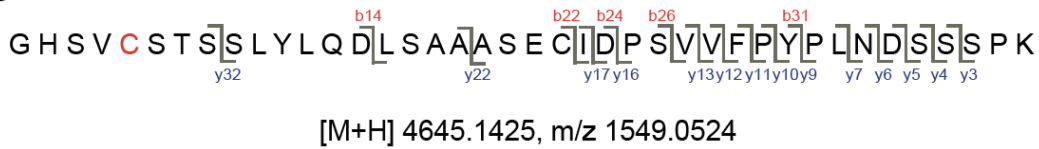
**a**



**b**



**c**



**Figure S1. Characterization of EN4 inhibition of c-MYC/MAX DNA binding and interaction of EN4 with c-MYC. (a)** Pure human c-MYC and MAX protein (0.2  $\mu$ g each) were pre-incubated for 30 min at 37  $^{\circ}$ C before addition of DMSO vehicle or EN4 (50  $\mu$ M) for 30 min at 37  $^{\circ}$ C prior to assessing DNA binding of the MYC/MAX complex to the E-box DNA consensus sequence. **(b)** Pure c-MYC (0.2  $\mu$ g) was pre-incubated with DMSO vehicle or EN4 (50  $\mu$ M) for 30 min at 37  $^{\circ}$ C prior to adding MAX (0.2  $\mu$ g) for 30 min at 37  $^{\circ}$ C prior to assessing DNA binding of the MYC/MAX complex to the E-box DNA consensus sequence. **(c)** MS/MS analysis of EN4 adduct on MYC C171. Pure human full-length MYC was incubated with EN4 (50  $\mu$ M) for 30 min. MYC was then subjected to



tryptic digestion and tryptic digests were analyzed by LC-MS/MS. Shown is the EN4 adduct on C171 on a fully tryptic peptide of c-MYC. Data in **(a, b)** are shown as individual replicate values and average  $\pm$  sem, n=4 biologically independent samples/group. Statistical significance was calculated with unpaired two-tailed Student's t-tests and is expressed as \*p<0.05 compared to vehicle-treated controls. This figure is related to **Figure 1** and **Figure 2**.

**Figure S2:**

```
sp|P05412|JUN_HUMAN      -----  
MTAKMETTFYDDALNASFLPSESGPYGYS---NPKILKQSM 38  
sp|P01106|MYC_HUMAN     -----MPLNVSFTNRNYDLDYDSVQPYFYCDEE-  
ENFYQQQQQSELQPPAPSEDIWKKFE 54  
sp|P04198|MYCN_HUMAN    MPSCSTSTMPGMICKNPDLEFDSLQPCFYDE--  
DDFYFGG----PDSTPPGEDIWKKFE 54  
sp|P12524|MYCL_HUMAN    -----MDYDSYQHYFYDYDCGEDFY-----  
RSTAPSEDIWKKFE 34  
      .: **  *      .:*:
```

```
sp|P05412|JUN_HUMAN      -----  
TLNLADPVGSLKPHLRAKNSDLLTSPDVGLLKLASPELERLIIQSSN 85  
sp|P01106|MYC_HUMAN     LLPTPPLSPSRRLGCLCSPSYVA-VTPFSLRGDN-----  
DGGGGSFSTADQLEMVTELLG 107  
sp|P04198|MYCN_HUMAN    LLPTPPLSPSRGFAEHSSEPPSWVTEMLL-----  
-----ENELWG 89  
sp|P12524|MYCL_HUMAN    LVSPPTSPPWGLGPGAGDPAPGIG-----  
----PPEPWP65  
      . :      :
```

```
sp|P05412|JUN_HUMAN      GHITTTPTPTQFLCPK-----NVTDEQEGFAEGFV-----  
RALAELHSQNTLPSVTS 132  
sp|P01106|MYC_HUMAN     GDM-VNQ---SFIQDPDDETFIKNII-  
IQDCMWSGFSAAAKLVSSEKLSYQAARKDSGSP 162  
sp|P04198|MYCN_HUMAN    SPA-EED---AFGLGGLGGLTPNPVI-  
LQDCMWSGFSAAREKLERAVSEKLGHRGPPTAG 144  
sp|P12524|MYCL_HUMAN    GGC-TGDEAESRGHSGWGRNYASII-  
RRDCMWSGFSAERLERAVSDRLAPGAPRGNPP 123  
      . : :::**
```

```
sp|P05412|JUN_HUMAN      AAQPVNGAGMVAPAVASVAGGSGSGGFSASL---H-----  
SEPPVY----- 170  
sp|P01106|MYC_HUMAN     N-----PARGHSV----  
CSTSSLYLQDLSAAASECIDPSVVFYPLNDSSSPK 206  
sp|P04198|MYCN_HUMAN    STAQSPGAGAASPAGRGHGGAAGAGRAGAALPAELA  
HPAAECVDPVVFPFPVVKREPAP 204  
sp|P12524|MYCL_HUMAN    KASA-----APDCT-----  
132
```

```
sp|P05412|JUN_HUMAN      --A---NLSNFPNGALS--SGGGAPSYGAAGLAFPA-----  
----- 199  
sp|P01106|MYC_HUMAN     SCASQD-SSAFSPSSDILLSSTESSPQGSPE-----  
--PLVLHEE-TPPT 247
```

```

sp|P04198|MYCN_HUMAN   VPAAPASAPAAGPAVAS--
GAGIAAPAGAPGVAPPRPGGRQTSGGDHKALSTSGEDTLD 262
sp|P12524|MYCL_HUMAN  -----PSLE-----AGNPAPAAPCPLGE-----
PKTQACSGS----- 159
      *      *

sp|P05412|JUN_HUMAN   -----QPQQQQQPPHHLPO
213
sp|P01106|MYC_HUMAN   TSSDSEEEQEDEEEEIDVVSVEKRQAPGKRSESGSPS---
----AGGHSKPPHSPLVLKR 299
sp|P04198|MYCN_HUMAN  SDDDEDDEEEDEEEIDVVTVEKRRSSSNTKAVTTFTIT
VRPKNAALGPGRAQSSELILKR 322
sp|P12524|MYCL_HUMAN  ----ESPDSSENEEIDVVTVEKRQSLGIRKPVT---ITVRAD-
-PLDPCM-----KHF 203

sp|P05412|JUN_HUMAN   QMPVQ-----HPRLQA-----LKE-----EP-
QTVPE 234
sp|P01106|MYC_HUMAN   CHVS-THQHNYAAPPSTRKDY-----PAAKR-VKL-----
---DSVR-VLRQ 335
sp|P04198|MYCN_HUMAN  CLPI-HQQHNYAAPSPYVESE----DAPPQKK-IK-----
--SEASPRPLKSV 362
sp|P12524|MYCL_HUMAN  HISIHQQQHNYAARFPPESCSQEEASERGPQEEVLER
DAAGEKEDEEEDIEIVSPPPVESE 263
      .      ::      .

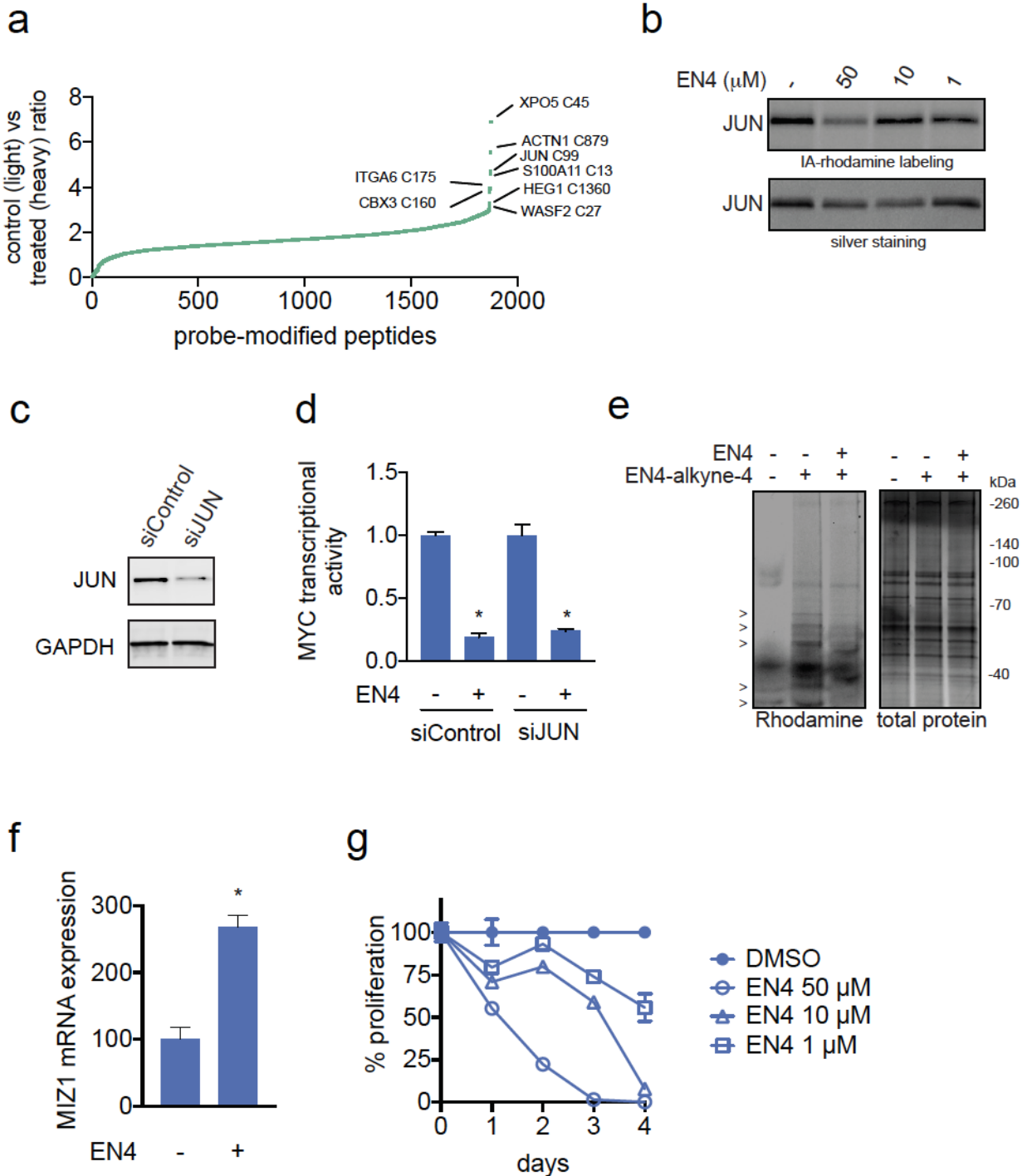
sp|P05412|JUN_HUMAN   MPGETPPLSPIDMESQERIKAEKRMNRNRIAASKCRKR-
-----KLER 276
sp|P01106|MYC_HUMAN   ISNNRKCTSPRSSDTEENVKRRTHNVLERQRRNELKRS
FFALRDQIPELENNEKAPKVVI 395
sp|P04198|MYCN_HUMAN  IPPKAKSLSPRNSDSEDSERRRNHNILERQRRNDLRS
SFLTLRDHVPELVKNEKAAKVVI 422
sp|P12524|MYCL_HUMAN  A-
AQSCHPKPVSSDTEEDVTKRKNHNFLERKRRNDLRSRFLALRDQVPTLASCASKAPKVVI
322
      :  *  :::  : : : *  .. :      * :

sp|P05412|JUN_HUMAN   IARLEEKVKTLLKAQNSELASTANMLREQVAQLKQKVMN
HVNSGCQLMLTQQLQTF 331
sp|P01106|MYC_HUMAN   LKKATAYILSVQAEEQKLISEEDLLRKRREQLKHKLE-
QLRNCA----- 439
sp|P04198|MYCN_HUMAN  LKKATEYVHSLQAEHQLLLEKEKQLARQQQLLKIE-
HARTC----- 464
sp|P12524|MYCL_HUMAN  LSKALEYLQALVGAEKRMATEKRQLRCRQQQLQKRIA-
YLTGY----- 364
      ::  : : : : :  * : * * : :

```

**Figure S2. Sequence alignment of human JUN, c-MYC, N-MYC and L-MYC.** Shown is a sequence alignment of human JUN, c-MYC, N-MYC, and L-MYC using CLUSTAL O (1.2.4). Highlighted in yellow C171 on c-MYC, which is not conserved between N-MYC and L-MYC. Highlighted in green is C99 of JUN and C117 of c-MYC. This figure is related to **Figure 2**.

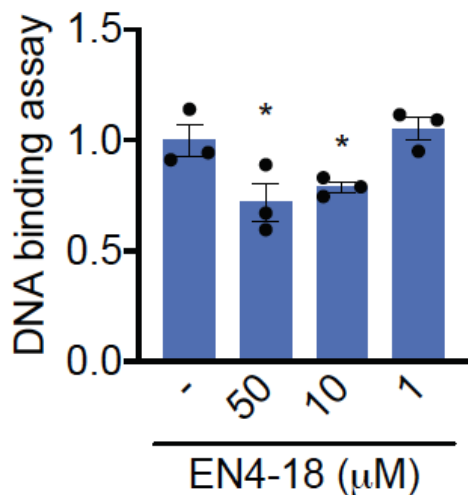
**Figure S3 (a-g):**



**Figure S3. Characterization of EN4 in cells.** (a) isoTOP-ABPP analysis of EN4 in 231MFP breast cancer cells. 231MFP cells were treated with DMSO vehicle or EN4 (50  $\mu\text{M}$ ) for 4 h. Cell lysates were subsequently labeled with IA-alkyne for 1 h, followed by appendage of a TEV protease-cleavable biotin-azide linker bearing isotopically light (for control) or heavy (for EN4-treated) tags by CuAAC. Control and treated proteomes from each replicate were combined in a 1:1 ratio and subjected to avidin-enrichment,

digestion by trypsin, and release of probe-modified peptides by TEV protease. Probe-modified peptides were analyzed by LC-LC-MS/MS and light to heavy ratios were quantified. Shown are those ratios with ratios >5 highlighted with the protein designation and cysteine number. Shown are average light/heavy ratios from n=3 biologically independent samples/group. **(b)** gel-based ABPP of EN4 against pure human JUN. Pure human JUN protein was pre-incubated with DMSO vehicle or EN4 for 30 min prior to labeling of protein with rhodamine-functionalized iodoacetamide (IA-rhodamine) (5  $\mu$ M) for 30 min. Proteins were separated by SDS/PAGE and visualized by in-gel fluorescence and protein loading was visualized by silver staining. **(c)** JUN knockdown in HEK293T cells shown by Western blotting alongside loading control GAPDH in cells transfected with siControl or pooled siJUN oligonucleotides for 48 h. Shown is a representative gel of n=3 biological replicates/group. **(d)** MYC luciferase reporter activity in HEK293T cells transiently transfected with siControl versus siJUN oligonucleotides treated with DMSO vehicle or EN4 (10  $\mu$ M) for 24 h. Values are reported relative to vehicle-treated controls in each group. **(e)** EN4-alkyne-4 labeling of 231MFP cells. 231MFP cells were pre-treated with DMSO vehicle or EN4 (50  $\mu$ M) for 2 h prior to treatment with DMSO or EN4-alkyne-4 (1  $\mu$ M) for 4 h. Resulting cell lysates were subjected to CuAAC to append rhodamine-azide to EN4-alkyne-4 labeled proteins. Proteins were separated by SDS/PAGE and visualized by rhodamine fluorescence (left) or for total protein by Simple Blue staining (right). Shown are representative gels from n=3 biological replicates/group. Arrows show EN4-alkyne-4 labeled proteins that are competed by EN4 treatment. **(f)** MIZ1 mRNA expression. 231MFP breast cancer cells were treated with DMSO vehicle or EN4 (50  $\mu$ M) and MIZ mRNA expression was assessed by qPCR. **(g)** 231MFP cell proliferation in cells treated with DMSO vehicle or EN4 once per day daily over 4 days, assessed by Hoechst staining. Data normalized to daily DMSO controls are shown. Data shown in **(d, f, g)** are average  $\pm$  sem, n=6 for **(d)**, n=3 for **(f)**, and n=5-6 for **(g)** biologically independent samples/group. Statistical significance was calculated with unpaired two-tailed Student's t-tests and is expressed as \*p<0.05 compared to vehicle-treated controls. This figure is related to **Figures 3 and 4**.

**Figure S4:**



**Figure S4. Characterization of EN4-18 in DNA binding assays.** Screening a cysteine-reactive covalent ligand library *in vitro* for compounds that would inhibit MYC/MAX binding to its E-box DNA consensus sequence. DMSO vehicle or EN4-18 were pre-incubated with pure human full-length MYC protein for 30 min before direct addition of MAX and then addition to DNA binding plates for 1 h. Data are shown as ratio relative to DMSO vehicle treated controls set to 1. Data shown are average ± sem, n=3 biologically independent samples/group. Statistical significance was calculated with unpaired two-tailed Student's t-tests is expressed as \*p<0.05. This figure is related to **Table S3**.

## CHAPTER 3

### Development of Potent Pyrazoline-Based Covalent Main Protease Inhibitors

(Adapted from published article by Patrick Moon, Lydia Boike, Dustin Dovala, Nathaniel J. Henning, Mark Knapp, Jessica N. Spradlin, Carl C. Ward, Helene Wolleb, Charlotte M. Zammit, Daniel Fuller, Gabrielle Blake, Jason P. Murphy, Feng Wang, Yipin Lu, Stephanie A. Moquin, Laura Tandeske, Matthew J. Hesse, Jeffrey M. McKenna, John A. Tallarico, Markus Schirle, F. Dean Toste, Daniel K. Nomura DOI: 10.1101/2022.03.05.483025)



## Development of Potent Pyrazoline-Based Covalent Main Protease Inhibitors

While vaccines are now being deployed for the current SARS-CoV-2 pandemic, we need more effective antiviral therapeutics that can effectively combat this COVID-19 pandemic and future coronaviruses that will inevitably arise. Furthermore, with the appearance of an increasing number of dangerous variants of SARS-CoV-2 and with all vaccines and antibody therapies targeting the viral Spike protein, which is prone to mutations, there is an increasing need for pan-coronavirus antiviral drugs that will be effective against any variant of the SARS-CoV-2 virus as well as any future coronaviruses. In particular, the discovery and design of covalent inhibitors against SARS-CoV-2 Mpro present an exciting therapeutic opportunity.

All coronaviruses have several highly conserved genes that provide a potential exploitable opening to develop an antiviral therapy that could be effective against not only SARS-CoV-2, but also all past and future coronaviruses. Among the various genes and proteins encoded by coronaviruses, one particularly “druggable” or relatively easy-to-drug target is the coronavirus Main Protease (Mpro) also known as 3C-like protease, an enzyme that is involved in cleaving a long peptide translated by the viral genome into its individual protein components that are then assembled into the virus to enable viral replication in the cell. The replicase gene of SARS-CoV-2 encodes two overlapping polyproteins—pp1a and pp1ab—that are required for viral replication and transcription.<sup>218</sup> Mpro is responsible for the majority of proteolytic processing events to release the functional polypeptides with at least 11 proteolytic cleavage events starting with the autolytic cleavage of Mpro itself from pp1a and pp1ab.<sup>218</sup> Inhibiting Mpro with a small-molecule antiviral could effectively stop the ability of the virus to replicate, providing immediate and complete therapeutic benefit. Unlike the viral Spike protein that currently all vaccines and antibody therapeutics target, which is highly prone to mutations that give rise to resistant variants, Mpro protein sequence is highly conserved across all former coronaviruses indicating that making a potent and effective Mpro inhibitor for one coronavirus would result in an efficacious drug for all future coronaviruses. Because Mpro is a critical protein necessary for coronavirus replication, mutations in the core active site domain of this protein would likely impair its catalytic activity and the ability of the virus to replicate, and thus Mpro inhibitor antivirals would be unlikely to run into resistance through viral mutations. There has been considerable effort over the past two years by many academic and industrial labs to discover Mpro inhibitors (**Fig. 1A**).<sup>224,3</sup> For example, Pfizer has now developed an effective orally bioavailable covalently-acting peptide-based Mpro inhibitor Paxlovid that been given emergency use authorization by the United States Food and Drug Administration (**Fig. 1A**).<sup>219</sup> However, these inhibitors may not be effective against Mpro enzymes from future variants of SARS-CoV-2 or other future coronaviruses. Thus, there is still significant room for discovering additional Mpro inhibitor scaffolds that may be tractable for drug discovery and development efforts.

### 3.1 Covalent Ligand Screening to Discover Pyrazoline-based SARS-CoV-2 Main Protease Inhibitors

### 3.1.1 Covalent Screening and Main Protease Activity Assay Development

Mpro is a cysteine protease that bears a catalytic cysteine (C145 for SARS-CoV-2 Mpro) that coordinates its catalytic protease activity.<sup>218</sup> As such, chemoproteomic platforms and cysteine-reactive covalent ligand discovery approaches are particularly attractive for developing potent, selective, non-peptidic, more drug-like, and covalently-acting Mpro inhibitors that irreversibly interact with the Mpro catalytic cysteine.

To rapidly identify cysteine-reactive inhibitors against SARS-CoV-2 Mpro, we screened a library of 582 acrylamides and chloroacetamides in a gel-based activity-based protein profiling (ABPP) screen, in which we competed these cysteine-reactive covalent ligands against the binding of a cysteine-reactive rhodamine-functionalized iodoacetamide probe (IA-rhodamine) using previously described gel-based ABPP approaches (**Fig. S1-S2**).<sup>225,13</sup> The most promising hits clustered together on pyrazoline-based chloroacetamide ligands EN71, EN82, EN216, and EN223 that demonstrated dose-responsive inhibition of Mpro IA-rhodamine labeling (**Fig. S3A**). We subsequently tested these four compounds in a FRET-based activity assay employing an Mpro peptide substrate to identify compounds that inhibited Mpro activity (**Fig. S3B**). The most promising hit to arise from this screen was a pyrazoline EN82 which displayed encouraging potency against the SARS-CoV-2 Mpro with a 50 % inhibitory concentration (IC<sub>50</sub>) value of 0.16 mM compared to 0.52-4.8 mM for EN216, EN71, and EN223 (**Fig. S3B**). The FRET-based peptide probe required significant concentrations of Mpro protein (115 nM) to obtain reliable Mpro activity readouts, and would thus hinder our ability to detect more potent inhibitors with IC<sub>50</sub> values <100 nM. For subsequent Mpro activity assays, we switched over to a rhodamine-based Mpro substrate activity assay which required significantly less protein. Using this rhodamine-based assay, EN82 showed potent IC<sub>50</sub> values against SARS-CoV-2 (0.091  $\mu$ M), SARS-CoV-1 (0.059  $\mu$ M) and MERS (0.56  $\mu$ M) (**Fig. 1B**).

### 3.1.2 Assessing Proteome-Wide Cysteine-Reactivity of Hit Compound EN82

We next performed chemoproteomic profiling of EN82 cysteine-reactivity in HEK293T cell lysate to assess whether this compound demonstrated some degree of selectivity or whether this compound was non-specific (**Fig. 2**). Using mass spectrometry-based activity-based protein profiling (MS-ABPP), in which HEK293T cell lysates spiked with pure Mpro protein were pre-treated with vehicle or EN82 (50  $\mu$ M), and subsequently labeled with an alkyne-functionalized cysteine-reactive iodoacetamide probe for subsequent quantitative chemoproteomic analysis of EN82-occupied cysteine sites, we observed Mpro C145 as the primary target of EN82 with 1 additional off-target HMOX2 C282 out of >1000 distinct quantified probe-modified cysteines. Thus, as an initial hit compound, EN82 showed a high degree of proteome-wide selectivity with only one potential off-target (**Fig. 2**). EN82 was a racemic mixture of two compounds. Upon enantioselective synthesis of each isomer, (*R*)-EN82 (0.53  $\mu$ M) proved more active than (*S*)-EN82 (2.7  $\mu$ M), consistent with a crystal structure obtained of (*R*)-EN82 bound to the catalytic cysteine C145 of SARS-CoV-2 Mpro (**Fig. 1B**). With EN82 as an

encouraging starting point, structure-activity relationship (SAR) studies around the central pyrazoline core were conducted to optimize for potency.

### 3.2 Exploration of 3,5-Disubstituted Pyrazoline SAR

We next devised an expedient synthesis of 3,5-disubstituted pyrazolines starting with an aldol condensation between commercially available aldehyde and acetophenone precursors to form the corresponding chalcone, followed by hydrazine condensation to access the unprotected pyrazoline core, and subsequent N-acylation to yield the final pyrazoline chloroacetamide in three steps (**Fig. 3A**). A number of nitrogen-containing heterocycles were introduced at the pyrazoline C5-carbon ( $Ar^1$ ) in an attempt to improve potency. Differentially substituted quinoxalines, quinolines, indoles, azaindoles, benzotriazoles and pyridines were tested at this position, but no significant breakthrough was achieved (**Fig. 3B**). Several of these derivatives showed comparable potency to EN82, suggesting there is some flexibility for C5-pyrazoline substituent. SAR at the pyrazoline C3-carbon ( $Ar^2$ ) was explored next (**Fig. 3C**). Inhibitors with larger 4-substituted arenes (LEB-2-187, LEB-2-175, LEB-2-174, LEB-2-173) showed decreased potency compared to smaller 3-substituted arenes (LEB-2-172, LEB-3-001, LEB-2-182, LEB-2-176, and LEB-3-004). This was consistent with the relatively shallow pocket in which the C3-arene is located in the EN82-bound Mpro crystal structure (**Fig. 1B**). Simple aliphatic tert-butyl and cyclopropyl substituents (HW-1-174, HW-1-185) at this position exhibited a significant loss of activity. Exploration of SAR was also examined using a 5-quinolinyl bearing pyrazoline scaffold (**Fig. 3C, bottom**), given that this scaffold displayed better overall solubility relative to the 5-quinoxalinyl derivatives. A consistent trend was observed with smaller 3- and 2-substituted arenes at the C3 position showing better activity. Ultimately, a 5-chloro-2-fluorophenyl C3-substituent (CMZ-53) proved optimal and provided a nearly three-fold improvement in potency compared to EN82.

### 3.3 Exploration of Trisubstituted Pyrazoline Inhibitors

We next explored the possibility of adding a substituent at the C4 position of the pyrazoline core. Examination of the crystal structure of EN82 bound to the SARS-CoV-2 Mpro revealed a small pocket occupied by an imidazole that should be accessible with cis-trisubstituted pyrazolines (**Fig. 1B**). Consistent with this premise, the trisubstituted cis-pyrazoline PM-2-020B showed improved potency compared to EN82 and its disubstituted analog EN23 (**Fig. 4B**). In contrast, trans-pyrazoline PM-2-020A showed an approximate 30-fold lower potency compared to cis-PM-02-020B, highlighting the profound impact of relative stereochemistry on the activity of these trisubstituted pyrazolines. A crystal structure of PM-2-020B bound to the SARS-CoV-2 Mpro was obtained which corroborated the initial hypothesis, with the C4-phenyl group nestled in the previously empty sub-pocket. Trisubstituted analogs with C5-heterocycles were next tested while holding the C4-phenyl group constant. A smaller 3-pyridyl C5-substituent (cis-PM-2-043B) was slightly less potent, while a 5-quinoxalinyl substituent (cis-PM-2-027B) proved even less potent, perhaps due to unfavorable steric clash between the

C4/C5 aryl substituents. Consistent with this hypothesis, a smaller isomeric 6-quinoxalinylnyl derivative (cis-HW-1-187B) proved to be our most potent inhibitor yet (IC<sub>50</sub> = 0.022 μM). Other heteroaryl and alkyl C4-substituents were also tested but showed no significant improvement (HW-1-198B, HW-1-176B, HW-2-01B).

A revised modular synthetic route was next designed for each pyrazoline substituent to originate from different building blocks. Ideally, the C4-substituent could be installed at a later stage via cross-coupling to avoid having to be introduced in the first aldol condensation step from 2-arylacetophenone derivatives which are not readily available. Initial aldol condensation, followed by mono-bromination and Suzuki coupling led to a tri-arylated chalcone intermediate, which then led to the final pyrazoline core via hydrazine condensation and acylation (**Fig. 3C**). With this route, the optimal 5-chloro-2-fluorophenyl C3-substituent could be installed, leading to our most potent 0.014 μM chloroacetamide inhibitor (HW-2-010B).

### 3.5 Exploration of Cysteine Reactive Warheads

Along with exploring the SAR around the pyrazoline core, concurrent efforts were aimed at identifying an alternative cysteine reactive warhead given chloroacetamides could potentially present metabolic stability issues (**Fig. 5**). A vinylsulfonamide was ultimately identified as a more promising alternative warhead,<sup>226,227</sup> culminating in the synthesis of PM-2-071B, a 0.091 μM inhibitor (**Fig. 5, Fig. 6**). Chiral SFC separation of PM-2-071B revealed that the (*S,S*)-enantiomer is most potent (0.035 μM), contrasting what is observed in the chloroacetamide series where the (*R*)-enantiomer of EN82 is most active (**Fig. 6**). This suggests a different binding mode may be at play with vinylsulfonamides relative to chloroacetamides. Other vinylsulfonamide derivatives were also tested; a modest improvement in potency was obtained with a 3-methoxy C4-substituent (PM-2-167B) while loss in potency was observed with a larger 3-benzyl substituent.

### 3.6 Testing Lead Compounds Against Coronavirus Panel

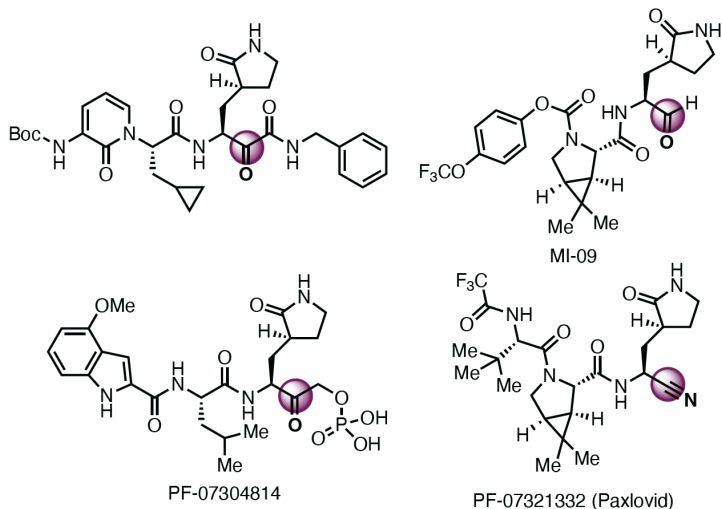
We also tested several of the di- and tri-substituted compounds, including PM-2-071, from our study against a panel of Mpro enzymes from other former coronaviruses, including SARS-CoV-2, SARS-CoV-1, HCoV-HKU1, HCoV-229E, MERS-CoV, PorCoV-HKU15, HCoV-NL63, avian infectious bronchitis virus (IBV), and HCoV-OC43. For these assays, we used an even more sensitive Mpro activity assay using an Agilent RapidFire-based substrate peptide activity assay that requires even less Mpro protein compared to the rhodamine-based activity assays to comparatively assess potencies of some of our best covalent ligands. In these assays, we show that most of the compounds tested inhibit Mpro from these other coronaviruses with IC<sub>50</sub> values in the nanomolar range (**Fig. S1, Fig. 7**). Most notably, we demonstrate that PM-2-071 shows IC<sub>50</sub>s <2 nM across SARS-CoV-2, HCoV-HKU1, and HCoV-OC43, and IC<sub>50</sub>s <1 mM for SARS-CoV-1, HCoV-NL63, IBV, and HCoV-229E Mpro enzymes (**Fig. 7**). Thus, this chemical scaffold may represent a promising starting point for further optimization of pan-coronavirus Mpro inhibitors.

In conclusion, in this study, we have discovered some of the most reported potent SARS-CoV-2 Mpro inhibitors based on pyrazoline-based chloroacetamides and vinyl sulfonamides, comparable with FDA-approved nirmatrelvir in standard *in vitro* assays. We further demonstrate that many of the inhibitors developed here also show exceptional potency against Mpro from previous coronaviruses, indicating the potential to develop a pan-coronavirus Mpro inhibitor. Exploration of SAR at the pyrazoline core revealed the importance of relative stereochemistry at the C4 and C5 carbons. We note that unfortunately our optimized inhibitors such as (S,S)-PM-2-071A appeared to have issues with solubility, metabolic stability, and cell permeability. We did not observe antiviral activity in cells with these series of compounds against SARS-CoV-2. As such, further medicinal chemistry optimization is required to optimize these potent inhibitors for their drug-like properties to promote antiviral efficacy. Nonetheless, we believe that this general scaffold and the exploration of cysteine-reactive warheads here represents a good foundational starting point for the development of more advanced non-peptidic and more drug-like pan-coronavirus Mpro inhibitors in the future.

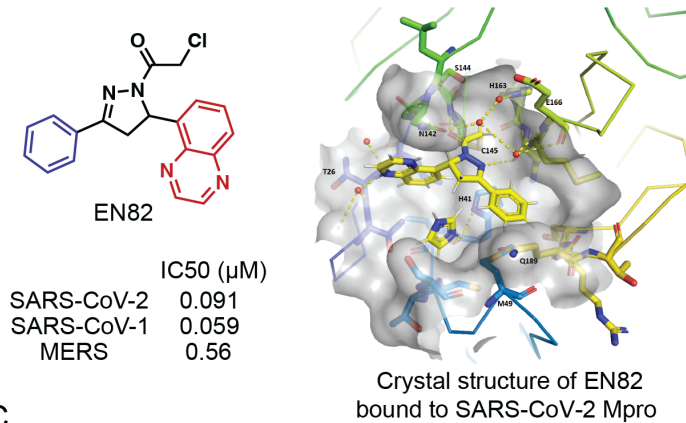
### 3.7 Figures and Tables

**Figure 1 (a-c):**

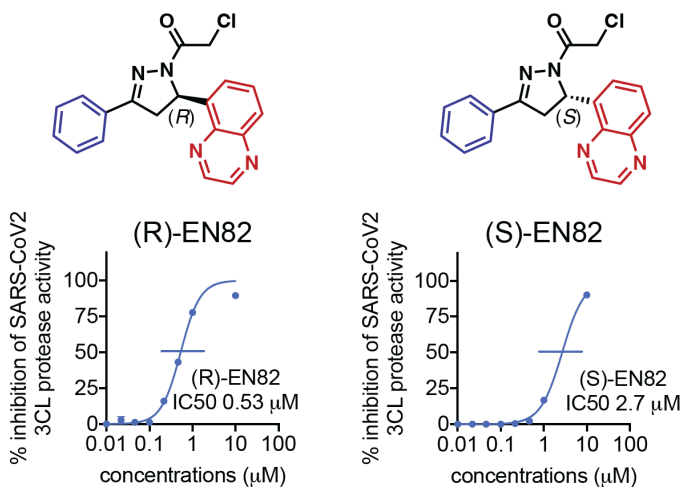
**A** Peptidomimetic SARS-CoV-2 Mpro (3CL<sup>pro</sup>) covalent inhibitors



**B** Promising Initial Hit Containing a Pyrazoline Core

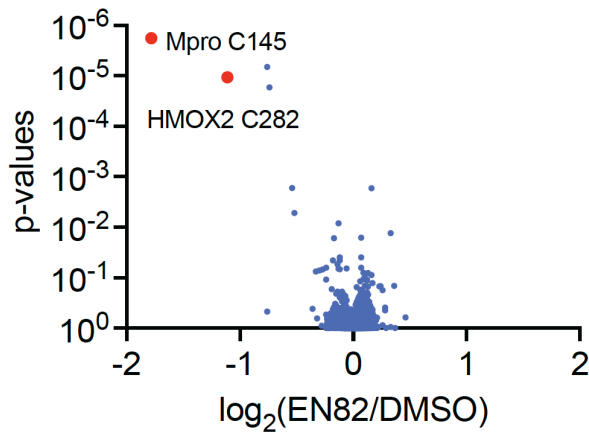


**C**



**Figure 1. Discovery of pyrazoline-based SARS-CoV-2 main protease inhibitors.** **(A)** Examples of peptidic SARS-CoV-2 Mpro inhibitors. **(B)** Structure of our top hit EN82 and potency of EN82 against SARS-CoV-2, SARS-CoV-1, and MERS-CoV Mpro. IC<sub>50</sub>s are derived from a rhodamine-based Mpro substrate peptide activity assay. On the left is a crystal structure of EN82 covalently bound to the catalytic C145 in the SARS-CoV-2 Mpro active site. **(C)** Potency of (*R*)- and (*S*)-EN82 against SARS-CoV-2 Mpro. IC<sub>50</sub>s were derived from averages of 3 biological replicates/group.

**Figure 2:**

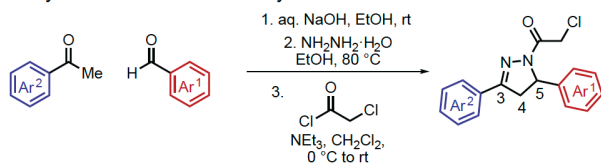


**Figure 2. Proteome-wide cysteine-reactivity of EN82 in HEK293T cell lysate.** Cell lysate was treated with DMSO or EN82 (50 mM), followed by labeling of proteomes with alkyne-functionalized iodoacetamide probe and subsequent MS-ABPP method. Shown in blue are >2-fold ratios with  $p < 0.05$  showing 41 potential off-targets of EN82 out of 7281 distinct cysteine sites quantified.

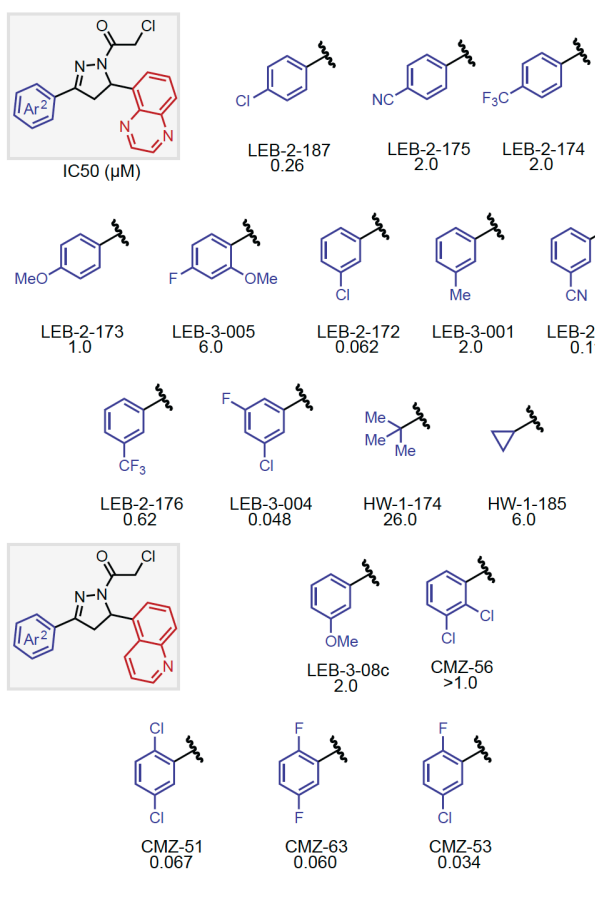


**Figure 3 (a-c):**

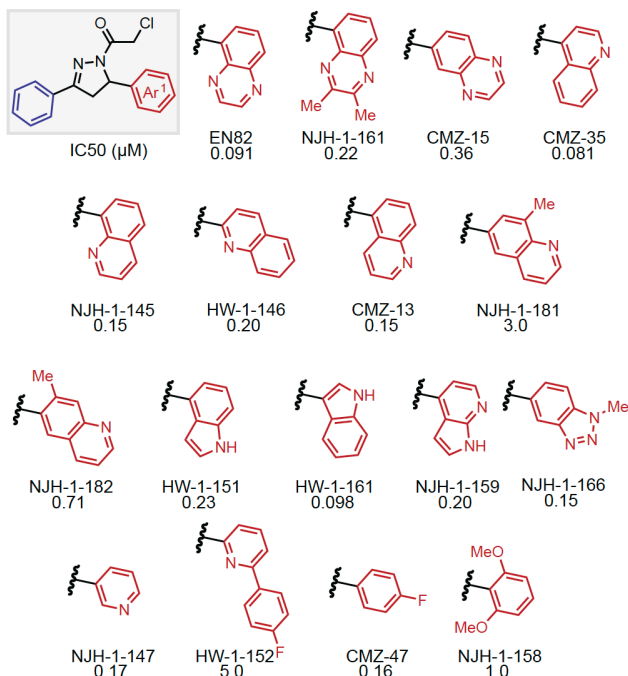
**A** Synthesis of Disubstituted Pyrazoline Scaffolds



**C** Exploration of SAR at Ar<sup>2</sup>

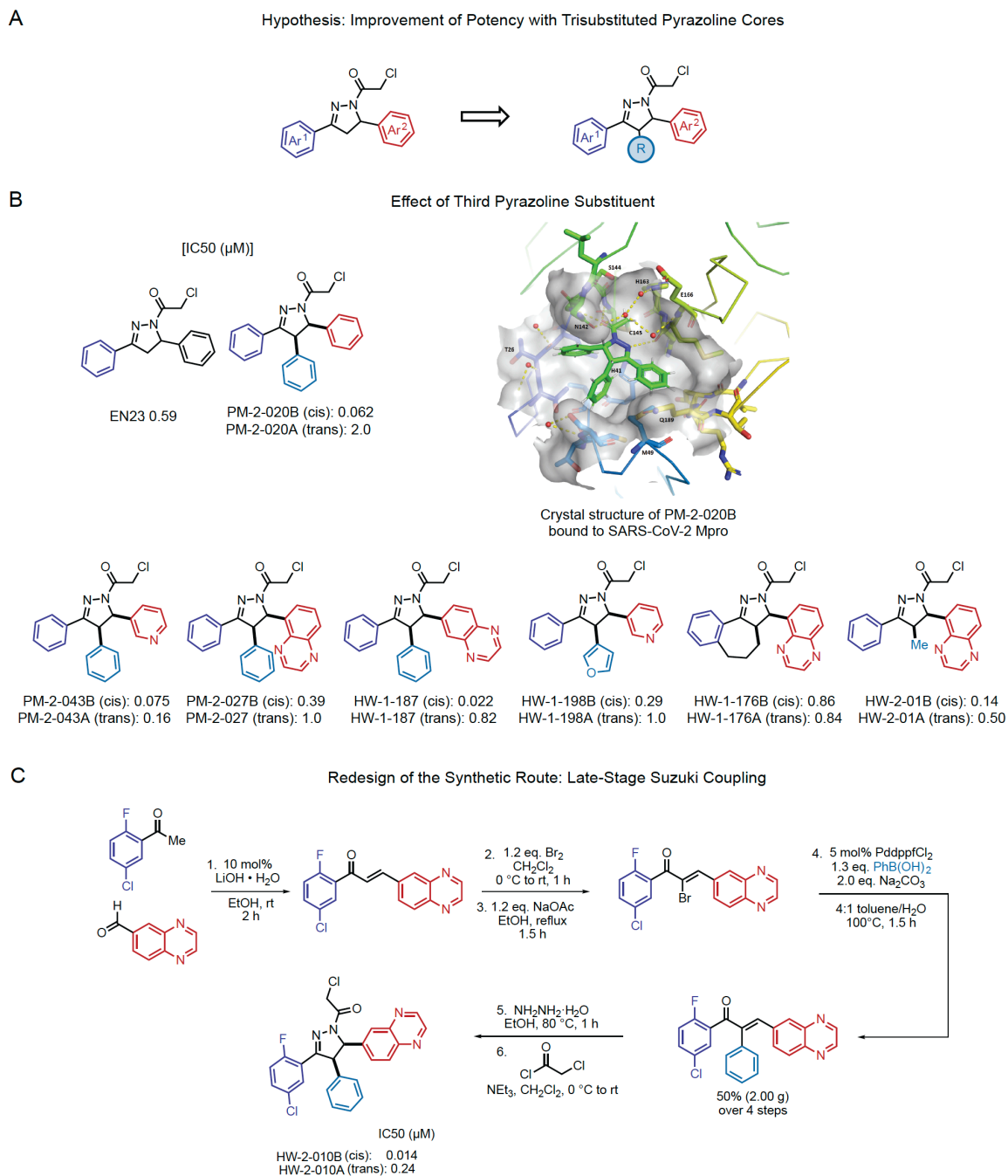


**B** Exploration of SAR at Ar<sup>1</sup>



**Figure 3. Exploration of 3,5-disubstituted pyrazoline SAR.** (A) Synthesis of disubstituted pyrazoline scaffolds. (B) Exploration of SAR at Ar<sup>1</sup>. (C) Exploration of SAR at Ar<sup>2</sup>. IC<sub>50</sub> values determined by rhodamine-based substrate activity assay for SARS-CoV-2 Mpro from n=3 biological replicates.

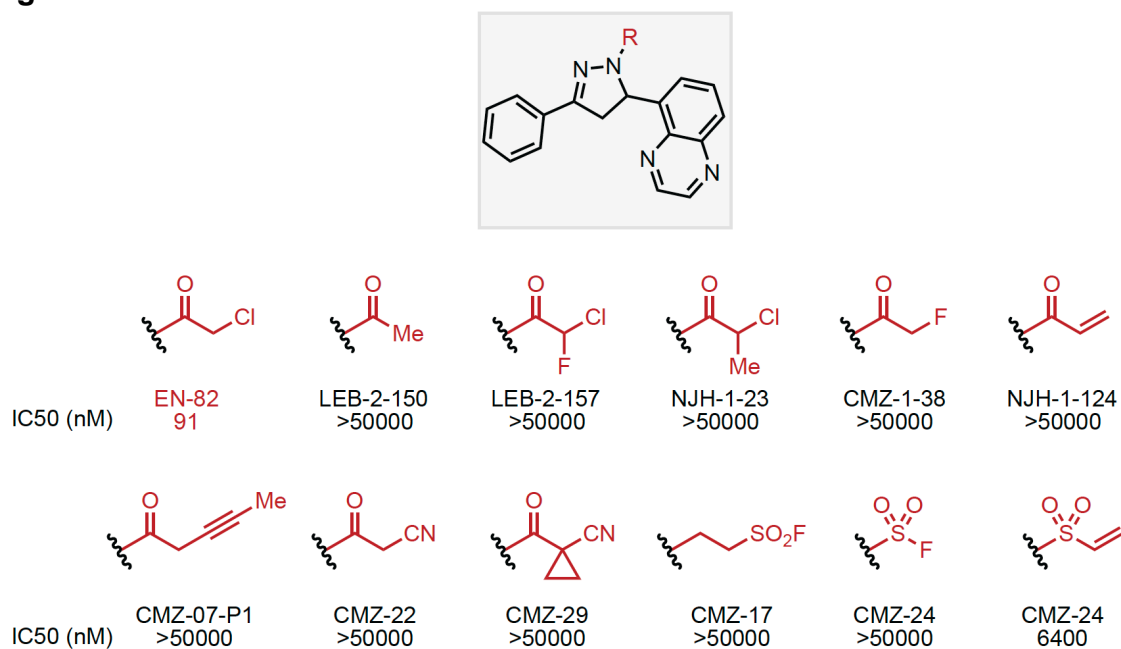
**Figure 4 (a-c):**



**Figure 4. Exploration of trisubstituted pyrazoline inhibitors.** (A) We hypothesized that potency would be improved through trisubstitution of the pyrazoline core. (B) Effect of third pyrazoline substituent on inhibitory potency of compounds against SARS-CoV-2 Mpro. Shown is the crystal structure of PM-2-20B covalently bound to the catalytic C145 of the SARS-CoV-2 Mpro active site. IC50 values determined by rhodamine-based substrate activity assay for SARS-CoV-2 Mpro from n=3 biological

replicates. **(C)** Redesign of the synthetic late-stage Suzuki coupling to yield HW-2-010B and HW-2-010A. Potencies against SARS-CoV-2 Mpro are shown from n=3 biological replicates.

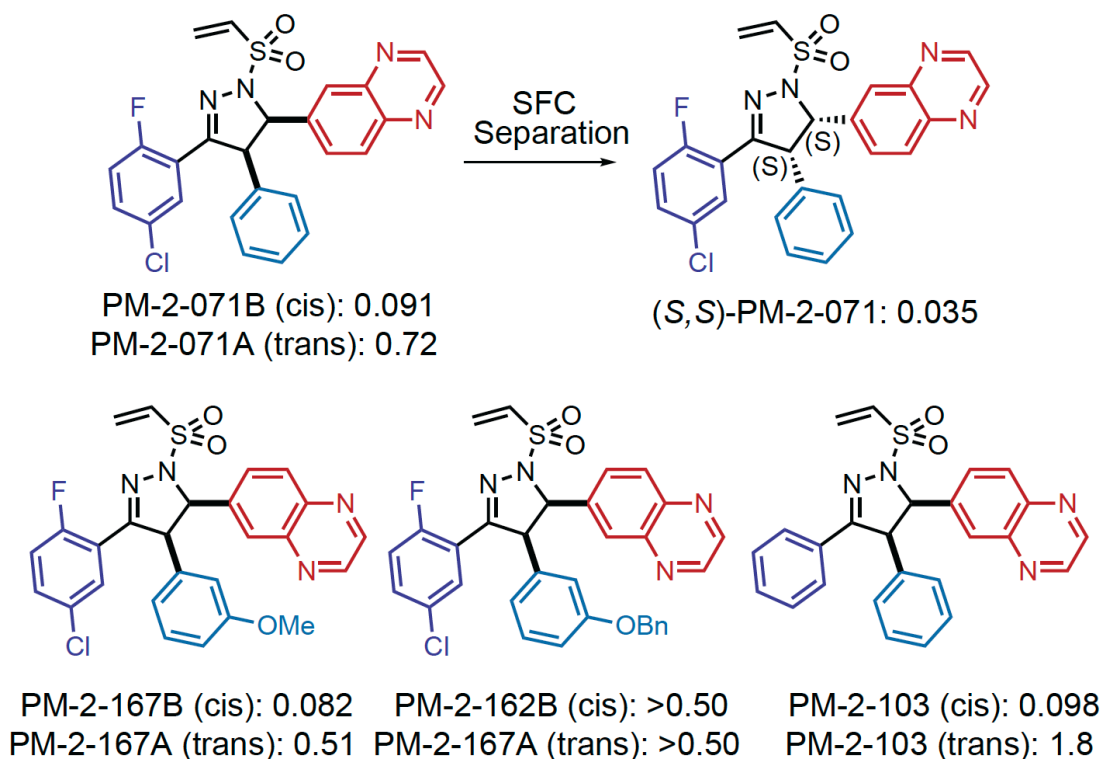
Figure 5:



**Figure 5. Exploration of Cysteine Reactive Warheads.** IC<sub>50</sub> values determined by rhodamine-based substrate activity assay for SARS-CoV-2 Mpro from n=3 biological replicates.

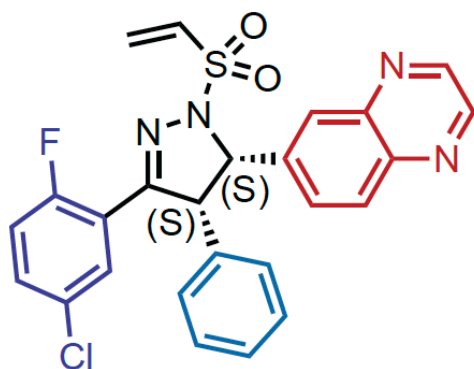
**Figure 6:**

Vinyl sulfonamide as the Cysteine-Reactive Warhead  
[IC<sub>50</sub> (μM)]



**Figure 6. Exploration of trisubstituted pyrazoline inhibitors with a vinyl sulfonamide cysteine-reactive warhead.** IC<sub>50</sub> values determined by rhodamine-based substrate activity assay for SARS-CoV-2 Mpro from n=3 biological replicates.

Figure 7:

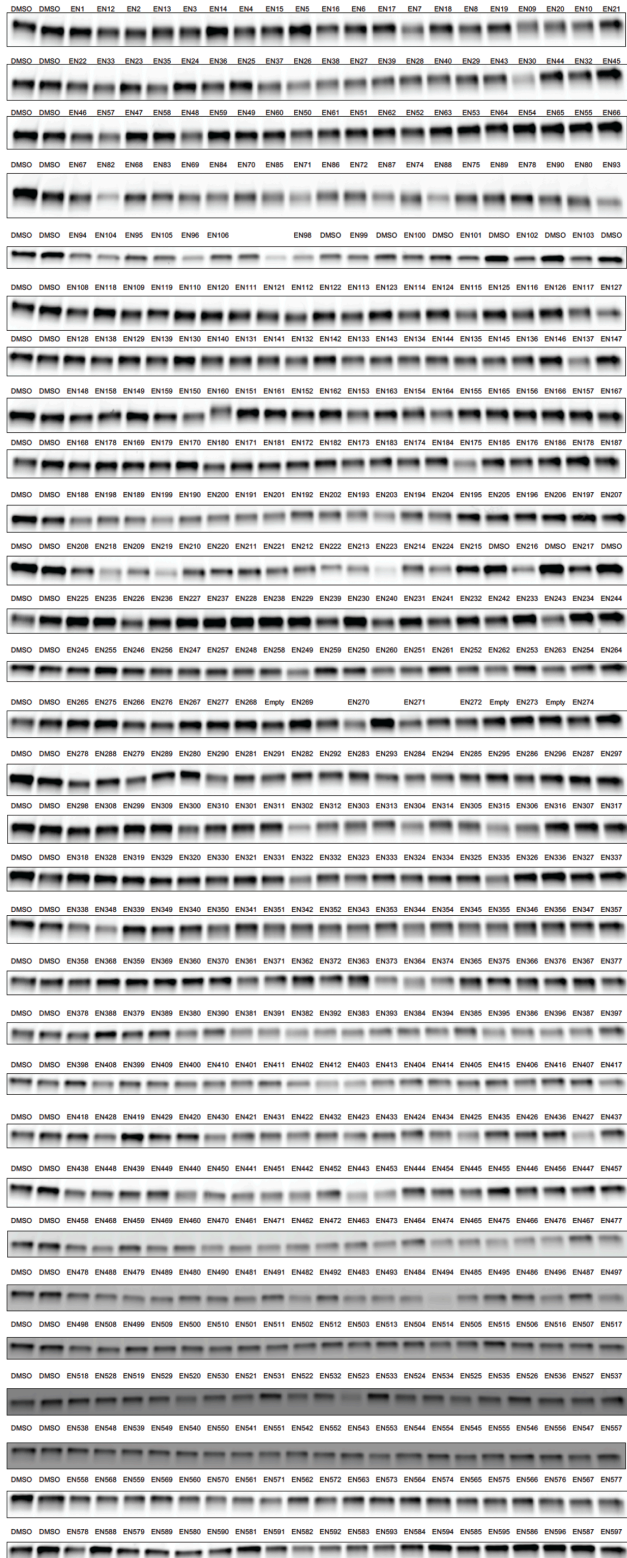


(S,S)-PM-2-071

	IC <sub>50</sub> (μM)
SARS-CoV-2	<0.0025
SARS-CoV-1	0.0117
HCoV-NL63	0.091
IBV	0.22
HCoV-HKU1	<0.0025
HCoV-OC43	<0.0025
HCoV-229E	0.012

Figure 7. Exploration of trisubstituted pyrazoline inhibitors with a vinyl sulfonamide cysteine-reactive warhead. IC<sub>50</sub> values determined by MS-based substrate activity assay for SARS-CoV-2 Mpro.

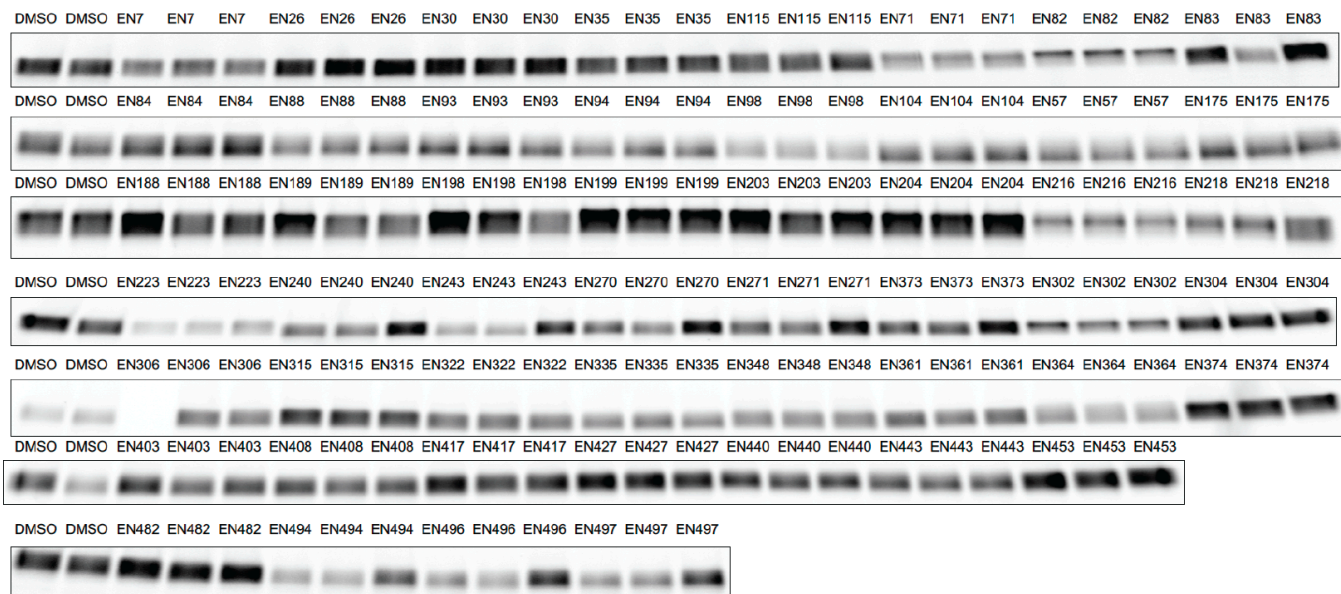
Figure S1:



**Figure S1. Gel-based ABPP screen of cysteine-reactive covalent ligands against Mpro enzyme.** Pure SARS-CoV-2 Mpro protein was pre-incubated with DMSO vehicle or cysteine-reactive covalent ligands (50  $\mu$ M, 30 min) prior to addition of a rhodamine-conjugated iodoacetamide (IA-rhodamine) probe (100 nM) for 1 hr after which reactions were quenched with SDS/PAGE loading buffer, resolved on SDS/PAGE and visualized by in-gel fluorescence.



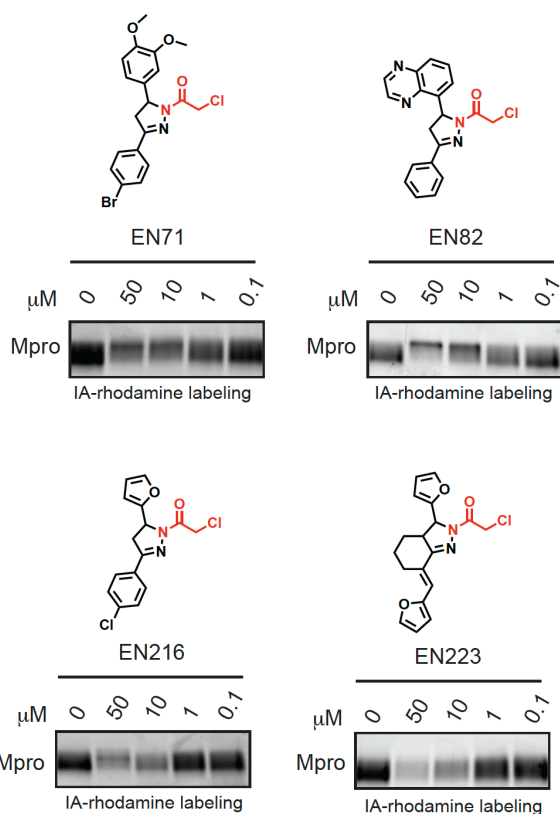
**Figure S2:**



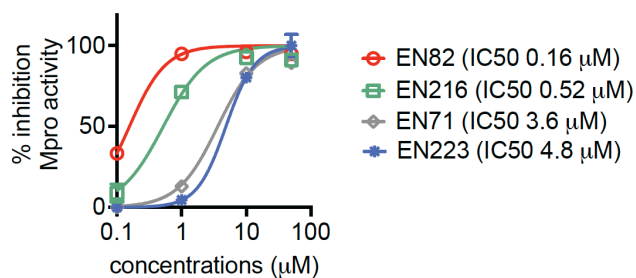
**Figure S2. Hit confirmation of initial gel-based ABPP screen.** Pure SARS-CoV-2 Mpro protein was pre-incubated with DMSO vehicle or cysteine-reactive covalent ligands (50 mM, 30 min) prior to addition of a rhodamine-conjugated iodoacetamide (IA-rhodamine) probe (100 nM) for 1hr min after which reactions were quenched with SDS/PAGE loading buffer, resolved on SDS/PAGE and visualized by in-gel fluorescence.

**Figure S3 (a-b):**

**A**

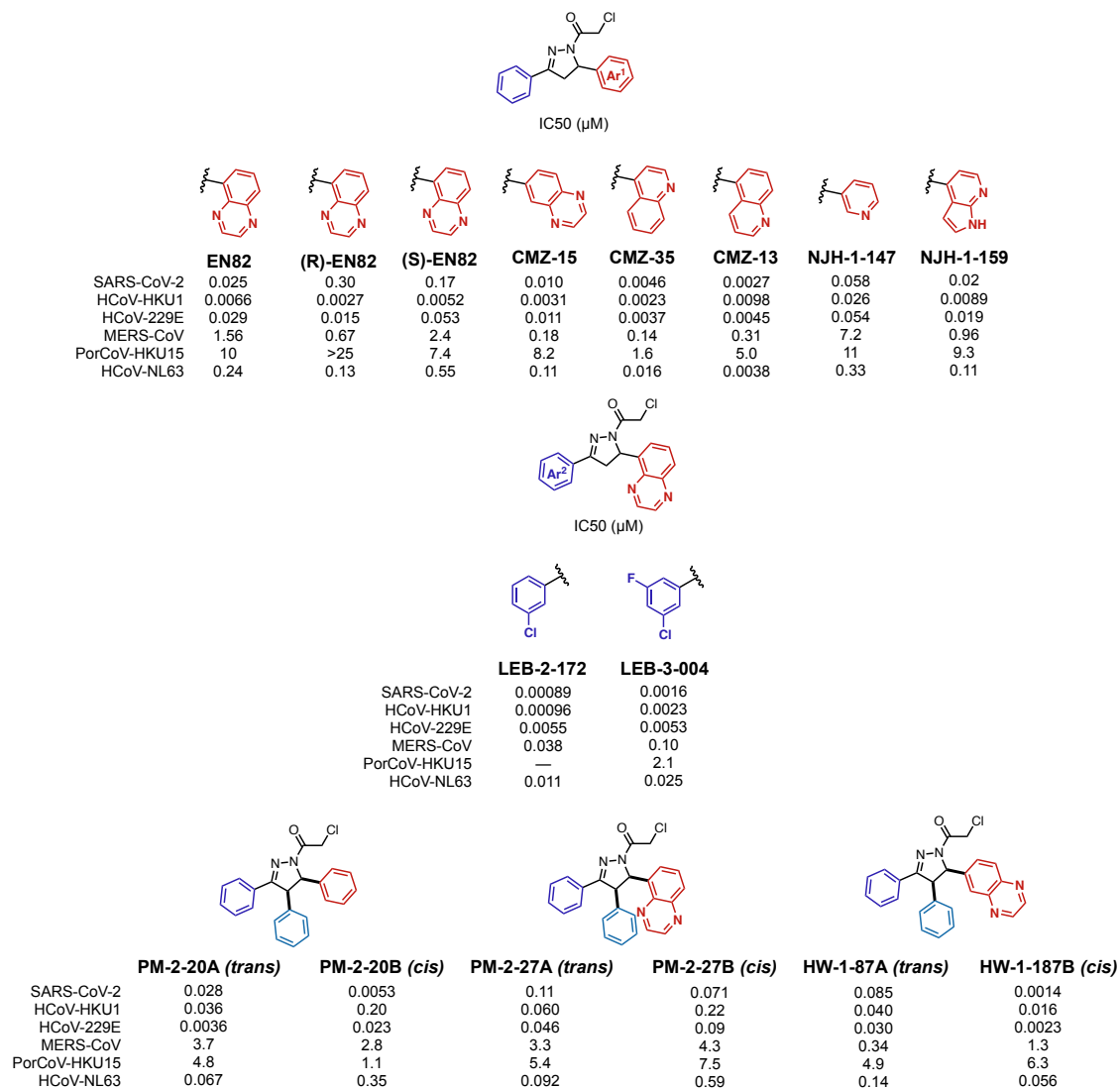


**B**



**Figure S3. Testing reproducible hit compounds in dose-response gel-based ABPP and Mpro substrate activity assays. (A)** Gel-based ABPP studies with reproducible hit compounds. Pure SARS-CoV-2 Mpro protein was pre-incubated with DMSO vehicle or cysteine-reactive covalent ligands (30 min) prior to addition of a rhodamine-conjugated iodoacetamide (IA-rhodamine) probe (100 nM) for 1 hr after which reactions were quenched with SDS/PAGE loading buffer, resolved on SDS/PAGE and visualized by in-gel fluorescence. **(B)** Mpro substrate activity assay using an FRET-based peptide probe. Activity assays in **(B)** are from  $n=3$  biological replicates/group and values are expressed as average  $\pm$  sem.

**Figure S4:**



**Figure S4. Potency (IC<sub>50</sub>, mM) of inhibitors against MPro enzymes from other coronaviruses.** The Mpro activity assays were performed with a MS-based substrate activity assay.

## CHAPTER 4

### Deubiquitinase-targeting Chimeras for Targeted Protein Stabilization

(Adapted from published paper by Nathaniel J. Henning, Lydia Boike, Jessica N. Spradlin, Carl C. Ward, Gang Liu, Erika Zhang, Bridget P. Belcher, Scott M. Brittain, Matthew J. Hesse, Dustin Dovala, Lynn M. McGregor, Rachel Valdez Misiolek, Lindsey W. Plasschaert, David J. Rowlands, Feng Wang, Andreas O. Frank, Daniel Fuller, Abigail R. Estes, Katelyn L. Randal, Anoohya Panidapu, Jeffrey M. McKenna, John A. Tallarico, Markus Schirle & Daniel K. Nomura DOI: 10.1038/s41589-022-00971-2)

## Deubiquitinase-targeting Chimeras for Targeted Protein Stabilization

Engaging the mostly undruggable proteome to uncover new disease therapies not only requires technological innovations that facilitate rapid discovery of ligandable hotspots across the proteome but also demands new therapeutic modalities that alter protein function through novel mechanisms.<sup>13,225</sup> Targeted protein degradation (TPD) tackles the undruggable proteome by targeting specific proteins for ubiquitination and proteasomal degradation. One major class of small-molecule effectors of TPD, proteolysis-targeting chimeras (PROTACs), use heterobifunctional molecules that consist of an E3 ligase recruiter linked to a protein-targeting ligand to induce the formation of ternary complexes that bring together an E3 ubiquitin ligase and the target protein as a neo-substrate.<sup>226,227,228</sup> PROTACs have enabled the targeted and specific degradation of numerous disease-causing proteins in cells.<sup>226,229</sup>

New approaches for TPD have also arisen that exploit endosomal and lysosomal degradation pathways with lysosome-targeting chimeras or autophagy with autophagy-targeting chimeras.<sup>230,231</sup> New approaches for chemically induced proximity beyond degradation have also been developed in recent years, including targeted phosphorylation with phosphorylation-inducing chimeric small molecules and targeted dephosphorylation, but no small-molecule-based induced proximity approaches exist for targeted deubiquitination and subsequent stabilization of proteins.<sup>232,233</sup>

Active ubiquitination and degradation of proteins is the root cause of several classes of diseases, including many tumor suppressors in cancer (for example, TP53, CDKN1A, CDN1C and BAX), and mutated and misfolded proteins, such as  $\Delta$ F508-CFTR in cystic fibrosis or glucokinase in pancreatic cells in maturity-onset diabetes of the young type 2. In these cases, a targeted protein stabilization (TPS) therapeutic strategy, rather than degradation, would be beneficial.<sup>235,236,237,137</sup>

Analogous to TPD, we hypothesized that TPS could be enabled by the discovery of a small-molecule recruiter of a deubiquitinase (DUB) that could be linked to a protein-targeting ligand to form a chimeric molecule, which would induce the deubiquitination and stabilization of proteins of interest. We call this heterobifunctional stabilizer a DUBTAC (**Fig. 1a**). In this study, we report the discovery of a covalent recruiter for the K48-ubiquitin chain-specific DUB OTUB1, which when linked to a protein-targeting ligand stabilizes an actively degraded target protein to demonstrate proof of concept for the DUBTAC platform.

### 4.1 Identifying Allosteric Ligandable Sites within DUBs

To enable the DUBTAC platform, our first goal was to identify a small-molecule recruiter that targeted an allosteric site on a DUB without inhibiting DUB function, as the recruitment of a functional DUB would be required to deubiquitinate and stabilize the target protein. While many DUBs possess well-defined active sites bearing a catalytic and highly nucleophilic cysteine, there have not yet been systematic evaluations of allosteric, non-catalytic and ligandable sites on DUBs that could be pharmacologically

targeted to develop a DUB recruiter. Chemoproteomic platforms, such as activity-based protein profiling (ABPP), have proven to be powerful approaches to map proteome-wide covalently ligandable sites. ABPP utilizes reactivity-based amino acid-specific chemical probes to profile reactive, functional and potentially ligandable sites directly in complex biological systems.<sup>13,137</sup> When used in a competitive manner, pretreatment with libraries of covalent small-molecule ligands can be used to screen for competition of probe binding to recombinant protein or complex proteomes to enable covalent ligand discovery against potential ligandable sites revealed by the reactivity-based probe.<sup>13,138</sup> Previous studies have shown that isotopic tandem orthogonal proteolysis-ABPP (isoTOP-ABPP) platforms for mapping sites of labeling with reactivity-based probes using quantitative proteomic approaches can identify hyperreactive, functional and ligandable cysteines.<sup>137,138</sup>

To identify DUB candidates that possess potential ligandable allosteric cysteines, we analyzed our research group's aggregate chemoproteomic data of proteome-wide sites modified by reactivity-based probes collected since the start of our laboratory. Specifically, we mined our collective chemoproteomic data of cysteine-reactive alkyne-functionalized iodoacetamide (IA-alkyne) probe labeling sites from 455 isoTOP-ABPP experiments in human cell line proteomes for total aggregate spectral counts identified for each probe-modified site across the DUB family. We postulated that probe-modified cysteines within DUBs that showed the highest spectral counts aggregated over all chemoproteomic datasets compared to those sites within the same DUB that showed lower spectral counts may represent more reactive and potentially more ligandable cysteines. Caveats to this premise include cysteines that might be located in regions within a protein sequence that do not yield suitable tryptic peptides with respect to ionization and compatibility with MS-based sequencing and labeling of surface-exposed cysteines that may not be part of binding pockets. However, we conjectured that the aggregate chemoproteomics data would still yield candidate allosteric ligandable sites within DUBs that could be prioritized for covalent ligand screening. We initially mined our aggregate chemoproteomic data for 66 members of the cysteine protease family of DUBs, including ubiquitin-specific proteases, ubiquitin C-terminal hydrolases, Machado-Josephin domain proteases and ovarian tumor proteases (OTU), as they encompass the majority of DUB superfamilies. Interestingly, we found probe-modified cysteines across all of these DUB enzymes (**Fig. 1b and Supplementary Table 1**).

Consistent with our aggregate chemoproteomic data of probe-modified sites being enriched in functional sites within DUBs, among the 40 DUBs that showed a total of >10 aggregate spectral counts of probe-modified peptides, 24 (60%) showed labeling of the DUB catalytic cysteine (**Fig. 1b**).

We next prioritized this list of 40 DUBs to identify suitable candidates for TPS. We prioritized DUBs where the dominant probe-modified cysteine was (1) located at an allosteric site and not the catalytic cysteine such that we could target the identified cysteine with a covalent ligand while retaining the catalytic activity of the DUB; (2) in a dominantly identified probe-labeled peptide compared to other probe-modified sites within the same DUB, which could indicate a high degree of reactivity and potential

covalent ligandability of the identified allosteric cysteine compared to the catalytic site and (3) frequently identified in chemoproteomics datasets, which would indicate the general accessibility of the cysteine in complex proteomes.

We found 10 DUBs where one probe-modified cysteine represented >50% of the spectral counts of all modified cysteines for the particular protein, of which 7 of these DUBs showed primary sites of probe modification that did not correspond to the catalytic cysteine (**Fig. 1c**). Of these 10 DUBs, OTUB1 C23 was captured with >1,000 total aggregate spectral counts compared to <500 aggregate spectral counts for the other DUBs (**Extended Data Fig. 1a**). In our aggregated chemoproteomic data, the tryptic peptide encompassing OTUB1 C23 was the dominant peptide labeled by IA-alkyne, with >1,500 total spectral counts compared to 15 spectral counts for the peptide encompassing the catalytic C91 and 115 spectral counts for C212 (**Fig. 1d**).

OTUB1 is a highly expressed DUB that specifically cleaves K48-linked polyubiquitin chains, the type of ubiquitin linkage that destines proteins for proteasome-mediated degradation, and C23 represents a known ubiquitin substrate recognition site that is distant from the catalytic C91 (**Extended Data Fig. 1b**).<sup>237,238,239,240,241</sup> Interestingly, C23 is also predicted to be in an intrinsically disordered region of the protein based on Predictor of Natural Disordered Regions (PONDR) analysis of OTUB1 (**Extended Data Fig. 1b**).<sup>242,243</sup> Given our analysis of chemoproteomic data, the properties of OTUB1 and the location of C23, we chose OTUB1 as our candidate DUB for covalent ligand screening using gel-based ABPP approaches to discover an OTUB1 recruiter.

## 4.2 Discovery of a Covalent Recruiter Against OTUB1

We performed a gel-based ABPP screen in which we screened 702 cysteine-reactive covalent ligands against labeling of pure OTUB1 protein with a rhodamine-functionalized cysteine-reactive iodoacetamide (IA-rhodamine) probe (**Fig. 2a, Extended Data Fig. 1b and Supplementary Table 2**). Through this screen, we identified the acrylamide EN523 (**1**) as a top hit (**Fig. 2b**). We confirmed that EN523 dose-responsively displaced IA-rhodamine labeling of OTUB1 without causing any protein aggregation or precipitation (**Fig. 2c**). We next performed liquid chromatography–tandem MS analysis (LC–MS/MS) of tryptic peptides from EN523 bound to OTUB1 and showed that EN523 selectively targets C23, with no detectable modification of the catalytic C91 (**Fig. 2d**). Following these data, we performed an in vitro reconstituted OTUB1 deubiquitination activity assay monitoring monoubiquitin release from diubiquitin and demonstrated that EN523 does not inhibit OTUB1 deubiquitination activity (**Fig. 2e**).<sup>244</sup> These studies were performed in the presence of OTUB1-stimulating ubiquitin-conjugating enzyme E2 D1 (UBE2D1), an E2 ubiquitin ligase that engages in a complex with OTUB1 to stimulate OTUB1 activity.

We next used NMR analysis to further characterize EN523 binding to OTUB1. A <sup>13</sup>C-heteronuclear multiple quantum coherence (HMQC) spectrum of OTUB1 revealed the presence of a homogenous and mostly folded protein with well-dispersed {U}-<sup>2</sup>H,

$^1\text{H}/^{13}\text{C}$ -methyl-Ile/Leu/Val/Ala (ILVA) methyl group resonances (**Extended Data Fig. 2a**). EN523 treatment of OTUB1 led to subtle but significant chemical shift perturbations (**Extended Data Fig. 2b**). As we did not assign any peaks in the OTUB1 spectrum, we could not determine the exact binding site of the small molecule. However, almost all affected resonances had stronger intensities than the average signal strengths and had chemical shifts that are close to the random coil values of the respective amino acids. These observations suggest that the amino acids giving rise to these resonances are located in unfolded sections of the protein (in agreement with our PONDR data), predicting that C23 (the site of EN523 binding) belongs to an intrinsically disordered region of the protein. Next, we investigated if the covalent binding of EN523 to C23 prevented the interaction of OTUB1 with the ubiquitin-loaded and free forms of UBE2D2. As the latter protein activates OTUB1 DUB activity, we wanted to confirm that EN523 did not interfere with this protein–protein interaction. We mixed OTUB1 with ubiquitinated or free UBE2D2 and compared the chemical shifts of OTUB1 residues in the presence and absence of EN523. Binding of either form of the conjugating enzyme to OTUB1 induced strong peak perturbations (**Extended Data Fig. 2c,d**). These shift differences were almost identical for samples with and without EN523. The only differences we detected were shift changes for peaks that are affected by compound binding. These results indicated that EN523 did not interfere with binding of UBE2D2 to OTUB1.

We also explored structure–activity relationships (SARs) of our OTUB1 recruiter (**2–12**). Consistent with the necessity of the reactive acrylamide warhead for interacting with C23 of OTUB1, a non-reactive acetamide version of EN523, NJH-2-080 (**2**), showed loss in binding against OTUB1 (**Extended Data Fig. 3**). Replacing the methylfuran substituent with a benzimidazolone, benzothiophene, benzofuran, phenyloxazole or imidazopyridine, but not methylimidazole, still retained potency against OTUB1 (**Extended Data Fig. 3**). We also explored preliminary SARs of the piperazinone core as well. Dimethyl and methylpiperazinone substitutions with a tert-butyl propionate extension from the furan still maintained potency against OTUB1. We also found that the (R)-methylpiperazinone derivative was more potent than the (S)-methylpiperazinone derivative of EN523, indicating that we may be able to achieve stereochemically specific interactions with OTUB1 (**Extended Data Fig. 3**). These data also demonstrated that extension off the furan may present an optimal exit vector for synthesis of DUBTACs. Overall, while the SAR showed room for flexibility within the EN523 core scaffold, we did not identify significantly more potent OTUB1 ligands, and thus we chose to further pursue EN523 as our OTUB1 recruiter for follow-up studies. An alkyne-functionalized probe of EN523 (NJH-2-075 (**13**)) was then synthesized with the goal of assessing whether this ligand engaged OTUB1 in cells (**Fig. 2f**). NJH-2-075 retained binding to OTUB1 *in vitro*, as shown by (1) gel-based ABPP demonstrating competition of NJH-2-075 against IA-rhodamine labeling of recombinant OTUB1 and (2) direct labeling of recombinant OTUB1 by NJH-2-075 visualized by copper-catalyzed azide-alkyne cycloaddition (CuAAC) of azide-functionalized rhodamine to NJH-2-075-labeled OTUB1 (monitored by *in-gel* fluorescence; **Fig. 2g,h**). We demonstrated NJH-2-075 engagement of OTUB1 in cells by enrichment of endogenous OTUB1, but not that of an unrelated protein vinculin, through NJH-2-075 compared to vehicle treatment in



HEK293T cells (**Fig. 2i**). Collectively, these data highlighted EN523 as a promising covalent OTUB1 ligand that targeted a non-catalytic and allosteric C23 on OTUB1 without inhibiting OTUB1 deubiquitination activity and engaged OTUB1 in cells.

### 4.3 Showing Proof of Concept of DUBTACs with Mutant CFTR

To demonstrate the feasibility of using EN523 as an OTUB1-recruiting module of a heterobifunctional DUBTAC, we identified the mutant  $\Delta$ F508-CFTR chloride channel as a proof-of-concept case where protein stabilization would be therapeutically desirable.  $\Delta$ F508, a frameshift mutation caused by deletion at codon 508 in exon 10 of CFTR, resulting in the absence of a phenylalanine residue, is the most common mutation that induces the cystic fibrosis phenotype.<sup>245</sup> This mutation causes the protein to become conformationally unstable, leading to K48 polyubiquitination and degradation before trafficking from the endoplasmic reticulum to the cell surface.<sup>245,246</sup> Previous studies have demonstrated the feasibility of stabilizing mutant CFTR not only by genetic and pharmacological inhibition of the cognate E3 ligase RNF5 but also through targeted recruitment of DUBs using a genetically encoded and engineered DUB targeted to CFTR using a CFTR-targeting nanobody.<sup>246,247,248</sup> Importantly for our work, suitable CFTR-targeting small-molecule ligands exist. Lumacaftor, a drug for cystic fibrosis developed by Vertex Pharmaceuticals, acts as a chemical chaperone for  $\Delta$ F508-CFTR and corrects its misfolding, leading to increased trafficking of  $\Delta$ F508-CFTR to the cell membrane and partial restoration of protein function.<sup>249</sup> Despite lumacaftor's chaperoning activity, the vast majority of  $\Delta$ F508-CFTR is still actively ubiquitinated and degraded, making the potential of a synergistic effect via DUBTAC-induced deubiquitination a therapeutically attractive option.

With this in mind, we synthesized DUBTACs linking the OTUB1 recruiter EN523 to the CFTR chaperone lumacaftor with two different C3 or C5 alkyl linkers, NJH-2-056 (**14**) and NJH-2-057 (**15**) (**Fig. 3a,b**). We confirmed that these two DUBTACs still engaged recombinant OTUB1 in vitro by gel-based ABPP (**Fig. 3c,d**). We used CFBE41o-4.7 human bronchial epithelial cells expressing  $\Delta$ F508-CFTR, a human cystic fibrosis bronchial epithelial cell line, as a model system to test our DUBTACs. We first showed that EN523 did not alter OTUB1 protein levels in these cells (**Extended Data Fig. 4a**). We also demonstrated that the alkyne-functionalized EN523 probe NJH-2-075 still engaged OTUB1 in this cell line (**Extended Data Fig. 4b**). The DUBTACs were then tested in these cells alongside lumacaftor or EN523 treatment alone. Treatment with NJH-2-056, lumacaftor, or EN523 did not alter mutant CFTR levels; however, we observed a robust and significant increase in CFTR protein levels with NJH-2-057 treatment (**Fig. 3e,f**). This stabilization was dose responsive and time dependent (**Extended Data Fig. 5**). We further confirmed that the stabilized protein was CFTR using three additional commercially available CFTR antibodies (**Extended Data Fig. 6**) and showed that the DUBTAC-stabilized CFTR band was attenuated following CFTR knockdown (**Extended Data Fig. 7**).

We also explored the dependence of CFTR stabilization on linker length and composition (**14–20**) (**Extended Data Fig. 8**). DUBTACs bearing C5 and C6 alkyl

linkers, but not C3 and C4 alkyl linkers, stabilized CFTR. Interestingly, none of the DUBTACs bearing PEG linkers were able to stabilize CFTR (**Extended Data Fig. 8**). We also made eight additional CFTR DUBTACs bearing more rigid heterocycle-containing linkers (**21–28**) to determine whether these compounds that may be less flexible and potentially more drug-like may perform better in stabilizing mutant CFTR (**Extended Data Fig. 9**). Interestingly, nearly all these DUBTACs increased CFTR levels, with substantially improved response from GL-03 (**23**) (bearing a fused azepane-pyrrolidine linker) compared to NJH2-075 (**Extended Data Fig. 9**). Consistent with the necessity of the cysteine-reactive warhead in binding to OTUB1, we also demonstrated that a non-reactive propionamide version of NJH-2-057, NJH-2-106 (**29**), was incapable of stabilizing CFTR (**Extended Data Fig. 10**). The alkyne probe NJH-2-075, which contains an identical linker to NJH-2-057 but exchanges an ethynylphenyl group for lumacaftor, also did not induce CFTR stabilization (**Extended Data Fig. 10**). The CFTR smear that we observed in the blot with NJH-2-057 treatment is consistent with previous studies investigating CFTR and in line with our observations with the proteasome inhibitor bortezomib that would maximally stabilize actively ubiquitinated and degraded mutant CFTR protein. This smear likely represents a combination of differential glycosylation states, other forms of ubiquitination on CFTR that may not be removed by OTUB1 (for example, K63 ubiquitin chains) and previously observed anomalous migration of CFTR on SDS-PAGE due to the presence of SDS-resistant ternary structures within the protein.<sup>250</sup> Based on the molecular weight of the darkest part of the CFTR blot >225 kDa, we conjectured that we are stabilizing the fully mature glycosylated form of mutant CFTR (**Fig. 3e**). To further validate our western blot data for CFTR stabilization and to assess the proteome-wide activity of NJH-2-057, we performed a tandem mass tags (TMT)-based quantitative proteomic analysis of NJH-2-057-treated CFBE41o-4.7 cells expressing  $\Delta$ F508-CFTR. Satisfyingly, the proteomic analysis showed CFTR among the most robustly stabilized proteins (ratio of 7.8 comparing NJH-2-057 to vehicle treatment; **Fig. 3g** and **Supplementary Table 3**). While there were additional proteins with significant changes in abundance levels, we only observed 21 proteins that were significantly stabilized by greater than fivefold with an adjusted P value of appear to be DUBTAC-specific changes because these changes were not detected with EN523 or lumacaftor treatment alone (**Supplementary Fig. 2 and Supplementary Table 3**) and may represent compensatory changes occurring from elevations in CFTR levels in cells or could represent changes resulting from off-targets of the DUBTAC. Interestingly, among proteins elevated along with CFTR were several protein chaperones, including heat shock 70-kDa protein 6, DnaJ heat shock protein family Hsp40 member B1, heat shock 50-kDa protein 1A and DNAJ homolog subfamily B member 4.

These changes could reflect potential compensatory on-target upregulation of protein chaperones in response to highly elevated levels of a relatively unstable mutant CFTR. Nonetheless, we did not observe widespread alterations in protein levels with DUBTAC treatment, suggesting that we were not substantially disrupting global protein turnover. Having identified NJH-2-057 as a DUBTAC that was capable of stabilizing mutant CFTR in cells, we next sought to confirm the formation of a ternary complex between CFTR, NJH-2-057 and OTUB1 in vitro using recombinant protein and native

MS-based approaches (**Fig. 3h,i**). While the highest intensity signals corresponded to unmodified OTUB1 and the  $\Delta$ F508-harboring CFTR nucleotide-binding domain used in this experiment, potentially indicating low levels of target engagement under these experimental conditions, we observed significant CFTR–OTUB1 complex formation with NJH-2-057 treatment but not with DMSO vehicle or EN523 treatment (**Fig. 3h,i**). The predominantly observed mass for this complex corresponded to OTUB1 and CFTR but not the combined masses of OTUB1, CFTR and NJH-2-057. This may be because the NJH-2-057 adduct on OTUB1 may be unstable to electrospray ionization and desolvation energy conditions required to observe the protein complex, as we were also not able to observe the NJH-2-057 mass adduct on OTUB1. However, minor peaks indicated the presence of adducts consistent with either full NJH2-057 or either MS-induced fragments or breakdown products (**Supplementary Fig. 3**). Nonetheless, given that this OTUB1–CFTR complex was only observed with DUBTAC treatment but not with DMSO or EN523 treatment, our data strongly suggest that the DUBTAC enables ternary complex formation.

To further confirm that the robust stabilization in mutant CFTR levels conferred by NJH-2-057 was due to the proposed on-target activity, we demonstrated that stabilization of CFTR was attenuated by pretreatment with either lumacaftor or EN523, indicating that stabilization by the DUBTAC was due to targets engaged by both lumacaftor and EN523 (**Fig. 4a,b**). These data also indicate the necessity for ternary complex formation to stabilize CFTR levels. To further verify that the CFTR stabilization was dependent on OTUB1, OTUB1 knockdown significantly attenuated mutant CFTR stabilization by NJH-2-057 (**Fig. 4c,d**). We next performed a competitive isoTOP-ABPP study to assess the overall proteome-wide selectivity and cysteine reactivity of NJH-2-057 treatment in CFBE41o-4.7 cells expressing  $\Delta$ F508-CFTR (**Supplementary Table 4**). Of 1,270 IA-alkyne probe-modified peptides quantified in two of three biological replicates, there were only five targets that showed isotopically light-to-heavy or control-to-NJH-2-057 treatment probe-modified peptide ratios of >4 with an adjusted P value of <0.05: VDAC2 C76, TUBB1 C201, RLF C744, VDAC2 C47 and VDAC3 C66 (**Supplementary Table 4**). Yang et al. previously showed that Nedd4 could ubiquitinate VDAC2 and VDAC3, and overexpression of Nedd4 significantly increased K48-linked ubiquitination of VDAC2 and VDAC3.<sup>251</sup>

However, none of these targets would be expected to directly influence the activity of our DUBTAC. OTUB1 C23 was captured in our isoTOP-ABPP experiment but only showed a ratio of 1.6, which would correspond to ~60% target occupancy (**Supplementary Table 3**). This likely indicates that the observed CFTR stabilization by NJH-2-057 is occurring through relatively low levels of OTUB1 occupancy in cells, which would also be in line with our in vitro labeling and native MS data. The activity of heterobifunctional molecules has also been reported previously in studies using covalent E3 ligase recruiters for targeted protein degradation applications showing that relatively minimal target occupancy of E3 ligases can still lead to robust degradation of target proteins due to the catalytic mechanism of action of the E3 ligases.<sup>148,150,151,152</sup> We conjectured that a similar catalytic effect in a DUBTAC also leads to robust stabilization of the target protein with partial OTUB1 occupancy. Having shown

stabilization of CFTR protein levels with our DUBTAC, we next sought to determine whether our DUBTAC-mediated increase in CFTR protein levels led to improved cell surface CFTR function (**Fig. 4e,f**). We measured transepithelial conductance in primary human cystic fibrosis donor bronchial epithelial cells bearing the  $\Delta F508$ -CFTR mutation. These cells were pretreated with vehicle, lumacaftor or NJH-2-057 for 24h before sequential treatments with a sodium channel inhibitor amiloride, a cAMP activator forskolin and a CFTR potentiator VX770 (ivacaftor) to fully activate CFTR function in cells. After chloride channel conductance was potentiated with VX770, the cells were treated with a CFTR inhibitor, CFTR(inh)-172, to show CFTR dependence of any increases observed in transepithelial conductance. The difference in conductance between potentiator VX770 and CFTR inhibitor treatment were quantified under the three different treatment conditions (vehicle, lumacaftor or DUBTAC treatment) to ascertain the effects that our DUBTAC had on CFTR-mediated conductance compared to lumacaftor or vehicle treatment (**Fig. 4e,f**). Our studies showed that treatment of these primary cells with NJH-2-057 led to significant improvement in CFTR-dependent transepithelial conductance compared to lumacaftor or vehicle treatment, indicating that our CFTR DUBTAC elevated not only CFTR protein levels but also functional CFTR at the cell surface, leading to improved CFTR function (**Fig. 4e,f**).

#### 4.4 Using DUBTACs to Stabilize WEE1

We next sought to show a second example of TPS with DUBTACs against another actively degraded target for which a well-validated ligand had been reported in the literature. We selected WEE1, a tumor suppressor kinase in non-malignant eukaryotic somatic cells that phosphorylates the cyclin-dependent kinase (CDK1)–cyclin B1 complex to inhibit cell cycle progression during S and G2 phases of mitosis and whose activity must be downregulated for mitotic progression to occur.<sup>252,253</sup> One mechanism through which WEE1 activity is suppressed is via ubiquitin-mediated proteasomal degradation.<sup>254</sup> Clinical WEE1 inhibitors, such as AZD1775, have been developed to be given in combination with DNA-damaging chemotherapy agents for inducing premature mitosis to exert anticancer effects, and WEE1 PROTACs using AZD1775 have also been developed to selectively degrade WEE1 in cancer cells.<sup>255</sup> We first confirmed previously reported results that treatment of HEP3B hepatoma cancer cell lines with a proteasome inhibitor bortezomib stabilized WEE1 levels, demonstrating that WEE1 was regulated by ubiquitin-mediated proteasomal degradation in this cell line (**Fig. 5a**).<sup>254</sup> We next synthesized four DUBTACs linking AZD1775 to our OTUB1 recruiter EN523 through no linker (LEB-03-153) (**30**), a C3 alkyl linker (LEB03-144) (**31**), a C5 alkyl linker (LEB-03-145) (**32**) or a PEG2 linker (LEB-03-146) (**33**) (**Fig. 5b**). LEB-03-144 and LEB-03-146 with the C3 alkyl linker and the PEG2 linker, respectively, showed significant WEE1 stabilization in HEP3B cells comparable to WEE1 levels observed with bortezomib treatment, whereas EN523 or AZD1775 treatment alone had no impact on WEE1 levels (**Fig. 5c,d**). While the therapeutic relevance of a WEE1 DUBTAC using a WEE1 inhibitor remains to be seen, these data show additional mechanistic proof of concept for TPS using DUBTACs.

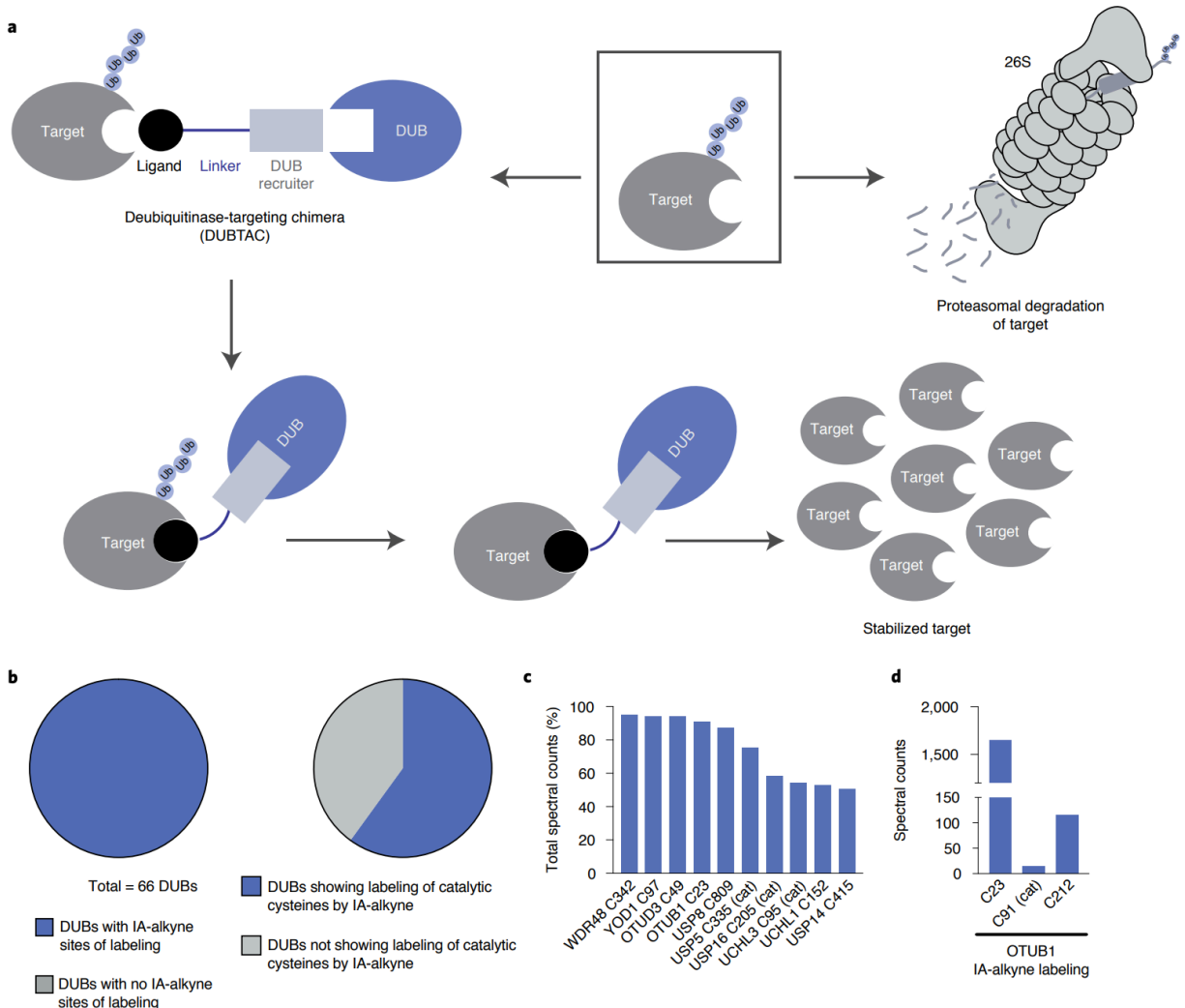
#### 4.5 Discussion

In this study, we discovered a covalent small-molecule recruiter EN523 for the K48-ubiquitin chain-specific DUB OTUB1. We demonstrated that this recruiter can be incorporated into fully synthetic heterobifunctional DUBTACs by linking a DUB recruiter to protein-targeting ligands to enable TPS of actively degraded target proteins in cells. We showed two successful examples of TPS with  $\Delta$ F508-CFTR and WEE1. For  $\Delta$ F508-CFTR, we also demonstrated that we not only heightened the levels of the mutant protein but also improved cell surface chloride channel conductance of CFTR with our DUBTAC in combination with the potentiator ivacaftor compared to lumacaftor and ivacaftor treatments. While we showed early validation of the DUBTAC platform here, there are many avenues for future exploration. These include further optimization of DUB recruiters against OTUB1 to improve their potency and proteome-wide selectivity as well as the discovery of new recruiters against other candidate DUBs. Much like with PROTACs, we observe significant dependency on linker length and composition for stabilizing target proteins with DUBTACs. Further exploration of this linker dependence with DUBTACs will be necessary for improving potency and kinetics of stabilization and bioavailability. In addition, elucidating the mechanism, structural underpinnings and kinetics in the formation of the ternary complex formed between the target protein and DUB and understanding how the target protein is deubiquitinated by the DUBTAC will be important. Using more advanced methods beyond native MS to monitor ternary complex formation, including AlphaLISA and time-resolved fluorescence resonance energy transfer, will be useful to achieve these goals.<sup>256</sup>

Given our initial proof of concept for CFTR and WEE1 stabilization with a DUBTAC, there are many promising areas that could benefit from targeted deubiquitination of actively ubiquitinated and degraded proteins to provide therapeutic benefit. Targets that could benefit from a DUBTAC that already possess protein-targeting ligands include stabilizing BAX levels in the mitochondria to induce apoptosis, stabilizing STING for immunooncology applications or stabilizing glucokinase in pancreatic cells for maturity-onset diabetes of the young type.<sup>257,258,259,260</sup> Other targets that would benefit from a DUBTAC would be various tumor suppressors that are actively ubiquitinated and degraded to maintain cancer cell proliferation.<sup>261</sup> There are also many other genetic disorders beyond cystic fibrosis where mutations can lead to protein destabilization and ubiquitin-mediated degradation that can also be stabilized by DUBTACs for therapeutic benefit. These include glucocerebrosidase mutations in Gaucher's disease or Parkinson's disease<sup>262</sup> and phenylalanine hydroxylase and fumarylacetoacetate hydroxylase mutations in phenylketonuria.<sup>263</sup> These disorders could directly benefit from TPS via DUBTACs. In diseases caused by haploinsufficiency where loss of one copy of a gene leads to disease pathology, DUBTACs could potentially slow down the turnover rate of the protein to increase the levels of the protein to alleviate the disease<sup>47</sup>. Overall, our study puts forth the discovery of DUB recruiters and shows proof of concept for the DUBTAC platform for TPS via induced proximity of a DUB with a target protein. In addition, our study underscores the utility of using chemoproteomics-enabled covalent ligand discovery platforms to facilitate development of unique induced proximity-based therapeutic modalities beyond TPD.

## 4.6 Figures and Tables

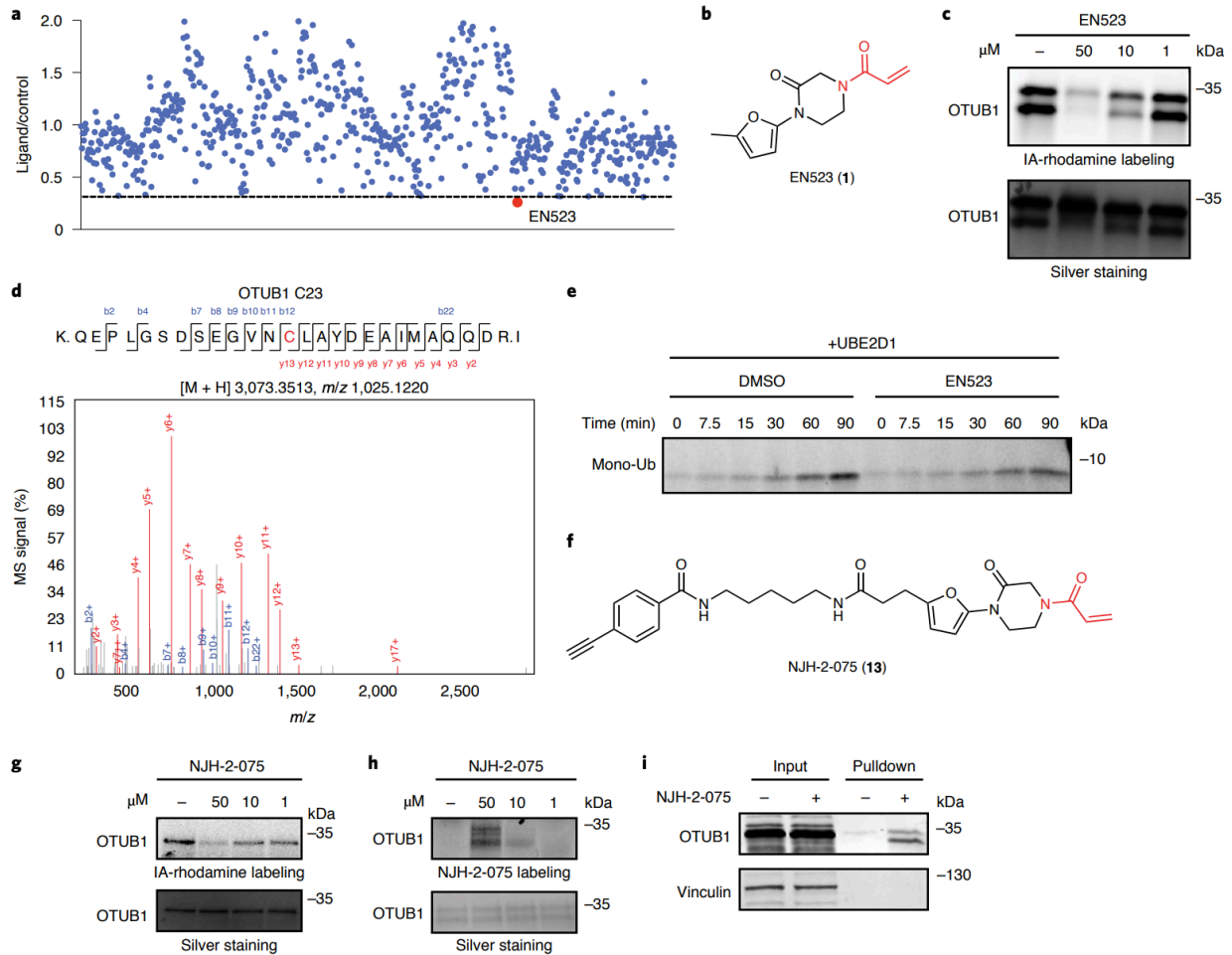
**Figure 1 (a-d):**



**Fig. 1. DUBTAC platform.** **a**, DUBTACs are heterobifunctional molecules consisting of a protein-targeting ligand linked to a DUB recruiter via a linker. DUBTACs are ideally used for stabilizing the levels of actively ubiquitinated proteins that are degraded by the proteasome. When treated in cells, a DUBTAC will induce the proximity of a DUB with a target protein to remove polyubiquitin chains to prevent the protein from undergoing proteasome-mediated degradation, thereby stabilizing and elevating the level of the actively degraded protein; Ub, ubiquitin. **b**, Of 65 DUBs mined in our research group's aggregate chemoproteomic datasets of cysteine-reactive probe labeling with IA-alkyne in various complex proteomes, we identified probe-modified cysteines across 100% of the 65 DUBs. This is shown in the first pie chart. Among the 65 DUBs that showed probe-modified cysteines, 40 of these DUBs showed >10 aggregate spectral counts across our chemoproteomic datasets. Twenty-four, or 60%, of the 40 DUBs showed labeling of the DUB catalytic or active-site cysteines. **c**, Mining the DUB data, we identified 10 DUBs wherein there was one probe-modified cysteine that represented

>50% of the total aggregate spectral counts for probe-modified cysteine peptides for the particular DUB. Seven of the 10 DUBs do not target a known catalytic cysteine, whereas three do target the catalytic cysteine (abbreviated 'cat'); USP, ubiquitin-specific proteases; UCH, ubiquitin C-terminal hydrolases; OTU, ovarian tumor proteases. **d**, Analysis of aggregate chemoproteomic data for OTUB1 IA-alkyne labeling showing that C23 is the dominant site labeled by IA-alkyne compared with the catalytic (cat) C91. Chemoproteomic data analysis of DUBs across aggregated datasets can be found in Supplementary Table 1.

**Figure 2 (a-i):**

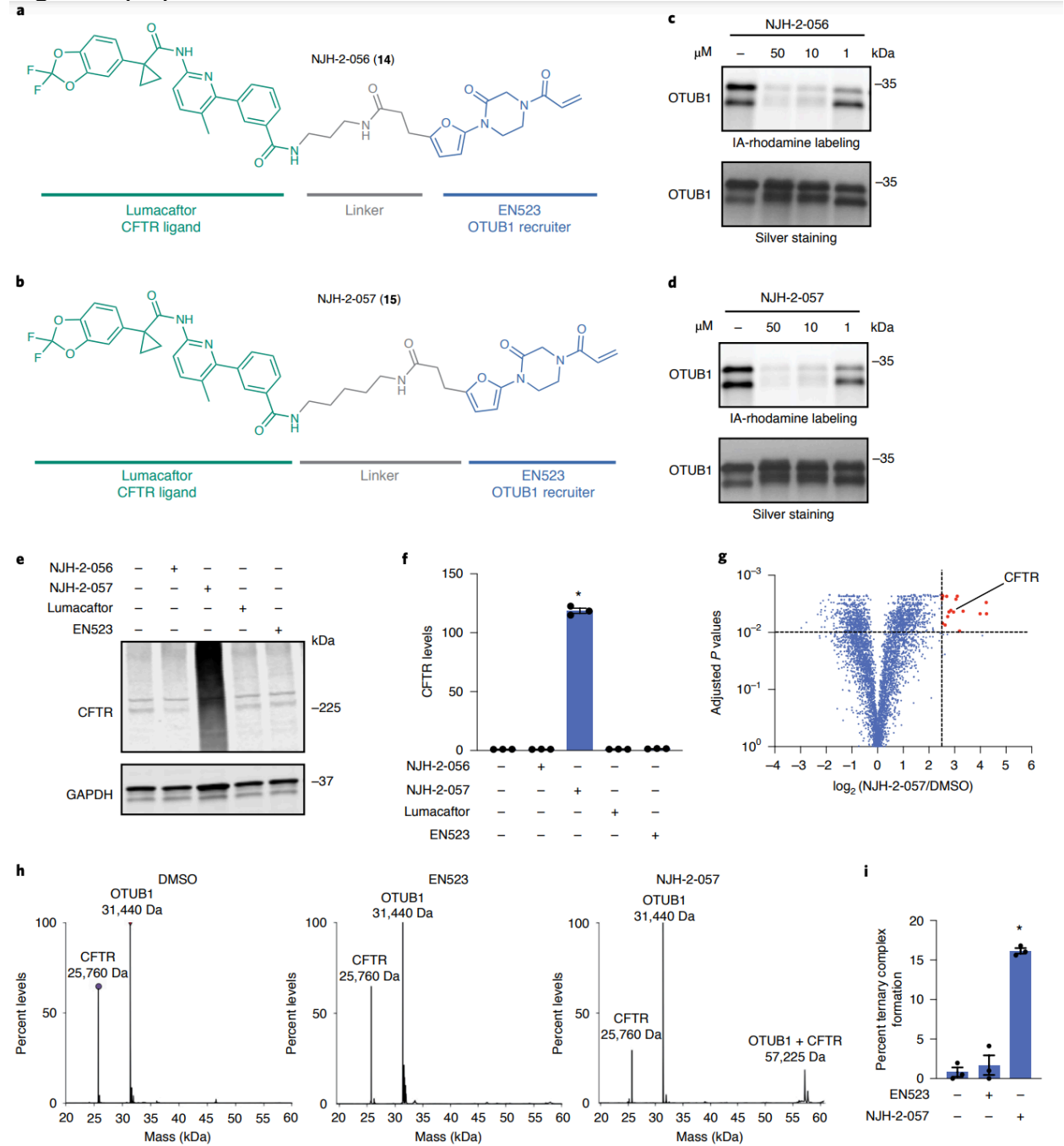


**Fig. 2 Discovery of covalent ligands that target OTUB1.** **a**, Covalent ligand screen of a cysteine-reactive library competed against IA-rhodamine labeling of recombinant OTUB1 to identify binders to OTUB1 by gel-based ABPP. Vehicle DMSO or cysteine-reactive covalent ligands (50 μM) were preincubated with OTUB1 for 30 min at room temperature before IA-rhodamine labeling (500 nM for 30 min at room temperature). OTUB1 was then separated by SDS-PAGE, and in-gel fluorescence was assessed and quantified. Gel-based ABPP data and quantification of in-gel fluorescence are shown in Extended Data Fig. 1b and Supplementary Table 2. EN523 annotated in red was the top hit that showed the greatest inhibition of OTUB1 IA-rhodamine labeling. **b**, Structure of EN523, with cysteine-reactive acrylamide highlighted in red. **c**, Gel-based ABPP confirmation showing dose-responsive inhibition of IA-rhodamine binding of OTUB1. Vehicle DMSO or EN523 were preincubated with recombinant OTUB1 for 30 min at 37 °C before IA-rhodamine labeling (500 nM for 30 min at room temperature). OTUB1 was then separated by SDS-PAGE, and in-gel fluorescence was assessed. The silver staining demonstrates protein loading. Shown is a representative gel of n=3 biologically independent samples per group. **d**, LC-MS/MS data showing an EN523-modified adduct on C23 of OTUB1. OTUB1 (10 μg) recombinant protein was incubated with EN523 (50 μM) for 30 min, after which the protein was precipitated and digested with



trypsin, and tryptic digests were analyzed by LC–MS/MS to identify modified sites. **e**, OTUB1 DUB activity monitored by cleavage of K48 diubiquitin. Recombinant OTUB1 was preincubated with DMSO or EN523 (50  $\mu$ M) for 1 h. After preincubation, OTUB1 was added to a mixture of diubiquitin and UBE2D1. The appearance of monoubiquitin (mono-Ub) was monitored by western blotting. **f**, Structure of the alkyne-functionalized EN523 probe NJH-2-075. **g**, Gel-based ABPP of NJH-2-075. Vehicle DMSO or NJH-2-075 were preincubated with OTUB1 for 30 min at 37 °C before IA-rhodamine labeling (500 nM for 30 min at room temperature). OTUB1 was then separated by SDS–PAGE, and in-gel fluorescence was assessed. Also shown is silver staining demonstrating protein loading. **h**, NJH-2-075 labeling of recombinant OTUB1. OTUB1 (0.5  $\mu$ g) was labeled with DMSO or NJH-2-075 for 1.5 h at 37 °C, after which rhodamine-azide was appended by CuAAC, OTUB1 was separated by SDS–PAGE and in-gel fluorescence was assessed. Also shown is silver staining demonstrating protein loading. **i**, NJH-2-075 engagement of OTUB1 in HEK293T cells. HEK293T cells were treated with DMSO vehicle or NJH-2-075 (50  $\mu$ M) for 2 h, after which cell lysates were subjected to CuAAC with biotin picolyl azide, and NJH-2-075-labeled proteins were subjected to avidin pulldown, elution, separation by SDS–PAGE and blotting for OTUB1 and vinculin. Both input lysate and pulldown levels are shown. Gels or blots shown in **c**, **e** and **g–i** are representative of  $n=3$  biologically independent samples per group. Raw gels and blots can be found in the source data.

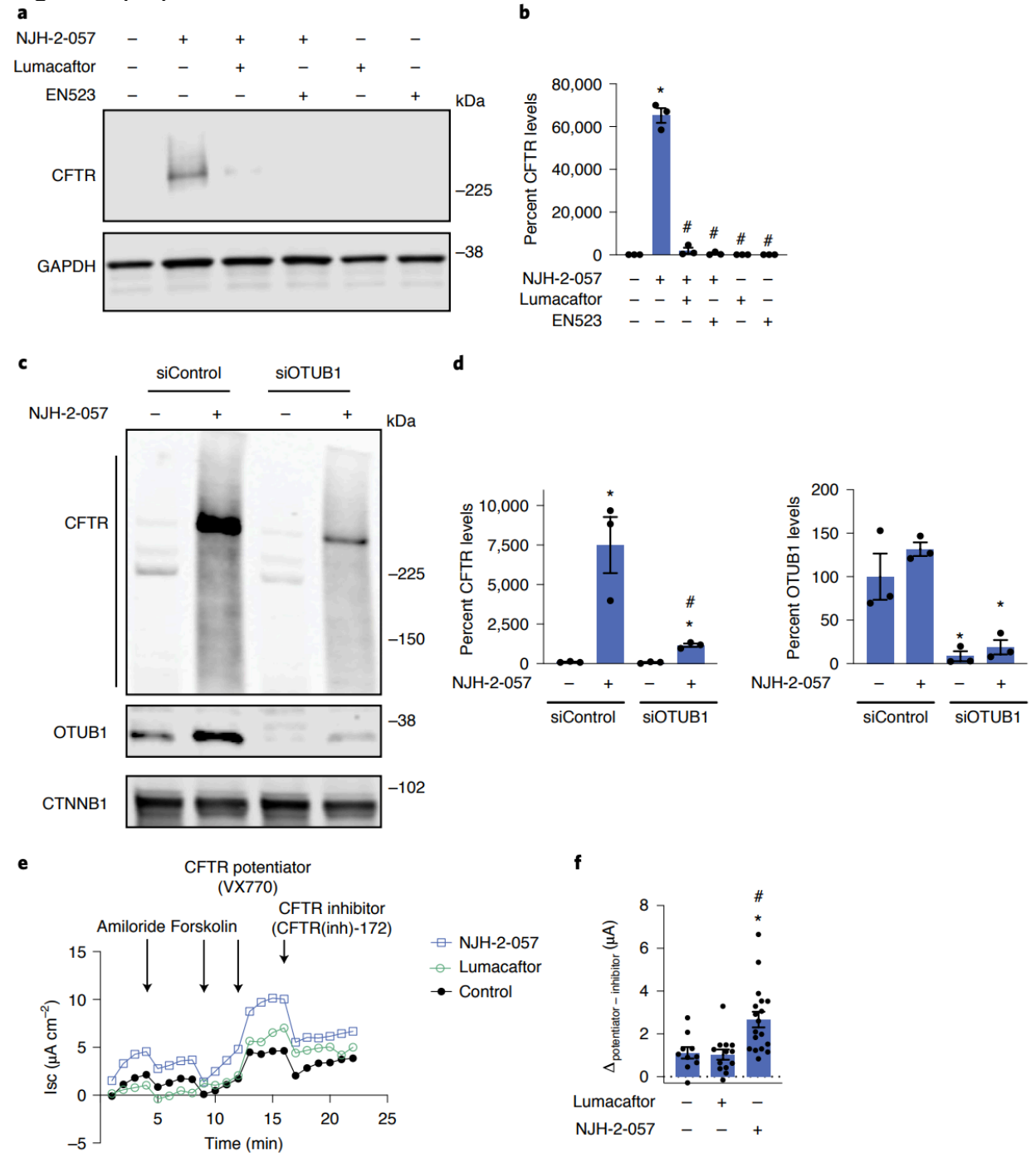
**Figure 3 (a-i):**



**Fig. 3 DUBTAC against mutant CFTR. a,b**, Structures of NJH-2-056 (**a**) and NJH-2-057 (**b**); these DUBTACs against mutant CFTR protein are formed by linking the CFTR ligand lumacaftor to the OTUB1 recruiter EN523 through C3 and C5 alkyl linkers, respectively. **c,d**, Gel-based ABPP analysis of NJH-2-056 (**c**) and NJH-2-057 (**d**) against OTUB1. Vehicle DMSO or DUBTACs were preincubated with recombinant OTUB1 for 30 min at 37 °C before addition of IA-rhodamine (100 nM) for 30 min at room temperature. OTUB1 was run on SDS-PAGE, and in-gel fluorescence was assessed. Protein loading was assessed by silver staining. **e**, Effect of DUBTACs on mutant CFTR levels. CFBE41o-4.7 cells expressing  $\Delta\text{F508}$ -CFTR were treated with vehicle DMSO,

NJH-2-056 (10  $\mu$ M), NJH-2-057 (10  $\mu$ M), lumacaftor (10  $\mu$ M) or EN523 (10  $\mu$ M) for 24 h, and mutant CFTR and loading control GAPDH levels were assessed by western blotting. **f**, Quantification of the experiment described in **e**. **g**, TMT-based quantitative proteomic profiling of NJH-2-057 treatment. CFBE41o-4.7 cells expressing  $\Delta$ F508-CFTR were treated with vehicle DMSO or NJH-2-057 (10  $\mu$ M) for 24 h. Data shown are from n= 3 biologically independent samples per group. Full data for this experiment can be found in Supplementary Table 3. **h**, Native MS analysis of DUBTAC-mediated ternary complex formation. OTUB1 (2  $\mu$ M) and the CFTR nucleotide-binding domain 1 (2  $\mu$ M) were incubated with DMSO vehicle, EN523 (50  $\mu$ M) or NJH-2-057 (50  $\mu$ M) in 150 mM ammonium acetate with MgCl<sub>2</sub> (100  $\mu$ M) and ATP (100  $\mu$ M). Representative mass spectra from n= 3 biologically independent samples per group are shown. **i**, Percentage of ternary complex formation assessed by measuring the CFTR–OTUB1 complex formed in the experiment described in **h**. Gels shown in **c**, **d** and **e** are representative of n= 3 biologically independent samples per group. Data in **f** and **i** show individual biological replicate values and average  $\pm$  s.e.m. from n= 3 biologically independent samples per group. Statistical significance was calculated with unpaired two-tailed Student's t-tests in **f** and **i** compared with vehicle-treated controls and is expressed as \*P<.05. Raw gels and blots, bar graph data and exact P values can be found in the source data.

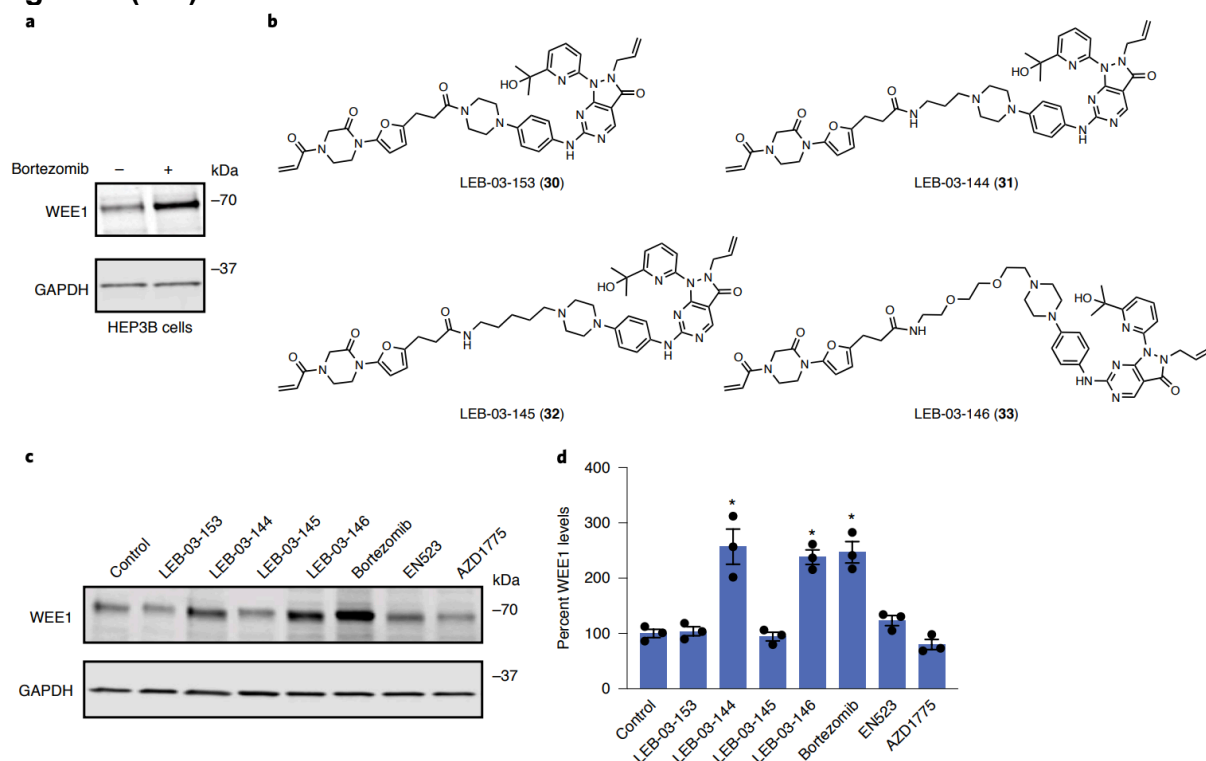
**Figure 4 (a-f):**



**Fig. 4 Characterizing the mechanism of the CFTR DUBTAC NJH-2-057.** a, Effect of lumacaftor or EN523 preincubation on NJH-2-057 DUBTAC-mediated stabilization of mutant CFTR levels. CFBE41o-4.7 cells expressing  $\Delta F508$ -CFTR were pretreated with vehicle DMSO, lumacaftor (100  $\mu$ M) or EN523 (100  $\mu$ M) for 1 h before treatment with NJH-2-057 (10  $\mu$ M) for 24 h. Mutant CFTR and loading control GAPDH levels were

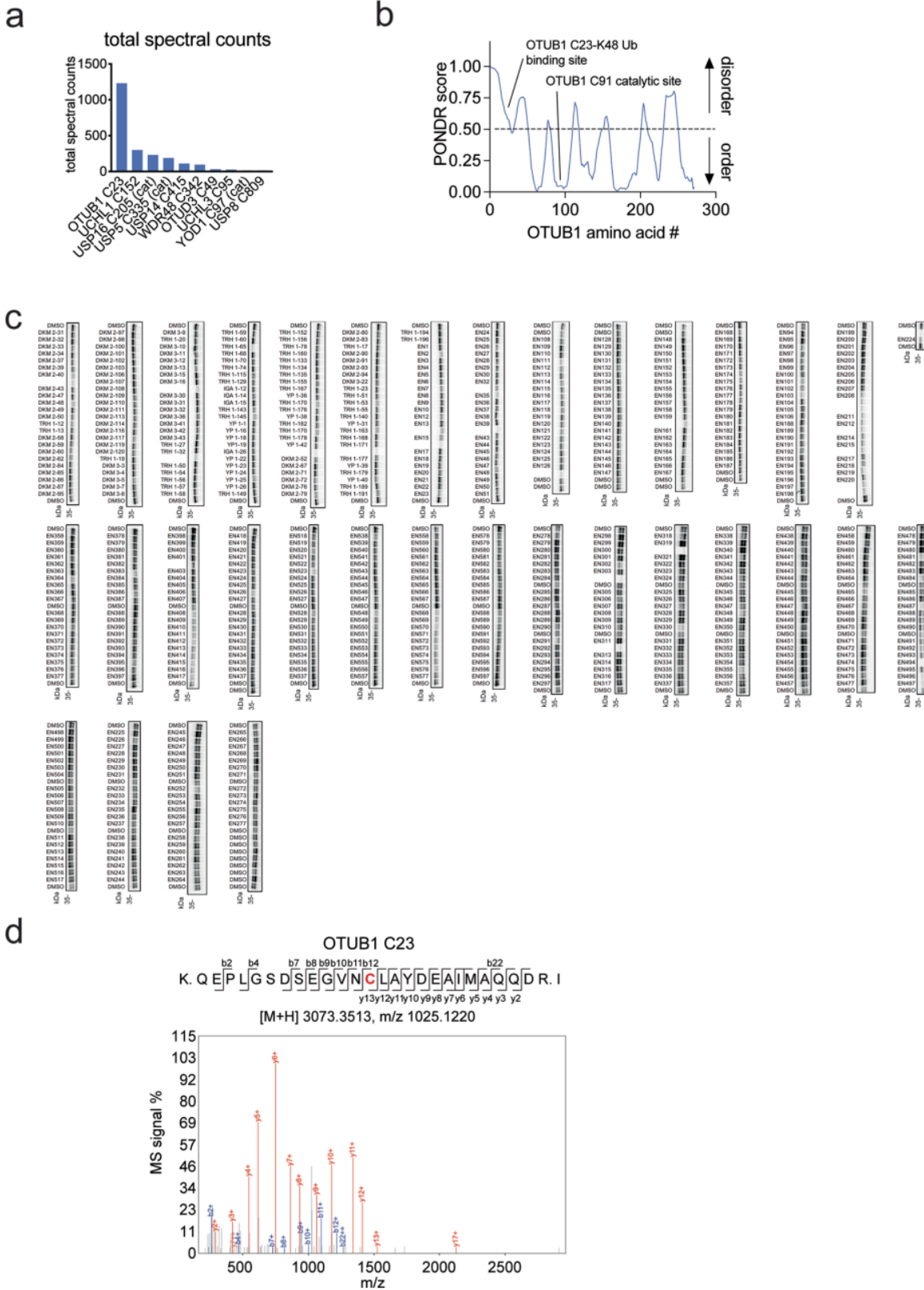
assessed by western blotting. **b**, Quantification of the experiment described in **a**. **c**, Effect of OTUB1 knockdown on NJH-2-057 DUBTAC-mediated mutant CFTR stabilization. CFBE41o-4.7 cells expressing  $\Delta F508$ -CFTR were transiently transfected with non-targeting short interfering RNA (siRNA; siControl) or siOTUB1 oligonucleotides for 48 h before treatment of cells with vehicle DMSO or NJH-2-057 (10  $\mu$ M) for 16 h. Mutant CFTR, OTUB1 and loading control GAPDH levels were assessed by western blotting. **d**, Levels of mutant CFTR and OTUB1 from the experiment described in **c**. **e**, Transepithelial conductance, expressed as short-circuit current ( $I_{sc}$ ), in primary human cystic fibrosis donor bronchial epithelial cells bearing the  $\Delta F508$ -CFTR mutation. Cells were treated with DMSO vehicle, NJH-2-057 (10  $\mu$ M) or lumacaftor (10  $\mu$ M) 24 h before the TECC24 assay in which cells received four additional sequential treatments with a sodium channel inhibitor amiloride (10  $\mu$ M), a cAMP activator forskolin (20  $\mu$ M), a CFTR potentiator VX770 (0.5  $\mu$ M) and a CFTR inhibitor CFTR(inh)-172 (30  $\mu$ M). Shown are the average values from conductance from a single donor. Experiments were conducted in primary cells from two donors. **f**, Changes in current between potentiator VX770 (ivacaftor) treatment and the CFTR inhibitor treatment in the experiment described in **e** in two primary human cystic fibrosis donor bronchial epithelial cells bearing the  $\Delta F508$ -CFTR mutation. Individual replicate data are shown in the bar graph from  $n=10$  biologically independent samples in the DMSO vehicle-treated group,  $n=13$  biologically independent samples in the lumacaftor-treated group and  $n=18$  biologically independent samples in the NJH-2-057-treated group. Gels shown in **a** and **c** are representative of  $n=3$  biologically independent samples per group. Data in **b** and **d** show individual biological replicate values and average  $\pm$  s.e.m. from  $n=3$  biologically independent samples per group. Statistical significance was calculated with unpaired two-tailed Student's t-tests in **b**, **d** and **f** and is expressed as  $*P<0.05$  compared with the NJH-2-057-treated group in **b** and compared with the NJH-2-057-treated siControl group for CFTR levels in **d**. Raw gels and blots, bar graph and line plot data and exact P values can be found in the source data.

**Figure 5 (a-d):**



**Fig. 5 WEE1 DUBTAC.** **a**, HEP3B cells were treated with DMSO vehicle or bortezomib (1  $\mu$ M) for 24 h. WEE1 and loading control GAPDH levels were assessed by western blotting. **b**, Structures of four WEE1 DUBTACs linking AZD1775 to the OTUB1 recruiter EN523 through four different linkers. **c**, HEP3B cells were treated with DMSO vehicle, the four DUBTACs, bortezomib, EN523 or AZD1775 at 1  $\mu$ M for 24 h. WEE1 and loading control GAPDH levels were assessed by western blotting. **d**, Quantitation of the data shown in **c**. Blots shown in **a** and **b** are representative blots from  $n=3$  biologically independent samples per group. Data in bar graphs show individual biological replicate values and average  $\pm$  s.e.m. from  $n=3$  biologically independent samples per group. Raw blots, bar graphs and exact P values can be found in the source data.

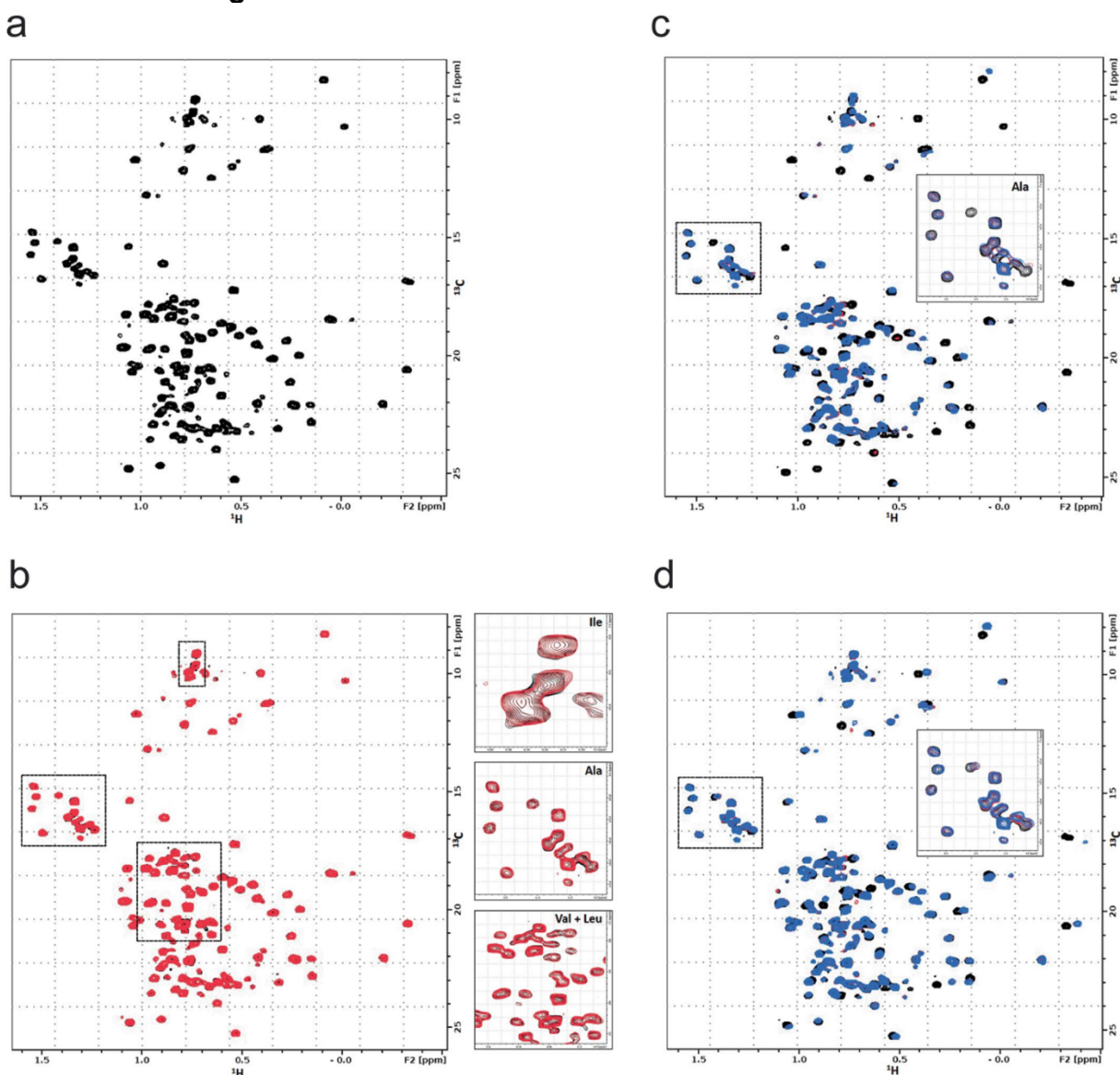
# Extended Data Figure 1 (a-d):



**Extended Data Fig. 1 Primary covalent ligand screen against OTUB1.** (a) Analysis of aggregate chemoproteomic data for DUBs. Top 10 candidate DUBs described in **Fig. 1c** for total aggregate spectral counts of the particular probe-modified cysteine found in our aggregate chemoproteomic data showing OTUB1 C23 appears far more frequently in chemoproteomic datasets compared to the other DUBs. (b) C23 belongs to an intrinsically disordered region within OTUB1 as assessed by PONDR. (c) Covalent ligand screen of cysteine-reactive libraries competed against IA-rhodamine labeling of recombinant OTUB1 to identify binders to OTUB1 by gel-based ABPP. Vehicle DMSO or cysteine-reactive covalent ligands (50  $\mu$ M) were pre-incubated with OTUB1 for 30 min at room temperature prior to IA-rhodamine labeling (500 nM, 30 min room temperature). OTUB1 was then separated by SDS/PAGE and in-gel fluorescence was assessed and quantified. This screen was performed with n = 1 biologically independent samples/group. Any hits showing >50 % loss of IA-rhodamine labeling were subsequently re-tested for reproducibility. EN523 was the only reproducible hit from this primary screen. Raw gels can be found in Source data. (d) Representative spectral file for labeling of OTUB1 C23 with EN523.



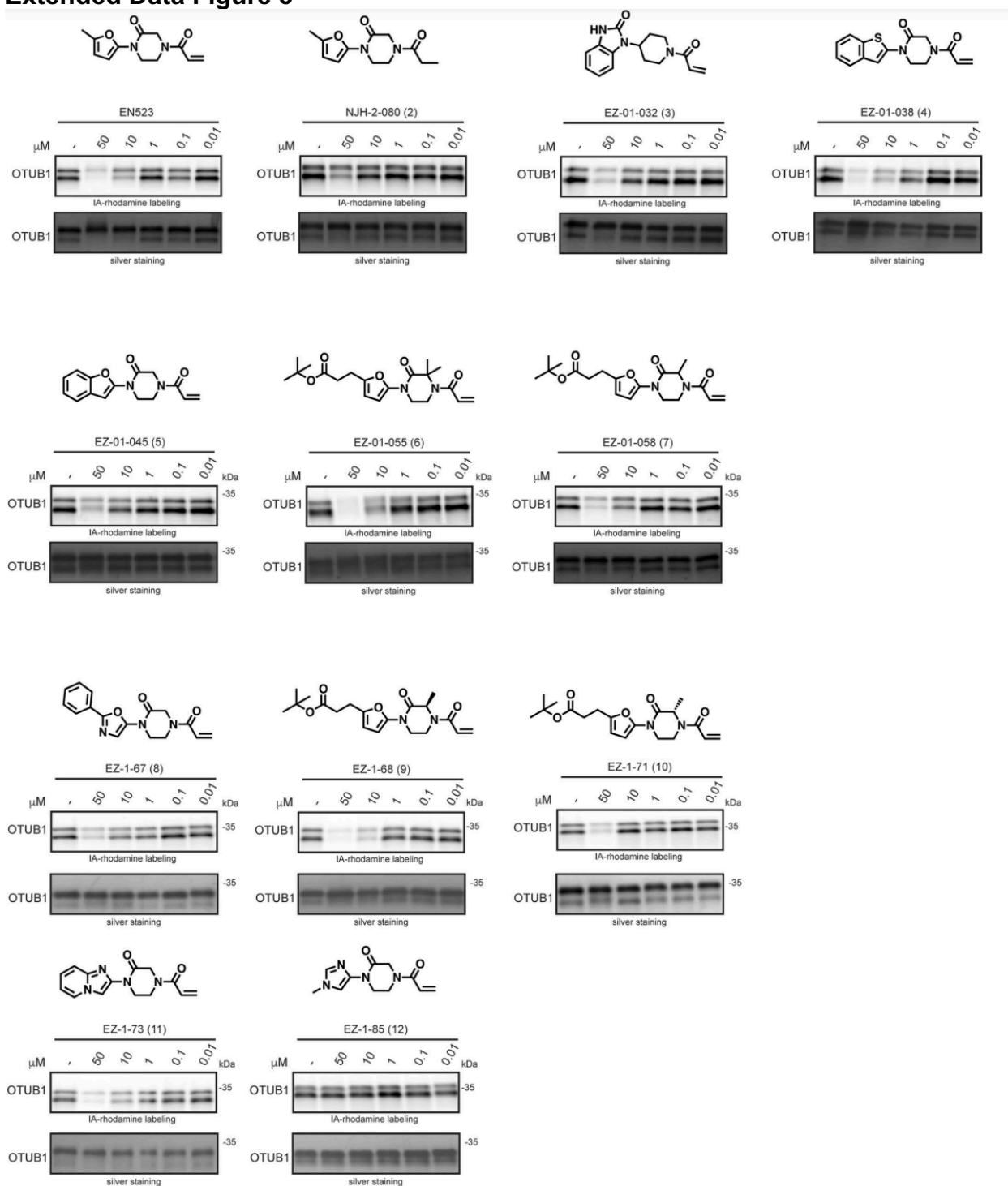
## Extended Data Figure 2



**Extended Data Fig. 2 NMR analysis of OTUB1, EN523, and UBE2D2.** (a)  $^{13}\text{C}$ -HMQC spectrum of OTUB1 labeled on methyl groups of isoleucine, alanine, valine and leucine residues. The presence of peaks with negative proton chemical shifts indicates that the protein is properly folded. (b) Overlay of HMQC spectra of apo-OTUB1 (black) and EN523-bound OTUB1 (red). While both spectra are mostly identical, we identified small but clear chemical shift perturbations of alanine, isoleucine, valine and leucine peaks. Some of these signal changes are shown in the respective blow-up boxes. (c) Overlay of HMQC spectra of apo-OTUB1 (black), UBE2D2 bound OTUB1 (red) and EN523/UBE2D2-bound OTUB1 (blue). The strong chemical shift perturbations (CSPs) are evidence of specific interactions between OTUB1 and the ubiquitylated ubiquitin-conjugating enzyme. The lack of significant differences between spectra recorded in the presence and absence of EN523 prove that the covalent ligand does not interfere with the protein-protein interaction. Differing peak shift pattern are only seen for peaks directly affected by compound binding (see inlay for blow-up of Ala region). (d) Overlay

of HMQC spectra of apo-OTUB1 (black), Ub-UBE2D2 bound OTUB1 (red) and EN523/Ub-UBE2D2-bound OTUB1 (blue). The strong CSPs are evidence of specific interactions between OTUB1 and the ubiquitin-conjugating enzyme. The lack of significant differences between spectra recorded in the presence and absence of EN523 prove that the covalent ligand does not interfere with the protein-protein interaction. Differing peak shift patterns are only seen for peaks directly affected by compound binding (see inlay for blow-up of Ala region).

### Extended Data Figure 3

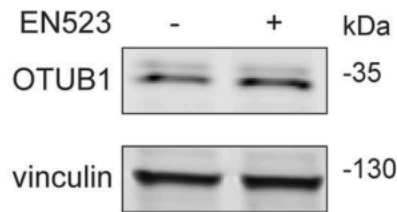


**Extended Data Fig. 3 Structure-activity relationships of EN523 analogs with OTUB1.** Gel-based ABPP analysis EN523 analogs against OTUB1. Vehicle DMSO or EN523 analogs were pre-incubated with recombinant OTUB1 for 30 min at 37 °C prior to IA-rhodamine labeling (100 nM, 30 min room temperature). OTUB1 was then separated by SDS/PAGE and in-gel fluorescence was assessed. Also shown is silver staining

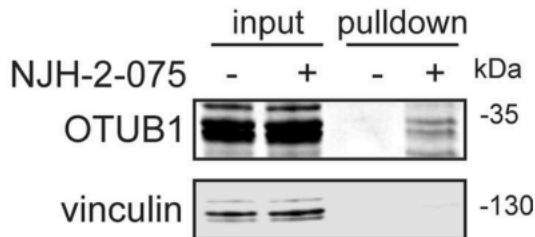
showing protein loading. Shown are representative gels of  $n = 3$  biologically independent samples/group. Raw gels can be found in Source data.

**Extended Data Figure 4 (a-b):**

**a**

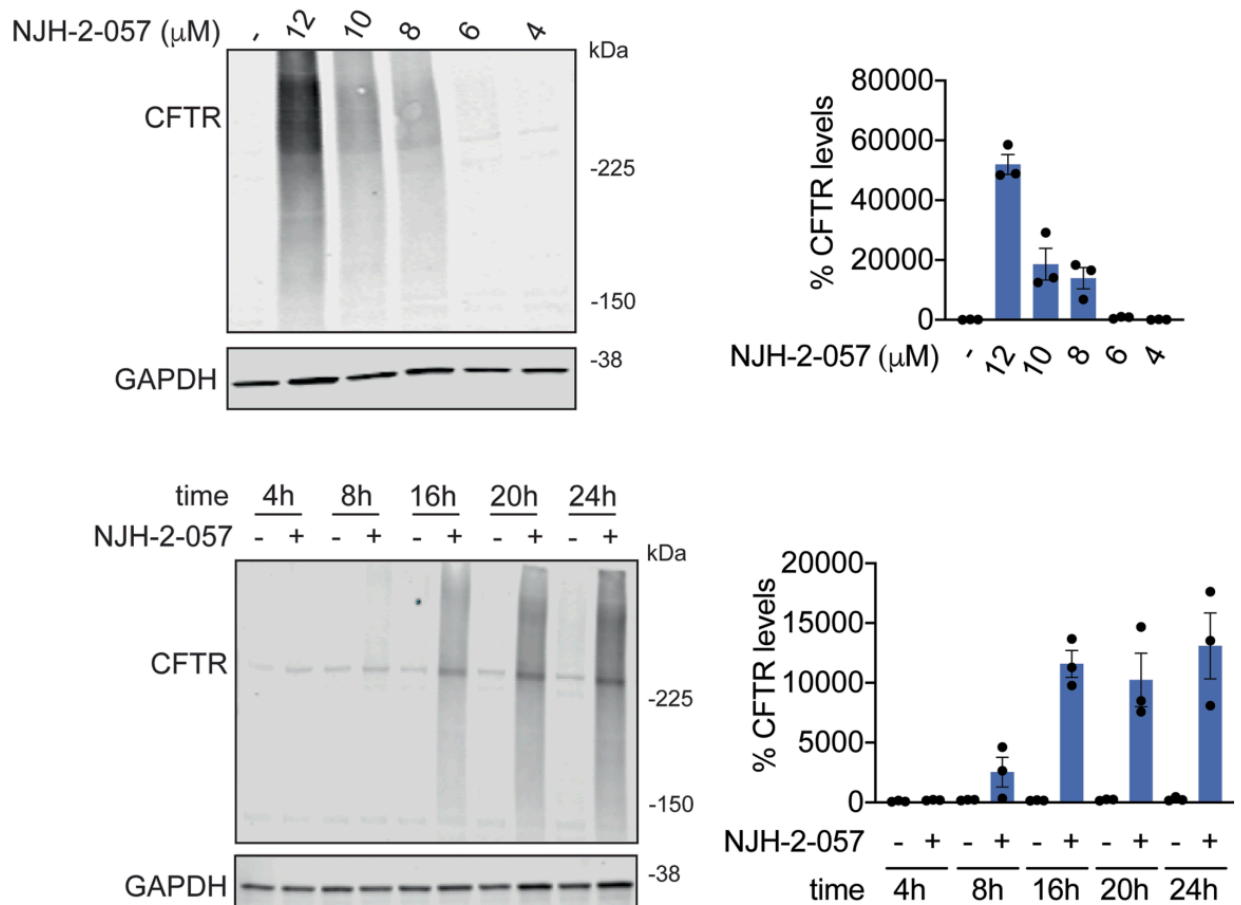


**b**



**Extended Data Fig. 4 EN523 does not alter OTUB1 levels and NJH-2-075 engages OTUB1 in CFBE41o-4.7 cells expressing  $\Delta$ F508-CFTR.** (a) CFBE41o-4.7 cells expressing  $\Delta$ F508-CFTR were treated with vehicle DMSO or EN523 (10  $\mu$ M) for 24 h and OTUB1 and loading control vinculin levels were assessed by Western blotting. (b) NJH-2-075 engagement of OTUB1 in CFBE41o-4.7 cells expressing  $\Delta$ F508-CFTR. Cells were treated with DMSO vehicle or NJH-2-075 (50  $\mu$ M) for 2 h, after which cell lysates were subjected to CuAAC with biotin picolyl azide and NJH-2-075 labeled proteins were subjected to avidin pulldown, eluted, separated by SDS/PAGE, and blotted for OTUB1 and vinculin. Both input lysate and pulldown levels are shown. Blots shown are representative blots from n = 3 biologically independent samples/group. Raw blots can be found in Source data

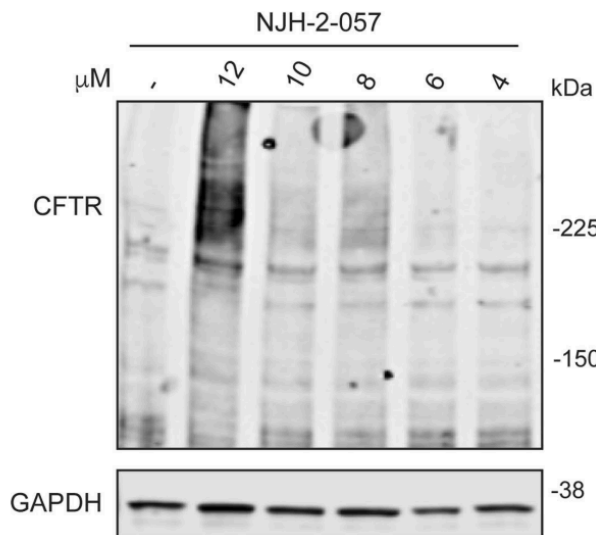
### Extended Data Figure 5



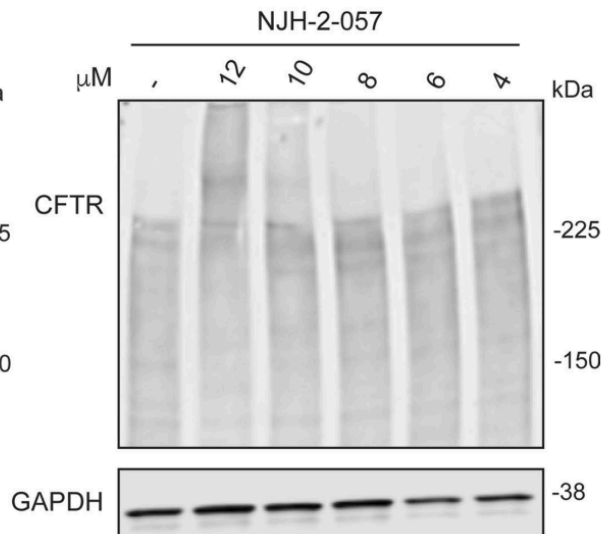
**Extended Data Fig. 5 | Effect of DUBTACs on mutant CFTR levels.** CFBE41o-4.7 cells expressing  $\Delta\text{F508}$ -CFTR were treated with vehicle DMSO or NJH-2-057 and CFTR and loading control GAPDH levels were assessed by Western blotting. For dose-response studies, NJH-2-057 was treated for 24 h. For time-course studies, NJH-2-057 was treated at 10  $\mu\text{M}$ . Dose-response and time-course data gels are representative of  $n = 3$  biologically independent samples/group and are quantified in the bar graphs to the right. Data in bar graphs show individual biological replicate values and average  $\pm$  sem from  $n = 3$  biologically independent samples/group. Raw blots and bar graph data can be found in Source data.

## Extended Data Figure 6

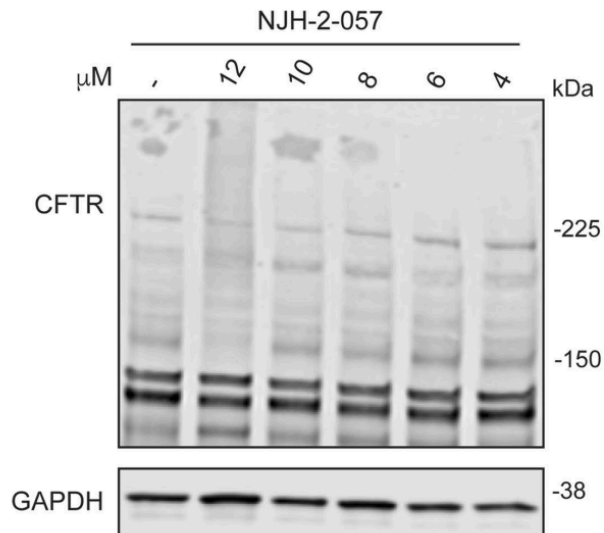
Sigma-Prestige CFTR antibody



Millipore CFTR antibody

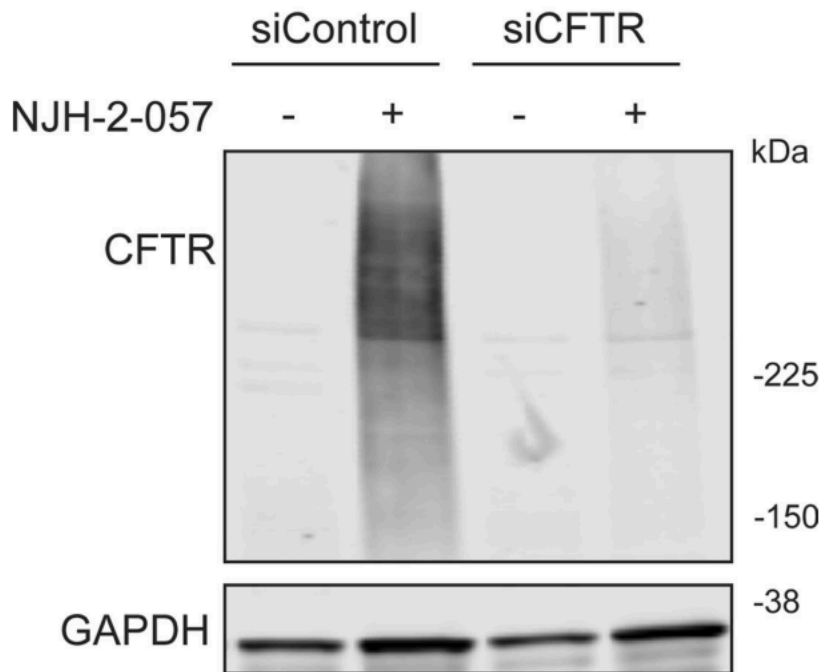


R&D 24-1 CFTR antibody



**Extended Data Fig. 6 Effect of DUBTACs on mutant CFTR levels.** CFBE41o-4.7 cells expressing  $\Delta$ F508-CFTR were treated with vehicle DMSO or NJH-2-057 and CFTR and loading control GAPDH levels were assessed by Western blotting using three different antibodies against CFTR from the ones used for the main figures. NJH-2-057 was treated for 24 h. Gels are representative of  $n = 3$  biologically independent samples/group. Raw blots can be found in Source data.

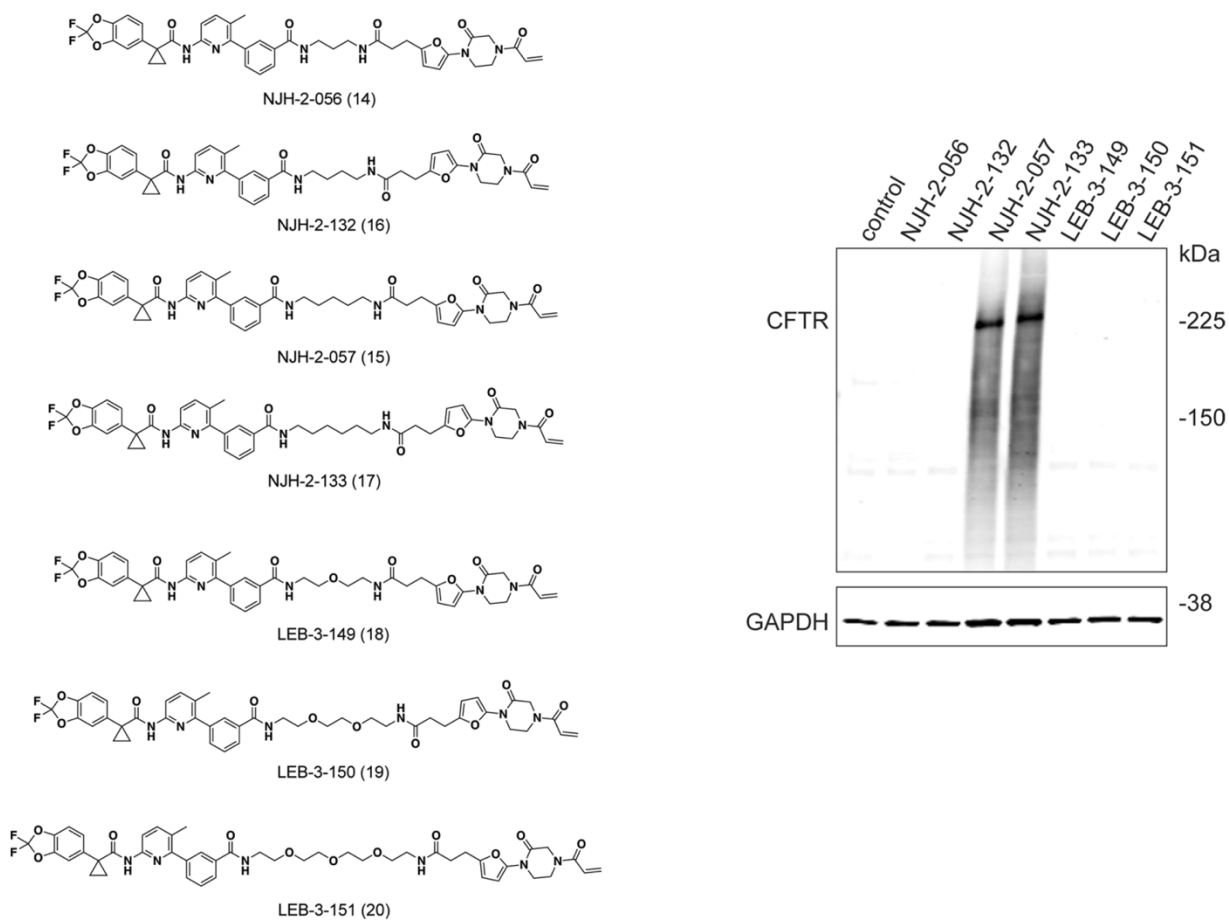
Extended Data Figure 7



**Extended Data Fig. 7 Effect of DUBTACs on mutant CFTR levels in siControl and siCFTR cells.** CFBE41o-4.7 cells expressing  $\Delta F508$ -CFTR were treated with vehicle DMSO or NJH-2-057 (10  $\mu$ M) for 24 h and CFTR and loading control GAPDH levels were assessed by Western blotting. Blot is representative of n = 3 biologically independent samples/group. Raw blots can be found in Source data.



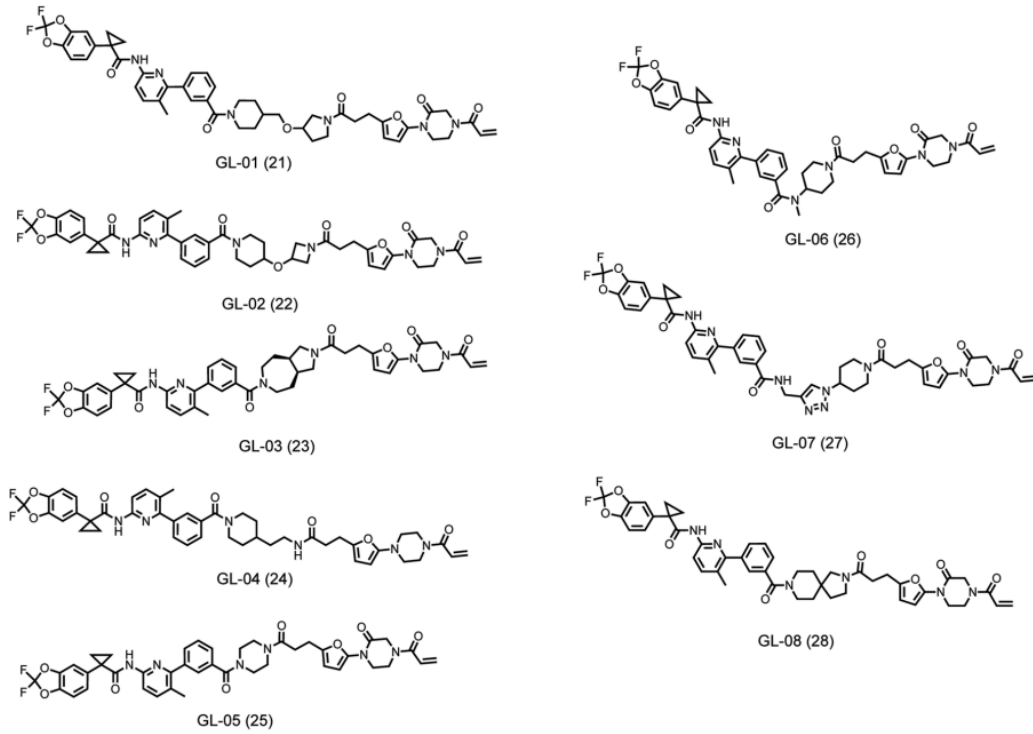
## Extended Data Figure 8



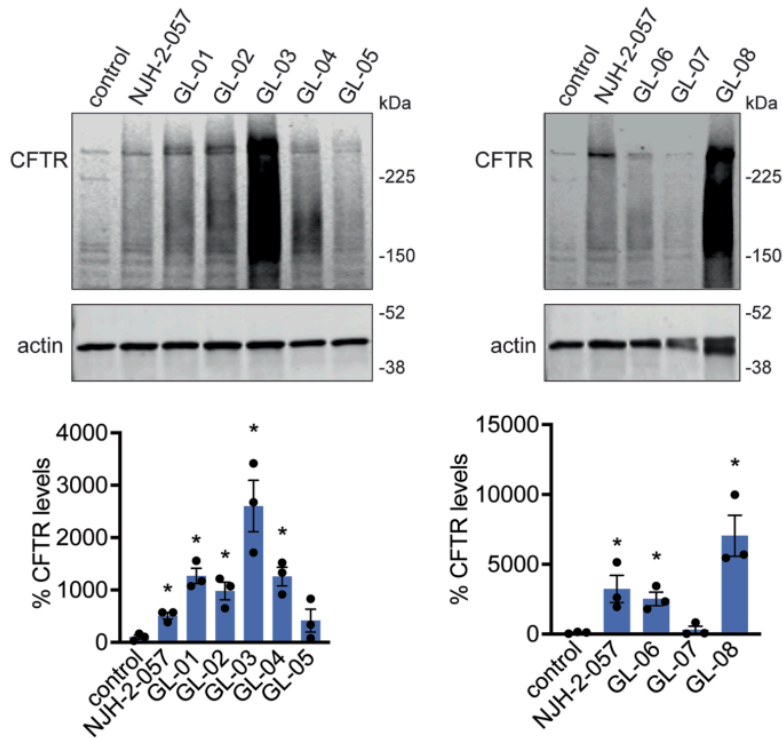
**Extended Data Fig. 8 Linker dependence of CFTR DUBTACs.** CFBE41o-4.7 cells expressing  $\Delta F508$ -CFTR were treated with vehicle DMSO or DUBTACs (10  $\mu$ M) for 24 h and CFTR and loading control GAPDH levels were assessed by Western blotting. Blot is representative of  $n = 3$  biologically independent samples/group. Raw blots can be found in Source data.

## Extended Data Figure 9

a



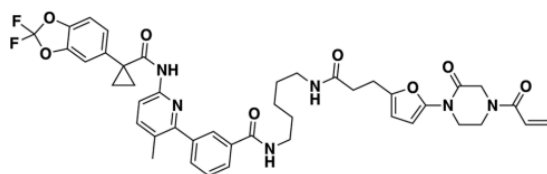
b



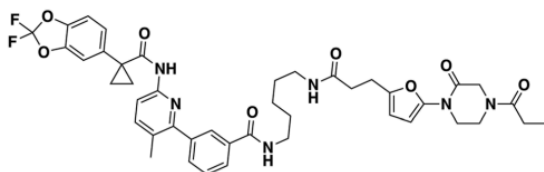
**Extended Data Fig. 9 CFTR DUBTACs with rigid linkers.** CFBE41o-4.7 cells expressing  $\Delta F508$ -CFTR were treated with vehicle DMSO or DUBTACs (10  $\mu$ M) for

24 h and CFTR and loading control actin levels were assessed by Western blotting. Blot is representative of n = 3 biologically independent samples/group. Bar graphs show quantification of CFTR levels shown as individual biological replicate data and average  $\pm$  sem with n = 3 biologically independent samples/group. Statistical significance was calculated with unpaired two-tailed Student's t-tests compared to vehicle-treated controls and is expressed as \*p < 0.05. Raw blots, bar graph, and exact p-values can be found in Source data.

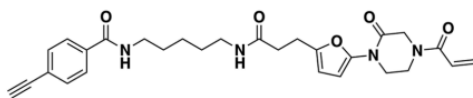
## Extended Data Figure 10



NJH-2-057

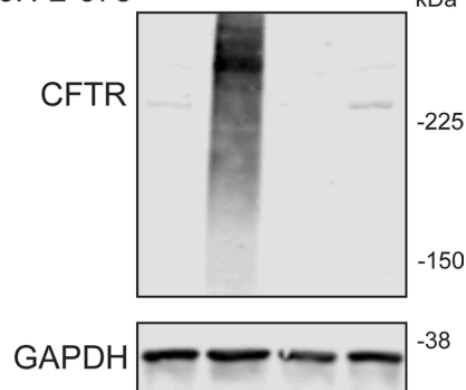


NJH-2-106 (29)



NJH-2-075

NJH-2-057	-	+	-	-	
NJH-2-106	-	-	+	-	
NJH-2-075	-	-	-	+	



**Extended Data Fig. 10 Negative control DUBTACs.** CFBE41o-4.7 cells expressing  $\Delta F508$ -CFTR were treated with vehicle DMSO or compounds (10  $\mu$ M) for 24 h and CFTR and loading control GAPDH levels were assessed by Western blotting. Blot is representative of  $n = 3$  biologically independent samples/ group. Raw blots can be found in Source data.

**Table S1 Chemoproteomic analysis of DUBs**

Available in supplementary downloadable files.

**Table S2 Structures of covalent ligands screened against OTUB1**

Available in supplementary downloadable files.

**Table S3 IsoTOP-ABPP analysis of NJH-2-057**

Available in supplementary downloadable files.

**Table S4 TMT-based quantitative proteomic profiling of NJH-2-057 treatment**

Available in supplementary downloadable files.

**CHAPTER 5**  
**CONCLUDING REMARKS**

## Concluding Remarks

In Chapter 2 (MYC) and in Chapter 4 (DUBTACs), chemoproteomics platforms enabled the discovery of both ligandable binding pockets and corresponding ligands against intrinsically disordered regions on MYC and OTUB1 respectively. In each of these projects, our searchable chemoproteomics database (comprised of >1000 cysteine reactive profiling experiments in complex biological samples) guided us in selecting potential ligandable hotspots on our proteins of interest prior to screening our covalent ligand libraries. Intrinsically disordered proteins do not possess well-defined binding pockets and structurally-enabled design isn't possible when it comes to designing ligands for these types of proteins. Therefore, taking advantage of the reactivity of allosteric cysteine hotspots on intrinsically disordered proteins can enable the discovery of tool covalent ligands without the need for well-studied binding pockets. While the ligands that were developed against MYC and OTUB1 (EN4 and EN523 respectively) were ultimately used for very different purposes, the project flow for ligand discovery was very similar. This chemoproteomics-enabled methodology for discovering ligands against "undruggable" protein targets is applicable to other ligand discovery efforts and can be prioritized in scenarios where structurally-enabled discovery proves difficult.

These projects showcase covalent ligands as allosteric inhibitors (Chapter 2 MYC), catalytic cysteine inhibitors (Chapter 3 Mpro), and as recruiters for targeted protein stabilization (Chapter 4 DUBTACs). As the field of covalent drug discovery continues to expand, covalent tool compounds serve as useful leads for developing clinically relevant molecules. These projects showcase synergy between covalent drug discovery and chemoproteomics platforms and indicate future successes in targeting disease relevant proteins with covalent compounds.

## References

1. Singh, J., Petter, R. C., Baillie, T. A. & Whitty, A. The resurgence of covalent drugs. *Nat Rev Drug Discov* **10**, 307–317 (2011).
2. Ghosh, A. K., Samanta, I., Mondal, A. & Liu, W. R. Covalent Inhibition in Drug Discovery. *ChemMedChem* **14**, 889–906 (2019).
3. Sutanto, F., Konstantinidou, M. & Dömling, A. Covalent inhibitors: a rational approach to drug discovery. *RSC Med. Chem.* **11**, 876–884 (2020).
4. De Cesco, S., Kurian, J., Dufresne, C., Mittermaier, A. K. & Moitessier, N. Covalent inhibitors design and discovery. *European Journal of Medicinal Chemistry* **138**, 96–114 (2017).
5. Zhang, T., Hatcher, J. M., Teng, M., Gray, N. S. & Kostic, M. Recent Advances in Selective and Irreversible Covalent Ligand Development and Validation. *Cell Chemical Biology* **26**, 1486–1500 (2019).
6. Lonsdale, R. & Ward, R. A. Structure-based design of targeted covalent inhibitors. *Chem. Soc. Rev.* **47**, 3816–3830 (2018).
7. Sagonowsky, E. The top 20 drugs by worldwide sales in 2020. *Fierce Pharma* <https://www.fiercepharma.com/special-report/top-20-drugs-by-2020-sales> (2021).
8. Moore, A. R., Rosenberg, S. C., McCormick, F. & Malek, S. RAS-targeted therapies: is the undruggable drugged? *Nat Rev Drug Discov* **19**, 533–552 (2020).
9. Huang, L., Guo, Z., Wang, F. & Fu, L. KRAS mutation: from undruggable to druggable in cancer. *Sig Transduct Target Ther* **6**, 386 (2021).
10. Vandyck, K. & Deval, J. Considerations for the discovery and development of 3-chymotrypsin-like cysteine protease inhibitors targeting SARS-CoV-2 infection. *Current Opinion in Virology* **49**, 36–40 (2021).
11. Lagoutte, R., Patouret, R. & Winssinger, N. Covalent inhibitors: an opportunity for rational target selectivity. *Current Opinion in Chemical Biology* **39**, 54–63 (2017).
12. Lu, W. *et al.* Fragment-based covalent ligand discovery. *RSC Chem. Biol.* **2**, 354–367 (2021).
13. Spradlin, J. N., Zhang, E. & Nomura, D. K. Reimagining Druggability Using Chemoproteomic Platforms. *Acc. Chem. Res.* **54**, 1801–1813 (2021).



14. Roberts, A. M., Ward, C. C. & Nomura, D. K. Activity-based protein profiling for mapping and pharmacologically interrogating proteome-wide ligandable hotspots. *Current Opinion in Biotechnology* **43**, 25–33 (2017).
15. Drewes, G. & Knapp, S. Chemoproteomics and Chemical Probes for Target Discovery. *Trends in Biotechnology* **36**, 1275–1286 (2018).
16. Moellering, R. E. & Cravatt, B. F. How Chemoproteomics Can Enable Drug Discovery and Development. *Chemistry & Biology* **19**, 11–22 (2012).
17. Maurais, A. J. & Weerapana, E. Reactive-cysteine profiling for drug discovery. *Current Opinion in Chemical Biology* **50**, 29–36 (2019).
18. Chandrasekharan, N. & Simmons, D. L. The cyclooxygenases. *Genome Biology* **5**, (2004).
19. Vane, J. R. Inhibition of Prostaglandin Synthesis as a Mechanism of Action for Aspirin-like Drugs. *Nature* **231**, 232–235 (1971).
20. Nicola, G., Tomberg, J., Pratt, R. F., Nicholas, R. A. & Davies, C. Crystal Structures of Covalent Complexes of  $\beta$ -Lactam Antibiotics with *Escherichia coli* Penicillin-Binding Protein 5: Toward an Understanding of Antibiotic Specificity. *Biochemistry* **49**, 8094–8104 (2010).
21. Dijkmans, A. C. *et al.* Fosfomycin: Pharmacological, Clinical and Future Perspectives. *Antibiotics* **6**, 24 (2017).
22. Hendlin, D. *et al.* Phosphonomycin, a New Antibiotic Produced by Strains of *Streptomyces*. *Science* **166**, 122–123 (1969).
23. Kahan, F. M., Kahan, J. S., Cassidy, P. J. & Kropp, H. The Mechanism of Action of Fosfomycin (Phosphonomycin). *Ann NY Acad Sci* **235**, 364–386 (1974).
24. Heck, A. M., Yanovski, J. A. & Calis, K. A. Orlistat, a New Lipase Inhibitor for the Management of Obesity. *Pharmacotherapy* **20**, 270–279 (2000).
25. Olbe, L., Carlsson, E. & Lindberg, P. A proton-pump inhibitor expedition: the case histories of omeprazole and esomeprazole. *Nat Rev Drug Discov* **2**, 132–139 (2003).
26. Savi, P. *et al.* Identification and Biological Activity of the Active Metabolite of Clopidogrel. *Thromb Haemost* **84**, 891–896 (2000).
27. Thomas, D. & Zalcborg, J. 5-FLUOROURACIL: A PHARMACOLOGICAL PARADIGM IN THE USE OF CYTOTOXICS. *Clin Exp Pharmacol Physiol* **25**, 887–895 (1998).

28. Danenberg, P. V., Langenbach, R. J. & Heidelberger, C. Structures of Reversible and Irreversible Complexes of Thymidylate Synthetase and Fluorinated Pyrimidine Nucleotides. *8*.
29. Xu, H., Faber, C., Uchiki, T., Racca, J. & Dealwis, C. Structures of eukaryotic ribonucleotide reductase I define gemcitabine diphosphate binding and subunit assembly. *Proceedings of the National Academy of Sciences* **103**, 4028–4033 (2006).
30. Curran, M. P. & McKeage, K. Bortezomib: A Review of its Use in Patients with Multiple Myeloma. *Drugs* **69**, 859–888 (2009).
31. Lynch, T. J., Okimoto, R. A., Supko, J. G. & Settleman, J. Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non–Small-Cell Lung Cancer to Gefitinib. *The New England Journal of Medicine* **11** (2004).
32. Paez, J. G. *et al.* EGFR Mutations in Lung Cancer: Correlation with Clinical Response to Gefitinib Therapy. *Science* **304**, 1497–1500 (2004).
33. Ou, S.-H. I. Second-generation irreversible epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (TKIs): A better mousetrap? A review of the clinical evidence. *Critical Reviews in Oncology/Hematology* **83**, 407–421 (2012).
34. Yu, H. A. *et al.* Analysis of Tumor Specimens at the Time of Acquired Resistance to EGFR-TKI Therapy in 155 Patients with EGFR -Mutant Lung Cancers. *Clin Cancer Res* **19**, 2240–2247 (2013).
35. Recondo, G., Facchinetti, F., Olaussen, K. A., Besse, B. & Friboulet, L. Making the first move in EGFR-driven or ALK-driven NSCLC: first-generation or next-generation TKI? *Nat Rev Clin Oncol* **15**, 694–708 (2018).
36. Yun, C.-H. *et al.* The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proceedings of the National Academy of Sciences* **105**, 2070–2075 (2008).
37. Soria, J.-C. *et al.* Afatinib versus erlotinib as second-line treatment of patients with advanced squamous cell carcinoma of the lung (LUX-Lung 8): an open-label randomised controlled phase 3 trial. *The Lancet Oncology* **16**, 897–907 (2015).
38. Yu, H. A. & Pao, W. Afatinib—new therapy option for EGFR-mutant lung cancer. *Nat Rev Clin Oncol* **10**, 551–552 (2013).
39. Tsou, H.-R. *et al.* Optimization of 6,7-Disubstituted-4-(arylamino)quinoline-3-carbonitriles as Orally Active, Irreversible Inhibitors of Human Epidermal Growth Factor Receptor-2 Kinase Activity. *J. Med. Chem.* **48**, 1107–1131 (2005).
40. Deeks, E. D. Neratinib: First Global Approval. *Drugs* **77**, 1695–1704 (2017).

41. Shirley, M. Dacomitinib: First Global Approval. *Drugs* **78**, 1947–1953 (2018).
42. Zhou, W. *et al.* Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. *Nature* **462**, 1070–1074 (2009).
43. Cross, D. A. E. *et al.* AZD9291, an Irreversible EGFR TKI, Overcomes T790M-Mediated Resistance to EGFR Inhibitors in Lung Cancer. *Cancer Discovery* **4**, 1046–1061 (2014).
44. Finlay, M. R. V. *et al.* Discovery of a Potent and Selective EGFR Inhibitor (AZD9291) of Both Sensitizing and T790M Resistance Mutations That Spares the Wild Type Form of the Receptor. *J. Med. Chem.* **57**, 8249–8267 (2014).
45. Walter, A. O. *et al.* Discovery of a Mutant-Selective Covalent Inhibitor of EGFR that Overcomes T790M-Mediated Resistance in NSCLC. *Cancer Discovery* **3**, 1404–1415 (2013).
46. Butterworth, S., Cross, D. A. E., Finlay, M. R. V., Ward, R. A. & Waring, M. J. The structure-guided discovery of osimertinib: the first U.S. FDA approved mutant selective inhibitor of EGFR T790M. *Med. Chem. Commun.* **8**, 820–822 (2017).
47. Jänne, P. A. *et al.* AZD9291 in EGFR Inhibitor-Resistant Non-Small-Cell Lung Cancer. *N Engl J Med* **372**, 1689–1699 (2015).
48. Mok, T. S. *et al.* Osimertinib or Platinum-Pemetrexed in EGFR T790M-Positive Lung Cancer. *n engl j med* **12** (2016).
49. Passaro, A., Jänne, P. A., Mok, T. & Peters, S. Overcoming therapy resistance in EGFR-mutant lung cancer. *Nat Cancer* **2**, 377–391 (2021).
50. Leonetti, A. *et al.* Resistance mechanisms to osimertinib in EGFR-mutated non-small cell lung cancer. *Br J Cancer* **121**, 725–737 (2019).
51. Thress, K. S. *et al.* Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M. *Nat Med* **21**, 560–562 (2015).
52. Jia, Y. *et al.* Overcoming EGFR(T790M) and EGFR(C797S) resistance with mutant-selective allosteric inhibitors. *Nature* **534**, 129–132 (2016).
53. To, C. *et al.* Single and Dual Targeting of Mutant EGFR with an Allosteric Inhibitor. *Cancer Discov* **9**, 926–943 (2019).
54. Pal Singh, S., Dammeijer, F. & Hendriks, R. W. Role of Bruton's tyrosine kinase in B cells and malignancies. *Mol Cancer* **17**, 57 (2018).

55. Burger, J. A. & Buggy, J. J. Bruton tyrosine kinase inhibitor ibrutinib (PCI-32765). *Leukemia & Lymphoma* **54**, 2385–2391 (2013).
56. Brooimans, R. A. *et al.* Identification of novel Bruton's tyrosine kinase mutations in 10 unrelated subjects with X linked agammaglobulinaemia. *Journal of Medical Genetics* **34**, 484–488 (1997).
57. Pan, Z. *et al.* Discovery of Selective Irreversible Inhibitors for Bruton's Tyrosine Kinase. *ChemMedChem* **2**, 58–61 (2007).
58. Honigberg, L. A. *et al.* The Bruton tyrosine kinase inhibitor PCI-32765 blocks B-cell activation and is efficacious in models of autoimmune disease and B-cell malignancy. *Proceedings of the National Academy of Sciences* **107**, 13075–13080 (2010).
59. Herman, S. E. M. *et al.* Bruton tyrosine kinase represents a promising therapeutic target for treatment of chronic lymphocytic leukemia and is effectively targeted by PCI-32765. *Blood* **117**, 6287–6296 (2011).
60. Byrd, J. C. *et al.* Targeting BTK with Ibrutinib in Relapsed Chronic Lymphocytic Leukemia. *N Engl J Med* **369**, 32–42 (2013).
61. Byrd, J. C. *et al.* Ibrutinib Treatment for First-Line and Relapsed/Refractory Chronic Lymphocytic Leukemia: Final Analysis of the Pivotal Phase Ib/II PCYC-1102 Study. *Clin Cancer Res* **26**, 3918–3927 (2020).
62. Wang, M. L. *et al.* Targeting BTK with Ibrutinib in Relapsed or Refractory Mantle-Cell Lymphoma. *N Engl J Med* **369**, 507–516 (2013).
63. Treon, S. P. *et al.* Ibrutinib in Previously Treated Waldenström's Macroglobulinemia. *N Engl J Med* **372**, 1430–1440 (2015).
64. Miklos, D. *et al.* Ibrutinib for chronic graft-versus-host disease after failure of prior therapy. *Blood* **130**, 2243–2250 (2017).
65. Davids, M. S. & Brown, J. R. Ibrutinib: a first in class covalent inhibitor of Bruton's tyrosine kinase. *Future Oncology* **10**, 957–967 (2014).
66. Advani, R. H. *et al.* Bruton Tyrosine Kinase Inhibitor Ibrutinib (PCI-32765) Has Significant Activity in Patients With Relapsed/Refractory B-Cell Malignancies. *JCO* **31**, 88–94 (2013).
67. Imbruvica. *FiercePharma* <https://www.fiercepharma.com/special-report/top-20-drugs-by-2020-sales-imbruvica>.

68. Barf, T. *et al.* Acabrutinib (ACP-196): A Covalent Bruton Tyrosine Kinase Inhibitor with a Differentiated Selectivity and In Vivo Potency Profile. *J Pharmacol Exp Ther* **363**, 240–252 (2017).
69. Flanagan, M. E. *et al.* Chemical and Computational Methods for the Characterization of Covalent Reactive Groups for the Prospective Design of Irreversible Inhibitors. *J. Med. Chem.* **57**, 10072–10079 (2014).
70. Ward, R. A. *et al.* Structure- and Reactivity-Based Development of Covalent Inhibitors of the Activating and Gatekeeper Mutant Forms of the Epidermal Growth Factor Receptor (EGFR). *J. Med. Chem.* **56**, 7025–7048 (2013).
71. Byrd, J. C. *et al.* Acabrutinib (ACP-196) in Relapsed Chronic Lymphocytic Leukemia. *N Engl J Med* **374**, 323–332 (2016).
72. Byrd, J. C. *et al.* Acabrutinib Versus Ibrutinib in Previously Treated Chronic Lymphocytic Leukemia: Results of the First Randomized Phase III Trial. *JCO* **39**, 3441–3452 (2021).
73. Narita, Y. *et al.* Phase I/II study of tirabrutinib, a second-generation Bruton's tyrosine kinase inhibitor, in relapsed/refractory primary central nervous system lymphoma. *Neuro-Oncology* **23**, 122–133 (2021).
74. Guo, Y. *et al.* Discovery of Zanubrutinib (BGB-3111), a Novel, Potent, and Selective Covalent Inhibitor of Bruton's Tyrosine Kinase. *J. Med. Chem.* **62**, 7923–7940 (2019).
75. Caldwell, R. D. *et al.* Discovery of Evobrutinib: An Oral, Potent, and Highly Selective, Covalent Bruton's Tyrosine Kinase (BTK) Inhibitor for the Treatment of Immunological Diseases. *J. Med. Chem.* **62**, 7643–7655 (2019).
76. Serafimova, I. M. *et al.* Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles. *Nat Chem Biol* **8**, 471–476 (2012).
77. Miller, R. M., Paavilainen, V. O., Krishnan, S., Serafimova, I. M. & Taunton, J. Electrophilic Fragment-Based Design of Reversible Covalent Kinase Inhibitors. *J. Am. Chem. Soc.* **135**, 5298–5301 (2013).
78. Bradshaw, J. M. *et al.* Prolonged and tunable residence time using reversible covalent kinase inhibitors. *Nat Chem Biol* **11**, 525–531 (2015).
79. Sanofi: Press Releases, Thursday, September 9, 2021.  
<https://www.sanofi.com/en/media-room/press-releases/2021/2021-09-09-07-00-00-2293920> <https://www.sanofi.com/media-room/press-releases/2021/2021-09-09-07-00-00-2293920>.

80. Zhang, D., Gong, H. & Meng, F. Recent Advances in BTK Inhibitors for the Treatment of Inflammatory and Autoimmune Diseases. *Molecules* **26**, 4907 (2021).
81. Zhao, Z. & Bourne, P. E. Progress with covalent small-molecule kinase inhibitors. *Drug Discovery Today* **23**, 727–735 (2018).
82. Liu, Q. *et al.* Developing Irreversible Inhibitors of the Protein Kinase Cysteine. *Chemistry & Biology* **20**, 146–159 (2013).
83. Thorarensen, A. *et al.* Design of a Janus Kinase 3 (JAK3) Specific Inhibitor 1-((2*S*,5*R*)-5-((7*H*-Pyrrolo[2,3-*d*]pyrimidin-4-yl)amino)-2-methylpiperidin-1-yl)prop-2-en-1-one (PF-06651600) Allowing for the Interrogation of JAK3 Signaling in Humans. *J. Med. Chem.* **60**, 1971–1993 (2017).
84. Goedken, E. R. *et al.* Tricyclic Covalent Inhibitors Selectively Target Jak3 through an Active Site Thiol. *Journal of Biological Chemistry* **290**, 4573–4589 (2015).
85. Tan, L. *et al.* Development of Selective Covalent Janus Kinase 3 Inhibitors. *J. Med. Chem.* **58**, 6589–6606 (2015).
86. Kempson, J. *et al.* Discovery of highly potent, selective, covalent inhibitors of JAK3. *Bioorganic & Medicinal Chemistry Letters* **27**, 4622–4625 (2017).
87. Telliez, J.-B. *et al.* Discovery of a JAK3-Selective Inhibitor: Functional Differentiation of JAK3-Selective Inhibition over pan-JAK or JAK1-Selective Inhibition. *ACS Chem. Biol.* **11**, 3442–3451 (2016).
88. Robinson, M. F. *et al.* Efficacy and Safety of PF-06651600 (Ritlecitinib), a Novel JAK3/TEC Inhibitor, in Patients With Moderate-to-Severe Rheumatoid Arthritis and an Inadequate Response to Methotrexate. *Arthritis Rheumatol* **72**, 1621–1631 (2020).
89. Kim, R. D. *et al.* First-in-Human Phase I Study of Fisogatinib (BLU-554) Validates Aberrant FGF19 Signaling as a Driver Event in Hepatocellular Carcinoma. *Cancer Discov* **9**, 1696–1707 (2019).
90. Tan, L. *et al.* Development of covalent inhibitors that can overcome resistance to first-generation FGFR kinase inhibitors. *Proc Natl Acad Sci USA* **111**, E4869–E4877 (2014).
91. Weiss, A. *et al.* FGF401, A First-In-Class Highly Selective and Potent FGFR4 Inhibitor for the Treatment of FGF19-Driven Hepatocellular Cancer. *Mol Cancer Ther* **18**, 2194–2206 (2019).
92. Kwiatkowski, N. *et al.* Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. *Nature* **511**, 616–620 (2014).

93. Zhang, T. *et al.* Covalent targeting of remote cysteine residues to develop CDK12 and CDK13 inhibitors. *Nat Chem Biol* **12**, 876–884 (2016).
94. Goebel, L., Müller, M. P., Goody, R. S. & Rauh, D. KRasG12C inhibitors in clinical trials: a short historical perspective. *RSC Med. Chem.* **11**, 760–770 (2020).
95. Janes, M. R. *et al.* Targeting KRAS Mutant Cancers with a Covalent G12C-Specific Inhibitor. *Cell* **172**, 578-589.e17 (2018).
96. Ostrem, J. M., Peters, U., Sos, M. L., Wells, J. A. & Shokat, K. M. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* **503**, 548–551 (2013).
97. Lito, P., Solomon, M., Li, L.-S., Hansen, R. & Rosen, N. Allele-specific inhibitors inactivate mutant KRAS G12C by a trapping mechanism. *Science* **351**, 604–608 (2016).
98. Patricelli, M. P. *et al.* Selective Inhibition of Oncogenic KRAS Output with Small Molecules Targeting the Inactive State. *Cancer Discov* **6**, 316–329 (2016).
99. Canon, J. *et al.* The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity. *Nature* **575**, 217–223 (2019).
100. Lanman, B. A. *et al.* Discovery of a Covalent Inhibitor of KRAS<sup>G12C</sup> (AMG 510) for the Treatment of Solid Tumors. *J. Med. Chem.* **63**, 52–65 (2020).
101. Skoulidis, F. *et al.* Sotorasib for Lung Cancers with KRAS p.G12C Mutation. *N Engl J Med* **384**, 2371–2381 (2021).
102. Fell, J. B. *et al.* Identification of the Clinical Development Candidate **MRTX849**, a Covalent KRAS<sup>G12C</sup> Inhibitor for the Treatment of Cancer. *J. Med. Chem.* **63**, 6679–6693 (2020).
103. Hallin, J. *et al.* The KRAS<sup>G12C</sup> Inhibitor MRTX849 Provides Insight toward Therapeutic Susceptibility of KRAS-Mutant Cancers in Mouse Models and Patients. *Cancer Discov* **10**, 54–71 (2020).
104. Mirati Therapeutics. Mirati Therapeutics' Adagrasib Receives Breakthrough Therapy Designation from U.S. Food and Drug Administration for Patients with Advanced Non-Small Cell Lung Cancer Harboring the KRAS G12C Mutation. <https://www.prnewswire.com/news-releases/mirati-therapeutics-adagrasib-receives-breakthrough-therapy-designation-from-us-food-and-drug-administration-for-patients-with-advanced-non-small-cell-lung-cancer-harboring-the-kras-g12c-mutation-301319824.html>.
105. COVID-19 Data Explorer. *Our World in Data*  
<https://ourworldindata.org/coronavirus-data-explorer>.

106. Commissioner, O. of the. Coronavirus (COVID-19) Update: FDA Authorizes First Oral Antiviral for Treatment of COVID-19. *FDA* <https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-first-oral-antiviral-treatment-covid-19> (2021).
107. Qiao, J. *et al.* SARS-CoV-2 M<sup>pro</sup> inhibitors with antiviral activity in a transgenic mouse model. *Science* **371**, 1374–1378 (2021).
108. Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R. & Hilgenfeld, R. Coronavirus Main Proteinase (3CLpro) Structure: Basis for Design of Anti-SARS Drugs. **300**, 6 (2003).
109. Hoffman, R. L. *et al.* Discovery of Ketone-Based Covalent Inhibitors of Coronavirus 3CL Proteases for the Potential Therapeutic Treatment of COVID-19. *J. Med. Chem.* **63**, 12725–12747 (2020).
110. Owen, D. R. *et al.* An oral SARS-CoV-2 M<sup>pro</sup> inhibitor clinical candidate for the treatment of COVID-19. *Science* eabl4784 (2021) doi:10.1126/science.abl4784.
111. Bandyopadhyay, A. & Gao, J. Targeting biomolecules with reversible covalent chemistry. *Current Opinion in Chemical Biology* **34**, 110–116 (2016).
112. Hammond, J. *et al.* Oral Nirmatrelvir for High-Risk, Nonhospitalized Adults with Covid-19. *N Engl J Med* NEJMoa2118542 (2022) doi:10.1056/NEJMoa2118542.
113. Mahase, E. Covid-19: Pfizer's paxlovid is 89% effective in patients at risk of serious illness, company reports. *BMJ* n2713 (2021) doi:10.1136/bmj.n2713.
114. de Leuw, P. & Stephan, C. Protease inhibitor therapy for hepatitis C virus-infection. *Expert Opinion on Pharmacotherapy* **19**, 577–587 (2018).
115. Hézode, C. *et al.* Telaprevir and Peginterferon with or without Ribavirin for Chronic HCV Infection. *The New England Journal of Medicine* **12** (2009).
116. Venkatraman, S. *et al.* Discovery of (1*R*,5*S*)-*N*-[3-Amino-1-(cyclobutylmethyl)-2,3-dioxopropyl]-3-[2(*S*)-[[[(1,1-dimethylethyl)amino]carbonyl]amino]-3,3-dimethyl-1-oxobutyl]-6,6-dimethyl-3-azabicyclo[3.1.0]hexan-2(*S*)-carboxamide (SCH 503034), a Selective, Potent, Orally Bioavailable Hepatitis C Virus NS3 Protease Inhibitor: A Potential Therapeutic Agent for the Treatment of Hepatitis C Infection. *J. Med. Chem.* **49**, 6074–6086 (2006).
117. Venkatraman, S. Discovery of boceprevir, a direct-acting NS3/4A protease inhibitor for treatment of chronic hepatitis C infections. *Trends in Pharmacological Sciences* **33**, 289–294 (2012).



118. Lin, C., Kwong, A. & Perni, R. Discovery and Development of VX-950, a Novel, Covalent, and Reversible Inhibitor of Hepatitis C Virus NS3.4A Serine Protease. *IDDT* **6**, 3–16 (2006).
119. Kwong, A. D., Kauffman, R. S., Hurter, P. & Mueller, P. Discovery and development of telaprevir: an NS3-4A protease inhibitor for treating genotype 1 chronic hepatitis C virus. *Nat Biotechnol* **29**, 993–1003 (2011).
120. Arasappan, A. *et al.* Discovery of Narlaprevir (SCH 900518): A Potent, Second Generation HCV NS3 Serine Protease Inhibitor. *ACS Med. Chem. Lett.* **1**, 64–69 (2010).
121. McCauley, J. A. & Rudd, M. T. Hepatitis C virus NS3/4a protease inhibitors. *Current Opinion in Pharmacology* **30**, 84–92 (2016).
122. Kwo, P. Y. *et al.* Efficacy of boceprevir, an NS3 protease inhibitor, in combination with peginterferon alfa-2b and ribavirin in treatment-naive patients with genotype 1 hepatitis C infection (SPRINT-1): an open-label, randomised, multicentre phase 2 trial. *The Lancet* **376**, 705–716 (2010).
123. Sherman, K. E. *et al.* Response-Guided Telaprevir Combination Treatment for Hepatitis C Virus Infection. *N Engl J Med* **365**, 1014–1024 (2011).
124. Ramsay, I. D., Lestner, J. M. & Barker, C. I. S. Antiviral drugs. in *Side Effects of Drugs Annual* vol. 35 503–550 (Elsevier, 2014).
125. Merck & Co., Inc. Merck Voluntarily Discontinuing VICTRELIS. (2015).
126. Lawitz, E. *et al.* Sofosbuvir and ledipasvir fixed-dose combination with and without ribavirin in treatment-naive and previously treated patients with genotype 1 hepatitis C virus infection (LONESTAR): an open-label, randomised, phase 2 trial. *The Lancet* **383**, 515–523 (2014).
127. Kim, K. B. & Crews, C. M. From epoxomicin to carfilzomib: chemistry, biology, and medical outcomes. *Nat. Prod. Rep.* **30**, 600 (2013).
128. Groll, M., Kim, K. B., Kairies, N., Huber, R. & Crews, C. M. Crystal Structure of Epoxomicin:20S Proteasome Reveals a Molecular Basis for Selectivity of  $\alpha'$ , $\beta'$ -Epoxyketone Proteasome Inhibitors. *J. Am. Chem. Soc.* **122**, 1237–1238 (2000).
129. Strelow, J. M. A Perspective on the Kinetics of Covalent and Irreversible Inhibition. *SLAS DISCOVERY: Advancing the Science of Drug Discovery* **22**, 3–20 (2017).
130. McWhirter, C. Kinetic mechanisms of covalent inhibition. in *Annual Reports in Medicinal Chemistry* vol. 56 1–31 (Elsevier, 2021).

131. Schnute, M. E. *et al.* Aminopyrazole Carboxamide Bruton's Tyrosine Kinase Inhibitors. Irreversible to Reversible Covalent Reactive Group Tuning. *ACS Med. Chem. Lett.* **10**, 80–85 (2019).
132. Maurer, T. S., Tabrizi-Fard, M. A. & Fung, H. Impact of mechanism-based enzyme inactivation on inhibitor potency: Implications for rational drug discovery. *J. Pharm. Sci.* **89**, 1404–1414 (2000).
133. Krippendorff, B.-F., Neuhaus, R., Lienau, P., Reichel, A. & Huisinga, W. Mechanism-Based Inhibition: Deriving  $K_i$  and  $k_{inact}$  Directly from Time-Dependent  $IC_{50}$  Values. *J. Biomol. Screen* **14**, 913–923 (2009).
134. Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-Based Protein Profiling: From Enzyme Chemistry to Proteomic Chemistry. *Annu. Rev. Biochem.* **77**, 383–414 (2008).
135. Liu, Y., Patricelli, M. P. & Cravatt, B. F. Activity-based protein profiling: The serine hydrolases. *Proceedings of the National Academy of Sciences* **96**, 14694–14699 (1999).
136. Weerapana, E., Speers, A. E. & Cravatt, B. F. Tandem orthogonal proteolysis-activity-based protein profiling (TOP-ABPP)—a general method for mapping sites of probe modification in proteomes. *Nat. Protoc.* **2**, 1414–1425 (2007).
137. Weerapana, E. *et al.* Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **468**, 790–795 (2010).
138. Backus, K. M. *et al.* Proteome-wide covalent ligand discovery in native biological systems. *Nature* **534**, 570–574 (2016).
139. Bar-Peled, L. *et al.* Chemical Proteomics Identifies Druggable Vulnerabilities in a Genetically Defined Cancer. *Cell* **171**, 696–709.e23 (2017).
140. Vinogradova, E. V. *et al.* An Activity-Guided Map of Electrophile-Cysteine Interactions in Primary Human T Cells. *Cell* **182**, 1009–1026.e29 (2020).
141. Blewett, M. M. *et al.* Chemical proteomic map of dimethyl fumarate-sensitive cysteines in primary human T cells. *Sci. Signal.* **9**, (2016).
142. Zaro, B. W. *et al.* Dimethyl Fumarate Disrupts Human Innate Immune Signaling by Targeting the IRAK4–MyD88 Complex. *J. I.* **202**, 2737–2746 (2019).
143. Chung, C. Y.-S. *et al.* Covalent targeting of the vacuolar H<sup>+</sup>-ATPase activates autophagy via mTORC1 inhibition. *Nat. Chem. Biol.* **15**, 776–785 (2019).

144. Grüner, B. M. *et al.* An in vivo multiplexed small-molecule screening platform. *Nat Methods* **13**, 883–889 (2016).
145. Boike, L. *et al.* Discovery of a Functional Covalent Ligand Targeting an Intrinsically Disordered Cysteine within MYC. *Cell Chemical Biology* S2451945620303421 (2020) doi:10.1016/j.chembiol.2020.09.001.
146. Yan, T. *et al.* SP3-FAIMS Chemoproteomics for High-Coverage Profiling of the Human Cysteinome. *ChemBioChem* **22**, 1841–1851 (2021).
147. Kuljanin, M. *et al.* Reimagining high-throughput profiling of reactive cysteines for cell-based screening of large electrophile libraries. *Nat Biotechnol* **39**, 630–641 (2021).
148. Spradlin, J. N. *et al.* Harnessing the anti-cancer natural product nimbolide for targeted protein degradation. *Nat Chem Biol* **15**, 747–755 (2019).
149. Luo, M. *et al.* Chemoproteomics-enabled discovery of covalent RNF114-based degraders that mimic natural product function. *Cell Chemical Biology* **28**, 559-566.e15 (2021).
150. Ward, C. C. *et al.* Covalent Ligand Screening Uncovers a RNF4 E3 Ligase Recruiter for Targeted Protein Degradation Applications. *ACS Chem. Biol.* **14**, 2430–2440 (2019).
151. Zhang, X., Crowley, V. M., Wucherpfennig, T. G., Dix, M. M. & Cravatt, B. F. Electrophilic PROTACs that degrade nuclear proteins by engaging DCAF16. *Nat Chem Biol* **15**, 737–746 (2019).
152. Zhang, X. *et al.* DCAF11 Supports Targeted Protein Degradation by Electrophilic Proteolysis-Targeting Chimeras. *J. Am. Chem. Soc.* **143**, 5141–5149 (2021).
153. Tong, B. *et al.* Bardoxolone conjugation enables targeted protein degradation of BRD4. *Sci Rep* **10**, 15543 (2020).
154. Henning, N. J. *et al.* Discovery of a Covalent FEM1B Recruiter for Targeted Protein Degradation Applications. *J. Am. Chem. Soc.* jacs.1c03980 (2022) doi:10.1021/jacs.1c03980.
155. Henning, N. J. *et al.* Deubiquitinase-targeting chimeras for targeted protein stabilization. *Nat Chem Biol* (2022) doi:10.1038/s41589-022-00971-2.
156. Belcher, B. P., Ward, C. C. & Nomura, D. K. Ligandability of E3 Ligases for Targeted Protein Degradation Applications. *Biochemistry* acs.biochem.1c00464 (2021) doi:10.1021/acs.biochem.1c00464.

157. Erlanson, D. A. *et al.* Site-directed ligand discovery. *Proceedings of the National Academy of Sciences* **97**, 9367–9372 (2000).
158. Erlanson, D. A., Wells, J. A. & Braisted, A. C. Tethering: Fragment-Based Drug Discovery. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 199–223 (2004).
159. Arkin, M. R. *et al.* Binding of small molecules to an adaptive protein-protein interface. *Proceedings of the National Academy of Sciences* **100**, 1603–1608 (2003).
160. Keedy, D. A. *et al.* An expanded allosteric network in PTP1B by multitemperature crystallography, fragment screening, and covalent tethering. *eLife* **7**, e36307 (2018).
161. Wolter, M. *et al.* Fragment-Based Stabilizers of Protein–Protein Interactions through Imine-Based Tethering. *Angew. Chem. Int. Ed.* **59**, 21520–21524 (2020).
162. Kathman, S. G., Xu, Z. & Statsyuk, A. V. A Fragment-Based Method to Discover Irreversible Covalent Inhibitors of Cysteine Proteases. *J. Med. Chem.* **57**, 4969–4974 (2014).
163. Kathman, S. G. *et al.* A Small Molecule That Switches a Ubiquitin Ligase From a Processive to a Distributive Enzymatic Mechanism. *J. Am. Chem. Soc.* **137**, 12442–12445 (2015).
164. Johansson, H. *et al.* Fragment-Based Covalent Ligand Screening Enables Rapid Discovery of Inhibitors for the RBR E3 Ubiquitin Ligase HOIP. *J. Am. Chem. Soc.* **141**, 2703–2712 (2019).
165. Resnick, E. *et al.* Rapid Covalent-Probe Discovery by Electrophile-Fragment Screening. *J. Am. Chem. Soc.* **141**, 8951–8968 (2019).
166. Dubiella, C. *et al.* Sulfopin is a covalent inhibitor of Pin1 that blocks Myc-driven tumors in vivo. *Nat Chem Biol* **17**, 954–963 (2021).
167. Shin, Y. *et al.* Discovery of *N*-(1-Acryloylazetidyl-3-yl)-2-(1*H*-indol-1-yl)acetamides as Covalent Inhibitors of KRAS<sup>G12C</sup>. *ACS Med. Chem. Lett.* **10**, 1302–1308 (2019).
168. Debaene, F., Da Silva, J. A., Pianowski, Z., Duran, F. J. & Winssinger, N. Expanding the scope of PNA-encoded libraries: divergent synthesis of libraries targeting cysteine, serine and metallo-proteases as well as tyrosine phosphatases. *Tetrahedron* **63**, 6577–6586 (2007).
169. Zimmermann, G. *et al.* A Specific and Covalent JNK-1 Ligand Selected from an Encoded Self-Assembling Chemical Library. *Chem. Eur. J.* **23**, 8152–8155 (2017).

170. Zambaldo, C., Daguer, J.-P., Saarbach, J., Barluenga, S. & Winssinger, N. Screening for covalent inhibitors using DNA-display of small molecule libraries functionalized with cysteine reactive moieties. *Med. Chem. Commun.* **7**, 1340–1351 (2016).
171. Chan, A. I., McGregor, L. M., Jain, T. & Liu, D. R. Discovery of a Covalent Kinase Inhibitor from a DNA-Encoded Small-Molecule Library × Protein Library Selection. *J. Am. Chem. Soc.* **139**, 10192–10195 (2017).
172. Guilinger, J. P. *et al.* Novel irreversible covalent BTK inhibitors discovered using DNA-encoded chemistry. *Bioorganic & Medicinal Chemistry* **42**, 116223 (2021).
173. London, N. *et al.* Covalent docking of large libraries for the discovery of chemical probes. *Nat Chem Biol* **10**, 1066–1072 (2014).
174. Shraga, A. *et al.* Covalent Docking Identifies a Potent and Selective MKK7 Inhibitor. *Cell Chemical Biology* **26**, 98-108.e5 (2019).
175. Nnadi, C. I. *et al.* Novel K-Ras G12C Switch-II Covalent Binders Destabilize Ras and Accelerate Nucleotide Exchange. *J. Chem. Inf. Model.* **58**, 464–471 (2018).
176. Hacker, S. M. *et al.* Global profiling of lysine reactivity and ligandability in the human proteome. *Nature Chem* **9**, 1181–1190 (2017).
177. Ward, C. C., Kleinman, J. I. & Nomura, D. K. NHS-Esters As Versatile Reactivity-Based Probes for Mapping Proteome-Wide Ligandable Hotspots. *ACS Chem. Biol.* **12**, 1478–1483 (2017).
178. Abbasov, M. E. *et al.* A proteome-wide atlas of lysine-reactive chemistry. *Nat. Chem.* **13**, 1081–1092 (2021).
179. Tamura, T. *et al.* Rapid labelling and covalent inhibition of intracellular native proteins using ligand-directed N-acyl-N-alkyl sulfonamide. *Nat Commun* **9**, 1870 (2018).
180. Vichinsky, E. *et al.* A Phase 3 Randomized Trial of Voxelotor in Sickle Cell Disease. *N Engl J Med* **381**, 509–519 (2019).
181. Metcalf, B. *et al.* Discovery of GBT440, an Orally Bioavailable R-State Stabilizer of Sickle Cell Hemoglobin. *ACS Med. Chem. Lett.* **8**, 321–326 (2017).
182. Garner, M. H., Bogardt, R. A. & Gurd, F. R. Determination of the pK values for the alpha-amino groups of human hemoglobin. *Journal of Biological Chemistry* **250**, 4398–4404 (1975).

183. Oksenberg, D. *et al.* GBT 440 increases haemoglobin oxygen affinity, reduces sickling and prolongs RBC half-life in a murine model of sickle cell disease. *Br J Haematol* **175**, 141–153 (2016).
184. Blair, H. A. Voxelotor: First Approval. *Drugs* **80**, 209–215 (2020).
185. Grimster, N. P. *et al.* Aromatic Sulfonyl Fluorides Covalently Kinetically Stabilize Transthyretin to Prevent Amyloidogenesis while Affording a Fluorescent Conjugate. *J. Am. Chem. Soc.* **135**, 5656–5668 (2013).
186. Dalton, S. E. *et al.* Selectively Targeting the Kinome-Conserved Lysine of PI3K $\delta$  as a General Approach to Covalent Kinase Inhibition. *J. Am. Chem. Soc.* **140**, 932–939 (2018).
187. Anscombe, E. *et al.* Identification and Characterization of an Irreversible Inhibitor of CDK2. *Chemistry & Biology* **22**, 1159–1164 (2015).
188. Cuesta, A., Wan, X., Burlingame, A. L. & Taunton, J. Ligand Conformational Bias Drives Enantioselective Modification of a Surface-Exposed Lysine on Hsp90. *J. Am. Chem. Soc.* **142**, 3392–3400 (2020).
189. Zhao, Q. *et al.* Broad-Spectrum Kinase Profiling in Live Cells with Lysine-Targeted Sulfonyl Fluoride Probes. *J. Am. Chem. Soc.* **139**, 680–685 (2017).
190. Hett, E. C. *et al.* Rational Targeting of Active-Site Tyrosine Residues Using Sulfonyl Fluoride Probes. *ACS Chem. Biol.* **10**, 1094–1098 (2015).
191. Zanon, P. R. A. *et al.* Profiling the proteome-wide selectivity of diverse electrophiles. **10**.
192. Amati, B., Alevizopoulos, K., and Vlach, J. Myc and the cell cycle. *Front. Biosci. J. Virtual Libr.* **3**, d250-268 (1998).
193. Bateman, L.A., Nguyen, T.B., Roberts, A.M., Miyamoto, D.K., Ku, W.-M., Huffman, T.R., Petri, Y., Heslin, M.J., Contreras, C.M., Skibola, C.F., *et al.* Chemoproteomics-enabled covalent ligand screen reveals a cysteine hotspot in reticulon 4 that impairs ER morphology and cancer pathogenicity. *Chem. Commun. Camb. Engl.* **53**, 7234–7237 (2017).
194. Berg, T., Cohen, S.B., Desharnais, J., Sonderegger, C., Maslyar, D.J., Goldberg, J., Boger, D.L., and Vogt, P.K. Small-molecule antagonists of Myc/Max dimerization inhibit Myc-induced transformation of chicken embryo fibroblasts. *Proc. Natl. Acad. Sci.* **99**, 3830–3835 (2002).
195. Chen, H., Liu, H., and Qing, G. Targeting oncogenic Myc as a strategy for cancer treatment. *Signal Transduct. Target. Ther.* **3**, 5 (2018).

196. Cimperman, P., Baranauskiene, L., Jachimoviciūte, S., Jachno, J., Torresan, J., Michailoviene, V., Matuliene, J., Sereikaite, J., Bumelis, V., and Matulis, D. A quantitative model of thermal stabilization and destabilization of proteins by ligands. *Biophys. J.* **95**, 3222–3231 (2008).
197. Clarke, D.J.B., Lachmann, A., Evangelista, J.E., and Ma'avan, A. Signature Commons Signature Search.
198. Conacci-Sorrell, M., Ngouenet, C., and Eisenman, R.N. Myc-nick: a cytoplasmic cleavage product of Myc that promotes alpha-tubulin acetylation and cell differentiation. *Cell* **142**, 480–493 (2010).
199. Dang, C.V., Reddy, E.P., Shokat, K.M., and Soucek, L. Drugging the “undruggable” cancer targets. *Nat. Rev. Cancer* **17**, 502–508 (2017).
200. Grossman, E.A., Ward, C.C., Spradlin, J.N., Bateman, L.A., Huffman, T.R., Miyamoto, D.K., Kleinman, J.I., and Nomura, D.K. Covalent Ligand Discovery against Druggable Hotspots Targeted by Anti-cancer Natural Products. *Cell Chem. Biol.* **24**, 1368-1376.e4 (2017).
201. Han, H., Jain, A.D., Truica, M.I., Izquierdo-Ferrer, J., Anker, J.F., Lysy, B., Sagar, V., Luan, Y., Chalmers, Z.R., Unno, K., *et al.* Small-Molecule MYC Inhibitors Suppress Tumor Growth and Enhance Immunotherapy. *Cancer Cell* **36**, 483-497.e15 (2019).
202. Horiuchi, D., Kusdra, L., Huskey, N.E., Chandriani, S., Lenburg, M.E., Gonzalez-Angulo, A.M., Creasman, K.J., Bazarov, A.V., Smyth, J.W., Davis, S.E., *et al.* MYC pathway activation in triple-negative breast cancer is synthetic lethal with CDK inhibition. *J. Exp. Med.* **209**, 679–696 (2012).
203. Hydring, P., Castell, A., and Larsson, L.-G. MYC Modulation around the CDK2/p27/SKP2 Axis. *Genes* **8** (2017).
204. Karin, M., Liu, Z. g, and Zandi, E. AP-1 function and regulation. *Curr. Opin. Cell Biol.* **9**, 240–246 (1997).
205. Kress, T.R., Sabò, A., and Amati, B. MYC: connecting selective transcriptional control to global RNA production. *Nat. Rev. Cancer* **15**, 593 (2015).
206. Martinez Molina, D., Jafari, R., Ignatushchenko, M., Seki, T., Larsson, E.A., Dan, C., Sreekumar, L., Cao, Y., and Nordlund, P. Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science* **341**, 84–87 (2013).
207. McKeown, M.R., and Bradner, J.E. Therapeutic strategies to inhibit MYC. *Cold Spring Harb. Perspect. Med.* **4** (2014).

208. Meyer, N., and Penn, L.Z. Reflecting on 25 years with MYC. *Nat. Rev. Cancer* **8**, 976–990 (2008).
209. Osorio, J. Cell cycle: repurposing MYC and E2F in the absence of RB. *Nat. Rev. Mol. Cell Biol.* **16**, 516–517 (2015).
210. Struntz, N.B., Chen, A., Deutzmann, A., Wilson, R.M., Stefan, E., Evans, H.L., Ramirez, M.A., Liang, T., Caballero, F., Wildschut, M.H.E., *et al.* Stabilization of the Max Homodimer with a Small Molecule Attenuates Myc-Driven Transcription. *Cell Chem. Biol.* **26**, 711-723.e14 (2019).
211. Whitfield, J.R., Beaulieu, M.-E., and Soucek, L. Strategies to Inhibit Myc and Their Clinical Applicability. *Front. Cell Dev. Biol.* **5**, 10 (2017).
212. Wolf, E., Lin, C.Y., Eilers, M., and Levens, D.L. Taming of the beast: shaping Myc-dependent amplification. *Trends Cell Biol.* **25**, 241–248 (2015).
213. Wolfer, A., Wittner, B.S., Irimia, D., Flavin, R.J., Lupien, M., Gunawardane, R.N., Meyer, C.A., Lightcap, E.S., Tamayo, P., Mesirov, J.P., *et al.* MYC regulation of a “poor-prognosis” metastatic cancer cell state. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 3698–3703 (2010).
214. Xu, J., Chen, Y., and Olopade, O.I. MYC and Breast Cancer. *Genes Cancer* **1**, 629–640 (2010).
215. Zeller, K.I., Jegga, A.G., Aronow, B.J., O’Donnell, K.A., and Dang, C.V. An integrated database of genes responsive to the Myc oncogenic transcription factor: identification of direct genomic targets. *Genome Biol.* **4**, R69 (2003).
216. Zörnig, M., and Evan, G.I. Cell cycle: On target with Myc. *Curr. Biol.* **6**, 1553–1556 (1996).
217. Reflecting on 25 years with MYC | *Nature Reviews Cancer*.
218. Jin, Z., *et al.* Structure of Mpro from SARS-CoV-2 and Discovery of Its Inhibitors. *Nature* **582**, 289–293 (2020).
219. von Delft, F., *et al.* A White-Knuckle Ride of Open COVID Drug Discovery. *Nature* **594**, 330–332 (2021).
220. Zaidman, D., *et al.* An Automatic Pipeline for the Design of Irreversible Derivatives Identifies a Potent SARS-CoV-2 Mpro Inhibitor. *Cell Chem. Biol.* **12**, 1795-1806.e5 (2021).



221. How Pfizer scientists transformed an old drug lead into a COVID-19 antiviral <https://cen.acs.org/pharmaceuticals/drug-discovery/How-Pfizer-scientists-transformed-an-old-drug-lead-into-a-COVID-19-antiviral/100/i3> (accessed 2022 -01 -19).
222. Shaqra, A.M., *et al.* Defining the Substrate Envelope of SARS-CoV-2 Main Protease to Predict and Avoid Drug Resistance. *bioRxiv* January 27, 2022, p 2022.01.25.477757. <https://doi.org/10.1101/2022.01.25.477757>.
223. Flynn, J.M., *et al.* Comprehensive Fitness Landscape of SARS-CoV-2 Mpro Reveals Insights into Viral Resistance Mechanisms. *bioRxiv* January 30, 2022, p 2022.01.26.477860. <https://doi.org/10.1101/2022.01.26.477860>.
224. Ettari, R., *et al.* Development of Novel Peptidomimetics Containing a Vinyl Sulfone Moiety as Proteasome Inhibitors. *ChemMedChem* **6**, 1228–1237 (2011).
225. Dixon, S. J. & Stockwell, B. R. Identifying druggable disease-modifying gene products. *Curr. Opin. Chem. Biol.* **13**, 549–555 (2009).
226. Nalawansa, D. A. & Crews, C. M. PROTACs: An Emerging Therapeutic Modality in Precision Medicine. *Cell Chem. Biol.* **27**, 998–1014 (2020).
227. Burslem, G. M. & Crews, C. M. Proteolysis-Targeting Chimeras as Therapeutics and Tools for Biological Discovery. *Cell* **181**, 102–114 (2020).
228. Schapira, M., Calabrese, M. F., Bullock, A. N. & Crews, C. M. Targeted protein degradation: expanding the toolbox. *Nat. Rev. Drug Discov.* **18**, 949–963 (2019).
229. Farnaby, W., Koegl, M., McConnell, D. B. & Ciulli, A. Transforming targeted cancer therapy with PROTACs: a forward-looking perspective. *Curr. Opin. Pharmacol* **57**, 175–183 (2021).
230. Banik, S. M. *et al.* Lysosome-targeting chimaeras for degradation of extracellular proteins. *Nature* **584**, 291–297 (2020).
231. Takahashi, D. *et al.* AUTACs: Cargo-Specific Degradors Using Selective Autophagy. *Mol. Cell* **76**, 797-810.e10 (2019).
232. Yamazoe, S. *et al.* Heterobifunctional Molecules Induce Dephosphorylation of Kinases—A Proof of Concept Study. *J. Med. Chem.* **63**, 2807–2813 (2020).
233. Siriwardena, S. U. *et al.* Phosphorylation-Inducing Chimeric Small Molecules. *J. Am. Chem. Soc.* **142**, 14052–14057 (2020).
234. Sabapathy, K. & Lane, D. P. Therapeutic targeting of p53: all mutants are equal, but some mutants are more equal than others. *Nat. Rev. Clin. Oncol.* **15**, 13–30 (2018).

235. Abbas, T. & Dutta, A. p21 in cancer: intricate networks and multiple activities. *Nat. Rev. Cancer* **9**, 400–414 (2009).
236. Li, B. & Dou, Q. P. Bax degradation by the ubiquitin/proteasome-dependent pathway: Involvement in tumor survival and progression. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 3850–3855 (2000).
237. Ward, C. L., Omura, S. & Kopito, R. R. Degradation of CFTR by the ubiquitin-proteasome pathway. *Cell* **83**, 121–127 (1995).
238. Wiener, R., Zhang, X., Wang, T. & Wolberger, C. The mechanism of OTUB1-mediated inhibition of ubiquitination. *Nature* **483**, 618–622 (2012).
239. Nakada, S. *et al.* Non-canonical inhibition of DNA damage-dependent ubiquitination by OTUB1. *Nature* **466**, 941–946 (2010).
240. Que, L. T., Morrow, M. E. & Wolberger, C. Comparison of Cross-Regulation by Different OTUB1:E2 Complexes. *Biochemistry* **59**, 921–932 (2020).
241. French, M. E., Koehler, C. F. & Hunter, T. Emerging functions of branched ubiquitin chains. *Cell Discov.* **7**, 1–10 (2021).
242. Dunker, A. K. *et al.* Intrinsically disordered protein. *J. Mol. Graph. Model.* **19**, 26–59 (2001).
243. Garner, null, Romero, null, Dunker, null, Brown, null & Obradovic, null. Predicting Binding Regions within Disordered Proteins. *Genome Inform. Workshop Genome Inform.* **10**, 41–50 (1999).
244. Wiener, R. *et al.* E2 ubiquitin-conjugating enzymes regulate the deubiquitinating activity of OTUB1. *Nat. Struct. Mol. Biol.* **20**, 1033–1039 (2013).
245. Riordan, J. R. CFTR function and prospects for therapy. *Annu. Rev. Biochem.* **77**, 701–726 (2008).
246. Kanner, S. A., Shuja, Z., Choudhury, P., Jain, A. & Colecraft, H. M. Targeted deubiquitination rescues distinct trafficking-deficient ion channelopathies. *Nat. Methods* **17**, 1245–1253 (2020).
247. Tomati, V. *et al.* Genetic Inhibition Of The Ubiquitin Ligase Rnf5 Attenuates Phenotypes Associated To F508del Cystic Fibrosis Mutation. *Sci. Rep.* **5**, 12138 (2015).
248. Sondo, E. *et al.* Pharmacological Inhibition of the Ubiquitin Ligase RNF5 Rescues F508del-CFTR in Cystic Fibrosis Airway Epithelia. *Cell Chem. Biol.* **25**, 891-905.e8 (2018).

249. Lopes-Pacheco, M. CFTR Modulators: The Changing Face of Cystic Fibrosis in the Era of Precision Medicine. *Front. Pharmacol.* **10**, 1662 (2019).
250. Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G. & Deber, C. M. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 1760–1765 (2009).
251. Yang, Y. et al. Nedd4 ubiquitylates VDAC2/3 to suppress erastin-induced ferroptosis in melanoma. *Nat. Commun.* **11**, 433 (2020).
252. Ghelli Luserna di Rorà, A., Cerchione, C., Martinelli, G. & Simonetti, G. A WEE1 family business: regulation of mitosis, cancer progression, and therapeutic target. *J. Hematol. Oncol. J Hematol Oncol* **13**, 126 (2020).
253. Smith, A., Simanski, S., Fallahi, M. & Ayad, N. G. Redundant ubiquitin ligase activities regulate wee1 degradation and mitotic entry. *Cell Cycle Georget. Tex* **6**, 2795–2799 (2007).
254. Hashimoto, O. et al. Inhibition of proteasome-dependent degradation of Wee1 in G2-arrested Hep3B cells by TGF beta 1. *Mol. Carcinog.* **36**, 171–182 (2003).
255. Li, Z. et al. Development and Characterization of a Wee1 Kinase Degradation. *Cell Chem. Biol.* **27**, 57-65.e9 (2020).
256. Casement, R., Bond, A., Craighan, C. & Ciulli, A. Mechanistic and structural features of PROTAC ternary complexes. *Methods Mol. Biol.* **2365**, 79–113 (2021).
257. Gavathiotis, E., Reyna, D. E., Bellairs, J. A., Leshchiner, E. S. & Walensky, L. D. Direct and selective small-molecule activation of proapoptotic BAX. *Nat. Chem. Biol.* **8**, 639–645 (2012).
258. Pryde, D. C. et al. The discovery of potent small molecule activators of human STING. *Eur. J. Med. Chem.* **209**, 112869 (2021).
259. Ramanjulu, J. M. et al. Design of amidobenzimidazole STING receptor agonists with systemic activity. *Nature* **564**, 439–443 (2018).
260. Zorn, J. A. & Wells, J. A. Turning enzymes ON with small molecules. *Nat. Chem. Biol.* **6**, 179–188 (2010).
261. Chen, L., Liu, S. & Tao, Y. Regulating tumor suppressor genes: post-translational modifications. *Signal Transduct. Target. Ther.* **5**, 1–25 (2020).
262. Riboldi, G. M. & Di Fonzo, A. B. GBA, Gaucher Disease, and Parkinson's Disease: From Genetic to Clinic to New Therapeutic Approaches. *Cells* **8**, E364 (2019).

263. Sarodaya, N., Suresh, B., Kim, K.-S. & Ramakrishna, S. Protein Degradation and the Pathologic Basis of Phenylketonuria and Hereditary Tyrosinemia. *Int. J. Mol. Sci.* **21**, E4996 (2020).
- 262–264. Bartha, I., di Iulio, J., Venter, J. C. & Telenti, A. Human gene essentiality. *Nat. Rev. Genet.* **19**, 51–62 (2018).
265. 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).
266. Isobe, Y. *et al.* Manumycin polyketides act as molecular glues between UBR7 and P53. *Nat. Chem. Biol.* **16**, 1189–1198 (2020).
267. Xu, T. *et al.* ProLuCID: An improved SEQUEST-like algorithm with enhanced sensitivity and specificity. *J. Proteomics* **129**, 16–24 (2015).
268. Käll, L., Canterbury, J. D., Weston, J., Noble, W. S. & MacCoss, M. J. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat. Methods* **4**, 923–925 (2007).

## APPENDIX 1 CHAPTER 2 METHODS

### Covalent Ligand Library

Compounds were purchased from Enamine LLC and structures of compounds are all listed in **Table S1**.

### MYC/MAX In Vitro DNA Binding Experiments

The protocol was adapted from the Abcam c-Myc Transcription Factor Assay kit (ab207200). 50 mM DMSO stock solutions of compound were diluted into dPBS then vortexed. 0.2 µg of c-MYC (Abcam, ab169901) was added to each Eppendorf tube then mixed gently. c-MYC was incubated with compound or DMSO control for 30 minutes at 37 °C. After incubation, 0.2 µg of MAX (Origene, TP306812) was added to each Eppendorf tube then mixed gently. Samples were added to 96-well plate containing the consensus sequence for MYC/MAX binding. Experiment proceeded following manufacturer's protocol for binding, primary antibody, and secondary antibody incubation steps. After addition of developing and stop solution, the OD<sub>450</sub> was read on a Molecular Devices VersaMax microplate reader. Data was normalized to DMSO control after subtracting blank wells.

### MYC Luciferase Reporter Experiments

MYC luciferase reporter was purchased from Qiagen (CCS-012L) and performed according to manufacturer's protocol. This reporter assay employed a MYC-responsive firefly luciferase reporter, encoding a luciferase reporter gene under the control of a basal promoter element, the TATA box, joined to tandem repeats of the MYC transcriptional E-box response element. The negative control construct was a non-inducible firefly luciferase under the control of the TATA box element without addition of transcriptional response elements. The positive control reporter was a construct for constitutively expressing GFP and firefly luciferase. Also included in the MYC reporter, positive and negative control is a constitutively expressing Renilla luciferase that acts as an internal control for normalization against cytotoxicity, transfection efficiency, and technical variability. In each set of assays performed, the MYC-reporter controls treated with vehicle and positive controls showed >100 x more signal than the negative control.

Prior to transfection, HEK293T cells were seeded into a 96-well plate (35,000 cells/well in 100 µL). Negative control, positive control, or Myc luciferase reporter were diluted into Opti-MEM I medium (1 µL DNA into 25 µL Opti-MEM I per each well). Lipofectamine 2000 (Invitrogen, 11668019) was diluted into Opti-MEM I medium (1 µL Lipofectamine 2000 into 25 µL Opti-MEM I per each well). DNA and diluted Lipofectamine 2000 were combined in a 1:1 ratio then incubated at room temperature for 25 minutes. 50 µL of DNA-Lipofectamine 2000 mixture was added to each well. 24 hours post-transfection, media was carefully aspirated from each well, and 50 µL of fresh media containing either DMSO or compound was added. After 24 hours of

compound treatment, Dual-Glo luciferase (Promega, E2920) workup was performed according to manufacturer's protocol. Firefly and Renilla luminescence were read on a SpectraMax i3 plate reader. Background luminescence was subtracted using a blank control then Firefly:Renilla was calculated for each individual well.

### **LC-MS/MS Analysis of MYC**

Purified c-MYC (5 µg, Origene, 301611) in 50 µL PBS was incubated 30 min at room temperature either with DMSO vehicle or EN4 (50 µM). The DMSO control was then treated with light iodoacetamide while the EN4 treated sample was incubated with heavy iodoacetamide for 1 h each at room temperature (200 µM final concentration, Sigma-Aldrich, 721328). The samples were precipitated by additional of 12.5 µl of 100% (w/v) trichloroacetic acid and the treated and control groups were combined pairwise (10 µg total) before cooling to -80 °C for 2 h. The combined sample was then spun for at max speed for 10 min at 4 °C, supernatant was carefully removed and the sample was washed with ice-cold 0.01 M HCl/90% acetone solution. The pellet was resuspended in 4 M urea containing 0.1% Protease Max (Promega, V2071) and diluted in 40 mM ammonium bicarbonate buffer. The samples were reduced with 10 mM TCEP at 60 °C for 30 min. The sample was then diluted 50% with PBS before sequencing grade trypsin (1 µg per sample, Promega, V5111) was added for an overnight incubation at 37 °C. The next day, the sample was centrifuged at 13,200 r.p.m. for 30 min. The supernatant was transferred to a new tube and acidified to a final concentration of 5% formic acid and stored at -80 °C until mass spectrometry analysis.

### **Western Blotting Experiments**

Antibodies to c-MYC (Cell Signaling Technology, 9402S), CDK4 (Abcam, ab199728), CCND2 (Abcam, ab207604), FLAG (Cell Signaling Technology, 14793S), and β-actin (Cell Signaling Technology, 4970S) were obtained from various commercial sources and dilutions were prepared per the recommended manufacturers' procedures. Proteins were resolved by SDS-PAGE and transferred to nitrocellulose membranes using the iBlot system (Invitrogen). Blots were blocked with 5% Milk in Tris-buffered saline containing Tween 20 (TBST) solution for 1 h at room temperature, washed in TBST and probed with primary antibody diluted in diluent, as recommended by the manufacturer, overnight at 4 °C. Following washes with TBST, the blots were incubated in the dark with secondary antibodies purchased from Li-Cor Biosciences and used at 1:10,000 dilution in 5% Milk in TBST at room temperature for 1 h. Blots were visualized using an Odyssey Li-Cor scanner after additional washes.

### **Gel-ABPP Experiments**

Gel-based ABPP methods were performed as previously described (Grossman et al., 2017; Spradlin et al., 2019). Recombinant pure human proteins were purchased from Origene or Abcam as indicated. Pure proteins (0.3 µg) were pre-treated with DMSO vehicle or EN4 for 30 min at 37 °C in an incubation volume of 50 µl PBS and were subsequently treated with alkyne-functionalized probes for 2 h at 37 °C. CuAAC

was performed for 1 h at room temperature to append rhodamine-azide (1  $\mu$ M final concentration) onto alkyne probe-labeled proteins. Samples were then diluted with 20  $\mu$ l of 4x reducing Laemmli SDS sample loading buffer (Alfa Aesar) and heated at 90 °C for 5 min. The samples were separated on precast 4–20% TGX gels (Bio-Rad Laboratories, Inc.). Before scanning by ChemiDoc MP (Bio-Rad Laboratories, Inc.), gels were fixed in a solution of 10% acetic acid, 30% ethanol for 1 h.

For gel-based ABPP experiments with JUN, recombinant pure human c-JUN protein (0.1  $\mu$ g per sample) was pretreated with either DMSO vehicle or EN4 at 37 °C for 1 h in 25  $\mu$ l PBS, and subsequently treated with 200 nM Tetramethylrhodamine-5-Iodoacetamide Dihydroiodide (IA-rhodamine) (Setareh Biotech 6222) at room temperature for 1 h. Samples were incubated with 10  $\mu$ l 4x Laemmli sample buffer, boiled at 95°C for 6 min, and separated by SDS/PAGE. Probe-labeled proteins were analyzed by in-gel rhodamine fluorescence using a ChemiDoc MP (Bio-Rad). Protein loading was assessed by Silver Staining.

### **Vectors for cMYC Overexpression**

The pCMV6-Entry-cMYC was purchased from Origene (RC201611). The wild-type mammalian expression plasmids contained a C-terminal FLAG tag. The corresponding pCMV6-Entry-eGFP vector was constructed via Gibson Assembly with primers: GATCTGCCGCCGCGATCGCCatggtgagcaagggcgag, TCGAGCGGCCGCGTACGCGTctgtacagctcgccatgcc to amplify the eGFP ORF with desired overlaps, and ACGCGTACGCGGCCG, GGCGATCGCGGCCG to linearize the pCMV6-Entry backbone.

### **Expression of Wild-Type and C171Y Mutant MYC in HEK293T Cells with MYC Luciferase Reporter Experiments**

HEK293T cells were seeded at 35,000 cells/well in a 96-well plate. Cells were transfected with the MYC luciferase reporter according to protocol described above using Lipofectamine 2000 (Invitrogen, 11668019) as a transfection reagent. Cells were also transfected with either FLAG-eGFP, FLAG-tagged wild-type c-MYC (Origene, RC201611) or C171Y cMYC mammalian expression plasmids using Lipofectamine 2000. Cells were transfected with 1  $\mu$ g/well of corresponding plasmid in a final ratio of 3:1 Lipofectamine 2000:DNA. The c-MYC C171Y mutant plasmid was generated using Q5 site-directed mutagenesis kit according to the manufacturer's instructions (New England Biolabs, E0554S). Sequencing was confirmed with Quintara Biosciences. 24 hours post-transfection, media was carefully aspirated from each well, and 50  $\mu$ l of fresh media containing either DMSO or compound was added. After 24 hours of compound treatment, Dual-Glo luciferase (Promega, E2920) detection was performed according to manufacturer's protocol. Firefly and Renilla luminescence were read on a SpectraMax i3 plate reader. Background luminescence was subtracted using a blank control then Firefly:Renilla was calculated for each individual well. Lysates were also obtained for confirmation FLAG-protein overexpression by Western blotting.

## Knockdown of c-JUN by RNA interference with MYC Luciferase Reporter Experiments

RNA interference was performed by using siRNA purchased from Sigma MISSION siRNA library (SI007 siRNA Fluor Universal Negative Control#1, and PDSIRNA2D SASI\_Hs01\_00150279 siJUN). HEK293T cells were seeded at 35,000 cells/well in a 96-well plate. On Day 0, cells were co-transfected with 150 ng siRNA and myc luciferase reporter per well according to previous protocol using Lipofectomine 2000 (Invitrogen, 11668019). On Day 1, 24 hours post-transfection, 75  $\mu$ L of fresh media containing vehicle DMSO or EN4 was added to cells in n=6 replicates. After 24 h of compound treatment, Dual-Glo luciferase (Promega, E2920) workup was performed according to manufacturer's protocol. Firefly and Renilla luminescence were read on a SpectraMax i3 plate reader. Background luminescence was subtracted using a blank control then Firefly:Renilla was calculated for each individual well.

## EN4-Alkyne-4 Labeling of Endogenous MYC

231 MFP cells were treated at 70% confluency with DMSO or 50  $\mu$ M EN4-Alkyne-4 *in situ* for 4 h. Cells were harvested, lysed via sonication, and the resulting lysate normalized to 3.2 mg/mL per sample. Following normalization, 100  $\mu$ L of each lysate sample was removed for Western blot analysis of input, and 500  $\mu$ L of each lysate sample was incubated for 1 h at room temperature with 10  $\mu$ L of 5 mM biotin picolyl azide (in water) (Sigma Aldrich 900912), 10  $\mu$ L of 50 mM TCEP (in water), 30  $\mu$ L of TBTA ligand (0.9mg/mL in DMSO:t-butanol = 1:4), and 10  $\mu$ L of 50 mM Copper (II) Sulfate (12.5 mg/mL in water). Proteins were precipitated, washed 3 x with cold MeOH, resolubilized in 1 mL of 1.2% SDS/PBS (w/v), heated for 5 min at 90  $^{\circ}$ C, and centrifuged to remove any insoluble components. 1 mL of each resolubilized sample was then transferred to 15 mL conical tubes containing 5 mL PBS with 100  $\mu$ L high capacity streptavidin resin (Thermo Scientific 20357) to give a final SDS concentration of 0.2%. Samples were incubated with the streptavidin beads at 4 $^{\circ}$ C overnight on a rotator.

The following day the samples were warmed to room temperature and washed with 0.2% SDS, then transferred to spin columns and further washed 3 x with 50  $\mu$ L PBS and 3 x with 500  $\mu$ L water to remove non-probe-labeled proteins. The washed beads were resuspended in 100  $\mu$ L PBS, transferred to 1.5 mL eppendorf low-adhesion tubes, combined with 30  $\mu$ L Laemmli Sample Buffer (4 x) and heated to 95 $^{\circ}$ C. Immunoprecipitated proteins in each sample were then analyzed by Western blotting to look for enriched c-MYC (Abcam ab32072) versus non-enriched control Vinculin (Cell Signaling Technology 4650S).

## IsoTOP-ABPP Chemoproteomic Experiments

IsoTOP-ABPP studies were done as previously reported (Grossman et al., 2017; Spradlin et al., 2019; Weerapana et al., 2010). Cells were lysed by probe sonication in PBS and protein concentrations were measured by BCA assay. Cells were treated for 4 h with either DMSO vehicle or EN4 (from 1,000x DMSO stock) before cell collection and



lysis. Proteomes were subsequently labeled with IA-alkyne labeling (100  $\mu$ M) for 1 h at room temperature. CuAAC was used by sequential addition of tris(2-carboxyethyl)phosphine (1 mM, Strem, 15-7400), tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (34  $\mu$ M, Sigma, 678937), copper(II) sulfate (1 mM, Sigma, 451657) and biotin-linker-azide—the linker functionalized with a tobacco etch virus (TEV) protease recognition sequence as well as an isotopically light or heavy valine for treatment of control or treated proteome, respectively. After CuAAC, proteomes were precipitated by centrifugation at 6,500g, washed in ice-cold methanol, combined in a 1:1 control:treated ratio, washed again, then denatured and resolubilized by heating in 1.2% SDS–PBS to 80 °C for 5 min. Insoluble components were precipitated by centrifugation at 6,500g and soluble proteome was diluted in 5 ml 0.2% SDS–PBS. Labeled proteins were bound to streptavidin-agarose beads (170  $\mu$ l resuspended beads per sample, Thermo Fisher, 20349) while rotating overnight at 4 °C. Bead-linked proteins were enriched by washing three times each in PBS and water, then resuspended in 6 M urea/PBS, and reduced in TCEP (1 mM, Strem, 15-7400), alkylated with iodoacetamide (18 mM, Sigma), before being washed and resuspended in 2 M urea/PBS and trypsinized overnight with 0.5  $\mu$ g/ $\mu$ L sequencing grade trypsin (Promega, V5111). Tryptic peptides were eluted off. Beads were washed three times each in PBS and water, washed in TEV buffer solution (water, TEV buffer, 100  $\mu$ M dithiothreitol) and resuspended in buffer with Ac-TEV protease (Invitrogen, 12575-015) and incubated overnight. Peptides were diluted in water and acidified with formic acid (1.2 M, Fisher, A117-50) and prepared for analysis.

### **IsoTOP-ABPP Mass Spectrometry Analysis**

Peptides from all chemoproteomic experiments were pressure-loaded onto a 250  $\mu$ m inner diameter fused silica capillary tubing packed with 4 cm of Aqua C18 reverse-phase resin (Phenomenex, 04A-4299), which was previously equilibrated on an Agilent 600 series high-performance liquid chromatograph using the gradient from 100% buffer A to 100% buffer B over 10 min, followed by a 5 min wash with 100% buffer B and a 5 min wash with 100% buffer A. The samples were then attached using a MicroTee PEEK 360  $\mu$ m fitting (Thermo Fisher Scientific p-888) to a 13 cm laser pulled column packed with 10 cm Aqua C18 reverse-phase resin and 3 cm of strong-cation exchange resin for isoTOP-ABPP studies. Samples were analyzed using an Q Exactive Plus mass spectrometer (Thermo Fisher Scientific) using a five-step Multidimensional Protein Identification Technology (MudPIT) program, using 0, 25, 50, 80 and 100% salt bumps of 500 mM aqueous ammonium acetate and using a gradient of 5–55% buffer B in buffer A (buffer A: 95:5 water:acetonitrile, 0.1% formic acid; buffer B 80:20 acetonitrile:water, 0.1% formic acid). Data were collected in data-dependent acquisition mode with dynamic exclusion enabled (60 s). One full mass spectrometry (MS1) scan (400–1,800 mass-to-charge ratio ( $m/z$ )) was followed by 15 MS2 scans of the  $n$ th most abundant ions. Heated capillary temperature was set to 200 °C and the nanospray voltage was set to 2.75 kV.

Data were extracted in the form of MS1 and MS2 files using Raw Extractor v.1.9.9.2 (Scripps Research Institute) and searched against the Uniprot human database using ProLuCID search methodology in IP2 v.3 (Integrated Proteomics

Applications, Inc.)(Xu et al., 2015). Cysteine residues were searched with a static modification for carboxyamino-methylation (+57.02146) and up to two differential modifications for methionine oxidation and either the light or heavy TEV tags (+464.28596 or +470.29977, respectively). Peptides were required to be fully tryptic peptides and to contain the TEV modification. ProLUCID data were filtered through DTASelect to achieve a peptide false-positive rate below 5%. Only those probe-modified peptides that were evident across two out of three biological replicates were interpreted for their isotopic light to heavy ratios. For those probe-modified peptides that showed ratios greater than two, we only interpreted those targets that were present across all three biological replicates, were statistically significant and showed good quality MS1 peak shapes across all biological replicates. Light versus heavy isotopic probe-modified peptide ratios are calculated by taking the mean of the ratios of each replicate paired light versus heavy precursor abundance for all peptide-spectral matches associated with a peptide. The paired abundances were also used to calculate a paired sample *t*-test *P* value in an effort to estimate constancy in paired abundances and significance in change between treatment and control. *P* values were corrected using the Benjamini–Hochberg method.

### **Thermal Shift Assays**

231MFP cells (80% confluent in 10 cm culture dishes) were treated with EN4 or vehicle (DMSO) for 120 min at 50  $\mu$ M. Cells were harvested and washed once with PBS, then suspended in 1 mL of PBS containing protease inhibitor cocktail (Pierce A32955) and also maintained with 50  $\mu$ M compound or DMSO. The cell suspension was allocated into eight 0.2 mL PCR tubes with 100  $\mu$ L volume per tube, and each tube designated a temperature (37.0  $^{\circ}$ C, 38.4  $^{\circ}$ C, 40.9  $^{\circ}$ C, 45.0  $^{\circ}$ C, 49.8  $^{\circ}$ C, 53.5  $^{\circ}$ C, 56.3  $^{\circ}$ C, and 58.0  $^{\circ}$ C). Samples were heated at their respective temperatures for 3 min in a 96-well thermal cycler, then incubated at room temperature (25  $^{\circ}$ C) for 3 min. After this incubation, cells were immediately snap-lysed in liquid nitrogen (3 freeze-thaw cycles). Cell debris along with precipitated and aggregated proteins were removed from lysate by centrifuging samples at 20,000 g for 20 min at 4  $^{\circ}$ C. Samples were then boiled for 7 min at 90  $^{\circ}$ C after addition of loading buffer. Lysate samples were analyzed by Western Blot analysis. Protein intensity was quantified through Image J software.

### **RNA Sequencing Analysis**

231MFP cells were seeded into 6 cm dishes and cells were treated with DMSO vehicle or EN4 (50  $\mu$ M) for 24 h. Cells were harvested by scraping and RNA was isolated using Qiagen RNeasy mini kit with Qiagen DNase max kit to remove any DNA contamination. mRNA was enriched for using Oligo dT beads from the KAPA mRNA Capture Kit (KK8581). Subsequent library preparation steps of fragmentation, adapter ligation and cDNA synthesis were done on the enriched mRNA using the KAPA RNA HyperPrep kit (KK8540). Truncated universal stub adapters were used for ligation, and indexed primers were used during PCR amplification to complete the adapters and to enrich the libraries for adapter-ligated fragments. Samples were checked for quality on an AATI (now Agilent) Fragment Analyzer. Samples were then transferred to the

Vincent J. Coates Genomics Sequencing Laboratory (GSL), another QB3-Berkeley Core Research Facility at UC Berkeley, where Illumina sequencing library molarity was measured with quantitative PCR with the Kapa Biosystems Illumina Quant qPCR Kits on a BioRad CFX Connect thermal cycler. Libraries were then pooled evenly by molarity and sequenced on an Illumina HiSeq4000 (150PE) lane. Raw sequencing data was converted into FASTQ format, sample specific files using the Illumina bcl2fastq2 software on the sequencing centers local Linux server system.

RNA-seq reads were quantified using Salmon (Patro et al., 2017) to generate Transcripts Per Million (TPM) normalized data that was then rolled into gene level summaries. Treated versus control groups were contrasted on the log transformed data after removal of low variance genes with a t-test. This was followed by Benjamini–Hochberg false discovery rate correction and log<sub>2</sub> fold change calculation for each. We performed signature enrichment analysis of genes with an FDR < 0.1 and Fold change ≤ 0.75 (567 genes) using the Signature Commons platform against the MSigDB Collection H hallmark gene set.

## **Survival and Proliferation Experiments**

Cell survival and proliferation assays were performed using Hoechst 33342 dye (Invitrogen, H3570) according to the manufacturer's protocol and as previously described. For survival assays, cells were seeded and treated with serum-free media. Cells were seeded into 96-well plates (40,000 for survival) in a volume of 150 µL and allowed to adhere overnight. Cells were treated with an additional 50 µL of media containing compound in DMSO. After the appropriate incubation period, the media was removed from each well, and 100 µL of staining solution containing 10% formalin and Hoechst 33342 dye was added to each well and incubated for 15 min in the dark at room temperature. Staining solution was then removed, and wells were washed with dPBS before imaging.

For proliferation experiments, 231MFP cells were seeded at 15,000 cells/well in a black clear-bottom 96 well plate (avoiding the outer edges of the plate) in serum-containing L-15 media. The following day, cells were treated with 150 µL media containing vehicle (DMSO) or compound (EN4) at designated concentrations. For each subsequent treatment day, media was aspirated and replaced with fresh media containing DMSO or respective treatment. Proliferation was analyzed via Hoechst stain protocol as previously above.

## **Tumor Xenograft Study**

C.B17 SCID female mice (6-8 weeks old) were injected subcutaneously in the flank with 231MFP cells ( $1 \times 10^6$  cells) suspended in serum-free media. Mice were exposed via intraperitoneal injection with either vehicle (18:1:1 PBS/ethanol/PEG40) or 50 mg/kg EN4 once per day starting 2 weeks after injection of cancer cells. Tumor size was assessed weekly by caliper measurements.

## Synthetic Methods of EN4 Analogs

### Preparation of 4-(acrylamidomethyl)benzoic acid

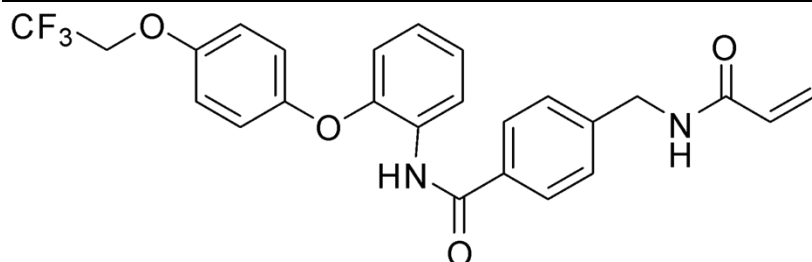
Step 1: methyl 4-(aminomethyl)benzoate (4.03 g, 20 mmol) and TEA (5.58 ml, 40.0 mmol) were stirred in acryloyl chloride (1.625 ml, 20.00 mmol) then acryloyl chloride (1.625 ml, 20.00 mmol) was added at 0 °C. The reaction mixture was stirred for 16 hr at ambient temperature, then it was diluted with EtOAc, washed with water, brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The resulting crude product was purified by flash column chromatography on silica gel eluting (10-100% EtOAc / Heptane) to give methyl 4-(acrylamidomethyl)benzoate as a white solid (3.03 g, 66%): **1H NMR** (400 MHz, Chloroform-d)  $\delta$  8.06 - 7.95 (m, 2H), 7.38 (d, *J* = 8.4 Hz, 2H), 6.37 (dd, *J* = 17.0, 1.4 Hz, 1H), 6.16 (dd, *J* = 17.0, 10.3 Hz, 1H), 5.72 (dd, *J* = 10.3, 1.4 Hz, 1H), 4.60 (d, *J* = 6.0 Hz, 2H), 3.93 (s, 3H); **LCMS** Rt 0.63 mins; *m/z* 220.1 (M+H)

Step 2: To a solution of methyl 4-(acrylamidomethyl)benzoate (500 mg, 2.28 mmol) in THF (5.70 ml) and water (5.70 ml) was added LiOH.H<sub>2</sub>O (383 mg, 9.12 mmol). The resulting solution was stirred at ambient temperature for 4 hr at which point HCl was added until the solution reached pH~2-3. The reaction mixture was then extracted with EtOAc (30 ml x2), the combined organic phases washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated to give 4-(acrylamidomethyl)benzoic acid as a white solid: **LCMS** Rt 0.49 mins; *m/z* 206.2 (M+H).

This material was used to prepare the following compounds according to general procedure A:

**Procedure A:** To 4-(acrylamidomethyl)benzoic acid (103 mg, 0.5 mmol), the corresponding aniline (0.500 mmol) in DMF (5 ml) was added EDC (115 mg, 0.600 mmol) and DMAP (61.1 mg, 0.500 mmol). The reaction mixture was stirred at 60 °C for 2 hr before being diluted with EtOAc (50 mL), washed with water, brine, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and the filtrate concentrated under reduced pressure. The crude product was purified by flash column chromatography on silica gel eluting (20-100% EtOAc/Heptane) to give the desired product

### 4-(acrylamidomethyl)-N-(2-(4-(2,2,2-trifluoroethoxy)phenoxy)phenyl)benzamide



**19 mg, 39 %**

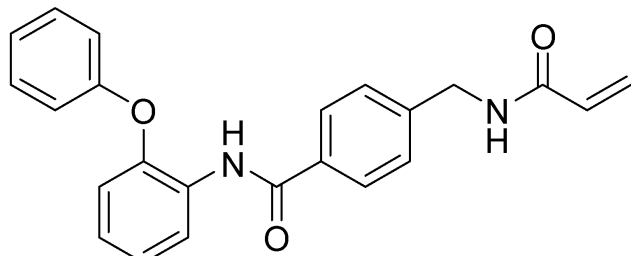
**1H NMR** (400 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  9.75 (s, 1H), 8.66 (t, *J* = 6.0 Hz, 1H), 7.81 (d, *J* = 8.3 Hz, 2H), 7.74 (dd, *J* = 7.5, 2.0 Hz, 1H), 7.35 (d, *J* = 8.3 Hz, 2H), 7.24 – 7.12 (m, 2H), 7.08 – 7.03 (m, 2H), 7.02 – 6.96 (m, 2H), 6.89 (dd, *J* = 7.8, 1.7 Hz, 1H), 6.28 (dd, *J* =

17.1, 10.1 Hz, 1H), 6.13 (dd,  $J = 17.1, 2.2$  Hz, 1H), 5.63 (dd,  $J = 10.1, 2.2$  Hz, 1H), 4.70 (q,  $J = 8.9$  Hz, 2H), 4.40 (d,  $J = 6.0$  Hz, 2H)

**<sup>19</sup>F NMR** (376 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  -72.59

**HRMS** calcd for C<sub>25</sub>H<sub>22</sub>F<sub>3</sub>N<sub>2</sub>O<sub>4</sub>(M+H)<sup>+</sup> 471.1532, found 471.1528

#### **4-(acrylamidomethyl)-N-(2-phenoxyphenyl)benzamide**



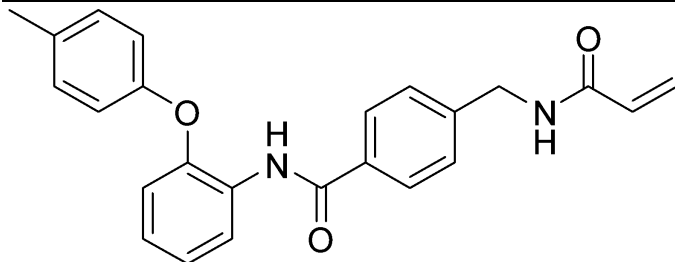
**83 mg, 45%**

**<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  9.78 (s, 1H), 8.68 (t,  $J = 6.0$  Hz, 1H), 7.77 – 7.69 (m, 3H), 7.33 (ddt,  $J = 8.4, 6.0, 3.0$  Hz, 4H), 7.28 – 7.17 (m, 2H), 7.11 – 7.04 (m, 1H), 7.03 – 6.99 (m, 1H), 6.99 – 6.93 (m, 2H), 6.28 (dd,  $J = 17.1, 10.1$  Hz, 1H), 6.13 (dd,  $J = 17.1, 2.2$  Hz, 1H), 5.64 (dd,  $J = 10.1, 2.2$  Hz, 1H), 4.39 (d,  $J = 6.0$  Hz, 2H).

**<sup>13</sup>C NMR** (101 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  163.53, 163.07, 155.15, 147.69, 141.47, 131.31, 129.90, 128.22, 127.99, 126.07, 125.48, 125.10, 124.86, 124.06, 122.29, 121.51, 118.01, 116.18, 40.21

**HRMS** calcd for C<sub>23</sub>H<sub>21</sub>N<sub>2</sub>O<sub>3</sub>(M+H)<sup>+</sup> 373.1552, found 373.1552

#### **4-(acrylamidomethyl)-N-(2-(p-tolyloxy)phenyl)benzamide**



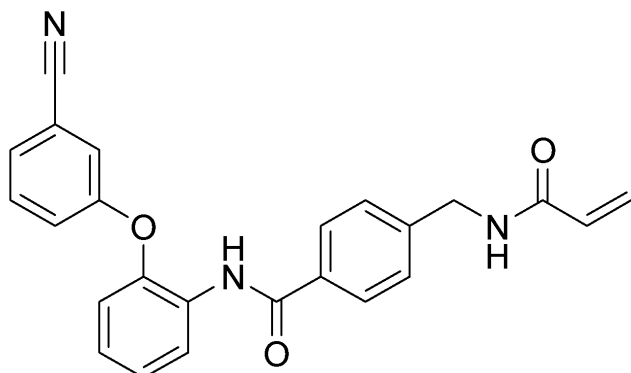
**94 mg, 48%**

**<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  9.74 (s, 1H), 8.68 (t,  $J = 6.0$  Hz, 1H), 7.78 (d,  $J = 8.3$  Hz, 2H), 7.74 (dd,  $J = 7.5, 2.1$  Hz, 1H), 7.34 (d,  $J = 8.4$  Hz, 2H), 7.23 – 7.12 (m, 4H), 6.93 (dd,  $J = 7.7, 1.8$  Hz, 1H), 6.91 – 6.86 (m, 2H), 6.28 (dd,  $J = 17.1, 10.1$  Hz, 1H), 6.13 (dd,  $J = 17.1, 2.2$  Hz, 1H), 5.64 (dd,  $J = 10.1, 2.2$  Hz, 1H), 4.40 (d,  $J = 6.0$  Hz, 2H), 2.25 (s, 3H)

**<sup>13</sup>C NMR** (101 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  165.08, 164.65, 154.35, 149.81, 143.08, 132.90, 132.28, 131.48, 130.16, 129.25, 127.67, 127.09, 126.44, 126.30, 125.63, 123.40, 118.85, 118.12, 41.80, 20.20

**HRMS** calcd for C<sub>24</sub>H<sub>23</sub>N<sub>2</sub>O<sub>3</sub>(M+H)<sup>+</sup> 387.1709, found 387.1708

#### 4-(acrylamidomethyl)-N-(2-(3-cyanophenoxy)phenyl)benzamide



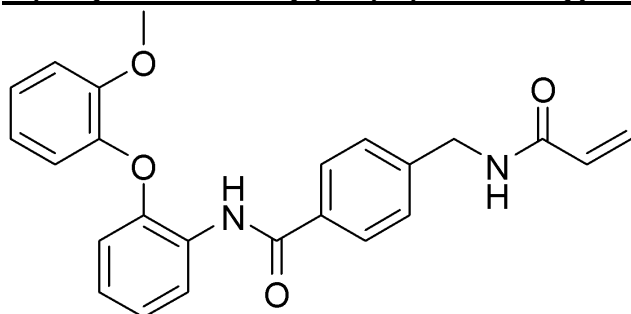
**69 mg, 35%**

**<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ 9.89 (s, 1H), 8.68 (t, *J* = 6.0 Hz, 1H), 7.74 – 7.66 (m, 3H), 7.55 – 7.48 (m, 2H), 7.38 – 7.24 (m, 6H), 7.17 – 7.11 (m, 1H), 6.28 (dd, *J* = 17.1, 10.1 Hz, 1H), 6.13 (dd, *J* = 17.1, 2.2 Hz, 1H), 5.64 (dd, *J* = 10.1, 2.2 Hz, 1H), 4.39 (d, *J* = 6.0 Hz, 2H)

**<sup>13</sup>C NMR** (101 MHz, DMSO-*d*<sub>6</sub>) δ 165.23, 164.65, 157.21, 148.59, 143.14, 132.67, 131.47, 131.19, 129.83, 127.62, 127.53, 127.04, 126.68, 125.64, 124.97, 122.49, 120.55, 120.46, 118.19, 112.28, 41.79

**HRMS** calcd for C<sub>24</sub>H<sub>20</sub>N<sub>3</sub>O<sub>3</sub>(M+H)<sup>+</sup> 398.1505, found 398.1503

#### 4-(acrylamidomethyl)-N-(2-(2-methoxyphenoxy)phenyl)benzamide



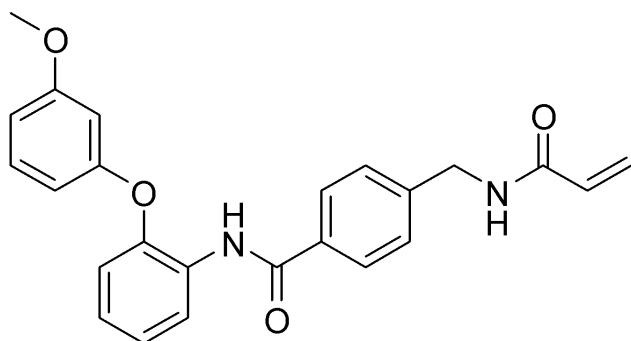
**108 mg, 53%**

**<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ 9.67 (s, 1H), 8.70 (t, *J* = 6.0 Hz, 1H), 7.89 (d, *J* = 8.3 Hz, 2H), 7.82 (dd, *J* = 7.3, 2.3 Hz, 1H), 7.38 (d, *J* = 8.3 Hz, 2H), 7.21 – 7.03 (m, 5H), 6.95 (ddd, *J* = 8.0, 6.7, 2.2 Hz, 1H), 6.73 – 6.68 (m, 1H), 6.29 (dd, *J* = 17.1, 10.1 Hz, 1H), 6.13 (dd, *J* = 17.1, 2.2 Hz, 1H), 5.64 (dd, *J* = 10.1, 2.2 Hz, 1H), 4.42 (d, *J* = 6.0 Hz, 2H), 3.73 (s, 3H)

**<sup>13</sup>C NMR** (101 MHz, DMSO-*d*<sub>6</sub>) δ 162.91, 162.60, 148.98, 148.07, 142.06, 141.12, 130.90, 129.42, 126.07, 125.56, 125.16, 123.74, 123.58, 123.34, 123.25, 120.41, 119.16, 118.97, 114.32, 111.15, 53.55, 39.76

**HRMS** calcd for C<sub>24</sub>H<sub>23</sub>N<sub>2</sub>O<sub>4</sub>(M+H)<sup>+</sup> 403.1658, found 403.1644

#### 4-(acrylamidomethyl)-N-(2-(3-methoxyphenoxy)phenyl)benzamide



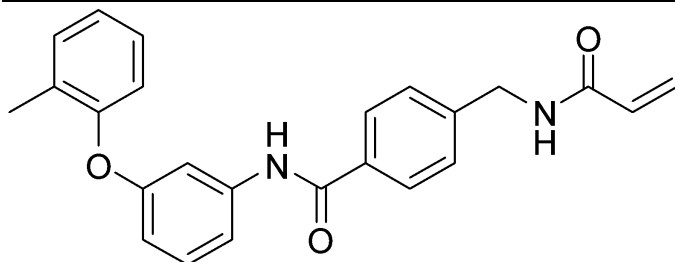
**79 mg, 39%**

**<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ 9.76 (s, 1H), 8.68 (t, *J* = 6.0 Hz, 1H), 7.79 – 7.72 (m, 3H), 7.32 (d, *J* = 8.3 Hz, 2H), 7.28 – 7.17 (m, 3H), 7.08 – 7.00 (m, 1H), 6.68 – 6.63 (m, 1H), 6.57 – 6.48 (m, 2H), 6.28 (dd, *J* = 17.1, 10.1 Hz, 1H), 6.13 (dd, *J* = 17.1, 2.2 Hz, 1H), 5.64 (dd, *J* = 10.1, 2.2 Hz, 1H), 4.39 (d, *J* = 6.0 Hz, 2H), 3.68 (s, 3H).

**<sup>13</sup>C NMR** (101 MHz, DMSO-*d*<sub>6</sub>) δ 165.12, 164.65, 160.47, 157.93, 149.15, 143.08, 132.88, 131.48, 130.27, 129.60, 127.65, 127.06, 126.74, 126.46, 125.64, 124.00, 119.79, 109.80, 108.75, 103.82, 55.15, 41.79

**HRMS** calcd for C<sub>24</sub>H<sub>23</sub>N<sub>2</sub>O<sub>4</sub>(M+H)<sup>+</sup> 403.1658, found 403.1653

**4-(acrylamidomethyl)-N-(3-(*o*-tolylloxy)phenyl)benzamide**



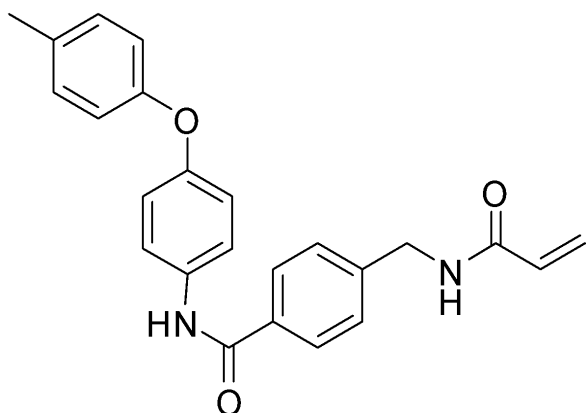
**92 mg, 47%**

**<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ 10.23 (s, 1H), 8.71 (t, *J* = 6.0 Hz, 1H), 7.88 (d, *J* = 8.3 Hz, 2H), 7.53 - 7.49 (m, 1H), 7.42 - 7.28 (m, 5H), 7.27 - 7.21 (m, 1H), 7.13 (td, *J* = 7.4, 1.0 Hz, 1H), 6.95 (d, *J* = 8.0 Hz, 1H), 6.72 - 6.61 (m, 1H), 6.29 (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), 4.42 (d, *J* = 6.0 Hz, 2H), 2.18 (s, 3H)

**<sup>13</sup>C NMR** (101 MHz, DMSO-*d*<sub>6</sub>) δ 164.01, 163.32, 156.26, 152.30, 141.80, 139.35, 131.97, 130.17, 130.14, 128.48, 128.01, 126.43, 126.18, 125.76, 124.30, 123.02, 118.58, 112.86, 110.77, 107.08, 40.49, 14.48

**HRMS** calcd for C<sub>24</sub>H<sub>23</sub>N<sub>2</sub>O<sub>3</sub>(M+H)<sup>+</sup> 387.1709, found 387.1705

**4-(acrylamidomethyl)-N-(4-(*p*-tolylloxy)phenyl)benzamide**



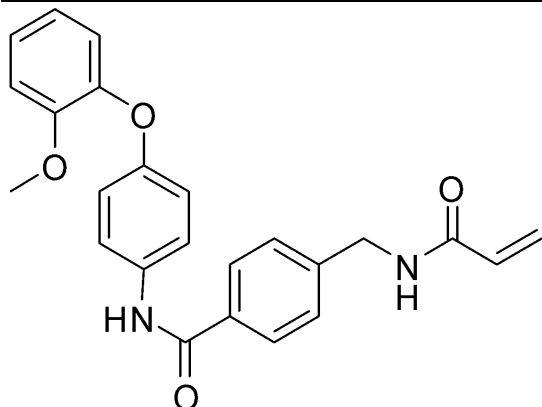
**93 mg, 48%**

**<sup>1</sup>H NMR** (400 MHz, DMSO-d<sub>6</sub>) δ 10.22 (s, 1H), 8.72 (t, *J* = 6.0 Hz, 1H), 7.92 (d, *J* = 8.3 Hz, 2H), 7.79 - 7.72 (m, 2H), 7.40 (d, *J* = 8.3 Hz, 2H), 7.18 (d, *J* = 8.1 Hz, 2H), 7.04 - 6.95 (m, 2H), 6.93 - 6.87 (m, 2H), 6.30 (dd, *J* = 17.1, 10.1 Hz, 1H), 6.14 (dd, *J* = 17.1, 2.2 Hz, 1H), 5.65 (dd, *J* = 10.1, 2.2 Hz, 1H), 4.43 (d, *J* = 6.0 Hz, 2H), 2.28 (s, 3H)

**<sup>13</sup>C NMR** (101 MHz, DMSO-d<sub>6</sub>) δ 165.04, 164.67, 154.80, 152.69, 143.02, 134.63, 133.42, 132.18, 131.50, 130.33, 127.71, 127.14, 125.64, 121.98, 118.66, 118.23, 41.85, 20.21

**HRMS** calcd for C<sub>24</sub>H<sub>23</sub>N<sub>2</sub>O<sub>3</sub>(M+H)<sup>+</sup> 387.1709, found 387.1702

**4-(acrylamidomethyl)-N-(4-(2-methoxyphenoxy)phenyl)benzamide**



**99 mg, 48%**

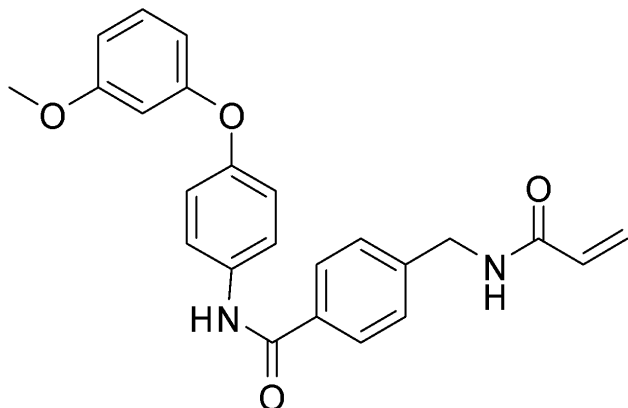
**<sup>1</sup>H NMR** (400 MHz, DMSO-d<sub>6</sub>) δ 10.17 (s, 1H), 8.71 (t, *J* = 6.0 Hz, 1H), 7.91 (d, *J* = 8.3 Hz, 2H), 7.71 - 7.65 (m, 2H), 7.39 (d, *J* = 8.3 Hz, 2H), 7.22 - 7.14 (m, 2H), 7.05 - 6.92 (m, 2H), 6.89 - 6.81 (m, 2H), 6.30 (dd, *J* = 17.1, 10.1 Hz, 1H), 6.14 (dd, *J* = 17.1, 2.2 Hz, 1H), 5.65 (dd, *J* = 10.1, 2.2 Hz, 1H), 4.43 (d, *J* = 6.0 Hz, 2H), 3.75 (s, 3H)

**<sup>13</sup>C NMR** (101 MHz, DMSO-d<sub>6</sub>) δ 162.79, 162.51, 151.51, 149.01, 141.89, 140.81, 131.55, 131.29, 129.35, 125.53, 124.97, 123.49, 123.07, 119.82, 118.93, 118.88, 114.23, 111.16, 53.40, 39.69

**HRMS** calcd for C<sub>24</sub>H<sub>23</sub>N<sub>2</sub>O<sub>4</sub>(M+H)<sup>+</sup> 403.1658, found 403.1648

**4-(acrylamidomethyl)-N-(4-(3-methoxyphenoxy)phenyl)benzamide**





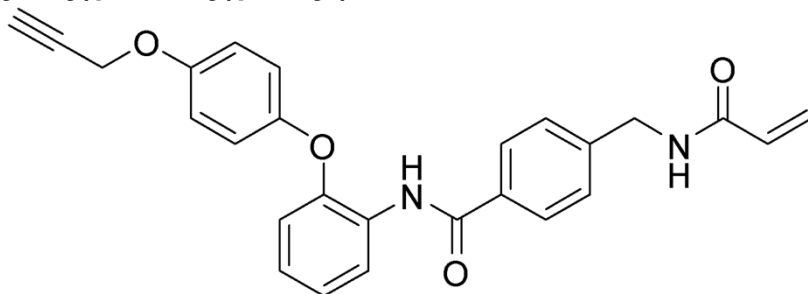
**121 mg, 58%**

**<sup>1</sup>H NMR** (400 MHz, DMSO-d<sub>6</sub>) δ 10.25 (s, 1H), 8.72 (t, *J* = 6.0 Hz, 1H), 7.96 - 7.89 (m, 2H), 7.83 - 7.76 (m, 2H), 7.45 - 7.37 (m, 2H), 7.27 (t, *J* = 8.2 Hz, 1H), 7.07 - 7.00 (m, 2H), 6.69 (ddd, *J* = 8.5, 2.4, 0.9 Hz, 1H), 6.56 (t, *J* = 2.3 Hz, 1H), 6.52 (ddd, *J* = 8.2, 2.3, 0.9 Hz, 1H), 6.30 (dd, *J* = 17.1, 10.1 Hz, 1H), 6.14 (dd, *J* = 17.1, 2.2 Hz, 1H), 5.65 (dd, *J* = 10.2, 2.2 Hz, 1H), 4.43 (d, *J* = 5.9 Hz, 2H), 3.73 (s, 3H)

**<sup>13</sup>C NMR** (101 MHz, DMSO-d<sub>6</sub>) δ 164.44, 164.01, 159.98, 157.86, 151.13, 142.40, 134.46, 132.75, 130.84, 129.78, 127.07, 126.49, 124.99, 121.29, 118.75, 109.12, 108.00, 103.26, 54.55, 41.19

**HRMS** calcd for C<sub>24</sub>H<sub>23</sub>N<sub>2</sub>O<sub>4</sub>(M+H)<sup>+</sup> 403.1658, found 403.1656

**Preparation of 4-(acrylamidomethyl)-N-(2-(4-(prop-2-yn-1-yloxy)phenoxy)phenyl)benzamide**



Step 1: To a suspension of K<sub>2</sub>CO<sub>3</sub> (10.04 g, 72.7 mmol) in DMAc (45.4 ml) was added hydroquinone (2 g, 18.16 mmol) and 3-bromoprop-1-yne (80% in toluene w/w) (2.023 ml, 18.16 mmol). The resulting dark solution was stirred at 40 °C for 4 hr when the reaction was partitioned between EtOAc and water, washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. The crude residue was purified by flash chromatography on silica gel eluting with (0-50% EtOAc / Heptanes) to give the desired product 4-(prop-2-yn-1-yloxy)phenol as a pale solid (533 mg, 20%); **<sup>1</sup>H NMR** (400 MHz, DMSO-d<sub>6</sub>) δ 9.01 (s, 1H), 6.83 - 6.76 (m, 2H), 6.71 - 6.64 (m, 2H), 4.65 (d, *J* = 2.4 Hz, 2H), 3.52 (t, *J* = 2.4 Hz, 1H); **LCMS** Rt 0.64 mins; m/z 149.0 (M+H)<sup>+</sup>

Step 2: A mixture of 4-(prop-2-yn-1-yloxy)phenol (548 mg, 3.70 mmol), 1-fluoro-2-nitrobenzene (390 μl, 3.70 mmol), K<sub>2</sub>CO<sub>3</sub> (767 mg, 5.55 mmol) and DMAc (3.70 ml) was stirred at 160° C for 2 hr under an inert atmosphere. The reaction was diluted with

EtOAc, washed with water and saturated brine, and dried over anhydrous sodium sulfate and filtered. The filtrate was concentrated under reduced pressure to give a residue which was purified by flash chromatography on silica gel eluting with (0-20% EtOAc / Heptanes) to give 1-nitro-2-(4-(prop-2-yn-1-yloxy)phenoxy)benzene as a pale oil (857 mg, 86%); **1H NMR** (400 MHz, Chloroform-d)  $\delta$  7.95 (dd,  $J$  = 8.2, 1.6 Hz, 1H), 7.49 (ddd,  $J$  = 8.8, 7.5, 1.7 Hz, 1H), 7.22 - 7.14 (m, 1H), 7.09 - 7.00 (m, 4H), 6.96 (dd,  $J$  = 8.4, 1.0 Hz, 1H), 4.72 (d,  $J$  = 2.4 Hz, 2H), 2.57 (t,  $J$  = 2.4 Hz, 1H)

Step 3: To a solution of 1-nitro-2-(4-(prop-2-yn-1-yloxy)phenoxy)benzene (857 mg, 3.18 mmol) in EtOH (18.100 ml) and water (1.80 ml) was added iron powder (978 mg, 17.51 mmol) and ammonium chloride (119 mg, 2.228 mmol). The resulting mixture was heated at reflux at 95 °C for 2 hr when the reaction mixture was passed through filtered through Celite™ rinsing the pad with MeOH and the filtrate evaporated *in vacuo*. The residue was partitioned between EtOAc / sat NaHCO<sub>3</sub> solution and the combined organic layers then dried over anhydrous magnesium sulphate, filtered and the filtrate evaporated to dryness to afford a light brown oil. The crude 2-(4-(prop-2-yn-1-yloxy)phenoxy)aniline was used to next step without further purification: **LCMS** Rt 0.94 mins; m/z 240.2 M+H

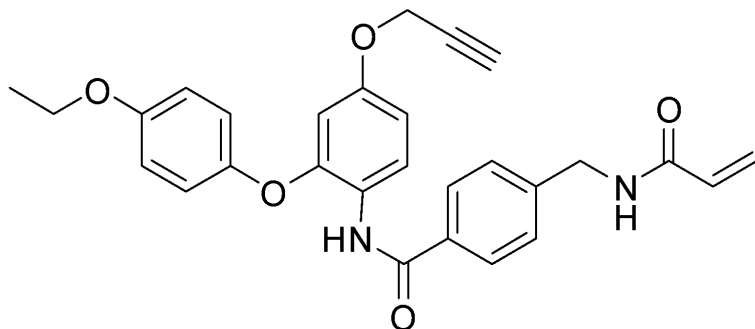
Step 4: To 4-(acrylamidomethyl)benzoic acid (103 mg, 0.5 mmol), 2-(4-(prop-2-yn-1-yloxy)phenoxy)aniline (120 mg, 0.500 mmol) in DMF (5.00 ml) was added EDC (115 mg, 0.600 mmol) and DMAP (61.1 mg, 0.500 mmol). The mixture was stirred at 60 °C for 2 hr before it was diluted with EtOAc (50 mL), the combined organics then washed with water, brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude product was purified by flash chromatography on silica gel eluting with (20-100% EtOAc/Heptane) to give a white solid which was then subjected to HPLC (C18 OBD 30 x 50mm; MeCN / H<sub>2</sub>O w 0.1% formic acid; 35-65% MeCN gradient) to give a 4-(acrylamidomethyl)-N-(2-(4-(prop-2-yn-1-yloxy)phenoxy)phenyl)benzamide as white solid (71 mg, 34%)

**1H NMR** (400 MHz, DMSO-d<sub>6</sub>)  $\delta$  9.76 (s, 1H), 8.70 (t,  $J$  = 6.0 Hz, 1H), 7.82 (d,  $J$  = 8.3 Hz, 2H), 7.74 (dd,  $J$  = 7.7, 1.9 Hz, 1H), 7.35 (d,  $J$  = 8.3 Hz, 2H), 7.22 - 7.11 (m, 2H), 6.98 (s, 4H), 6.87 (dd,  $J$  = 7.9, 1.7 Hz, 1H), 6.29 (dd,  $J$  = 17.1, 10.1 Hz, 1H), 6.13 (dd,  $J$  = 17.1, 2.2 Hz, 1H), 5.64 (dd,  $J$  = 10.1, 2.2 Hz, 1H), 4.75 (d,  $J$  = 2.4 Hz, 2H), 4.40 (d,  $J$  = 6.0 Hz, 2H), 3.56 (t,  $J$  = 2.4 Hz, 1H)

**13C NMR** (101 MHz, DMSO-d<sub>6</sub>)  $\delta$  165.02, 164.57, 153.15, 150.33, 150.27, 143.00, 132.85, 131.39, 128.77, 127.60, 127.04, 126.21, 126.17, 125.56, 122.96, 119.71, 117.97, 115.91, 79.17, 78.12, 55.65, 41.73

**HRMS** calcd for C<sub>26</sub>H<sub>23</sub>N<sub>2</sub>O<sub>4</sub>(M+H)<sup>+</sup> 427.1658, found 427.1652

#### **Preparation of 4-(acrylamidomethyl)-N-(2-(4-ethoxyphenoxy)-4-(prop-2-yn-1-yloxy)phenyl) benzamide**



Step 1: To a solution of 2,4-difluoro-1-nitrobenzene (1.992 ml, 18.17 mmol) in DMF (91 ml) was added propargyl alcohol (1.586 ml, 27.2 mmol) and K<sub>2</sub>CO<sub>3</sub> (7.53 g, 54.5 mmol). The solution was stirred at 40 °C for 18 hr when the reaction mixture was diluted with EtOAc, washed with water, brine, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and the filtrate concentrated *in vacuo*. The residue was purified by flash chromatography on silica gel (0-20% EtOAc/Heptane) to give the undesired regioisomer (1.1g, 30%) and the desired product 2-fluoro-1-nitro-4-(prop-2-yn-1-yloxy)benzene (545 mg, 16%): **<sup>1</sup>H NMR** (400 MHz, Chloroform-d) δ 8.24 - 8.07 (m, 1H), 6.93 - 6.78 (m, 2H), 4.81 (d, J = 2.4 Hz, 2H), 2.64 (t, J = 2.4 Hz, 1H)

Step 2: A mixture of 4-ethoxyphenol (370 mg, 2.67 mmol), 2-fluoro-1-nitro-4-(prop-2-yn-1-yloxy)benzene (522 mg, 2.67 mmol), K<sub>2</sub>CO<sub>3</sub> (555 mg, 4.01 mmol) in DMA (2.68 ml) was stirred at 160 °C for 2 hr under an inert atmosphere. The reaction mixture was then diluted with EtOAc, washed with water and saturated brine, and the combined organics dried over anhydrous sodium sulfate and filtered. The filtrate was evaporated under reduced pressure to yield a crude product which was then purified by flash chromatography on silica gel eluting (0-20% EtOAc / Heptanes) to give 2-(4-ethoxyphenoxy)-1-nitro-4-(prop-2-yn-1-yloxy)benzene as a light pale solid (720 mg, 86%): **<sup>1</sup>H NMR** (400 MHz, Chloroform-d) δ 8.08 (d, J = 9.2 Hz, 1H), 7.08 - 7.01 (m, 2H), 6.97 - 6.90 (m, 2H), 6.70 (dd, J = 9.2, 2.6 Hz, 1H), 6.45 (d, J = 2.5 Hz, 1H), 4.66 (d, J = 2.3 Hz, 2H), 4.06 (q, J = 7.0 Hz, 2H), 2.55 (t, J = 2.3 Hz, 1H), 1.45 (t, J = 7.0 Hz, 3H); **LCMS** Rt 1.15 mins; m/z 314.1 (M+H)

Step 3: To a solution of 2-(4-ethoxyphenoxy)-1-nitro-4-(prop-2-yn-1-yloxy)benzene (710 mg, 2.27 mmol) in EtOH (12.900 ml) and water (1.29 ml) was added iron powder (696 mg, 12.46 mmol) and ammonium chloride (85 mg, 1.586 mmol). The resulting mixture was heated at reflux at 90 °C for 18 hr. The reaction mixture was cooled to ambient temperature, filtered through Celite™ rinsing the pad with MeOH and the filtrate evaporated *in vacuo*. The residue was partitioned between EtOAc / sat NaHCO<sub>3</sub> solution and the combined organic layers then dried over anhydrous magnesium sulphate, filtered and the filtrate evaporated to dryness to afford a light brown oil. The crude product was purified by flash chromatography on silica gel eluting (0-50% EtOAc / Heptanes) to give 2-(4-ethoxyphenoxy)-4-(prop-2-yn-1-yloxy)aniline as a light brown oil (477 mg, 74%): **<sup>1</sup>H NMR** (400 MHz, Chloroform-d) δ 6.99 - 6.94 (m, 2H), 6.90 - 6.85 (m, 2H), 6.78 (d, J = 8.6 Hz, 1H), 6.61 (dd, J = 8.6, 2.7 Hz, 1H), 6.47 (d, J = 2.7 Hz, 1H), 4.56 (d, J = 2.4 Hz, 2H), 4.03 (q, J = 7.0 Hz, 2H), 3.83

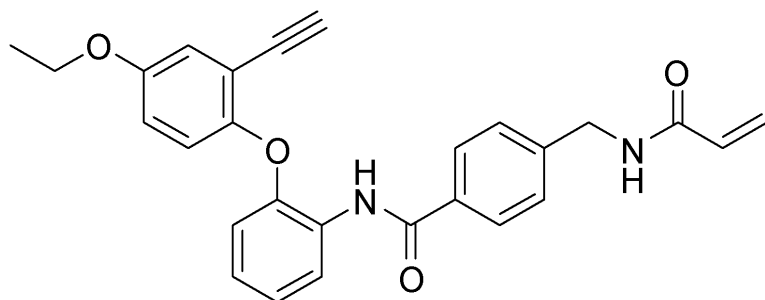
(br s, 2H), 2.49 (t,  $J = 2.4$  Hz, 1H), 1.44 (t,  $J = 7.0$  Hz, 3H); **LCMS** Rt 0.89 mins;  $m/z$  284.1 (M+H).

Step 4: 4-(acrylamidomethyl)benzoic acid (103 mg, 0.5 mmol) and 2-(4-ethoxyphenoxy)-4-(prop-2-yn-1-yloxy)aniline (142 mg, 0.500 mmol) were dissolved in DMF (5 ml) to give a clear solution, then EDC (115 mg, 0.600 mmol) and DMAP (61.1 mg, 0.500 mmol) were added. The mixture was stirred at 60 °C for 2 hr before the mixture was diluted with EtOAc (50 mL), washed with water, brine, and the combined organics dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated. The crude product was purified by flash chromatography on silica gel eluting (20-100% EtOAc/Heptane) to give 4-(acrylamidomethyl)-N-(2-(4-ethoxyphenoxy)-4-(prop-2-yn-1-yloxy)phenyl) benzamide as a white solid (153 mg, 63%): **<sup>1</sup>H NMR** (400 MHz, DMSO-d<sub>6</sub>)  $\delta$  9.67 (s, 1H), 8.69 (t,  $J = 6.0$  Hz, 1H), 7.81 (d,  $J = 8.2$  Hz, 2H), 7.53 (d,  $J = 8.8$  Hz, 1H), 7.34 (d,  $J = 8.3$  Hz, 2H), 6.99 - 6.85 (m, 4H), 6.77 (dd,  $J = 8.8, 2.8$  Hz, 1H), 6.45 (d,  $J = 2.7$  Hz, 1H), 6.29 (dd,  $J = 17.1, 10.1$  Hz, 1H), 6.13 (dd,  $J = 17.1, 2.2$  Hz, 1H), 5.64 (dd,  $J = 10.1, 2.2$  Hz, 1H), 4.74 (d,  $J = 2.3$  Hz, 2H), 4.40 (d,  $J = 6.0$  Hz, 2H), 3.97 (q,  $J = 7.0$  Hz, 2H), 3.58 (t,  $J = 2.3$  Hz, 1H), 1.30 (t,  $J = 7.0$  Hz, 3H)

**<sup>13</sup>C NMR** (101 MHz, DMSO-d<sub>6</sub>)  $\delta$  164.43, 163.94, 154.79, 154.00, 151.29, 148.66, 142.23, 132.25, 130.78, 127.19, 126.94, 126.35, 124.92, 121.46, 119.24, 114.66, 108.10, 104.48, 78.27, 77.75, 62.60, 54.97, 41.10, 13.93

**HRMS** calcd for C<sub>28</sub>H<sub>27</sub>N<sub>2</sub>O<sub>5</sub>(M+H)<sup>+</sup> 471.1920, found 471.1915

#### Preparation of 4-(acrylamidomethyl)-N-(2-(4-ethoxy-2-ethynylphenoxy)phenyl)benzamide



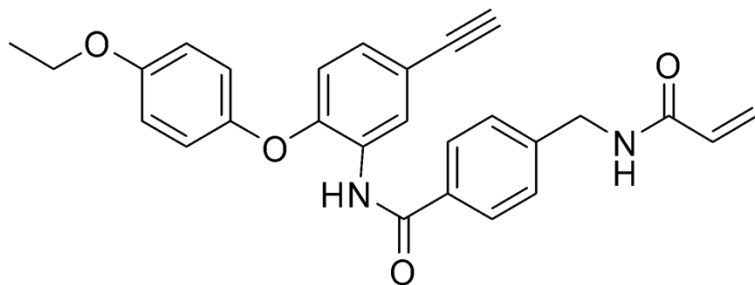
Step1: A mixture of 5-ethoxy-2-hydroxybenzaldehyde (405 mg, 2.44 mmol), 1-fluoro-2-nitrobenzene (257  $\mu$ l, 2.44 mmol), K<sub>2</sub>CO<sub>3</sub> (505 mg, 3.66 mmol) in DMA (2.44 ml) was stirred at 160° C for 2 hr under an inert atmosphere. The reaction mixture was cooled to room temperature, diluted with water and extracted with ethyl acetate (3 x 25 ml). The combined organic phases were washed with water, brine, and dried over anhydrous sodium sulfate, filtered and the filtrate evaporated under reduced pressure. The residue obtained was purified by flash column chromatography on silica gel eluting (0-30% EtOAc/Heptane) to give the desired compound 5-ethoxy-2-(2-nitrophenoxy)benzaldehyde as a yellow solid (690 mg, 80%): **<sup>1</sup>H NMR** (400 MHz, Chloroform-d)  $\delta$  10.39 (s, 1H), 8.02 (dd,  $J = 8.2, 1.6$  Hz, 1H), 7.55 (ddd,  $J = 8.4, 7.5, 1.7$  Hz, 1H), 7.44 (d,  $J = 3.2$  Hz, 1H), 7.30 - 7.23 (m, 1H), 7.17 (dd,  $J = 9.0, 3.2$  Hz, 1H), 7.04 - 6.93 (m, 2H), 4.11 (q,  $J = 7.0$  Hz, 2H), 1.46 (t,  $J = 7.0$  Hz, 3H); **LCMS** Rt 1.08 mins;  $m/z$  288.0 (M+H)

Step 2: To a solution of 5-ethoxy-2-(2-nitrophenoxy)benzaldehyde (690 mg, 2.402 mmol) in MeOH (2.40 ml) was added K<sub>2</sub>CO<sub>3</sub> (1095 mg, 7.93 mmol) and dimethyl (1-diazo-2-oxopropyl)phosphonate (923 mg, 4.80 mmol). The reaction was stirred at ambient temperature for 18 hr before the reaction was diluted with EtOAc, washed with water, brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated to give a solid 4-ethoxy-2-ethynyl-1-(2-nitrophenoxy)benzene (680 mg, 100%): **1H NMR** (400 MHz, Chloroform-d)  $\delta$  7.96 (dd, *J* = 8.1, 1.6 Hz, 1H), 7.47 (ddd, *J* = 8.9, 7.5, 1.6 Hz, 1H), 7.19 - 7.12 (m, 1H), 7.09 - 7.02 (m, 2H), 6.94 (dd, *J* = 9.0, 3.0 Hz, 1H), 6.88 - 6.82 (m, 1H), 4.05 (q, *J* = 7.0 Hz, 2H), 3.13 (s, 1H), 1.45 (t, *J* = 7.0 Hz, 3H); **LCMS** Rt 1.15 mins; *m/z* 284.1 (M+H)

Step 3: To a solution of 4-ethoxy-2-ethynyl-1-(2-nitrophenoxy)benzene (680 mg, 2.4 mmol) in EtOH (14.5 ml) and water (1.45 ml) was added iron (0.737 g, 13.20 mmol) and ammonium chloride (0.090 g, 1.680 mmol). The resulting mixture was heated at reflux at 90 °C for 2 hr before it was cooled, the crude reaction mixture was then filtered through Celite™ and rinsing the pad with MeOH. The combined organic solvent was removed *in vacuo* and the resulting residue dissolved in EtOAc, washed with sat. NaHCO<sub>3</sub> solution, the combined organic phases washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated to give a pale solid 2-(4-ethoxy-2-ethynylphenoxy)aniline in quantitative yield {**LCMS**: Rt 1.05 mins; 254.2 (M+H)}. The crude product was used without further purification.

Step 4: To 4-(acrylamidomethyl)benzoic acid (103 mg, 0.5 mmol), 2-(4-ethoxy-2-ethynylphenoxy)aniline (127 mg, 0.500 mmol) was added DMF (5 ml) giving a clear solution, then EDC (115 mg, 0.600 mmol) and DMAP (61.1 mg, 0.500 mmol) were added. The mixture was stirred at 60 °C for 2 hr when LC-MS indicated all the acid was consumed. The reaction mixture was diluted with EtOAc (50 mL), washed with water, brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated. Purification by flash silica gel chromatography eluting with (20-100% EtOAc / Heptane) gave 4-(acrylamidomethyl)-N-(2-(4-ethoxy-2-ethynylphenoxy)phenyl)benzamide as a white solid (75 mg, 33%): **1H NMR** (400 MHz, DMSO-d<sub>6</sub>)  $\delta$  9.69 (s, 1H), 8.70 (t, *J* = 6.0 Hz, 1H), 7.88 (d, *J* = 8.3 Hz, 2H), 7.80 (dd, *J* = 7.4, 2.1 Hz, 1H), 7.38 (d, *J* = 8.3 Hz, 2H), 7.19 - 7.09 (m, 2H), 7.05 (t, *J* = 1.7 Hz, 1H), 6.99 (d, *J* = 1.7 Hz, 2H), 6.83 - 6.69 (m, 1H), 6.29 (dd, *J* = 17.1, 10.1 Hz, 1H), 6.13 (dd, *J* = 17.1, 2.2 Hz, 1H), 5.64 (dd, *J* = 10.1, 2.2 Hz, 1H), 4.42 (d, *J* = 6.0 Hz, 2H), 4.26 (s, 1H), 4.01 (q, *J* = 7.0 Hz, 2H), 1.30 (t, *J* = 7.0 Hz, 3H) **13C NMR** (101 MHz, DMSO-d<sub>6</sub>)  $\delta$  160.24, 159.65, 150.16, 145.17, 141.45, 136.95, 128.93, 125.12, 123.83, 122.87, 122.37, 122.13, 118.76, 118.39, 115.80, 115.34, 113.70, 112.04, 110.82, 110.33, 77.10, 73.42, 58.87, 37.98, 9.54 **HRMS** calcd for C<sub>27</sub>H<sub>25</sub>N<sub>2</sub>O<sub>4</sub>(M+H)<sup>+</sup> 441.1814, found 441.1805

### **Preparation of 4-(acrylamidomethyl)-N-(2-(4-ethoxyphenoxy)-5-ethynylphenyl)benzamide**



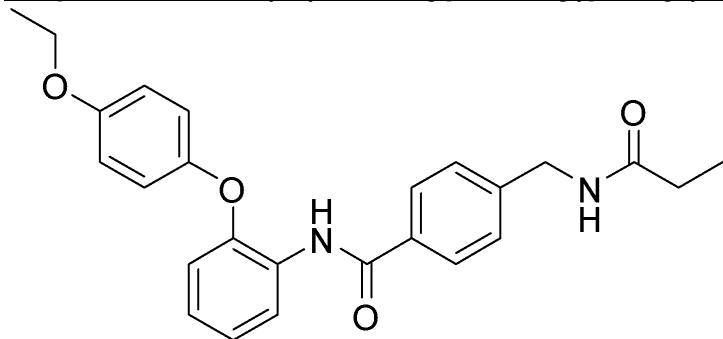
Step 1: A mixture of 4-ethoxyphenol (418 mg, 3.03 mmol), 4-ethynyl-1-fluoro-2-nitrobenzene (500 mg, 3.03 mmol),  $K_2CO_3$  (628 mg, 4.54 mmol) and DMA (3.02 ml) was stirred at 160 °C for 2 hours under  $N_2$  atmosphere. The reaction mixture was cooled to room temperature then diluted with water, and extracted with ethyl acetate (x3). The combined organic layers were washed with water and saturated brine, dried over anhydrous sodium sulfate, filtered and the filtrate evaporated *in vacuo*. The residue obtained was purified by flash column chromatography on silica gel (0-50% EtOAc / Heptane) to give 1-(4-ethoxyphenoxy)-4-ethynyl-2-nitrobenzene as a yellow solid (703 mg, 82%): **1H NMR** (400 MHz, Chloroform- $d$ )  $\delta$  8.06 (d,  $J$  = 2.0 Hz, 1H), 7.54 (dd,  $J$  = 8.7, 2.1 Hz, 1H), 7.06 - 7.02 (m, 2H), 6.97 - 6.91 (m, 2H), 6.86 (d,  $J$  = 8.7 Hz, 1H), 4.06 (q,  $J$  = 7.0 Hz, 2H), 3.13 (s, 1H), 1.45 (t,  $J$  = 7.0 Hz, 3H); **LCMS**  $m/z$  283.0

Step 2: To a solution of 1-(4-ethoxyphenoxy)-4-ethynyl-2-nitrobenzene (703 mg, 2.482 mmol) in EtOH (Volume: 14.100 mL, Ratio: 10) and Water (Volume: 1.410 mL, Ratio: 1) was added iron powder (762 mg, 13.65 mmol) and ammonium chloride (93 mg, 1.737 mmol). The resulting mixture was heated at reflux at 90 °C for 18 hr. The reaction mixture was filtered through Celite™ and the pad rinsed with MeOH. The filtrate was reduced *in vacuo* and the residue then partitioned between EtOAc/sat  $NaHCO_3$  solution. The combined organic layers dried over anhydrous magnesium sulphate and evaporated to dryness, to afford 2-(4-ethoxyphenoxy)-5-ethynylaniline as a light brown oil which was used in the next step without further purification (609 mg, 89%): **LCMS** Rt 1.09 mins;  $m/z$  254.2

Step 3: To 4-(acrylamidomethyl)benzoic acid (103 mg, 0.5 mmol), 2-(4-ethoxyphenoxy)-5-ethynylaniline (127 mg, 0.500 mmol) and DMF (5 ml) was added EDC (115 mg, 0.600 mmol) and DMAP (61.1 mg, 0.500 mmol). The reaction mixture was stirred at 60 °C for 2 hr. The reaction mixture was diluted with EtOAc (50 mL), washed with water, brine, dried over  $Na_2SO_4$ , filtered and concentrated to a crude product. Purification by flash silica gel chromatography eluting with (20-100% EtOAc / Heptane) gave the desired product 4-(acrylamidomethyl)-N-(2-(4-ethoxyphenoxy)-5-ethynylphenyl)benzamide as a white solid (77 mg, 33%): **1H NMR** (400 MHz, DMSO- $d_6$ )  $\delta$  9.82 (s, 1H), 8.70 (t,  $J$  = 6.0 Hz, 1H), 7.90 - 7.83 (m, 3H), 7.37 (d,  $J$  = 8.3 Hz, 2H), 7.27 (dd,  $J$  = 8.5, 2.1 Hz, 1H), 7.05 - 6.97 (m, 2H), 6.99 - 6.90 (m, 2H), 6.77 (d,  $J$  = 8.5 Hz, 1H), 6.29 (dd,  $J$  = 17.1, 10.1 Hz, 1H), 6.13 (dd,  $J$  = 17.1, 2.3 Hz, 1H), 5.64 (dd,  $J$  = 10.1, 2.2 Hz, 1H), 4.41 (d,  $J$  = 6.0 Hz, 2H), 4.15 (s, 1H), 3.99 (q,  $J$  = 6.9 Hz, 2H), 1.31 (t,  $J$  = 6.9 Hz, 3H) **13C NMR** (101 MHz, DMSO- $d_6$ )  $\delta$  164.10, 163.53, 154.05, 150.27, 147.49, 142.19, 131.56, 130.35, 128.50, 127.79, 127.33, 126.65, 126.03, 124.51, 119.66, 115.99, 114.49, 114.39, 81.77, 78.86, 62.22, 40.69, 13.50

**HRMS** calcd for C<sub>27</sub>H<sub>25</sub>N<sub>2</sub>O<sub>4</sub>(M+H)<sup>+</sup> 441.1814, found 441.1809

**Preparation of N-(2-(4-ethoxyphenoxy)phenyl)-4-(propionamidomethyl)benzamide**



Step 1: To a solution of 4-(((tert-butoxycarbonyl)amino)methyl)benzoic acid (13.06 g, 52.0 mmol) in toluene (200 ml) under inert atmosphere, was added anhydrous DMF (0.5 ml), pyridine (25.2 ml, 312 mmol) and oxalyl chloride (8.64 ml, 99 mmol), and the resulting mixture was stirred for 5 hr, with the formation of a precipitate being observed. This precipitate was filtered off and washed with toluene (x2). The combined filtrates were dried over Na<sub>2</sub>SO<sub>4</sub> and evaporated under reduced pressure, to obtain the corresponding tert-butyl (4-(chlorocarbonyl)benzyl)carbamate as a light pale solid. The crude product was used to next step without further purification.

Step 2: To tert-butyl (4-(chlorocarbonyl)benzyl)carbamate (1471 mg, 2.181 mmol), 2-(4-ethoxyphenoxy)aniline (500 mg, 2.181 mmol) in DCM (10.9 ml) was added TEA (0.912 mL, 6.54 mmol) and DMAP (133 mg, 1.090 mmol). The reaction mixture was stirred for 16 hr before being diluted with DCM, washed by sat. NaHCO<sub>3</sub> solution, brine, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure to give a crude product. The crude product was purified by flash column chromatography on silica gel eluting (0-40% EtOAc / Heptane) to give tert-butyl (4-((2-(4-ethoxyphenoxy)phenyl)carbamoyl)benzyl) carbamate as a white solid (690 mg, 69%): **<sup>1</sup>H NMR** (400 MHz, Chloroform-d) δ 8.65 - 8.57 (m, 2H), 7.84 (d, J = 8.2 Hz, 2H), 7.40 (d, J = 8.1 Hz, 2H), 7.14 (t, J = 7.8 Hz, 1H), 7.07 - 6.99 (m, 3H), 6.96 - 6.90 (m, 2H), 6.84 - 6.78 (m, 1H), 4.95 (s, 1H), 4.39 (d, J = 4.3 Hz, 2H), 4.05 (q, J = 7.0 Hz, 2H), 1.48 (s, 9H), 1.45 (t, J = 7.0 Hz, 3H); **LCMS** Rt 1.20 mins; m/z 463.2 (M+H)

Step 3: To a solution of tert-butyl (4-((2-(4-ethoxyphenoxy)phenyl)carbamoyl)benzyl) carbamate (690 mg, 1.492 mmol) in MeOH (7.46 ml) was added 4N HCl in 1,4-dioxane (7.46 ml, 29.8 mmol). The resulting solution was stirred at ambient temperature for 18 hr before the solvent was removed under reduced pressure to give a white solid, 4-(aminomethyl)-N-(2-(4-ethoxyphenoxy)phenyl)benzamide hydrochloride. The crude product was used to next step without further purification: **LCMS** Rt 0.81 mins; m/z 363.2 (M+H).

Step 4: To crude 4-(aminomethyl)-N-(2-(4-ethoxyphenoxy)phenyl)benzamide hydrochloride (120 mg, 0.3 mmol) in DCM (Volume: 1.5 ml) at 0 °C was added TEA (84 μl, 0.600 mmol) followed by propionyl chloride (26.2 μl, 0.300 mmol). The reaction

mixture was stirred for 16 hr at ambient temperature before it was diluted with EtOAc, washed with water, brine, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure to give the crude product. The crude product was purified by flash column chromatography on silica gel eluting (30-100% EtOAc/Heptane) to give N-(2-(4-ethoxyphenoxy)phenyl)-4-(propionamidomethyl)benzamide as a white solid (103 mg, 82%): **1H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.61 (dd, *J* = 8.3, 1.4 Hz, 2H), 7.83 (d, *J* = 8.2 Hz, 2H), 7.39 (d, *J* = 8.2 Hz, 2H), 7.19 – 7.10 (m, 1H), 7.08 – 6.99 (m, 3H), 6.96 – 6.89 (m, 2H), 6.81 (dd, *J* = 8.1, 1.3 Hz, 1H), 5.85 (s, 1H), 4.52 (d, *J* = 5.9 Hz, 2H), 4.05 (q, *J* = 7.0 Hz, 2H), 2.30 (q, *J* = 7.6 Hz, 2H), 1.45 (t, *J* = 7.0 Hz, 3H), 1.22 (t, *J* = 7.6 Hz, 3H).

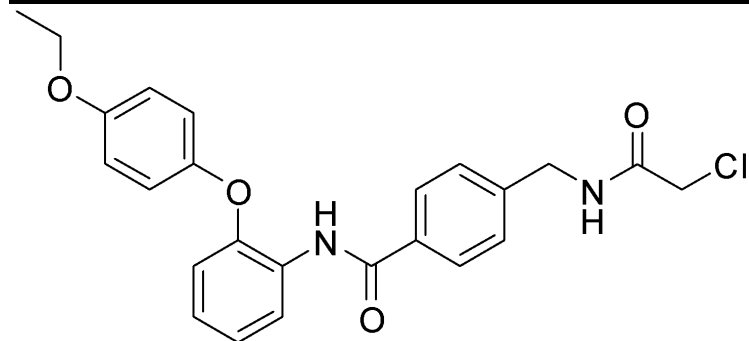
**13C NMR** (101 MHz, Chloroform-*d*)  $\delta$  173.74, 165.02, 155.75, 149.18, 147.10, 142.59, 134.20, 129.14, 128.06, 127.52, 124.09, 123.37, 120.54, 116.23, 115.66, 63.96, 43.13, 29.70, 14.89, 9.86

**HRMS** calcd for C<sub>25</sub>H<sub>27</sub>N<sub>2</sub>O<sub>4</sub>(M+H)<sup>+</sup> 419.1971, found 419.1973

The following compounds were prepared using a general procedure B (as used to prepare N-(2-(4-ethoxyphenoxy)phenyl)-4-(propionamidomethyl)benzamide: step 4) from tert-butyl (4-(chlorocarbonyl)benzyl)carbamate and the necessary aniline with subsequent deprotection.

**Procedure B:** To the corresponding crude 4-(aminomethyl)-benzamide hydrochloride (0.3 mmol) in DCM (1.5 ml) at 0 °C was added TEA (0.60 mmol) followed by either crotonyl chloride, acryloyl chloride or chloro acetyl chloride (0.300 mmol). The reaction mixture was stirred for 16 hr at ambient temperature before it was diluted with EtOAc, washed with water, brine, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure to give the crude product. The crude product was then purified by flash column chromatography on silica gel eluting (20-100% EtOAc / Heptane) to give the desired product

#### 4-((2-chloroacetamido)methyl)-N-(2-(4-ethoxyphenoxy)phenyl)benzamide



**73 mg, 56%**

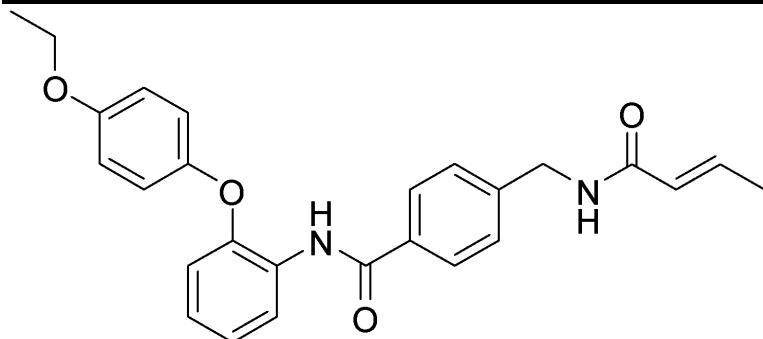
**1H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.61 (dd, *J* = 8.2, 1.3 Hz, 2H), 7.87 (d, *J* = 8.2 Hz, 2H), 7.42 (d, *J* = 8.1 Hz, 2H), 7.19 – 7.11 (m, 1H), 7.07 – 7.00 (m, 3H), 6.97 (s, 1H), 6.95 – 6.89 (m, 2H), 6.81 (dd, *J* = 8.2, 1.3 Hz, 1H), 4.59 (d, *J* = 6.0 Hz, 2H), 4.15 (s, 2H), 4.05 (q, *J* = 7.0 Hz, 2H), 1.45 (t, *J* = 7.0 Hz, 3H).



**13C NMR** (101 MHz, Chloroform-*d*)  $\delta$  166.02, 164.90, 155.77, 149.15, 147.12, 141.34, 134.55, 129.09, 128.05, 127.66, 124.13, 123.37, 120.57, 120.54, 116.20, 115.66, 63.96, 43.39, 42.62, 14.89

**HRMS** calcd for C<sub>24</sub>H<sub>24</sub>ClN<sub>2</sub>O<sub>4</sub>(M+H)<sup>+</sup> 439.1425, found 439.1421

**(E)-4-(but-2-enamidomethyl)-N-(2-(4-ethoxyphenoxy)phenyl)benzamide**



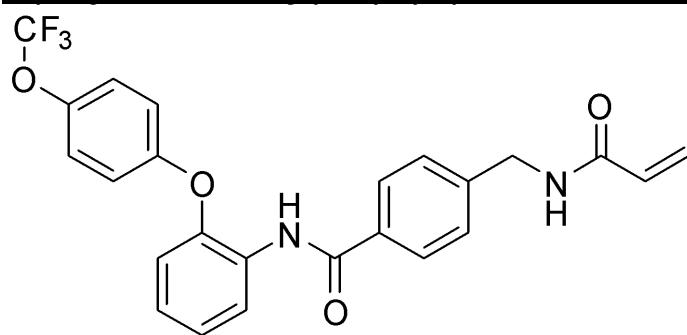
**66 mg, 46%**

**1H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.65 – 8.54 (m, 2H), 7.83 (d, *J* = 8.2 Hz, 2H), 7.41 (d, *J* = 8.3 Hz, 2H), 7.15 (td, *J* = 7.8, 1.4 Hz, 1H), 7.07 – 6.99 (m, 3H), 6.95 – 6.90 (m, 3H), 6.81 (dd, *J* = 8.2, 1.4 Hz, 1H), 5.85 (dd, *J* = 15.2, 1.5 Hz, 1H), 5.77 (s, 1H), 4.59 (d, *J* = 6.0 Hz, 2H), 4.05 (q, *J* = 7.0 Hz, 2H), 1.90 (dd, *J* = 7.0, 1.7 Hz, 3H), 1.45 (t, *J* = 7.0 Hz, 3H).

**13C NMR** 101 MHz, Chloroform-*d*)  $\delta$  165.91, 165.05, 155.73, 149.19, 147.08, 142.56, 140.99, 134.17, 129.16, 128.05, 127.50, 124.54, 124.08, 123.38, 120.53, 120.51, 116.26, 115.66, 63.96, 43.10, 17.84, 14.89

**HRMS** calcd for C<sub>26</sub>H<sub>27</sub>N<sub>2</sub>O<sub>4</sub>(M+H)<sup>+</sup> 431.1971, found 431.1968

**4-(acrylamidomethyl)-N-(2-(4-(trifluoromethoxy)phenoxy)phenyl)benzamide**



**86 mg, 57%**

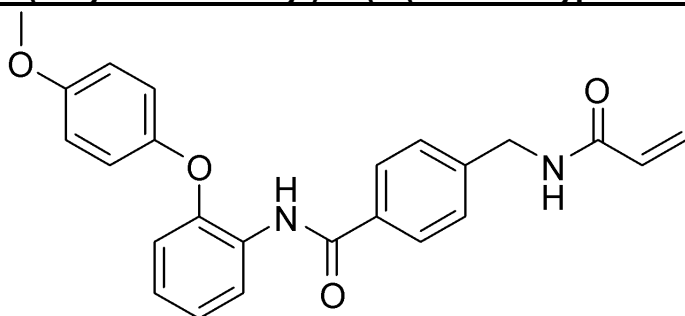
**1H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.61 (dd, *J* = 8.2, 1.5 Hz, 1H), 8.40 (s, 1H), 7.77 (d, *J* = 8.2 Hz, 2H), 7.41 (s, 1H), 7.39 (s, 1H), 7.23 (t, *J* = 7.7 Hz, 3H), 7.15 – 7.05 (m, 3H), 6.93 (dd, *J* = 8.1, 1.3 Hz, 1H), 6.37 (dd, *J* = 17.0, 1.3 Hz, 1H), 6.15 (dd, *J* = 17.0, 10.3 Hz, 1H), 5.98 (s, 1H), 5.72 (dd, *J* = 10.3, 1.3 Hz, 1H), 4.59 (d, *J* = 6.0 Hz, 2H)

**13C NMR** (101 MHz, Chloroform-*d*)  $\delta$  165.51, 165.05, 154.84, 145.50, 145.07 (d, *J* = 2.1 Hz), 142.41, 134.04, 130.31, 129.87, 128.13, 127.47, 127.44, 124.78, 124.47, 122.98, 121.17, 120.46 (q, *J* = 257.1 Hz), 119.51, 117.99, 43.20

**19F NMR** (376 MHz, Chloroform-*d*)  $\delta$  -58.26

**HRMS** calcd for C<sub>24</sub>H<sub>20</sub>F<sub>3</sub>N<sub>2</sub>O<sub>4</sub>(M+H)<sup>+</sup> 457.1375, found 457.1378

#### 4-(acrylamidomethyl)-N-(2-(4-methoxyphenoxy)phenyl)benzamide



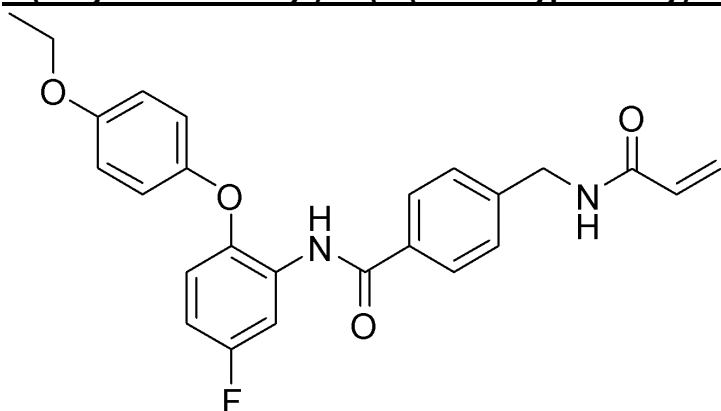
**105 mg, 73%**

**<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.66 – 8.56 (m, 2H), 7.84 (d, *J* = 8.2 Hz, 2H), 7.42 (d, *J* = 8.2 Hz, 2H), 7.19 – 7.10 (m, 1H), 7.08 – 7.00 (m, 3H), 6.97 – 6.89 (m, 2H), 6.81 (dd, *J* = 8.2, 1.2 Hz, 1H), 6.38 (dd, *J* = 16.9, 1.3 Hz, 1H), 6.16 (dd, *J* = 16.9, 10.3 Hz, 1H), 5.96 (s, 1H), 5.73 (dd, *J* = 10.3, 1.3 Hz, 1H), 4.61 (d, *J* = 6.0 Hz, 2H), 3.84 (s, 3H).

**<sup>13</sup>C NMR** (101 MHz, Chloroform-*d*)  $\delta$  165.49, 165.02, 156.37, 149.30, 147.09, 142.22, 134.27, 130.36, 129.13, 128.13, 127.54, 127.38, 124.12, 123.41, 120.56, 116.24, 115.07, 55.72, 43.23

**HRMS** calcd for C<sub>24</sub>H<sub>23</sub>N<sub>2</sub>O<sub>4</sub>(M+H)<sup>+</sup> 403.1658, found 403.1651.

#### 4-(acrylamidomethyl)-N-(2-(4-ethoxyphenoxy)-5-fluorophenyl)benzamide



**116 mg, 77%**

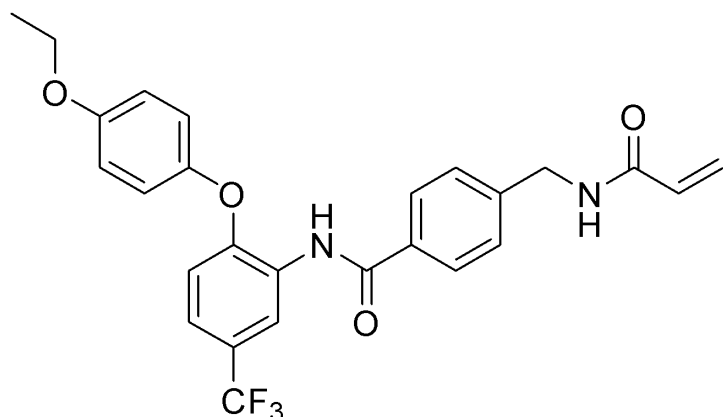
**<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.57 (s, 1H), 8.45 (dd, *J* = 10.6, 2.9 Hz, 1H), 7.80 (d, *J* = 8.3 Hz, 2H), 7.41 (d, *J* = 8.3 Hz, 2H), 7.01 – 6.95 (m, 2H), 6.93 – 6.88 (m, 2H), 6.80 (dd, *J* = 9.0, 5.2 Hz, 1H), 6.73 (ddd, *J* = 9.0, 7.7, 3.0 Hz, 1H), 6.37 (dd, *J* = 17.0, 1.3 Hz, 1H), 6.15 (dd, *J* = 17.0, 10.3 Hz, 1H), 5.95 (s, 1H), 5.72 (dd, *J* = 10.3, 1.4 Hz, 1H), 4.60 (d, *J* = 6.0 Hz, 2H), 4.04 (q, *J* = 7.0 Hz, 2H), 1.44 (t, *J* = 7.0 Hz, 3H)

**<sup>13</sup>C NMR** (101 MHz, Chloroform-*d*)  $\delta$  165.50, 164.98, 158.42 (d, *J* = 240.3 Hz), 155.69, 149.56, 142.68 (d, *J* = 2.8 Hz), 142.53, 133.80, 130.31, 130.20, 128.17, 127.54, 127.45, 119.88, 117.47 (d, *J* = 9.5 Hz), 115.72, 110.07 (d, *J* = 23.7 Hz), 107.95 (d, *J* = 29.9 Hz), 63.98, 43.20, 14.87

**<sup>19</sup>F NMR** (376 MHz, Chloroform-*d*)  $\delta$  -117.09

**HRMS** calcd for C<sub>25</sub>H<sub>24</sub>FN<sub>2</sub>O<sub>4</sub> (M+H)<sup>+</sup> 435.1720, found 435.1721

#### 4-(acrylamidomethyl)-N-(2-(4-ethoxyphenoxy)-5-(trifluoromethyl)phenyl)benzamide



**125 mg, 76%**

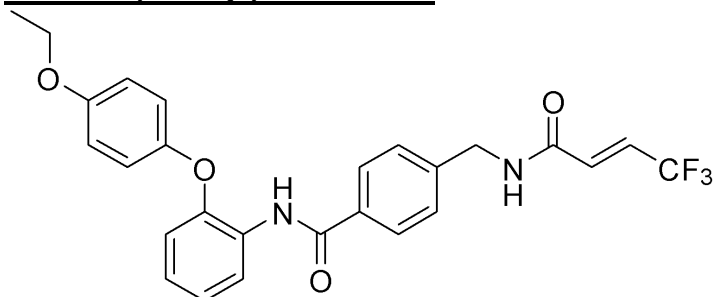
**<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.98 (d, *J* = 2.1 Hz, 1H), 8.69 (s, 1H), 7.88 (d, *J* = 8.3 Hz, 2H), 7.45 (d, *J* = 8.3 Hz, 2H), 7.31 – 7.23 (m, 1H), 7.08 – 7.02 (m, 2H), 6.99 – 6.93 (m, 2H), 6.81 (d, *J* = 8.5 Hz, 1H), 6.37 (dd, *J* = 17.0, 1.3 Hz, 1H), 6.16 (dd, *J* = 17.0, 10.3 Hz, 1H), 5.95 (s, 1H), 5.73 (dd, *J* = 10.3, 1.3 Hz, 1H), 4.62 (d, *J* = 6.0 Hz, 2H), 4.07 (q, *J* = 7.0 Hz, 2H), 1.46 (t, *J* = 7.0 Hz, 3H)

**<sup>13</sup>C NMR** (101 MHz, Chloroform-*d*)  $\delta$  165.53, 165.13, 156.49, 149.86, 147.78, 142.64, 133.74, 130.30, 128.83, 128.20, 127.58, 127.45, 125.00 (q, *J* = 32.7 Hz), 124.04 (q, *J* = 271.8 Hz), 121.44, 121.02 (q, *J* = 3.7 Hz), 117.57 (q, *J* = 3.8 Hz), 115.85, 114.75, 63.99, 43.21, 14.85

**<sup>19</sup>F NMR** (376 MHz, Chloroform-*d*)  $\delta$  -61.84

**HRMS** calcd for C<sub>26</sub>H<sub>24</sub>F<sub>3</sub>N<sub>2</sub>O<sub>4</sub> (M+H)<sup>+</sup> 485.1688, found 485.1684

**Preparation of (*E*)-N-(2-(4-ethoxyphenoxy)phenyl)-4-((4,4,4-trifluorobut-2-enamido)methyl) benzamide**



Step 1: (*E*)-4,4,4-trifluorobut-2-enoic acid (56.0 mg, 0.40 mmol), crude 4-(aminomethyl)-*N*-(2-(4-ethoxyphenoxy)phenyl)benzamide hydrochloride (160 mg, 0.40 mmol) and DIPEA (210  $\mu$ l, 1.20 mmol) were stirred in THF (2 ml) at 0 °C before EDC (115 mg, 0.600 mmol) was added. The reaction mixture was stirred for 16 hr at ambient temperature before the mixture was diluted with EtOAc, washed with water, brine, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure to give the crude product. The crude product was purified by flash column chromatography on silica gel eluting (0-80% EtOAc/heptane) to give (*E*)-*N*-(2-(4-ethoxyphenoxy)phenyl)-4-((4,4,4-trifluorobut-2-enamido)methyl) benzamide as a white solid (59 mg, 30%): **<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.61 – 8.54 (m, 2H), 7.80 (d, *J* = 8.2 Hz, 2H), 7.37 (d, *J* = 8.2 Hz, 2H), 7.18 – 7.11 (m, 1H), 7.07 – 6.97 (m, 3H), 6.95 – 6.78 (m, 4H), 6.56 (dd, *J* = 15.4,

1.9 Hz, 1H), 6.38 (s, 1H), 4.61 (d,  $J = 5.9$  Hz, 2H), 4.04 (q,  $J = 7.0$  Hz, 2H), 1.44 (t,  $J = 7.0$  Hz, 3H)

**<sup>13</sup>C NMR** (101 MHz, Chloroform-*d*)  $\delta$  165.05, 162.54, 155.79, 149.10, 147.17, 141.32, 134.47, 130.31 (q,  $J = 5.8$  Hz), 129.19 (q,  $J = 35.2$  Hz), 128.94, 128.17, 127.59, 124.29, 123.36, 122.41 (q,  $J = 270.0$  Hz), 120.55, 116.23, 115.68, 63.97, 43.58, 14.88

**<sup>19</sup>F NMR** (376 MHz, Chloroform-*d*)  $\delta$  -65.07

**HRMS** calcd for C<sub>26</sub>H<sub>24</sub>F<sub>3</sub>N<sub>2</sub>O<sub>4</sub> (M+H)<sup>+</sup> 485.1688, found 485.1690

## QUANTIFICATION AND STATISTICAL ANALYSIS

For covalent ligand screening and dose-response studies of covalent ligands in DNA binding and MYC transcriptional assays, EN4-alkyne MYC enrichment studies, thermal shift assays, proliferation and survival assays, tumor xenograft studies, and quantitation of Western blots, we quantified the data as described in the Methods section and statistical analysis was performed using a Student's unpaired two-tailed *t*-test.

For isoTOP-ABPP data analysis, data were extracted in the form of MS1 and MS2 files using Raw Extractor v.1.9.9.2 (Scripps Research Institute) and searched against the Uniprot human database using ProLuCID search methodology in IP2 v.3 (Integrated Proteomics Applications, Inc.)(Xu et al., 2015). Cysteine residues were searched with a static modification for carboxyamino-methylation (+57.02146) and up to two differential modifications for methionine oxidation and either the light or heavy TEV tags (+464.28596 or +470.29977, respectively). Peptides were required to be fully tryptic peptides and to contain the TEV modification. ProLUCID data were filtered through DTASelect to achieve a peptide false-positive rate below 5%. Only those probe-modified peptides that were evident across two out of three biological replicates were interpreted for their isotopic light to heavy ratios. For those probe-modified peptides that showed ratios greater than two, we only interpreted those targets that were present across all three biological replicates, were statistically significant and showed good quality MS1 peak shapes across all biological replicates. Light versus heavy isotopic probe-modified peptide ratios are calculated by taking the mean of the ratios of each replicate paired light versus heavy precursor abundance for all peptide-spectral matches associated with a peptide. The paired abundances were also used to calculate a paired sample *t*-test *P* value in an effort to estimate constancy in paired abundances and significance in change between treatment and control. *P* values were corrected using the Benjamini–Hochberg method.

For quantification and statistical analysis of the RNA sequencing data, we performed Contrast analysis to perform *t*-tests to generate *p*-values, Benjamini-Hochberg false discovery rates (FDRs) and log<sub>2</sub> fold change for each gene comparing treated versus control groups. We performed signature enrichment analysis of genes with an FDR < 0.1 and Fold change ≤ 0.75 (567 genes) using the MSigDB Collection H hallmark gene set.

## APPENDIX 2

### CHAPTER 3 METHODS

#### Production of authentic SARS-CoV-2 Main Proteases

The coding sequence for SARS-CoV-2 main protease was codon-optimized for *E. coli* and synthesized by Integrated DNA Technologies. The sequence was amplified by PCR and cloned into the pGEX6P-1 vector, downstream of GST and an HRV 3C protease cleavage site, using the Gibson Assembly Master Mix kit (New England BioLabs, Inc). To ensure authentic termini, the amino acids AVLQ were added to the N-terminus of the main protease by addition of their coding sequence to the 5' end of the gene product. This sequence reconstitutes the NSP4/5 cleavage site, resulting in auto-cleavage by the main protease protein product and removal of the GST tag. We also added a GP-6xHis tag (to enable IMAC purification) on the C-terminus (the GP completes a non-consensus 3C cleavage site along with the C-terminus of the main protease which allows for cleavage of the his tag after purification, resulting in an authentic C-terminus).

Hi Control BL21(DE3) cells were transformed with the expression plasmid using standard techniques. We used Hi Control cells as we observed expression of the main protease was toxic in other standard *E. coli* cell lines. A single colony was used to start an overnight culture in LB + carbenicillin media. This culture was used to inoculate 2 x 1 L cultures in Terrific Broth, supplemented with 50 mM sodium phosphate pH 7.0 and 100 µg/mL carbenicillin. These cultures grew in Fernbach flasks at 37 °C while shaking at 225 rpm, until the OD600 reached approximately 2.0, at which point the temperature was reduced to 20 °C and 0.5 mM IPTG (final) was added to each culture. The cells were allowed to grow overnight.

The next day, the cultures were centrifuged at 6,000 x *g* for 20 minutes at 4 °C, and the resulting cell pellets were resuspended in IMAC\_A buffer (50 mM Tris pH 8.0, 400 mM NaCl, 1 mM TCEP). Cells were lysed with two passes through a cell homogenizer (Microfluidics model M-110P) at 18,000 psi. The lysate was clarified with centrifugation at 42,000 x *g* for 30 minutes and the cleared lysate was loaded onto 3 x 5 mL HiTrap Ni-NTA columns (GE) pre-equilibrated with IMAC\_A buffer, using an AKTA Pure FPLC. After loading, the columns were washed with IMAC\_A buffer until the A280 levels reached a sustained baseline. The protein was then eluted with a linear gradient with IMAC\_B buffer (50 mM Tris pH 8.0, 400 mM NaCl, 500 mM imidazole, 1 mM TCEP) across 25 column volumes, while 2 mL fractions were collected automatically. Peak fractions were analyzed by SDS-PAGE and those containing SARS-CoV-2 main protease were pooled. Importantly, auto-cleavage of the N-terminal GST tag was observed and the eluted protein had a mass consistent with SARS-CoV-2 main protease along with the C-terminal GP-6xHis tag, as determined by ESI-LC/MS.

Pooled fractions were treated with HRV 3C protease (also known as “PreScission” protease) while dialyzing against IMAC\_A buffer at room temperature (2 x 2 L dialyses). Room temperature dialysis was important as we observed a tendency for

the main protease protein to precipitate with prolonged exposure to 4 °C. Cleavage of the C-terminal GP-6xHis tag was confirmed after 2 hours by ESI-LC/MS. The dialyzed and cleaved protein was then re-run through a 5 mL HiTrap Ni-NTA column pre-equilibrated with IMAC\_A buffer. The main protease eluted in the flow-through as expected.

The protein was then concentrated to approximately 5 mL and loaded onto a Superdex 75 16/60 column pre-equilibrated with SEC Buffer (25 mM HEPES pH 7.5, 150 mM NaCl, 1 mM TCEP). The protein was run through the column at 1 mL/min and eluted as one large peak well in the included volume (at ~75 mL). Fractions from this peak were analyzed by SDS-PAGE and pure fractions were pooled and concentrated to 10 mg/mL, aliquotted, and stored at -80 °C. Final yield was typically in the realm of 60-70 mg/L of culture.

### **Gel-Based ABPP Screens**

Recombinant Mpro (100 nM) was pre-treated with either DMSO vehicle or covalent ligand at room temperature for 30 min in 25 µL of PBS, and subsequently treated with Tetramethylrhodamine-5-iodoacetamide dihydroiodide (IA-rhodamine) (500 nM) (ThermoFisher Scientific) at room temperature for 1 h. The reaction was stopped by addition of 4×reducing Laemmli SDS sample loading buffer (Alfa Aesar). After boiling at 95 °C for 5 min, the samples were separated on precast 4–20% Criterion TGX gels (Bio-Rad). Probe-labeled proteins were analyzed by in-gel fluorescence using a ChemiDoc MP (Bio-Rad).

### **SARS CoV2 MPro Activity Assay using a Fluorescent Substrate Peptide Probe (FRET-based assay)**

Compounds were made up in DMSO to 50X the desired screening concentration. DMSO was used as a solvent control. MPro protein was diluted in assay buffer (Tris buffered saline with 1 mM EDTA) to a concentration of 115 nM and was aliquoted to each well of a 96-well plate. Each well was treated with compound or vehicle and the plate was incubated for 30 min at room temperature. During the compound incubation the quenched fluorescent peptide probe (7-methoxycoumarin-4-ylacetyl) MCA-ABLQSGFR-Lys(2,4,-dinitrophenyl (Dnp))-Lys-NH was added from a 80 mM stock solution to a final concentration of 10 mM. Values were read-out on a Tecan Spark plate-reader.

### **SARS CoV2 MPro Activity Assay using a Fluorescent Substrate Peptide Probe (Rhodamine-based assay)**

Compounds were made up in DMSO to 50X the desired screening concentration. DMSO was used as a solvent control. MPro protein was diluted in assay buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM EDTA, 0,01% pluronic acid F127) to a concentration of 30 nM and 24.5 µL of diluted protein was aliquoted to each well of a black 384 well plate (Corning 384-Well, Flat-Bottom Microplate). Each well was treated

with 0.5  $\mu\text{L}$  of compound or vehicle and the plate was incubated for 30 min at room temperature. During the compound incubation the peptide probe KTSAVLQ-(Rhodamine-110 (Rh-110))-gammaGlu (Biosynton) was diluted from 5mM DMSO stock into assay buffer. After pre-incubation 5  $\mu\text{L}$  of 75  $\mu\text{M}$  Rh-110 probe was added to each well. RFU value was immediately measured on a Tecan Spark plate reader with an excitation wavelength of 488 nm and an emission wavelength of 535 nm at 30 °C for 30 min.

### **MS-Based ABPP Study**

We performed MS-based ABPP studies as previously described.<sup>148</sup>

### **General Synthetic Methods**

All non-aqueous reactions were performed under an inert atmosphere of dry nitrogen in flame dried glassware sealed with a rubber septum unless stated otherwise. Nitrogen was supplied through a glass manifold. Reactions were stirred magnetically and monitored by thin layer chromatography (TLC). Analytical thin layer chromatography was performed using MERCK Silica Gel 60 F254 TLC glass plates and visualized by ultraviolet light (UV). Additionally, TLC plates were stained with aqueous potassium permanganate ( $\text{KMnO}_4$ ) [1.5 g  $\text{KMnO}_4$ , 200 mL  $\text{H}_2\text{O}$ , 10 g  $\text{K}_2\text{CO}_3$ , 1.25 mL 10%  $\text{NaOH}$ ]. Concentration under reduced pressure was performed by rotator evaporation at 40 °C at the appropriate pressure. Chromatographic purification was performed as flash chromatography on MERCK silica gel 60 Å (230 x 400 mesh) at 0.2–0.5 bar overpressure. Purified compounds were dried further under high vacuum (0.01-0.1 mbar). Yields refer to the purified compound.

### **Chemicals**

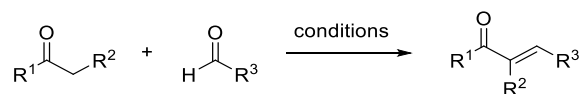
All chemicals and solvents were used as received from the commercial supplier without further purification unless mentioned otherwise.  $\text{CH}_2\text{Cl}_2$  was purified by passage through an activated alumina column under an atmosphere of dry argon.

### **Analytics**

Nuclear Magnetic Resonance (NMR) spectra were recorded on BRUKER AV (600 MHz and 300 MHz), AVB (400 MHz), AVQ (400 MHz) and NEO (500 MHz) spectrometers. Measurements were carried out at ambient temperature. Chemical shifts ( $\delta$ ) are reported in ppm with the residual solvent signal as internal standard (chloroform at 7.26 and 77.00 ppm for  $^1\text{H}$  NMR and  $^{13}\text{C}$  NMR spectroscopy, respectively). The data is reported as (s = singlet, d = doublet, t = triplet, q = quartet, p = quintet, m = multiplet or unresolved, br = broad signal, coupling constant(s) in Hz, integration).  $^{13}\text{C}$  NMR spectra were recorded with broadband  $^1\text{H}$  decoupling. Mass spectrometry (MS) analyses were obtained at the Catalysis Center at the College of Chemistry, University of California, Berkeley.

## General Procedures

### Synthesis of $\alpha,\beta$ -Unsaturated Ketones



#### General Procedure A

The ketone (1.0 equiv) and aldehyde (2.0 equiv) were taken up in EtOH (0.6 M) and cooled to 0 °C. 40% NaOH (10 equiv) was added dropwise, and the resulting reaction mixture was stirred at ambient temperature for 4 h. The precipitated solid was filtered, washed with cold water and cold EtOH and dried under high vacuum. Where necessary, the crude product was recrystallized from hot EtOH to afford the corresponding  $\alpha,\beta$ -unsaturated ketone.

#### General Procedure B

LiOH·H<sub>2</sub>O (10 mol%) was added to a solution of the ketone (1.0 equiv) in EtOH (1 M) and the resulting mixture was stirred for 10 min at ambient temperature. The aldehyde (1.0 equiv) was then added, and the reaction mixture was stirred at ambient temperature for 30 min to 1 h before it was concentrated under reduced pressure. Purification by column chromatography afforded the corresponding  $\alpha,\beta$ -unsaturated ketone.

#### General Procedure C

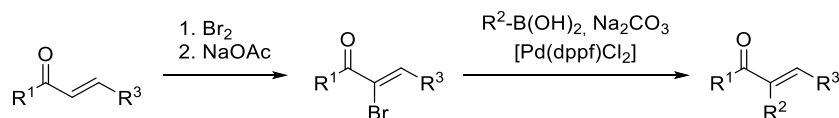
A mixture of the aldehyde (1.0 equiv) and deoxybenzoin (3.0 equiv) in toluene (0.4 M) was treated with acetic acid (0.9 equiv) and piperidine (0.2 equiv). Powdered 4 Å molecular sieves (500 mg/mmol) was added, and the resulting reaction mixture was stirred at reflux temperature for 7 to 16 h. It was then allowed to cool to ambient temperature, quenched with NaHCO<sub>3</sub> solution (sat. aqueous) and extracted with EtOAc (3 x). The combined organic phases were washed with NaCl solution (sat. aqueous), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. Purification by column chromatography afforded the corresponding  $\alpha,\beta$ -unsaturated ketone.

#### General Procedure D

A mixture of the aldehyde (1.0 equiv) and propiophenone (1.0 equiv) in EtOH (0.3 M) was heated gently until both starting materials dissolved. A solution of NaOH (1.2 equiv) in EtOH/H<sub>2</sub>O (1:1 v/v, 0.3 M) was added dropwise, and the resulting reaction mixture was stirred at ambient temperature for 1 h, then at 60 °C for 16 h. The reaction mixture was allowed to cool to ambient temperature, diluted with water and extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x). The combined organic phases were washed with NaCl solution (sat. aqueous), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. Purification by column chromatography afforded the corresponding  $\alpha,\beta$ -unsaturated ketone.



## General Procedure E

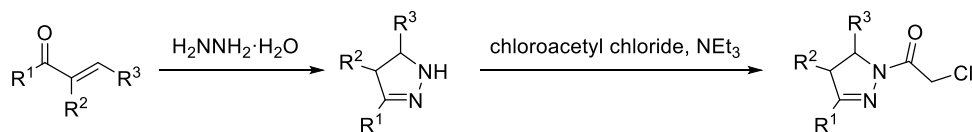


A solution of the chalcone (1.0 equiv) in  $\text{CH}_2\text{Cl}_2$  (0.1 M) was cooled to  $0\text{ }^\circ\text{C}$ . A solution of bromine (1.2 equiv) in  $\text{CH}_2\text{Cl}_2$  (1 M) was added dropwise over 3 min, and the resulting reaction mixture was stirred at  $0\text{ }^\circ\text{C}$  for 15 min, then at ambient temperature for 15 min. It was then quenched with  $\text{Na}_2\text{S}_2\text{O}_3$  solution (sat. aqueous) and the resulting mixture was extracted with  $\text{CH}_2\text{Cl}_2$  (3 x). The combined organic phases were washed with  $\text{NaCl}$  solution (sat. aqueous), dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated under reduced pressure.

The crude dibromide was dissolved in  $\text{EtOH}$  (0.3 M) and sodium acetate (1.2 equiv) was added. The resulting reaction mixture was stirred at reflux temperature for 1.5 h before the solvent was removed under reduced pressure. The residue was partitioned between  $\text{CH}_2\text{Cl}_2$  and  $\text{NaHCO}_3$  solution (sat. aqueous), the phases were separated and the aqueous phase was extracted with  $\text{CH}_2\text{Cl}_2$  (2 x). The combined organic phases were washed with  $\text{NaCl}$  solution (sat. aqueous), dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated under reduced pressure.

A mixture of the crude vinyl bromide (1.0 equiv), the boronic acid (1.3 equiv), sodium carbonate (2.0 equiv) and  $\text{Pd(dppf)Cl}_2$  (5 mol%) was taken up in  $\text{toluene}/\text{H}_2\text{O}$  (4:1 v/v, 0.2 M). The resulting suspension was sparged with nitrogen for 10 min. The flask was sealed and the reaction mixture was stirred at  $100\text{ }^\circ\text{C}$  for 1 h. It was then allowed to cool to ambient temperature and partitioned between  $\text{CH}_2\text{Cl}_2$  and 1 M  $\text{NaOH}$ . The phases were separated and the aqueous phase was extracted with  $\text{CH}_2\text{Cl}_2$  (2 x). The combined organic phases were washed with  $\text{NaCl}$  solution (sat. aqueous), dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated under reduced pressure. Purification by column chromatography afforded the corresponding  $\alpha,\beta$ -unsaturated ketone.

## Synthesis of Pyrazoline Chloroacetamides

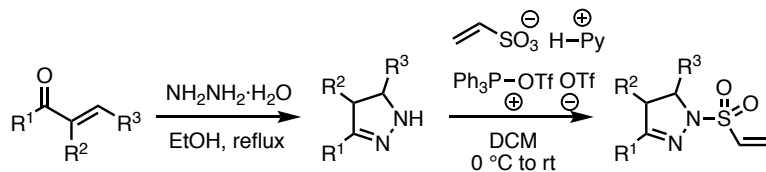


## General Procedure F

Hydrazine monohydrate (2.0 equiv) was added to a suspension of the  $\alpha,\beta$ -unsaturated ketone (1.0 equiv) in  $\text{EtOH}$  (0.3 M). The resulting reaction mixture was stirred at reflux temperature for 2.5 to 4 h before it was concentrated under reduced pressure. The crude pyrazoline was then dissolved in  $\text{CH}_2\text{Cl}_2$  (0.2 M) and cooled to  $0\text{ }^\circ\text{C}$ . Triethylamine (3.0 equiv) was added dropwise, followed by chloroacetyl chloride (1.5 equiv). The resulting reaction mixture was stirred at ambient temperature for 30 min before it was diluted with  $\text{CH}_2\text{Cl}_2$ . The organic phase was sequentially washed with

NaHCO<sub>3</sub> solution (sat. aqueous) and NaCl solution (sat. aqueous), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. Purification by column chromatography afforded the corresponding chloroacetamide.

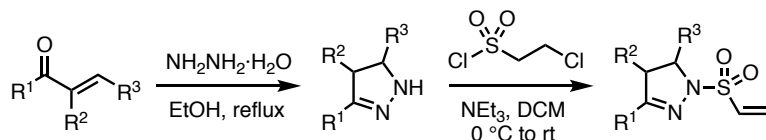
### Synthesis of Pyrazoline Vinyl Sulfonamides



### General Procedure G

**Step 1:** Hydrazine monohydrate (2.0 equiv) was added to a suspension of the  $\alpha,\beta$ -unsaturated ketone (1.0 equiv) in EtOH (0.3 M) at rt. The reaction mixture was then stirred at reflux for 2h, then allowed to cool to rt and concentrated under reduced pressure to give the crude pyrazoline. This was used in the next step without further purification.

**Step 2:** A solution of triphenylphosphine oxide (2.2 equiv) in dichloromethane (0.25 M) was degassed at rt. Trifluoromethanesulfonic anhydride (1.0 equiv) was added at rt and the resultant mixture was stirred at this temperature for 15 min. In a separate flask, pyridine (1.0 equiv) was added to a solution of vinyl sulfonic acid (1.0 equiv) in dichloromethane (0.15 M) at rt and the mixture was then concentrated *in vacuo* to give a white solid. A solution of the vinyl sulfonic acid-pyridinium salt in dichloromethane (0.30 M) was added to the reaction mixture at rt and the resultant mixture was stirred at this temperature for 30 min. A solution of the crude pyrazoline (1.0 equiv) and triethylamine (2.0 equiv) in dichloromethane (0.20 M) was then added to the reaction mixture at 0 °C and the resultant mixture was allowed to warm to rt and stirred at rt for 16 h. The mixture was then diluted with dichloromethane and washed sequentially with 2.0 M aq HCl, sat aq NaHCO<sub>3</sub> and brine, then dried and concentrated under reduced pressure. Purification *via* flash column chromatography using both normal-phase (gradient elution, EtOAc in hexane) and reverse-phase (gradient elution, acetonitrile in water) conditions afforded the vinyl sulfonamides.



### General Procedure H

Hydrazine monohydrate (2.0 equiv) was added to a suspension of the  $\alpha,\beta$ -unsaturated ketone (1.0 equiv) in EtOH (0.3 M). The resulting reaction mixture was stirred at reflux temperature for 1 to 4 h before it was concentrated under reduced pressure. The crude pyrazoline was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (0.2 M) and cooled to 0 °C. Triethylamine (3.0 equiv) was added dropwise, followed by 2-chloroethanesulfonyl

chloride (1.2–1.5 equiv). The resulting reaction mixture was stirred at room temperature 1 h before it was diluted with CH<sub>2</sub>Cl<sub>2</sub>. The organic phase was sequentially washed with 1M HCl, sat. aq. NaHCO<sub>3</sub> and brine, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. Purification by column chromatography or preparative TLC afforded the corresponding vinyl sulfonamide.

## **Stereoselective Synthesis of Pyrazoline Chloroacetamides**

### **General Procedure I**

(E)-1-phenyl-3-(quinoxalin-5-yl)prop-2-en-1-one (1.0 eq.), tert-butyl carbazate (1.1 eq.), potassium phosphate (1.3 eq.), indicated catalyst (10 mol%) were combined in THF (0.5 M) under nitrogen and stirred at 0 °C for 18h. The mixture was diluted with EtOAc, filtered to remove salts, concentrated and purified by silica gel chromatography to provide the stereo-enriched pyrazoline.

### **General Procedure J**

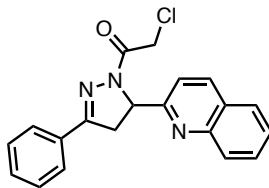
Boc-protected pyrazoline (1.0 eq) was dissolved in DCM (0.2 M), cooled to 0 °C, and 4.0 M HCl in dioxane (10.0 eq.) was added under nitrogen. The mixture was stirred at room temperature for several hours (2-3) until starting material was consumed, then additional DCM was added (reducing HCl to 0.1 M). At 0 °C, chloroacetyl chloride (3.0 eq.) was added followed by TEA (13.0 eq.), and the reaction allowed to warm to room temperature and stirred for 1h. Water was added and the mixture extracted with DCM. Extracts were combined, washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated under vacuum and purified by silica gel chromatography.

## **Formylation of Aryl Bromides**

### **General Procedure K**

Aryl bromide (1.0 Eq.) was dissolved in THF (to 0.2 M) and cooled to -78 °C. n-BuLi (2.5 M in hexanes) was added dropwise, the reaction turned red-brown, and was stirred for 20 minutes at -78 °C. Then, DMF (2.0 Eq.) was added dropwise. The reaction was removed from the -78 °C bath and allowed to stir for 2h at room temperature. To quench the reaction the flask was cooled to -78 °C quickly, sat. NH<sub>4</sub>Cl (2 mL) was added dropwise, and the reaction allowed to warm to rt again. Water was added and the mixture was extracted with ethyl acetate. Combined organic extracts were washed with brine, dried over sodium sulfate, concentrated, and purified by silica gel chromatography to provide the resulting aldehyde.

## **Characterization Data**



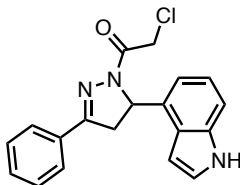
HW-01-146

**2-Chloro-1-(3-phenyl-5-(quinolin-2-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one.**

General Procedure B was followed starting from acetophenone (0.30 mL, 2.6 mmol) and quinoline-2-carbaldehyde (445 mg, 2.83 mmol). Purification by column chromatography (EtOAc/hexane, 15:85) afforded the corresponding chalcone (326 mg) as a white solid.

General Procedure F was followed starting from the above product (0.20 g, 0.77 mmol). Purification by column chromatography (EtOAc/hexane, 30:70) afforded the corresponding chloroacetamide (158 mg, 29% over three steps) as a white solid.

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>): δ 8.15 (d, *J* = 8.5 Hz, 1H), 8.01 (dd, *J* = 8.5, 1.0 Hz, 1H), 7.80 (ddd, *J* = 6.9, 4.0, 1.6 Hz, 3H), 7.68 (ddd, *J* = 8.4, 6.9, 1.5 Hz, 1H), 7.55–7.49 (m, 2H), 7.49–7.43 (m, 3H), 5.88 (dd, *J* = 10.9, 6.1 Hz, 1H), 4.64 (d, *J* = 2.1 Hz, 2H), 3.89 – 3.71 (m, 2H); **<sup>13</sup>C NMR** (151 MHz, CD<sub>2</sub>Cl<sub>2</sub>) δ 164.5, 159.6, 156.5, 148.3, 137.4, 131.3, 131.1, 130.1, 129.5, 129.2, 128.0, 128.0, 127.3, 127.0, 119.9, 62.7, 42.9, 40.6. **HRMS** (ESI): exact mass calculated for C<sub>20</sub>H<sub>17</sub>ClN<sub>3</sub>O [(M+H)<sup>+</sup>] 350.1055, found 350.1048.



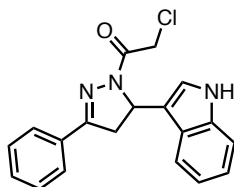
HW-01-151

**1-(5-(1H-Indol-4-yl)-3-phenyl-4,5-dihydro-1H-pyrazol-1-yl)-2-chloroethan-1-one.**

General Procedure B was followed starting from acetophenone (0.30 mL, 2.6 mmol) and indole-4-carbaldehyde (373 mg, 2.57 mmol). Purification by column chromatography (EtOAc/hexane, 30:70) afforded the corresponding chalcone (375 mg) as a yellow solid.

General Procedure F was followed starting from the above product (0.20 g, 0.81 mmol). Purification by column chromatography (EtOAc/hexane, 30:70) afforded the corresponding chloroacetamide (100 mg, 20% over three steps) as a yellowish solid.

**<sup>1</sup>H NMR** (600 MHz, CDCl<sub>3</sub>): δ 8.27 (s, br, 1H), 7.83–7.71 (m, 2H), 7.52–7.39 (m, 3H), 7.32 (dt, *J* = 8.2, 1.0 Hz, 1H), 7.19 (dd, *J* = 3.3, 2.5 Hz, 1H), 7.14 (t, *J* = 7.7 Hz, 1H), 7.02 (d, *J* = 7.2 Hz, 1H), 6.43 (ddd, *J* = 3.2, 1.9, 0.9 Hz, 1H), 5.94 (dd, *J* = 11.9, 5.4 Hz, 1H), 4.74–4.51 (m, 2H), 3.85 (dd, *J* = 17.8, 12.0 Hz, 1H), 3.38 (dd, *J* = 17.8, 5.4 Hz, 1H); **<sup>13</sup>C NMR** (151 MHz, CDCl<sub>3</sub>): δ 164.0, 155.8, 136.4, 132.0, 131.0, 130.7, 128.8, 126.8, 124.4, 124.3, 122.1, 117.4, 111.0, 100.2, 59.9, 42.3, 41.5; **HRMS** (ESI): exact mass calculated for C<sub>19</sub>H<sub>17</sub>ClN<sub>3</sub>O [(M+H)<sup>+</sup>] 338.1055, found 338.1026.



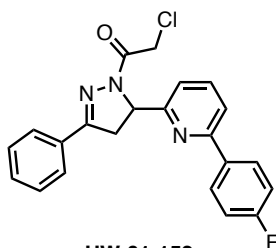
HW-01-161

**1-(5-(1*H*-Indol-3-yl)-3-phenyl-4,5-dihydro-1*H*-pyrazol-1-yl)-2-chloroethan-1-one.**

General Procedure B was followed starting from acetophenone (0.10 mL, 0.86 mmol) and indole-3-carbaldehyde (124 mg, 0.857 mmol). The stirring time was prolonged to 22 h at ambient temperature followed by 1 h at 40 °C. Purification by column chromatography (EtOAc/hexane, 35:65) afforded the corresponding chalcone (22 mg) as a yellow solid.

General Procedure F was followed starting from the above product (22 mg, 89 μmol). Purification by column chromatography (EtOAc/hexane, 40:60) afforded the corresponding chloroacetamide (19 mg, 6% over three steps) as a white solid.

**<sup>1</sup>H NMR** (600 MHz, CDCl<sub>3</sub>): δ 8.48 (s, br, 1H), 7.84–7.77 (m, 2H), 7.53–7.43 (m, 3H), 7.34 (ddd, *J* = 8.9, 8.0, 1.0 Hz, 2H), 7.20–7.07 (m, 2H), 7.01 (ddd, *J* = 8.0, 7.0, 1.0 Hz, 1H), 5.92 (dd, *J* = 11.8, 4.8 Hz, 1H), 4.63 (d, *J* = 13.8 Hz, 1H), 4.57 (d, *J* = 13.8 Hz, 1H), 3.76 (dd, *J* = 17.8, 11.8 Hz, 1H), 3.48 (dd, *J* = 17.8, 4.8 Hz, 1H); **<sup>13</sup>C NMR** (151 MHz, CDCl<sub>3</sub>): δ 164.1, 156.1, 136.9, 131.0, 130.8, 128.9, 126.9, 124.2, 123.5, 122.2, 119.8, 118.5, 114.6, 111.8, 54.6, 42.5, 40.4; **HRMS** (ESI): exact mass calculated for C<sub>19</sub>H<sub>17</sub>ClN<sub>3</sub>O [(M+H)<sup>+</sup>] 338.1055, found 338.1032.



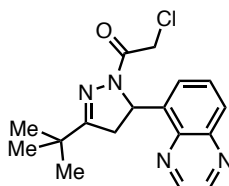
HW-01-152

**2-Chloro-1-(5-(6-(4-fluorophenyl)pyridin-2-yl)-3-phenyl-4,5-dihydro-1*H*-pyrazol-1-yl)-ethan-1-one.**

General Procedure B was followed starting from acetophenone (50 μL, 0.43 mmol) and 6-(4-fluorophenyl)picolinaldehyde (86 mg, 0.43 mmol). Purification by column chromatography (EtOAc/hexane, 15:85) afforded the corresponding chalcone (108 mg) as a white solid.

General Procedure F was followed starting from the above product (60 mg, 0.20 mmol). Purification by column chromatography (EtOAc/hexane, 30:70) afforded the corresponding chloroacetamide (57 mg, 61% over three steps) as a white solid.

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>): δ 7.98–7.88 (m, 2H), 7.84–7.76 (m, 2H), 7.72 (t, *J* = 7.7 Hz, 1H), 7.59 (d, *J* = 7.8 Hz, 1H), 7.51–7.40 (m, 3H), 7.32 (d, *J* = 7.6 Hz, 1H), 7.08 (t, *J* = 8.7 Hz, 2H), 5.76 (dd, *J* = 10.9, 5.5 Hz, 1H), 4.63 (d, *J* = 1.1 Hz, 2H), 3.84–3.63 (m, 2H); **<sup>13</sup>C NMR** (151 MHz, CD<sub>2</sub>Cl<sub>2</sub>) δ 164.4, 164.0 (d, *J* = 247.8 Hz), 158.8, 156.6, 156.4, 138.1, 135.6 (d, *J* = 3.2 Hz), 131.5, 131.1, 129.2, 129.1 (d, *J* = 8.4 Hz), 127.3, 120.5, 119.5, 115.8 (d, *J* = 21.6 Hz), 62.1, 42.9, 40.7. **HRMS** (ESI): exact mass calculated for C<sub>22</sub>H<sub>18</sub>ClFN<sub>3</sub>O [(M+H)<sup>+</sup>] 394.1117, found 394.1099.

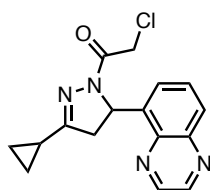


HW-1-174

**1-(3-(tert-Butyl)-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)-2-chloroethan-1-one.** General Procedure A was followed starting from pinacolone (20  $\mu$ L, 0.16 mmol) and quinoxaline-5-carbaldehyde (50 mg, 0.32 mmol). The crude  $\alpha,\beta$ -unsaturated ketone (27 mg) was used in the following transformation without further purification.

General Procedure F was followed starting from the above product (26 mg, 0.11 mmol). Purification by column chromatography (EtOAc/hexane, 40:60) afforded the corresponding chloroacetamide (18 mg, 35% over three steps) as a white solid.

**$^1\text{H NMR}$**  (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.88 (d,  $J = 1.8$  Hz, 1H), 8.83 (d,  $J = 1.8$  Hz, 1H), 8.04 (d,  $J = 8.6$  Hz, 1H), 7.81–7.66 (m, 1H), 7.48 (d,  $J = 7.1$  Hz, 1H), 6.46 (dd,  $J = 11.7, 4.9$  Hz, 1H), 4.69–4.43 (m, 2H), 3.63 (dd,  $J = 18.1, 11.7$  Hz, 1H), 2.79 (dd,  $J = 18.1, 5.0$  Hz, 1H), 1.20 (s, 9H);  **$^{13}\text{C NMR}$**  (151 MHz,  $\text{CD}_2\text{Cl}_2$ )  $\delta$  168.4, 164.1, 145.6, 144.4, 143.9, 140.5, 139.5, 130.2, 129.5, 125.9, 56.91, 42.8, 42.0, 34.5, 28.1. **HRMS** (ESI): exact mass calculated for  $\text{C}_{17}\text{H}_{20}\text{ClN}_4\text{O}$  [(M+H) $^+$ ] 331.1320, found 331.1343.

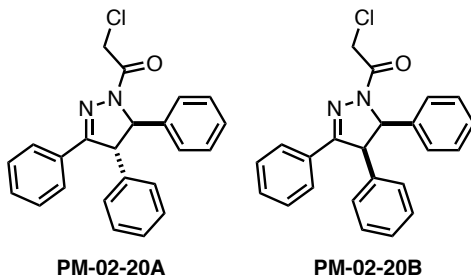


HW-01-185

**2-Chloro-1-(3-cyclopropyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one.** General Procedure A was followed starting from cyclopropyl methyl ketone (20  $\mu$ L, 0.20 mmol) and quinoxaline-5-carbaldehyde (64 mg, 0.40 mmol). Upon completion of the reaction, the reaction mixture was diluted with water (10 mL) and extracted with diethyl ether (3 x 20 mL). The combined organic phases were washed with NaCl solution (20 mL, sat. aqueous), dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated under reduced pressure. Purification by column chromatography (EtOAc/hexane, 30:70) afforded the corresponding  $\alpha,\beta$ -unsaturated ketone (53 mg) as a white solid.

General Procedure F was followed starting from the above product (23 mg, 0.10 mmol). Purification by column chromatography (EtOAc/hexane, 50:50) afforded the corresponding chloroacetamide (16 mg, 50% over three steps) as a white solid.

**$^1\text{H NMR}$**  (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.87 (d,  $J = 1.8$  Hz, 1H), 8.82 (d,  $J = 1.8$  Hz, 1H), 8.03 (dd,  $J = 8.5, 1.4$  Hz, 1H), 7.73 (dd,  $J = 8.5, 7.2$  Hz, 1H), 7.56–7.42 (m, 1H), 6.47 (dd,  $J = 11.7, 4.9$  Hz, 1H), 4.69–4.36 (m, 2H), 3.49 (dd,  $J = 18.0, 11.6$  Hz, 1H), 2.57 (dd,  $J = 18.0, 4.9$  Hz, 1H), 1.81 (ddd,  $J = 13.3, 8.4, 5.0$  Hz, 1H), 1.04–0.66 (m, 4H);  **$^{13}\text{C NMR}$**  (151 MHz,  $\text{CD}_2\text{Cl}_2$ )  $\delta$  163.6, 163.5, 145.6, 144.4, 143.9, 140.4, 139.3, 130.2, 129.5, 125.8, 56.3, 42.9, 42.7, 11.7, 7.4, 7.0. **HRMS** (ESI): exact mass calculated for  $\text{C}_{16}\text{H}_{15}\text{ClN}_4\text{NaO}$  [(M+MeCN+Na) $^+$ ] 378.1082, found 378.1055.



PM-02-20A

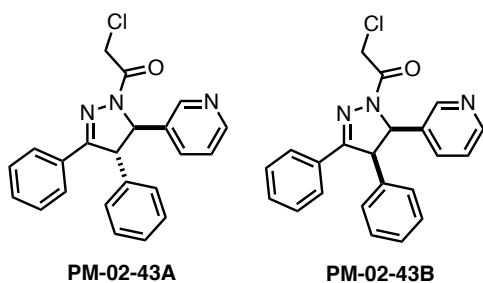
PM-02-20B

**2-chloro-1-(3,4,5-triphenyl-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one.** General Procedure C was followed starting from deoxybenzoin (235 mg, 1.2 mmol) and benzaldehyde (0.12 mL, 1.2 mmol). Purification by column chromatography (EtOAc/hexane, 10:90) afforded the corresponding  $\alpha,\beta$ -unsaturated ketone (126 mg) as a mixture with benzaldehyde. This mixture was used in the following transformation without further purification.

General Procedure F was followed starting from the above product (120 mg, 0.42 mmol). Purification by column chromatography (EtOAc/hexane, 0:100 to 20:80) afforded the *trans*-substituted chloroacetamide **PM-02-20A** (13 mg, 2% over three steps) as a colorless film and the *cis*-substituted chloroacetamide **PM-02-20B** (8 mg, 1% over three steps) as a colorless film.

**PM-02-20A [trans]:**  $^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.71 – 7.63 (m, 2H), 7.40 – 7.27 (m, 9H), 7.25 – 7.10 (m, 4H), 5.36 (d,  $J = 3.6$  Hz, 1H), 4.68 (d,  $J = 2.7$  Hz, 2H), 4.55 (d,  $J = 3.5$  Hz, 1H).  $^{13}\text{C NMR}$  (126 MHz,  $\text{CDCl}_3$ )  $\delta$  164.2, 157.2, 140.2, 139.3, 130.6, 130.2, 129.7, 129.3, 128.8, 128.3, 128.2, 127.6, 127.1, 125.5, 70.7, 61.5, 42.3. **HRMS** (ESI): exact mass calculated for  $\text{C}_{23}\text{H}_{20}\text{ClN}_2\text{O}$  [(M+H) $^+$ ] 375.1259, found 375.1253.

**PM-02-20B [cis]:**  $^1\text{H NMR}$  (300 MHz,  $\text{C}_6\text{D}_6$ )  $\delta$  7.56 (dd,  $J = 6.9, 2.8$  Hz, 2H), 6.96 (t,  $J = 3.4$  Hz, 3H), 6.89 – 6.69 (m, 5H), 6.69 – 6.60 (m, 3H), 6.45 (dd,  $J = 6.3, 2.7$  Hz, 2H), 5.46 (d,  $J = 11.8$  Hz, 1H), 4.54 – 4.40 (m, 2H), 4.34 (d,  $J = 11.7$  Hz, 1H).  $^{13}\text{C NMR}$  (151 MHz,  $\text{C}_6\text{D}_6$ )  $\delta$  164.4, 155.8, 136.2, 135.3, 131.2, 130.1, 129.9, 128.6, 127.2, 127.2, 127.2, 66.9, 57.2, 42.2. **HRMS** (ESI): exact mass calculated for  $\text{C}_{23}\text{H}_{20}\text{ClN}_2\text{O}$  [(M+H) $^+$ ] 375.1259, found 375.1280.



**2-chloro-1-(3,4-diphenyl-5-(pyridin-3-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one.**

General procedure E was followed starting from 1-phenyl-3-(pyridin-3-yl)prop-2-en-1-one (630 mg, 3.0 mmol, 1 equiv.). The intermediate product 2-bromo-1-phenyl-3-(pyridin-3-yl)prop-2-en-1-one was obtained as a light orange oil (896 mg, 99%).  $^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.86 (d,  $J = 2.3$  Hz, 1H), 8.65 (dd,  $J = 4.8, 1.6$  Hz, 1H), 8.38 (dt,  $J = 7.8, 1.9$  Hz, 1H), 7.87 – 7.80 (m, 2H), 7.65 (s, 1H), 7.65 – 7.58 (m, 1H), 7.51 (m, 2H), 7.46 – 7.37 (m, 1H).

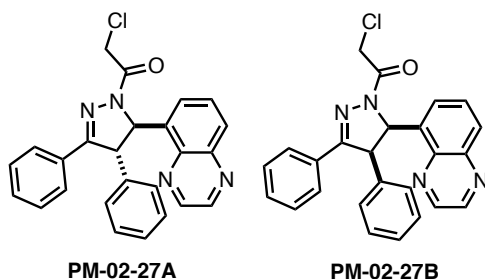
A mixture of 2-bromo-1-phenyl-3-(pyridin-3-yl)prop-2-en-1-one (58 mg, 0.20 mmol, 1.0 equiv), phenylboronic acid (32 mg, 0.26 mmol, 1.3 equiv), sodium carbonate (43 mg, 0.40 mmol, 2.0 equiv) and  $\text{Pd}(\text{dppf})\text{Cl}_2$  (7 mg, 10  $\mu\text{mol}$ , 5 mol%) was taken up in toluene/ $\text{H}_2\text{O}$  (4:1 v/v, 1 mL). The resulting suspension was sparged with nitrogen for 10 min. The flask was sealed and the reaction mixture was stirred at 100  $^\circ\text{C}$  for 1 h. The reaction mixture was then allowed to cool to ambient temperature and partitioned between  $\text{CH}_2\text{Cl}_2$  (20 mL) and 1 M NaOH (10 mL). The phases were separated and the aqueous phase was extracted with  $\text{CH}_2\text{Cl}_2$  (2 x 20 mL). The combined organic phases were washed with NaCl solution (20 mL, sat. aqueous), dried over  $\text{Na}_2\text{SO}_4$ , filtered and

concentrated under reduced pressure. Purification by column chromatography (EtOAc/hexane, 40:60 to 100:0) afforded the corresponding phenyl-substituted  $\alpha,\beta$ -unsaturated ketone (41 mg, 72%).  $^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.54 – 8.32 (m, 2H), 7.96 (dt,  $J = 8.4, 1.2$  Hz, 1H), 7.85 (dt,  $J = 7.0, 1.3$  Hz, 1H), 7.59 – 7.50 (m, 1H), 7.50 – 7.40 (m, 2H), 7.40 – 7.22 (m, 6H), 7.14 (d,  $J = 2.7$  Hz, 1H), 7.06 (dd,  $J = 8.6, 4.3$  Hz, 1H).

General Procedure F was followed starting from the above product (41 mg, 0.14 mmol). Purification by column chromatography (7% to 9% MeOH/DCM) afforded the *trans*-substituted chloroacetamide **PM-02-43A** (18 mg, 25% over three steps) as a colorless film. Further purification (EtOAc/hexane, 20:80 to 100:0) afforded the *cis*-substituted chloroacetamide **PM-02-43B** (8 mg, 11% over three steps) as a colorless film.

**PM-02-43A [trans]** :  $^1\text{H NMR}$  (600 MHz,  $\text{C}_6\text{D}_6$ )  $\delta$  8.61 (d,  $J = 2.3$  Hz, 1H), 8.45 (dd,  $J = 4.8, 1.6$  Hz, 1H), 7.56 – 7.48 (m, 2H), 7.11 (dt,  $J = 8.0, 2.0$  Hz, 1H), 6.94 (ddt,  $J = 11.0, 8.6, 4.4$  Hz, 6H), 6.83 – 6.77 (m, 2H), 6.66 (dd,  $J = 7.9, 4.7$  Hz, 1H), 5.31 (d,  $J = 3.9$  Hz, 1H), 4.39 – 4.26 (m, 2H), 4.19 (d,  $J = 4.0$  Hz, 1H).  $^{13}\text{C NMR}$  (151 MHz,  $\text{C}_6\text{D}_6$ )  $\delta$  164.0, 156.2, 149.9, 148.2, 139.4, 136.2, 132.8, 130.5, 130.5, 129.9, 128.8, 128.3, 127.8, 127.2, 123.8, 68.9, 61.3, 41.8. **HRMS** (ESI): exact mass calculated for  $\text{C}_{22}\text{H}_{19}\text{ClN}_3\text{O}$  [(M+H) $^+$ ] 376.1211, found 376.1196.

**PM-02-43B [cis]** :  $^1\text{H NMR}$  (500 MHz,  $\text{C}_6\text{D}_6$ )  $\delta$  8.29 (d,  $J = 2.5$  Hz, 1H), 8.15 (dd,  $J = 4.8, 1.7$  Hz, 1H), 7.56 – 7.48 (m, 2H), 7.04 – 6.89 (m, 3H), 6.78 (dt,  $J = 7.9, 2.1$  Hz, 1H), 6.65 (dd,  $J = 5.2, 2.8$  Hz, 3H), 6.44 – 6.34 (m, 3H), 5.30 (d,  $J = 11.6$  Hz, 1H), 4.47 – 4.36 (m, 2H), 4.30 (d,  $J = 11.7$  Hz, 1H).  $^{13}\text{C NMR}$  (126 MHz,  $\text{C}_6\text{D}_6$ )  $\delta$  164.8, 155.8, 149.1, 148.8, 134.9, 133.7, 131.7, 130.9, 130.3, 129.6, 128.7, 128.6, 128.6, 127.5, 122.5, 64.8, 56.9, 42.2. **HRMS** (ESI): exact mass calculated for  $\text{C}_{22}\text{H}_{19}\text{ClN}_3\text{O}$  [(M+H) $^+$ ] 376.1211, found 376.1211.



**2-chloro-1-(3,4-diphenyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one.** General Procedure C was followed starting from deoxybenzoin (39 mg, 0.20 mmol) and quinoxaline-6-carbaldehyde (32 mg, 0.20 mmol). Purification by column chromatography (EtOAc/hexane, 1:4) afforded the corresponding  $\alpha,\beta$ -unsaturated ketone (22 mg) as a mixture with quinoxaline-6-carbaldehyde. This mixture was used in the following transformation without further purification.

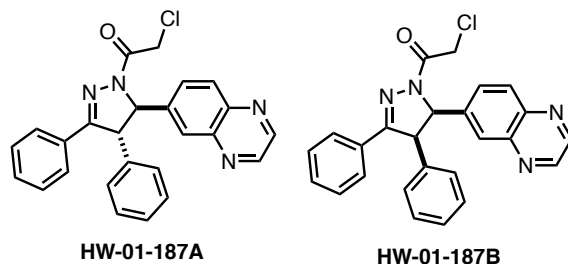
General Procedure F was followed starting from the above product (22 mg, 0.066 mmol). Purification by column chromatography (EtOAc/hexane, 0:100 to 40:60) afforded the *trans*-substituted chloroacetamide **PM-02-27A** (10 mg, 12% over three steps) as a colorless film and the *cis*-substituted chloroacetamide **PM-02-27B** (1.4 mg, 2% over three steps) as a colorless film.

**PM-02-27A [trans]:**  $^1\text{H NMR}$  (300 MHz,  $\text{CD}_2\text{Cl}_2$ )  $\delta$  8.87 (dd,  $J = 22.7, 1.8$  Hz, 2H), 8.08 (d,  $J = 8.4$  Hz, 1H), 7.81 – 7.61 (m, 3H), 7.51 (d,  $J = 7.3$  Hz, 1H), 7.46 – 7.13 (m, 8H), 6.47 (d,  $J = 3.0$  Hz, 1H), 4.92 – 4.71 (m, 2H), 4.61 (d,  $J = 3.1$  Hz, 1H).  $^{13}\text{C NMR}$  (151



MHz, CD<sub>2</sub>Cl<sub>2</sub>) δ 164.3, 158.4, 145.8, 144.5, 144.1, 140.6, 139.5, 137.8, 130.9, 130.8, 130.2, 129.9, 129.6, 129.0, 128.3, 128.0, 127.8, 125.8, 66.8, 61.3, 42.7. **HRMS** (ESI): exact mass calculated for C<sub>25</sub>H<sub>20</sub>ClN<sub>4</sub>O [(M+H)<sup>+</sup>] 427.1320, found 427.1293.

**PM-02-27B [cis]:** <sup>1</sup>H NMR (300 MHz, C<sub>6</sub>D<sub>6</sub>) δ 8.26 – 8.16 (m, 2H), 7.71 (d, *J* = 8.4 Hz, 1H), 7.60 (dd, *J* = 6.8, 2.9 Hz, 1H), 7.24 (d, *J* = 7.2 Hz, 1H), 7.05 – 6.87 (m, 5H), 6.43 (m, 4H), 5.00 (d, *J* = 11.7 Hz, 1H), 4.68 (d, *J* = 12.8 Hz, 1H), 4.40 (d, *J* = 12.8 Hz, 1H). <sup>13</sup>C NMR (126 MHz, C<sub>6</sub>D<sub>6</sub>) δ 165.2, 156.2, 144.8, 143.4, 143.2, 140.3, 135.9, 135.3, 131.2, 130.1, 129.3, 129.2, 128.9, 128.6, 128.6, 127.7, 127.6, 127.1, 62.6, 57.1, 42.6. **HRMS** (ESI): exact mass calculated for C<sub>25</sub>H<sub>20</sub>ClN<sub>4</sub>O [(M+H)<sup>+</sup>] 427.1320, found 427.1338.

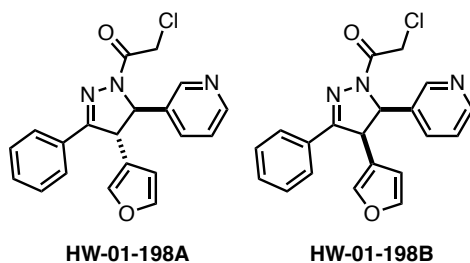


**2-Chloro-1-(3,4-diphenyl-5-(quinoxalin-6-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one.** General Procedure C was followed starting from deoxybenzoin (118 mg, 0.599 mmol) and quinoxaline-6-carbaldehyde (32 mg, 0.20 mmol). Purification by column chromatography (EtOAc/hexane, 40:60) afforded the corresponding α,β-unsaturated ketone (57 mg) containing impurities. This product was used in the following transformation without further purification. <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 8.83 – 8.71 (m, 2H), 8.02 (d, *J* = 7.5 Hz, 2H), 7.96 – 7.82 (m, 2H), 7.74 – 7.29 (m, 10H).

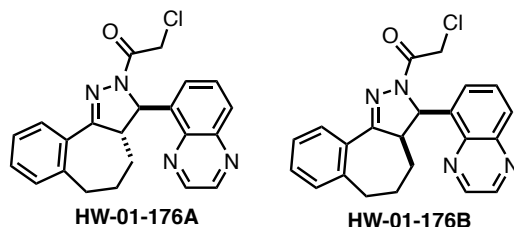
General Procedure F was followed starting from the above product (57 mg). Purification by column chromatography (EtOAc/hexane, 45:55) afforded the *trans*-substituted chloroacetamide **HW-01-187A** (33 mg, 39% over three steps) as a white solid and the *cis*-substituted chloroacetamide **HW-01-187B** (6 mg, 7% over three steps) as a colorless film.

**HW-01-187A [trans]:** <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>): δ 8.84 (s, 2H), 8.15 (d, *J* = 8.7 Hz, 1H), 7.97 (d, *J* = 2.0 Hz, 1H), 7.68 (ddd, *J* = 8.8, 4.9, 1.6 Hz, 3H), 7.42–7.27 (m, 6H), 7.24–7.16 (m, 2H), 5.61 (d, *J* = 3.8 Hz, 1H), 4.75 (d, *J* = 13.4 Hz, 1H), 4.71 (d, *J* = 13.4 Hz, 1H), 4.64 (d, *J* = 3.8 Hz, 1H); <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>): δ 164.3, 156.9, 145.4, 145.2, 143.1, 142.7, 142.1, 138.7, 130.9, 130.7, 129.8, 129.8, 128.7, 128.3, 127.7, 127.6, 127.0, 125.8, 70.3, 61.5, 42.0; **HRMS** (ESI): exact mass calculated for C<sub>25</sub>H<sub>20</sub>ClN<sub>4</sub>O [(M+MeCN+H)<sup>+</sup>] 468.1585, found 468.1574.

**HW-01-187B [cis]:** <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>): δ 8.75 (d, *J* = 1.8 Hz, 1H), 8.73 (d, *J* = 1.8 Hz, 1H), 7.74–7.70 (m, 2H), 7.64–7.60 (m, 2H), 7.37–7.33 (m, 1H), 7.31–7.27 (m, 2H), 7.20 (dd, *J* = 8.8, 2.0 Hz, 1H), 6.87 (s, 3H), 6.78 (s, 2H), 6.08 (d, *J* = 11.7 Hz, 1H), 5.34 (d, *J* = 11.7 Hz, 1H), 4.80 (d, *J* = 13.5 Hz, 1H), 4.67 (d, *J* = 13.5 Hz, 1H); <sup>13</sup>C NMR (151 MHz, CD<sub>2</sub>Cl<sub>2</sub>) δ 165.4, 157.0, 145.6, 145.5, 142.8, 142.4, 138.7, 134.7, 130.8, 130.7, 130.0 (2C), 129.2 (2C), 128.9, 128.6, 128.1, 127.7, 66.8, 57.7, 43.0. **HRMS** (ESI): exact mass calculated for C<sub>25</sub>H<sub>19</sub>ClN<sub>4</sub>NaO [(M+MeCN+Na)<sup>+</sup>] 490.1405, found 490.1400.



**2-Chloro-1-(4-(furan-3-yl)-3-phenyl-5-(pyridin-3-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one.** General procedure E was followed starting from 1-phenyl-3-(pyridin-3-yl)prop-2-en-1-one (630 mg, 3.0 mmol, 1 equiv.). The intermediate product 2-bromo-1-phenyl-3-(pyridin-3-yl)prop-2-en-1-one was obtained as a light orange oil (896 mg, 99%). A mixture of 2-bromo-1-phenyl-3-(pyridin-3-yl)prop-2-en-1-one (58 mg, 0.20 mmol, 1.0 equiv), 3-furanylboronic acid (29 mg, 0.26 mmol, 1.3 equiv), sodium carbonate (43 mg, 0.40 mmol, 2.0 equiv) and Pd(dppf)Cl<sub>2</sub> (7 mg, 10 μmol, 5 mol%) was taken up in toluene/H<sub>2</sub>O (4:1 v/v, 1 mL). The resulting suspension was sparged with nitrogen for 10 min. The flask was sealed and the reaction mixture was stirred at 100 °C for 1 h. More portions of 3-furanylboronic acid (29 mg, 0.26 mmol, 1.3 equiv), sodium carbonate (22 mg, 0.20 mmol, 1.0 equiv) and Pd(dppf)Cl<sub>2</sub> (7 mg, 10 μmol, 5 mol%) were then added, the mixture was sparged again with nitrogen for 10 min, the flask was sealed and stirring was continued at 100 °C for 2 h. The reaction mixture was then allowed to cool to ambient temperature and partitioned between CH<sub>2</sub>Cl<sub>2</sub> (20 mL) and 1 M NaOH (10 mL). The phases were separated and the aqueous phase was extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 x 20 mL). The combined organic phases were washed with NaCl solution (20 mL, sat. aqueous), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. Purification by column chromatography (EtOAc/hexane, 40:60) afforded the corresponding furyl-substituted α,β-unsaturated ketone (46 mg, d.r. = 2:1). General Procedure F was followed starting from the above product (43 mg, 0.16 mmol). Purification by column chromatography (EtOAc/hexane, 50:50 to 75:25) afforded the *trans*-substituted chloroacetamide **HW-01-198A** (27 mg, 39% over three steps) as a yellowish oil. Further purification by pTLC (EtOAc/hexane, 80:20) afforded the *cis*-substituted chloroacetamide **HW-01-198B** (4 mg, 6% over three steps) as a colorless film. **HW-01-198A [trans]:** <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 8.56 (dd, *J* = 5.0, 1.8 Hz, 2H), 7.73 (dd, *J* = 8.1, 1.7 Hz, 2H), 7.53 (dt, *J* = 7.9, 2.0 Hz, 1H), 7.44–7.27 (m, 6H), 6.26 (dd, *J* = 1.9, 0.9 Hz, 1H), 5.40 (d, *J* = 3.6 Hz, 1H), 4.63 (s, 2H), 4.54 (d, *J* = 3.6 Hz, 1H); <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>): δ 164.3, 156.3, 149.5, 147.3, 144.6, 139.5, 135.3, 133.3, 131.0, 129.6, 128.8, 127.4, 124.0, 122.9, 108.8, 67.3, 51.4, 41.8; **HRMS** (ESI): exact mass calculated for C<sub>20</sub>H<sub>17</sub>ClN<sub>3</sub>O<sub>2</sub> [(M+MeCN+H)<sup>+</sup>] 407.1269, found 407.1246. **HW-01-198B [cis]:** <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>): δ 8.39 (dd, *J* = 4.9, 1.6 Hz, 1H), 8.32 (d, *J* = 2.3 Hz, 1H), 7.70–7.66 (m, 2H), 7.42–7.38 (m, 1H), 7.37–7.32 (m, 2H), 7.30 (d, *J* = 8.0 Hz, 1H), 7.13 (dd, *J* = 7.9, 4.8 Hz, 1H), 7.03 (d, *J* = 1.5 Hz, 2H), 5.78 (d, *J* = 11.4 Hz, 1H), 5.59 (s, 1H), 5.19 (d, *J* = 11.4 Hz, 1H), 4.73 (d, *J* = 13.5 Hz, 1H), 4.57 (d, *J* = 13.4 Hz, 1H); <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>): δ 165.0, 156.1, 148.0, 147.6, 143.7, 141.4, 135.1, 134.9, 130.8, 129.8, 128.7, 127.5, 123.2, 119.0, 110.5, 63.8, 47.9, 42.1; **HRMS** (ESI): exact mass calculated for C<sub>20</sub>H<sub>17</sub>ClN<sub>3</sub>O<sub>2</sub> [(M+MeCN+H)<sup>+</sup>] 407.1269, found 407.1242.

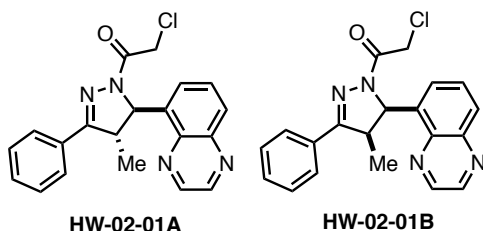


**2-Chloro-1-(3-(quinoxalin-5-yl)-3a,4,5,6-tetrahydrobenzo[6,7]cyclohepta[1,2-c]pyrazol-2(3H)-yl)ethan-1-one.** General Procedure A was followed starting from 1-benzosuberone (30  $\mu$ L, 0.20 mmol) and quinoxaline-5-carbaldehyde (63 mg, 0.40 mmol). The crude  $\alpha,\beta$ -unsaturated ketone (42 mg) was used in the following transformation without further purification.

General Procedure F was followed starting from the above product (41 mg, 0.14 mmol). Purification by column chromatography (EtOAc/hexane, 40:60) afforded the *trans*-substituted chloroacetamide **HW-01-176A** (9 mg, 17% over three steps) as a colorless film and the *cis*-substituted chloroacetamide **HW-01-176B** (5 mg, 9% over three steps) as a colorless film.

**HW-01-176A [trans]:**  $^1\text{H NMR}$  (600 MHz,  $\text{CD}_2\text{Cl}_2$ )  $\delta$  8.86 (m, 2H), 8.03 (dd,  $J = 8.5, 1.4$  Hz, 1H), 7.82 – 7.71 (m, 2H), 7.58 – 7.49 (m, 1H), 7.28 (m, 2H), 7.17 (dd,  $J = 7.4, 1.4$  Hz, 1H), 6.23 (d,  $J = 4.4$  Hz, 1H), 4.74 (d,  $J = 13.6$  Hz, 1H), 4.56 (d,  $J = 13.6$  Hz, 1H), 3.21 (dt,  $J = 12.3, 5.0$  Hz, 1H), 2.95 – 2.80 (m, 2H), 2.74 (m, 1H), 2.12 – 1.93 (m, 2H), 1.74 (m, 1H).  $^{13}\text{C NMR}$  (151 MHz,  $\text{CD}_2\text{Cl}_2$ )  $\delta$  164.3, 162.5, 145.7, 144.5, 143.9, 142.7, 140.8, 138.8, 132.2, 130.7, 130.5, 130.1, 129.5, 129.0, 126.8, 125.9, 64.5, 57.9, 42.8, 36.7, 34.2, 26.2. **HRMS** (ESI): exact mass calculated for  $\text{C}_{22}\text{H}_{20}\text{ClN}_4\text{O}$  [(M+H) $^+$ ] 391.1320, found 391.1312.

**HW-01-176B [cis]:**  $^1\text{H NMR}$  (600 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.88 (d,  $J = 6.7$  Hz, 2H), 8.03 (td,  $J = 7.6, 6.9, 1.5$  Hz, 2H), 7.71 (t,  $J = 7.8$  Hz, 1H), 7.46 (d,  $J = 7.2$  Hz, 1H), 7.33 (dtd,  $J = 18.5, 7.5, 1.6$  Hz, 2H), 7.16–7.11 (m, 1H), 7.02 (d,  $J = 11.3$  Hz, 1H), 4.73 (d,  $J = 13.0$  Hz, 1H), 4.57 (d,  $J = 13.0$  Hz, 1H), 4.03 (ddd,  $J = 13.0, 11.3, 3.5$  Hz, 1H), 3.00 (ddd,  $J = 14.9, 10.6, 4.4$  Hz, 1H), 2.65 (ddd,  $J = 14.7, 6.2, 3.9$  Hz, 1H), 1.70 (ddp,  $J = 21.4, 10.7, 3.6$  Hz, 2H), 1.32 (ddd,  $J = 13.5, 6.4, 3.3$  Hz, 1H), 0.97 (ddt,  $J = 13.4, 9.6, 6.7$  Hz, 1H);  $^{13}\text{C NMR}$  (151 MHz,  $\text{CDCl}_3$ ):  $\delta$  164.0, 162.0, 145.0, 144.2, 143.0, 141.4, 140.8, 135.0, 130.9, 130.8, 130.5, 129.8, 129.3, 128.9, 126.8, 126.7, 59.4, 50.0, 42.1, 34.2, 25.8, 24.7; **HRMS** (ESI): exact mass calculated for  $\text{C}_{22}\text{H}_{20}\text{ClN}_4\text{O}$  [(M+H) $^+$ ] 391.1320, found 391.1295.

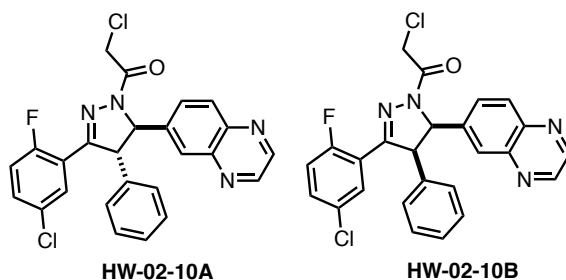


**2-Chloro-1-(4-methyl-3-phenyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one.** General Procedure D was followed starting from propiophenone (50  $\mu$ L, 0.37 mmol) and quinoxaline-5-carbaldehyde (59 mg, 0.37 mmol). Purification by column chromatography (EtOAc/hexane, 25:75) afforded the corresponding  $\alpha,\beta$ -unsaturated ketone (49 mg).

General Procedure F was followed starting from the above product (46 mg, 0.17 mmol). Purification by column chromatography (EtOAc/hexane, 35:65 to 40:60) afforded the *trans*-substituted chloroacetamide **HW-02-01A** (13 mg, 10% over three steps) as a white solid. Further purification by pTLC (EtOAc/hexane, 70:30) afforded the *cis*-substituted chloroacetamide **HW-02-01B** (2.5 mg, 2% over three steps, containing 9% of the *trans* product **xx**) as a white solid.

**HW-02-01A [trans]:**  $^1\text{H NMR}$  (600 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.90 (d,  $J = 1.8$  Hz, 1H), 8.88 (d,  $J = 1.8$  Hz, 1H), 8.04 (dd,  $J = 8.4, 1.3$  Hz, 1H), 7.75–7.71 (m, 2H), 7.69 (dd,  $J = 8.4, 7.2$  Hz, 1H), 7.49–7.35 (m, 4H), 6.38 (d,  $J = 2.9$  Hz, 1H), 4.78 (d,  $J = 13.2$  Hz, 1H), 4.63 (d,  $J = 13.1$  Hz, 1H), 3.55 (qd,  $J = 7.2, 3.0$  Hz, 1H), 1.64 (d,  $J = 7.2$  Hz, 3H);  $^{13}\text{C NMR}$  (151 MHz,  $\text{CDCl}_3$ ):  $\delta$  164.3, 160.8, 144.9, 143.9, 143.4, 140.6, 137.5, 130.6, 130.0, 129.1, 128.8, 127.2, 125.1, 64.5, 50.0, 41.9, 18.9; **HRMS** (ESI): exact mass calculated for  $\text{C}_{20}\text{H}_{18}\text{ClN}_4\text{O}$  [(M+H) $^+$ ] 365.1164, found 365.1144.

**HW-02-01B [cis]:**  $^1\text{H NMR}$  (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.91 (d,  $J = 1.8$  Hz, 1H), 8.88 (d,  $J = 1.8$  Hz, 1H), 8.07 (dd,  $J = 8.5, 1.4$  Hz, 1H), 7.80–7.75 (m, 3H), 7.55 (dt,  $J = 7.2, 0.9$  Hz, 1H), 7.45 (dd,  $J = 5.0, 1.9$  Hz, 3H), 6.81 (d,  $J = 11.3$  Hz, 1H), 4.81 (d,  $J = 13.4$  Hz, 1H), 4.54 (d,  $J = 13.4$  Hz, 1H), 4.38 (dq,  $J = 11.2, 7.5$  Hz, 1H), 0.64 (d,  $J = 7.5$  Hz, 3H); **HRMS** (ESI): exact mass calculated for  $\text{C}_{20}\text{H}_{18}\text{ClN}_4\text{O}$  [(M+H) $^+$ ] 365.1164, found 365.1140.



**2-Chloro-1-(3-(5-chloro-2-fluorophenyl)-4-phenyl-5-(quinoxalin-6-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one.** General procedure E was followed starting from 1-(5-chloro-2-fluorophenyl)-3-(quinoxalin-6-yl)prop-2-en-1-one (2.15g, 6.9 mmol, 1 equiv.). The intermediate product 2-bromo-1-(5-chloro-2-fluorophenyl)-3-(quinoxalin-6-yl)prop-2-en-1-one was obtained as a beige solid (2.66g, 99%).  $^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.90 (s, 2H), 8.63 (s, 1H), 8.18 (m, 2H), 7.96 (s, 1H), 7.64 – 7.45 (m, 2H), 7.17 (t,  $J = 8.8$ , 1H).

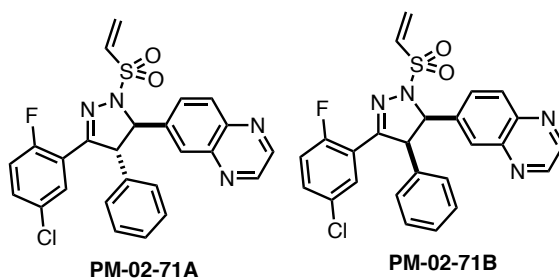
The vinyl bromide from above (2.66 g, 6.8 mmol, 1.0 equiv), phenylboronic acid (1.08 g, 8.8 mmol, 1.3 equiv), sodium carbonate (1.4 g, 0.14 mmol, 2.0 equiv) and  $\text{Pd}(\text{dppf})\text{Cl}_2$  (277 mg, 0.34 mmol, 5 mol%) was taken up in toluene/ $\text{H}_2\text{O}$  (4:1 v/v, 34 mL). The resulting suspension was sparged with nitrogen for 10 min. The flask was sealed and the reaction mixture was stirred at 100 °C for 1.5 h. The reaction mixture was then allowed to cool to ambient temperature and partitioned between EtOAc (100 mL) and 1 M NaOH (30 mL). The phases were separated and the aqueous phase was extracted with EtOAc (3 x 70 mL). The combined organic phases were washed with NaCl solution (50 mL, sat. aqueous), dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated under reduced pressure. Purification by column chromatography (EtOAc/hexane, 4:1 to 1:1 Hexane/EtOAc) afforded the corresponding phenyl-substituted  $\alpha,\beta$ -unsaturated ketone (2.0 g, 76%) as a light yellow oil.  $^1\text{H NMR}$  (500 MHz,  $\text{CDCl}_3$ ) *major isomer*,  $\delta$  8.82 – 8.77 (m, 2H), 7.90 (d,  $J = 1.9$  Hz, 1H), 7.83 (d,  $J = 8.8$  Hz, 1H), 7.59 (dd,  $J = 5.8, 2.7$  Hz, 1H), 7.52 (s, 1H), 7.50

– 7.43 (m, 2H), 7.39 (dt,  $J = 4.9, 2.6$  Hz, 3H), 7.34 (dd,  $J = 8.9, 2.1$  Hz, 1H), 7.27 (d,  $J = 2.1$  Hz, 1H), 7.07 (t,  $J = 8.9$  Hz, 1H).

General Procedure F was followed starting from the corresponding chalcone (200 mg, 0.51 mmol). Purification by column chromatography (EtOAc/hexane, 50:50) afforded the *trans*-substituted chloroacetamide **HW-02-10A** (103 mg, 42% over two steps) as a yellow oil. Further purification by pTLC (EtOAc/hexane, 70:30) and column chromatography (MeOH/CH<sub>2</sub>Cl<sub>2</sub>, 1:99) afforded the *cis*-substituted chloroacetamide **HW-01-10B** (22 mg, 9% over three steps) as a colorless film.

**HW-02-10A** [*trans*]: <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>): δ 8.86 (s, 2H), 8.17 (d,  $J = 8.7$  Hz, 1H), 7.96 (d,  $J = 2.0$  Hz, 1H), 7.93 (dd,  $J = 6.2, 2.7$  Hz, 1H), 7.67 (dd,  $J = 8.7, 2.1$  Hz, 1H), 7.38–7.28 (m, 4H), 7.15–7.11 (m, 2H), 6.93 (dd,  $J = 10.6, 8.8$  Hz, 1H), 5.63 (d,  $J = 4.0$  Hz, 1H), 4.76 (dd,  $J = 4.0, 2.9$  Hz, 1H), 4.74 (d,  $J = 13.4$  Hz, 1H), 4.66 (d,  $J = 13.4$  Hz, 1H); <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>): δ 164.5, 159.1 (d,  $J = 254.5$  Hz), 153.1 (d,  $J = 3.9$  Hz), 145.5, 145.2, 143.1, 142.7, 141.9, 138.0, 132.2 (d,  $J = 8.9$  Hz), 130.9, 129.9 (d,  $J = 3.3$  Hz), 129.6, 129.1 (d,  $J = 3.2$  Hz), 128.4, 127.7, 127.0, 125.8, 119.5 (d,  $J = 12.9$  Hz), 118.1 (d,  $J = 24.3$  Hz), 70.2, 62.6 (d,  $J = 6.1$  Hz), 41.9; HRMS (ESI): exact mass calculated for C<sub>25</sub>H<sub>18</sub>Cl<sub>2</sub>FN<sub>4</sub>O [(M+MeCN+H)<sup>+</sup>] 520.1101, found 520.1101.

**HW-01-10B** [*cis*]: <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>): δ 8.75 (d,  $J = 1.9$  Hz, 1H), 8.73 (d,  $J = 1.8$  Hz, 1H), 7.89 (dd,  $J = 6.1, 2.7$  Hz, 1H), 7.73 (d,  $J = 8.7$  Hz, 1H), 7.69 (d,  $J = 2.0$  Hz, 1H), 7.29 (ddd,  $J = 8.8, 4.2, 2.7$  Hz, 1H), 7.17 (dd,  $J = 8.7, 2.0$  Hz, 1H), 6.90–6.82 (m, 4H), 6.76–6.65 (m, 2H), 6.09 (d,  $J = 12.0$  Hz, 1H), 5.46 (dd,  $J = 12.0, 2.6$  Hz, 1H), 4.78 (d,  $J = 13.5$  Hz, 1H), 4.62 (d,  $J = 13.5$  Hz, 1H); <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>): δ –112.4; <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>): δ 165.2, 158.9 (d,  $J = 254.3$  Hz), 153.5 (d,  $J = 3.4$  Hz), 144.9, 144.7, 142.1, 141.9, 137.9, 133.2, 132.0 (d,  $J = 8.8$  Hz), 129.9 (d,  $J = 3.0$  Hz), 129.2, 129.1 (d,  $J = 3.3$  Hz), 128.9, 128.2, 127.5, 126.9, 120.1 (d,  $J = 13.5$  Hz), 118.0 (d,  $J = 23.9$  Hz), 66.2, 58.6 (d,  $J = 5.3$  Hz), 42.1; HRMS (ESI): exact mass calculated for C<sub>25</sub>H<sub>18</sub>Cl<sub>2</sub>FN<sub>4</sub>O [(M+MeCN+H)<sup>+</sup>] 520.1101, found 520.1088.



**6-(3-(5-chloro-2-fluorophenyl)-4-phenyl-1-(vinylsulfonyl)-4,5-dihydro-1H-pyrazol-5-yl)quinoxaline.** General procedure E was followed starting from 1-(5-chloro-2-fluorophenyl)-3-(quinoxalin-6-yl)prop-2-en-1-one (2.15g, 6.9 mmol, 1 equiv.). The intermediate product 2-bromo-1-(5-chloro-2-fluorophenyl)-3-(quinoxalin-6-yl)prop-2-en-1-one was obtained as a beige solid (2.66g, 99%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 8.90 (s, 2H), 8.63 (s, 1H), 8.18 (m, 2H), 7.96 (s, 1H), 7.64 – 7.45 (m, 2H), 7.17 (t,  $J = 8.8, 1$ Hz).

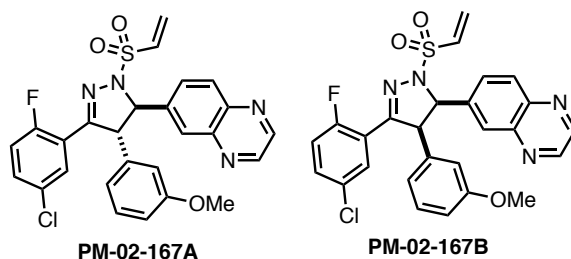
The vinyl bromide from above (2.66 g, 6.8 mmol, 1.0 equiv), phenylboronic acid (1.08 g, 8.8 mmol, 1.3 equiv), sodium carbonate (1.4 g, 0.14 mmol, 2.0 equiv) and Pd(dppf)Cl<sub>2</sub> (277 mg, 0.34 mmol, 5 mol%) was taken up in toluene/H<sub>2</sub>O (4:1 v/v, 34 mL). The resulting suspension was sparged with nitrogen for 10 min. The flask was sealed and the reaction mixture was stirred at 100 °C for 1.5 h. The reaction mixture was then allowed

to cool to ambient temperature and partitioned between EtOAc (100 mL) and 1 M NaOH (30 mL). The phases were separated and the aqueous phase was extracted with EtOAc (3 x 70 mL). The combined organic phases were washed with NaCl solution (50 mL, sat. aqueous), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. Purification by column chromatography (EtOAc/hexane, 4:1 to 1:1 Hexane/EtOAc) afforded the corresponding phenyl-substituted  $\alpha,\beta$ -unsaturated ketone (2.0 g, 76%) as a light yellow oil. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) *major isomer*,  $\delta$  8.82 – 8.77 (m, 2H), 7.90 (d, *J* = 1.9 Hz, 1H), 7.83 (d, *J* = 8.8 Hz, 1H), 7.59 (dd, *J* = 5.8, 2.7 Hz, 1H), 7.52 (s, 1H), 7.50 – 7.43 (m, 2H), 7.39 (dt, *J* = 4.9, 2.6 Hz, 3H), 7.34 (dd, *J* = 8.9, 2.1 Hz, 1H), 7.27 (d, *J* = 2.1 Hz, 1H), 7.07 (t, *J* = 8.9 Hz, 1H).

General Procedure H was followed starting from the above product (267 mg, 0.69 mmol, 1.0 equiv.). Purification by column chromatography (EtOAc:hexane, 50:50 to 100:0) afforded the *trans*-substituted vinyl sulfonamide **PM-02-71A** (73 mg, 21%) as a white solid and the *cis*-substituted vinylsulfonamide **PM-02-71B** (68 mg, 20%) as a white solid.

**PM-02-71A [trans]** : <sup>1</sup>H NMR (600 MHz, CD<sub>2</sub>Cl<sub>2</sub>)  $\delta$  8.87 – 8.83 (m, 2H), 8.17 (d, *J* = 8.7 Hz, 1H), 7.94 (d, *J* = 2.1 Hz, 1H), 7.87 (dd, *J* = 6.1, 2.7 Hz, 1H), 7.80 (dd, *J* = 8.7, 2.1 Hz, 1H), 7.31 (m, 4H), 7.05 – 7.03 (m, 2H), 6.91 (dd, *J* = 10.4, 8.8 Hz, 1H), 6.77 (dd, *J* = 16.6, 10.0 Hz, 1H), 6.40 (d, *J* = 16.7 Hz, 1H), 6.25 (d, *J* = 10.0 Hz, 1H), 5.13 (d, *J* = 8.6 Hz, 1H), 4.83 (dd, *J* = 8.6, 3.1 Hz, 1H). <sup>13</sup>C NMR (151 MHz, CD<sub>2</sub>Cl<sub>2</sub>)  $\delta$  159.1 (d, *J* = 253.5 Hz), 154.5 (d, *J* = 3.6 Hz), 146.2, 146.0, 143.3, 142.3, 138.1, 132.8, 132.5 (d, *J* = 8.8 Hz), 131.7, 131.0, 130.2 (d, *J* = 3.1 Hz), 129.9 (d, *J* = 3.4 Hz), 129.8, 128.7, 128.5, 128.3, 127.8, 120.2 (d, *J* = 13.8 Hz), 118.3 (d, *J* = 24.2 Hz), 75.2, 64.6 (d, *J* = 5.0 Hz). HRMS (ESI): exact mass calculated for C<sub>25</sub>H<sub>18</sub>ClFN<sub>4</sub>O<sub>2</sub>S [(M+MeCN+H)<sup>+</sup>] 534.1161, found 534.1164.

**PM-02-71B [cis]** : <sup>1</sup>H NMR (600 MHz, CD<sub>2</sub>Cl<sub>2</sub>)  $\delta$  8.77 – 8.71 (m, 2H), 8.06 (dd, *J* = 6.3, 2.7 Hz, 1H), 7.88 (d, *J* = 2.0 Hz, 1H), 7.69 (d, *J* = 8.7 Hz, 1H), 7.40 (dd, *J* = 8.7, 2.0 Hz, 1H), 7.35 (ddd, *J* = 8.8, 4.3, 2.7 Hz, 1H), 7.01 – 6.93 (m, 4H), 6.86 – 6.78 (m, 2H), 6.74 (dd, *J* = 16.7, 10.0 Hz, 1H), 6.40 (d, *J* = 16.7 Hz, 1H), 6.30 (d, *J* = 9.9 Hz, 1H), 5.46 (d, *J* = 10.8 Hz, 1H), 5.18 (dd, *J* = 10.8, 2.5 Hz, 1H). <sup>13</sup>C NMR (151 MHz, CD<sub>2</sub>Cl<sub>2</sub>)  $\delta$  159.6 (d, *J* = 252.2 Hz), 156.0 (d, *J* = 3.5 Hz), 145.7 (d, *J* = 8.5 Hz), 142.7, 142.6, 138.1, 133.0, 132.8, 132.8, 132.5, 131.7, 130.3, 129.9, 129.7 (d, *J* = 3.3 Hz), 129.5, 128.9, 128.8, 128.7, 128.3, 120.1, 118.5 (d, *J* = 24.3 Hz), 71.0, 60.4 (d, *J* = 7.0 Hz). HRMS (ESI): exact mass calculated for C<sub>25</sub>H<sub>18</sub>ClFN<sub>4</sub>O<sub>2</sub>S [(M+MeCN+H)<sup>+</sup>] 534.1161, found 534.1159.



**6-(3-(5-chloro-2-fluorophenyl)-4-(3-methoxyphenyl)-1-(vinylsulfonyl)-4,5-dihydro-1H-pyrazol-5-yl)quinoxaline.** General procedure E was followed starting from 1-(5-chloro-2-fluorophenyl)-3-(quinoxalin-6-yl)prop-2-en-1-one (2.15g, 6.9 mmol, 1 equiv.). The intermediate product 2-bromo-1-(5-chloro-2-fluorophenyl)-3-(quinoxalin-6-yl)prop-2-

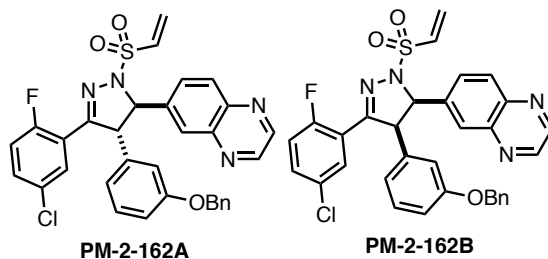
en-1-one was obtained as a beige solid (2.66g, 99%). **<sup>1</sup>H NMR** (300 MHz, CDCl<sub>3</sub>) δ 8.90 (s, 2H), 8.63 (s, 1H), 8.18 (m, 2H), 7.96 (s, 1H), 7.64 – 7.45 (m, 2H), 7.17 (t, *J* = 8.8, 1H).

The vinyl bromide from above (200 mg, 0.51 mmol, 1.0 equiv), 3-methoxyphenylboronic acid (58 mg, 0.61 mmol, 1.2 equiv), sodium carbonate (162 mg, 1.5 mmol, 2.0 equiv) and Pd(dppf)Cl<sub>2</sub> (19 mg, 0.026 mmol, 5 mol%) was taken up in toluene/H<sub>2</sub>O (4:1 v/v, 2.5 mL). The resulting suspension was sparged with nitrogen for 10 min. The flask was sealed and the reaction mixture was stirred at 100 °C for 2 h. The reaction mixture was then allowed to cool to ambient temperature and partitioned between CH<sub>2</sub>Cl<sub>2</sub> (20 mL) and 1 M NaOH (10 mL). The phases were separated and the aqueous phase was extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 x 20 mL). The combined organic phases were washed with NaCl solution (20 mL, sat. aqueous), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. Purification by column chromatography (EtOAc/hexane, 10:90 to 60:40) afforded the corresponding phenyl-substituted α,β-unsaturated ketone (174 mg, 81%). **<sup>1</sup>H NMR** (300 MHz, CDCl<sub>3</sub>) *major isomer* δ 8.75 (s, 2H), 7.89 (t, *J* = 2.0 Hz, 1H), 7.80 (d, *J* = 8.8 Hz, 1H), 7.54 (dd, *J* = 5.8, 2.7 Hz, 1H), 7.46 (s, 1H), 7.41 (ddd, *J* = 8.8, 4.4, 2.7 Hz, 1H), 7.33 (dd, *J* = 8.9, 2.0 Hz, 1H), 7.30 – 7.20 (m, 1H), 7.04 (t, *J* = 8.9 Hz, 1H), 6.92 – 6.87 (m, 1H), 6.84 – 6.73 (m, 2H), 3.70 (s, 3H).

General Procedure H was followed starting from the above product (35 mg, 0.08 mmol, 1.0 equiv.). Purification by column chromatography (EtOAc:hexane, 30:70 to 100:0) afforded the *trans*-substituted vinyl sulfonamide **PM-02-167A** (22 mg, 54%) as a white solid and the *cis*-substituted vinylsulfonamide **PM-02-167B** (6 mg, 14%) as a white solid.

**PM-02-167A [trans]** : **<sup>1</sup>H NMR** (600 MHz, CD<sub>2</sub>Cl<sub>2</sub>) δ 8.86 (q, *J* = 1.9 Hz, 2H), 8.18 (d, *J* = 8.7 Hz, 1H), 7.97 (d, *J* = 2.0 Hz, 1H), 7.86 (dd, *J* = 6.2, 2.7 Hz, 1H), 7.81 (dd, *J* = 8.7, 2.1 Hz, 1H), 7.32 (ddd, *J* = 8.8, 4.3, 2.7 Hz, 1H), 7.24 (t, *J* = 7.9 Hz, 1H), 6.93 (dd, *J* = 10.4, 8.8 Hz, 1H), 6.83 (ddd, *J* = 8.3, 2.6, 0.9 Hz, 1H), 6.78 (dd, *J* = 16.7, 10.0 Hz, 1H), 6.63 (dt, *J* = 7.6, 1.2 Hz, 1H), 6.57 (t, *J* = 2.2 Hz, 1H), 6.41 (d, *J* = 16.6 Hz, 1H), 6.25 (d, *J* = 9.9 Hz, 1H), 5.17 (d, *J* = 8.5 Hz, 1H), 4.80 (dd, *J* = 8.5, 3.0 Hz, 1H), 3.71 (s, 3H). **<sup>13</sup>C NMR** (151 MHz, CD<sub>2</sub>Cl<sub>2</sub>) δ 160.7, 159.2 (d, *J* = 253.4 Hz), 154.4 (d, *J* = 4.0 Hz), 146.1, 145.9, 143.3 (d, *J* = 5.4 Hz), 142.4, 139.6, 132.9, 132.5 (d, *J* = 9.0 Hz), 131.6, 131.0, 130.8, 130.2 (d, *J* = 3.6 Hz), 129.8 (d, *J* = 4.0 Hz), 128.6, 127.8, 120.4, 120.3 (d, *J* = 13.9 Hz), 118.3 (d, *J* = 24.0 Hz), 114.4, 113.6, 75.0, 64.5 (d, *J* = 5.6 Hz), 55.6. **HRMS** (ESI): exact mass calculated for C<sub>26</sub>H<sub>21</sub>ClFN<sub>4</sub>O<sub>3</sub>S [(M+MeCN+H)<sup>+</sup>] 564.1267, found 564.1267.

**PM-02-167B [cis]** : **<sup>1</sup>H NMR** (400 MHz, CD<sub>2</sub>Cl<sub>2</sub>) δ 8.76 (d, *J* = 5.9 Hz, 2H), 8.04 (dd, *J* = 6.3, 2.7 Hz, 1H), 7.90 (s, 1H), 7.73 (d, *J* = 8.8 Hz, 1H), 7.45 (dd, *J* = 8.7, 2.0 Hz, 1H), 7.35 (dt, *J* = 7.3, 3.1 Hz, 1H), 7.03 – 6.94 (m, 1H), 6.90 (t, *J* = 8.0 Hz, 1H), 6.73 (dd, *J* = 16.6, 10.0 Hz, 1H), 6.54 – 6.23 (m, 5H), 5.45 (d, *J* = 10.8 Hz, 1H), 5.15 (dd, *J* = 10.9, 2.5 Hz, 1H), 3.48 (s, 3H). **<sup>13</sup>C NMR** (151 MHz, CD<sub>2</sub>Cl<sub>2</sub>) δ 160.1, 159.6 (d, *J* = 254.2 Hz), 155.9 (d, *J* = 3.8 Hz), 145.7, 145.6, 142.7, 142.6, 138.1, 134.5, 132.8 (d, *J* = 8.9 Hz), 132.4, 131.7, 130.3 (d, *J* = 3.2 Hz), 129.9, 129.7 (d, *J* = 3.2 Hz), 128.8, 128.66, 121.8, 120.2 (d, *J* = 12.8 Hz), 118.5 (d, *J* = 24.3 Hz), 115.4, 113.6, 70.9, 60.4 (d, *J* = 6.9 Hz), 55.5. **HRMS** (ESI): exact mass calculated for C<sub>26</sub>H<sub>20</sub>ClFN<sub>4</sub>NaO<sub>3</sub>S [(M+MeCN+Na)<sup>+</sup>] 586.1087, found 586.1109.



**6-(4-(3-(benzyloxy)phenyl)-3-(5-chloro-2-fluorophenyl)-1-(vinylsulfonyl)-4,5-dihydro-1H-pyrazol-5-yl)quinoxaline.** General procedure E was followed starting from 1-(5-chloro-2-fluorophenyl)-3-(quinoxalin-6-yl)prop-2-en-1-one (2.15g, 6.9 mmol, 1 equiv.). The intermediate product 2-bromo-1-(5-chloro-2-fluorophenyl)-3-(quinoxalin-6-yl)prop-2-en-1-one was obtained as a beige solid (2.66g, 99%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 8.90 (s, 2H), 8.63 (s, 1H), 8.18 (m, 2H), 7.96 (s, 1H), 7.64 – 7.45 (m, 2H), 7.17 (t, *J* = 8.8, 1H).

The vinyl bromide from above (200 mg, 0.51 mmol, 1.0 equiv), 3-benzyloxyphenylboronic acid (139 mg, 0.61 mmol, 1.2 equiv), sodium carbonate (162 mg, 1.5 mmol, 2.0 equiv) and Pd(dppf)Cl<sub>2</sub> (19 mg, 0.026 mmol, 5 mol%) was taken up in toluene/H<sub>2</sub>O (4:1 v/v, 2.5 mL). The resulting suspension was sparged with nitrogen for 10 min. The flask was sealed and the reaction mixture was stirred at 100 °C for 2 h. The reaction mixture was then allowed to cool to ambient temperature and partitioned between CH<sub>2</sub>Cl<sub>2</sub> (20 mL) and 1 M NaOH (10 mL). The phases were separated and the aqueous phase was extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 x 20 mL). The combined organic phases were washed with NaCl solution (20 mL, sat. aqueous), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. Purification by column chromatography (EtOAc/hexane, 10:90 to 60:40) afforded the corresponding phenyl-substituted α,β-unsaturated ketone (203 mg, 80%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 8.81 (s, 2H), 7.92 (d, *J* = 1.9 Hz, 1H), 7.82 (d, *J* = 8.8 Hz, 1H), 7.58 (dd, *J* = 5.8, 2.7 Hz, 1H), 7.49 (s, 1H), 7.40 – 7.27 (m, 8H), 7.09 (t, *J* = 8.9 Hz, 2H), 6.92 – 6.83 (m, 2H), 5.00 (s, 2H).

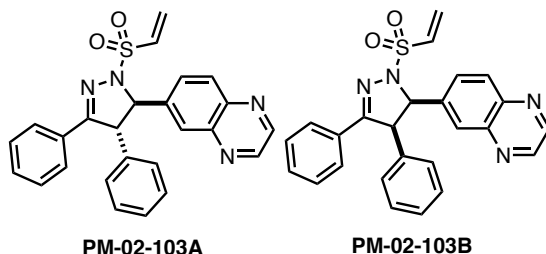
General Procedure H was followed starting from the above product (41 mg, 0.08 mmol, 1.0 equiv.). Purification by column chromatography (EtOAc:hexane, 20:80 to 100:0) afforded the *trans*-substituted vinyl sulfonamide **PM-02-162A** (28 mg, 59%) as a white solid and the *cis*-substituted vinylsulfonamide **PM-02-162B** (9 mg, 20%) as a white solid.

**PM-02-162A [trans] :** <sup>1</sup>H NMR (600 MHz, CD<sub>2</sub>Cl<sub>2</sub>) δ 8.85 (m, 2H), 8.17 (d, *J* = 8.7 Hz, 1H), 7.96 (d, *J* = 2.0 Hz, 1H), 7.85 (dd, *J* = 6.1, 2.7 Hz, 1H), 7.79 (dd, *J* = 8.7, 2.1 Hz, 1H), 7.38 – 7.35 (m, 4H), 7.34 – 7.29 (m, 2H), 7.27 – 7.21 (m, 1H), 6.95 – 6.89 (m, 2H), 6.75 (dd, *J* = 16.6, 9.9 Hz, 1H), 6.66 – 6.63 (m, 2H), 6.40 (d, *J* = 16.6 Hz, 1H), 6.22 (d, *J* = 9.9 Hz, 1H), 5.16 (d, *J* = 8.5 Hz, 1H), 4.98 (d, *J* = 1.6 Hz, 2H), 4.80 (dd, *J* = 8.5, 3.0 Hz, 1H). <sup>13</sup>C NMR (151 MHz, CD<sub>2</sub>Cl<sub>2</sub>) δ 159.8, 159.2 (d, *J* = 253.8 Hz), 154.3 (d, *J* = 3.9 Hz), 146.1, 145.9, 143.3 (d, *J* = 1.9 Hz), 142.3, 139.7, 137.1, 132.9, 132.5 (d, *J* = 9.4 Hz), 131.6, 131.0, 130.9, 130.2 (d, *J* = 3.8 Hz), 129.8 (d, *J* = 3.8 Hz), 128.9, 128.5, 128.5, 127.9, 127.8, 120.7, 120.2 (d, *J* = 14.0 Hz), 118.3 (d, *J* = 24.0 Hz), 115.1, 114.8, 75.0, 70.5, 64.5 (d, *J* = 5.5 Hz). HRMS (ESI): exact mass calculated for C<sub>32</sub>H<sub>24</sub>ClFN<sub>4</sub>NaO<sub>3</sub>S [(M+Na)<sup>+</sup>] 621.1134, found 621.1110.

**PM-02-162B [cis] :** <sup>1</sup>H NMR (600 MHz, CD<sub>2</sub>Cl<sub>2</sub>) δ 8.78 – 8.73 (m, 2H), 8.02 (dd, *J* = 6.3, 2.7 Hz, 1H), 7.90 (d, *J* = 2.0 Hz, 1H), 7.69 (d, *J* = 8.8 Hz, 1H), 7.38 – 7.27 (m, 6H), 7.23



(d,  $J = 7.4$  Hz, 2H), 6.99 (dd,  $J = 10.7, 8.8$  Hz, 1H), 6.73 (dd,  $J = 16.7, 10.0$  Hz, 1H), 6.57 (dd,  $J = 8.4, 2.5$  Hz, 1H), 6.45 – 6.37 (m, 3H), 6.29 (d,  $J = 10.0$  Hz, 1H), 5.45 (d,  $J = 10.8$  Hz, 1H), 5.13 (dd,  $J = 10.9, 2.4$  Hz, 1H), 4.85 – 4.62 (m, 2H).  $^{13}\text{C}$  NMR (151 MHz,  $\text{CD}_2\text{Cl}_2$ )  $\delta$  159.6 (d,  $J = 254.3$  Hz), 159.2, 155.8 (d,  $J = 4.4$  Hz), 145.7, 145.6, 142.7, 142.6, 138.1, 137.2, 134.5, 132.8 (d,  $J = 9.3$  Hz), 132.4, 131.7, 130.3 (d,  $J = 3.5$  Hz), 129.9, 129.8, 129.6 (d,  $J = 4.0$  Hz), 128.9, 128.8, 128.6 (d,  $J = 3.3$  Hz), 128.3, 127.7, 122.1 (d,  $J = 5.6$  Hz), 120.1 (d,  $J = 13.1$  Hz), 118.5 (d,  $J = 24.5$  Hz), 114.8, 71.0, 70.3, 60.4 (d,  $J = 7.1$  Hz). **HRMS** (ESI): exact mass calculated for  $\text{C}_{32}\text{H}_{24}\text{ClFN}_4\text{NaO}_3\text{S}$  [(M+Na) $^+$ ] 621.1134, found 621.1121.



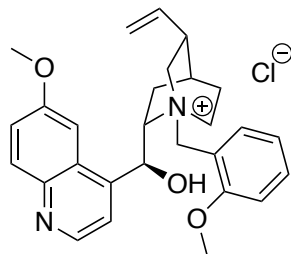
**6-(3,4-diphenyl-1-(vinylsulfonyl)-4,5-dihydro-1H-pyrazol-5-yl)quinoxaline.** General Procedure C was followed starting from deoxybenzoin (2.9 g, 15 mmol) and quinoxaline-6-carbaldehyde (790 mg, 5.0 mmol). Purification by column chromatography (EtOAc/hexane, 1:4 to 1:0) afforded the corresponding  $\alpha,\beta$ -unsaturated ketone (1.5 g, 89%).  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.83 – 8.71 (m, 2H), 8.02 (d,  $J = 7.5$  Hz, 2H), 7.96 – 7.82 (m, 2H), 7.74 – 7.29 (m, 10H).

General Procedure H was followed starting from the above product (108 mg, 0.32 mmol, 1.0 equiv.). Purification by column chromatography (EtOAc:hexane, 40:60 to 100:0) afforded the *trans*-substituted vinyl sulfonamide **PM-02-103A** (70 mg, 49% over three steps) as a white solid and the *cis*-substituted vinylsulfonamide **PM-02-103B** (19 mg, 13% over three steps) as a white solid.

**PM-02-103A [trans]** :  $^1\text{H}$  NMR (600 MHz,  $\text{CD}_2\text{Cl}_2$ )  $\delta$  8.85 (d,  $J = 1.7$  Hz, 2H), 8.16 (d,  $J = 8.7$  Hz, 1H), 7.95 (d,  $J = 2.0$  Hz, 1H), 7.80 (dd,  $J = 8.7, 2.0$  Hz, 1H), 7.64 – 7.58 (m, 2H), 7.38 – 7.32 (m, 4H), 7.32 – 7.25 (m, 2H), 7.13 (dd,  $J = 8.0, 1.6$  Hz, 2H), 6.76 (dd,  $J = 16.6, 10.0$  Hz, 1H), 6.37 (d,  $J = 16.7$  Hz, 1H), 6.20 (d,  $J = 9.9$  Hz, 1H), 5.13 (d,  $J = 8.0$  Hz, 1H), 4.76 (d,  $J = 8.0$  Hz, 1H).  $^{13}\text{C}$  NMR (151 MHz,  $\text{CD}_2\text{Cl}_2$ )  $\delta$  157.9, 146.1, 145.9, 143.3, 143.3, 142.7, 139.2, 133.0, 131.1, 131.0, 130.9, 130.3, 130.0, 128.9, 128.6, 128.6, 128.3, 128.1, 127.7, 74.9, 63.6. **HRMS** (ESI): exact mass calculated for  $\text{C}_{25}\text{H}_{21}\text{N}_4\text{O}_2\text{S}$  [(M+H) $^+$ ] 441.1380, found 441.1390.

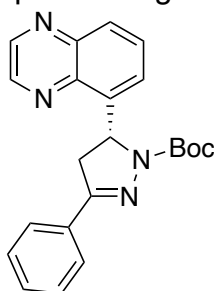
**PM-02-103B [cis]** :  $^1\text{H}$  NMR (600 MHz,  $\text{CD}_2\text{Cl}_2$ )  $\delta$  8.75 (d,  $J = 9.1$  Hz, 2H), 7.90 (d,  $J = 1.9$  Hz, 1H), 7.77 – 7.67 (m, 3H), 7.44 (dd,  $J = 8.7, 2.0$  Hz, 1H), 7.42 – 7.38 (m, 1H), 7.35 (t,  $J = 7.5$  Hz, 2H), 7.05 – 6.94 (m, 3H), 6.89 (s, 2H), 6.77 (dd,  $J = 16.7, 10.0$  Hz, 1H), 6.38 (d,  $J = 16.7$  Hz, 1H), 6.27 (d,  $J = 10.0$  Hz, 1H), 5.43 (d,  $J = 10.5$  Hz, 1H), 5.07 (d,  $J = 10.6$  Hz, 1H).  $^{13}\text{C}$  NMR (151 MHz,  $\text{CD}_2\text{Cl}_2$ )  $\delta$  159.9, 145.7, 145.6, 142.7, 142.6, 138.5, 133.7, 132.6, 132.2, 131.7, 131.2, 130.4, 130.0, 129.5, 129.1, 129.0, 128.7, 128.2, 128.0, 71.4, 59.2. **HRMS** (ESI): exact mass calculated for  $\text{C}_{25}\text{H}_{21}\text{N}_4\text{O}_2\text{S}$  [(M+H) $^+$ ] 441.1380, found 441.1400.

**(R)-EN082:**

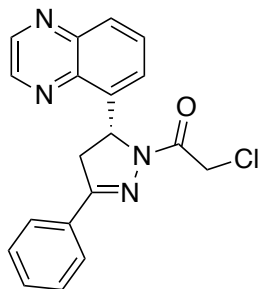


Catalyst A

**Catalyst A (1S,2S,4S,5R)-2-((R)-hydroxy(6-methoxyquinolin-4-yl)methyl)-1-(2-methoxybenzyl)-5-vinylquinuclidin-1-ium chloride** (for (*R*)-EN082): Thionyl chloride (1.0 mL, 14 mmol) was added dropwise to a solution of 2-methoxybenzyl alcohol (967 mg, 7.0 mmol) in DCM (20 mL) at 0 °C, and the solution was then stirred at rt for 4h. Volatiles were then evaporated to provide the benzyl chloride. Quinidine (324 mg, 1.0 mmol) was added to a solution of 2-methoxybenzyl chloride (313 mg, 2.0 mmol) in toluene (5 mL) and the mixture was stirred at 80 °C for 19h. After cooling to rt, the mixture was concentrated and purified by silica gel chromatography (0-4% MeOH/DCM) to provide the title compound (150 mg, 0.31 mmol, 31%) as a maroon foam. **LC/MS** calc. 445.2, found 455.2. **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ 8.82 (d, *J* = 4.6 Hz, 1H), 8.02 (d, *J* = 9.2 Hz, 1H), 7.77 (d, *J* = 4.6 Hz, 1H), 7.67 (d, *J* = 7.7 Hz, 1H), 7.58 (t, *J* = 7.9 Hz, 1H), 7.49 (d, *J* = 9.3 Hz, 1H), 7.39 (s, 1H), 7.25 (d, *J* = 8.3 Hz, 1H), 7.15 (t, *J* = 7.4 Hz, 1H), 6.98 (d, *J* = 3.6 Hz, 1H), 6.56 (s, 1H), 6.05 (ddd, *J* = 17.5, 10.7, 7.3 Hz, 1H), 5.77 (s, 1H), 5.29 – 5.17 (m, 2H), 5.06 (d, *J* = 12.6 Hz, 1H), 4.75 (d, *J* = 12.5 Hz, 1H), 4.21 (t, *J* = 10.5 Hz, 1H), 4.09 (s, 3H), 3.90 (s, 3H), 3.47 (t, *J* = 11.6 Hz, 1H), 2.95 (q, *J* = 10.1 Hz, 1H), 2.65 (d, *J* = 9.2 Hz, 1H), 2.38 (t, *J* = 11.6 Hz, 1H), 2.09 (s, 1H), 1.88 (s, 1H), 1.82 – 1.71 (m, 2H), 1.10 – 0.97 (m, 1H). Procedure followed from Mahé *et al.* *Enantioselective Phase-Transfer Catalysis: Synthesis of Pyrazolines*. *Angewandte Chemie International Edition* **2010**, 49 (39), 7072–7075. <https://doi.org/10.1002/anie.201002485>.



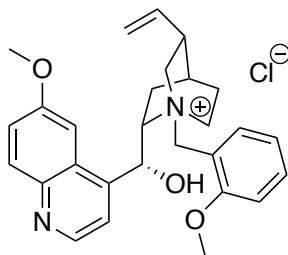
**tert-butyl (R)-3-phenyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazole-1-carboxylate**: Following general procedure I and Catalyst A, (*E*)-1-phenyl-3-(quinoxalin-5-yl)prop-2-en-1-one (130 mg, 0.50 mmol) were combined in THF (1 mL) under nitrogen and stirred at 0 °C for 18 h. The mixture was diluted with EtOAc, filtered to remove salts, concentrated and purified by silica gel chromatography (0-35% EtOAc/Hex) to provide the pyrazoline (61 mg, 0.16 mmol, 31%) as a white solid. **LC/MS** calc. 375.15, found 375.2. **<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*) δ 8.97 – 8.91 (m, 2H), 8.09 (d, *J* = 8.4 Hz, 1H), 7.84 – 7.75 (m, 3H), 7.69 (d, *J* = 7.3 Hz, 1H), 7.44 – 7.39 (m, 3H), 6.65 (dd, *J* = 12.2, 5.4 Hz, 1H), 4.03 (dd, *J* = 17.3, 12.0 Hz, 1H), 3.17 (dd, *J* = 17.3, 5.4 Hz, 1H), 1.30 (s, 9H).



(R)-EN82

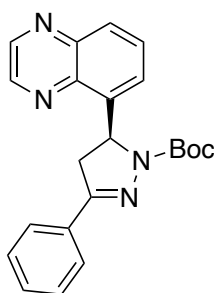
**(R)-EN82** - **(R)-2-chloro-1-(3-phenyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one** - Following general procedure J, tert-butyl (R)-3-phenyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazole-1-carboxylate (40 mg, 0.11 mmol) was converted to **(R)-EN82** (22 mg, 0.064 mmol, 58%) as a white solid. **HRMS** calc. 351.1007, found 351.1007. **<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.92 (dd, *J* = 16.1, 1.7 Hz, 2H), 8.09 (d, *J* = 8.5 Hz, 1H), 7.84 – 7.71 (m, 3H), 7.59 (d, *J* = 7.2 Hz, 1H), 7.52 – 7.42 (m, 3H), 6.72 (dd, *J* = 11.9, 5.1 Hz, 1H), 4.81 (d, *J* = 13.4 Hz, 1H), 4.64 (d, *J* = 13.5 Hz, 1H), 4.07 (dd, *J* = 17.9, 11.9 Hz, 1H), 3.25 (dd, *J* = 17.9, 5.1 Hz, 1H). **<sup>13</sup>C NMR** (151 MHz, DMSO)  $\delta$  163.8, 156.8, 146.3, 145.3, 143.1, 139.7, 139.5, 131.2, 131.2, 130.5, 129.2, 129.0, 127.4, 125.8, 56.6, 43.0, 42.3.

**(S)-EN082:**



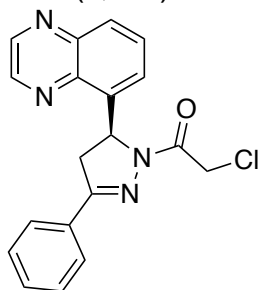
Catalyst B

**Catalyst B** **(1S,2S,4S,5R)-2-((S)-hydroxy(6-methoxyquinolin-4-yl)methyl)-1-(2-methoxybenzyl)-5-vinylquinuclidin-1-ium chloride (for S-EN082)** was synthesized from quinine, otherwise identically to Catalyst A, according to Mahé *et al.* Enantioselective Phase-Transfer Catalysis: Synthesis of Pyrazolines. *Angewandte Chemie International Edition* **2010**, 49 (39), 7072–7075. <https://doi.org/10.1002/anie.201002485>.



**tert-butyl (S)-3-phenyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazole-1-carboxylate:**

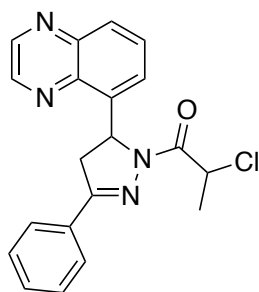
General Procedure I was followed from (E)-1-phenyl-3-(quinoxalin-5-yl)prop-2-en-1-one (130 mg, 0.50 mmol), using Catalyst B to provide the title compound (122 mg, 0.32 mmol, 65%) as a white solid. **LC/MS** calc. 375.17, found 375.2. **<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.94 (d, *J* = 10.0 Hz, 2H), 8.09 (d, *J* = 8.5 Hz, 1H), 7.84 – 7.76 (m, 3H), 7.69 (d, *J* = 7.3 Hz, 1H), 7.45 – 7.35 (m, 3H), 6.64 (dd, *J* = 12.2, 5.3 Hz, 1H), 4.02 (dd, *J* = 18.0, 12.7 Hz, 1H), 3.17 (d, *J* = 18.0 Hz, 1H), 1.30 (s, 9H).



(S)-EN82

**(S)-EN082 - (S)-2-chloro-1-(3-phenyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one** - Following general procedure J, tert-butyl (R)-3-phenyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazole-1-carboxylate (40 mg, 0.11 mmol) was converted to **(S)-EN082** (21.3 mg, 0.061 mmol, 47%) as a white solid. **HRMS** calc. 351.1007, found 351.1002. **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  9.04 (d, *J* = 3.9 Hz, 2H), 8.06 (d, *J* = 8.5 Hz, 1H), 7.85 – 7.78 (m, 3H), 7.56 (d, *J* = 7.2 Hz, 1H), 7.50 – 7.43 (m, 3H), 6.56 (dd, *J* = 11.9, 4.9 Hz, 1H), 4.92 (d, *J* = 13.9 Hz, 1H), 4.78 (d, *J* = 13.9 Hz, 1H), 4.08 (dd, *J* = 18.2, 12.0 Hz, 1H), 3.25 (dd, *J* = 18.1, 5.0 Hz, 1H). **<sup>13</sup>C NMR** (151 MHz, DMSO)  $\delta$  163.8, 156.8, 146.3, 145.3, 143.1, 139.7, 139.5, 131.2, 131.2, 130.6, 129.3, 129.0, 127.4, 125.8, 56.6, 43.0, 42.3.

**NJH-01-123:**



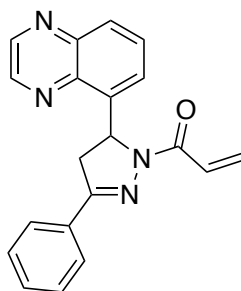
NJH-01-123

**2-chloro-1-(3-phenyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)propan-1-one:**

(E)-1-phenyl-3-(quinoxalin-5-yl)prop-2-en-1-one (55mg, 0.20 mmol) was converted via General Procedure F and purified by silica gel chromatography (0-25% EtOAc/Hex) to give **NJH-01-123** (43 mg, 0.12 mmol, 59%) as a cream-colored solid. **HRMS** calc. 365.1164, found 365.1173. **<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.93 (dd, *J* = 11.0, 1.9 Hz, 2H), 8.09 (dd, *J* = 8.5, 1.4 Hz, 1H), 7.83 – 7.79 (m, 2H), 7.77 (d, *J* = 7.7 Hz, 1H), 7.65 (d, *J* = 7.1 Hz, 1H), 7.50 – 7.43 (m, 3H), 6.77 (dd, *J* = 11.9, 4.9 Hz, 1H), 5.60 (q, *J* = 6.8

Hz, 1H), 4.07 (dd,  $J = 17.9, 11.9$  Hz, 1H), 3.22 (dd,  $J = 17.9, 4.9$  Hz, 1H), 1.79 (d,  $J = 6.8$  Hz, 3H).

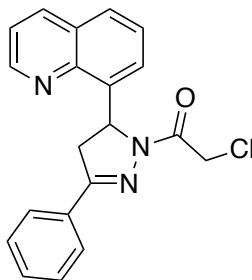
**NJH-01-124:**



NJH-01-124

**1-(3-phenyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)prop-2-en-1-one (NJH-01-124):** (E)-1-phenyl-3-(quinoxalin-5-yl)prop-2-en-1-one (55 mg, 0.20 mmol) was converted via General Procedure F and purified by silica gel chromatography (0-50% EtOAc/Hex) to give **NJH-01-124** (31 mg, 0.094 mmol, 47%) as a pale yellow solid. **HRMS** calc. 329.1397, found 329.1403.  **$^1\text{H NMR}$**  (400 MHz, Chloroform- $d$ )  $\delta$  8.93 (d,  $J = 7.1$  Hz, 2H), 8.07 (d,  $J = 8.6$  Hz, 1H), 7.86 – 7.70 (m, 3H), 7.60 – 7.39 (m, 5H), 6.82 (dd,  $J = 11.8, 4.9$  Hz, 1H), 6.51 (d,  $J = 17.2$  Hz, 1H), 5.86 (d,  $J = 9.7$  Hz, 1H), 4.06 (dd,  $J = 17.9, 12.0$  Hz, 1H), 3.21 (dd,  $J = 17.8, 5.0$  Hz, 1H).  **$^{13}\text{C NMR}$**  (151 MHz, DMSO)  $\delta$  162.5, 156.1, 146.3, 145.2, 143.1, 140.0, 139.7, 131.5, 131.0, 130.6, 129.2, 128.8, 128.5, 127.3, 125.7, 56.3, 42.1.

**NJH-01-145:**



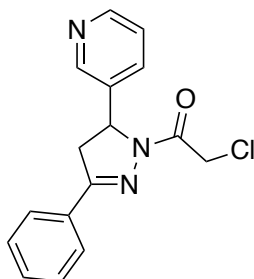
NJH-01-145

**2-chloro-1-(3-phenyl-5-(quinolin-8-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (NJH-01-145):** General Procedure A was followed starting from quinoline-8-carbaldehyde (157 mg, 1.0 mmol). After stirring at room temperature for 15 h, the solution was diluted with water and extracted with EtOAc (3x5mL). Extracts were combined, washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , concentrated and purified by silica gel chromatography (0-60% EtOAc/Hex) to obtain the chalcone (85 mg, 0.33 mmol, 33%) as an orange oil.

This chalcone ((E)-1-phenyl-3-(quinolin-8-yl)prop-2-en-1-one) (81 mg, 0.31 mmol) was converted via General Procedure F and purified by silica gel chromatography (0-50% EtOAc/Hex) to give **NJH-01-145** (35 mg, 0.10 mmol, 32%) as a white solid. **HRMS** calc. 350.1055, found 350.1057.  **$^1\text{H NMR}$**  (400 MHz, Chloroform- $d$ )  $\delta$  8.98 (dd,  $J = 4.2, 1.8$  Hz, 1H), 8.21 (dd,  $J = 8.3, 1.8$  Hz, 1H), 7.79 (d,  $J = 8.2$  Hz, 3H), 7.57 – 7.39 (m, 6H), 6.78 (dd,

$J = 11.7, 4.9$  Hz, 1H), 4.75 (dd,  $J = 73.5, 13.2$  Hz, 2H), 4.10 (dd,  $J = 18.0, 11.8$  Hz, 1H), 3.26 (dd,  $J = 17.9, 4.9$  Hz, 1H).  $^{13}\text{C}$  NMR (151 MHz, DMSO)  $\delta$  163.7, 156.8, 150.4, 144.9, 138.7, 137.0, 131.3, 131.1, 129.2, 128.8, 128.1, 127.4, 126.8, 124.8, 122.1, 57.1, 43.0, 42.5.

#### NJH-01-147:

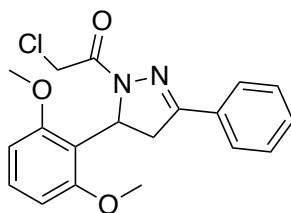


NJH-01-147

**2-chloro-1-(3-phenyl-5-(pyridin-3-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (NJH-01-147):** General Procedure A was followed starting from pyridine-3-carbaldehyde (96  $\mu\text{L}$ , 1.0 mmol). After stirring at room temperature for 15h, the solution was diluted with water and extracted with EtOAc (3x5mL). Extracts were combined, washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , concentrated and purified by silica gel chromatography (0-70% EtOAc/Hex) to obtain the chalcone (50 mg, 0.24 mmol, 24%) as a clear colorless oil.

This chalcone (E)-1-phenyl-3-(pyridin-3-yl)prop-2-en-1-one (50 mg, 0.24 mmol) was converted via General Procedure F and purified by silica gel chromatography (0-80% EtOAc/Hex) to give **NJH-01-147** (21 mg, 0.070 mmol, 29%) as a white solid. **HRMS** calc. 300.0898, found 300.0917.  $^1\text{H}$  NMR (400 MHz, Chloroform- $d$ )  $\delta$  8.62 (d,  $J = 2.4$  Hz, 1H), 8.59 (d,  $J = 4.8$  Hz, 1H), 7.80 (d,  $J = 7.7$  Hz, 2H), 7.61 (d,  $J = 7.9$  Hz, 1H), 7.56 – 7.47 (m, 3H), 7.32 (d,  $J = 4.9$  Hz, 1H), 5.67 (dd,  $J = 11.9, 4.8$  Hz, 1H), 4.63 (dd,  $J = 22.3, 13.8$  Hz, 2H), 3.90 (dd,  $J = 17.9, 11.8$  Hz, 1H), 3.29 (dd,  $J = 17.9, 4.9$  Hz, 1H).  $^{13}\text{C}$  NMR (151 MHz, DMSO)  $\delta$  163.9, 156.3, 149.2, 148.1, 137.3, 133.9, 131.3, 131.0, 129.3, 127.5, 124.3, 58.5, 42.9, 42.1.

#### NJH-01-158:

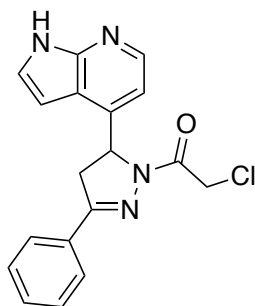


NJH-01-158

**2-chloro-1-(5-(2,6-dimethoxyphenyl)-3-phenyl-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (NJH-01-158):** General Procedure A was followed starting from 2,6-dimethoxybenzaldehyde (332mg, 2.0 mmol). After stirring at rt 15h, the mixture was diluted with water and extracted with EtOAc (3x5mL). Extracts were combined, washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , concentrated and purified by silica gel chromatography (0-30% EtOAc/Hex) to obtain the chalcone (525 mg, 1.96 mmol, 98%) as a colorless oil. This chalcone (E)-3-(2,6-dimethoxyphenyl)-1-phenylprop-2-en-1-one (100 mg, 0.37

mmol) was converted via General Procedure F and purified by silica gel chromatography (0-50% EtOAc/Hex) to give **NJH-01-158** (89 mg, 0.25 mmol, 67%) as a white solid. **HRMS** calc. 359.1157, found 359.1139. **<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*) δ 7.85 – 7.73 (m, 2H), 7.54 – 7.42 (m, 3H), 7.23 (t, *J* = 8.3 Hz, 1H), 6.58 (d, *J* = 8.3 Hz, 2H), 6.21 (dd, *J* = 12.6, 6.2 Hz, 1H), 4.58 (d, *J* = 2.1 Hz, 2H), 3.80 (s, 6H), 3.60 (dd, *J* = 17.3, 12.6 Hz, 1H), 3.27 (dd, *J* = 17.3, 6.3 Hz, 1H). **<sup>13</sup>C NMR** (151 MHz, DMSO) δ 162.7, 158.4, 156.0, 131.8, 130.7, 129.4, 129.2, 127.1, 117.2, 105.2, 56.5, 51.7, 42.9.

#### **NJH-01-159:**

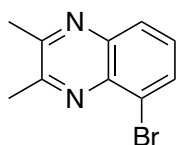


NJH-01-159

**2-chloro-1-(3-phenyl-5-(1H-pyrrolo[2,3-b]pyridin-4-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (NJH-01-159):** 1H-pyrrolo[2,3-b]pyridine-4-carbaldehyde: n-BuLi (3.0 mL, 7.62 mmol, 2.5 M in hexanes) was added dropwise to a solution of 4-bromo-1H-pyrrolo[2,3-b]pyridine (500 mg, 2.54 mmol) in THF (20 mL) at -78 °C and stirred at that temperature for 30 min. DMF (988 μL, 12.7 mmol) was added and the reaction was stirred at rt for 1h before being cooled to -78 °C and quenched with sat. aq. NH<sub>4</sub>Cl. The mixture was then warmed to rt, diluted with water and extracted with EtOAc (3x10 mL). Extracts were combined, washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated and purified by silica gel chromatography (0-40% EtOAc/Hex) to obtain the aldehyde (178 mg, 1.22 mmol, 48%) as a yellow solid.

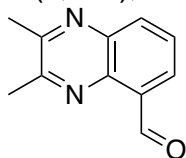
General Procedure A was followed starting from 1H-pyrrolo[2,3-b]pyridine-4-carbaldehyde (75 mg, 0.51 mmol). After stirring at rt 4h, water (15 mL) was added and the precipitate filtered to provide the chalcone (93 mg, 0.38 mmol, 74%) as a bright yellow solid. This chalcone (E)-1-phenyl-3-(1H-pyrrolo[2,3-b]pyridin-4-yl)prop-2-en-1-one was converted via General Procedure F and purified by silica gel chromatography (0-60% EtOAc/Hex) to give **NJH-01-159** (15 mg, 0.044 mmol, 22%) as a yellow solid. **LC/MS** calc. 339.1, found 339.1. **<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*) δ 9.55 (s, 1H), 8.31 (d, *J* = 4.9 Hz, 1H), 7.81 (d, *J* = 7.8 Hz, 2H), 7.55 – 7.46 (m, 3H), 7.35 (s, 1H), 7.04 (d, *J* = 5.0 Hz, 1H), 6.45 (d, *J* = 3.3 Hz, 1H), 5.94 (dd, *J* = 12.1, 5.4 Hz, 1H), 4.67 (s, 2H), 3.94 (dd, *J* = 17.8, 12.1 Hz, 1H), 3.39 (dd, *J* = 17.8, 5.4 Hz, 1H).

#### **NJH-01-161:**

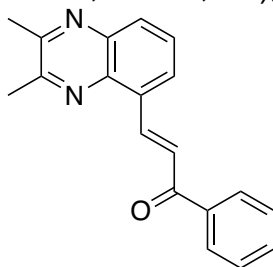


**5-bromo-2,3-dimethylquinoxaline:** 3-bromobenzene-1,2-diamine (1.0 g, 5.35 mmol) and diacetyl (517 mL, 5.89 mmol) were added to EtOH (20 mL), and the solution was

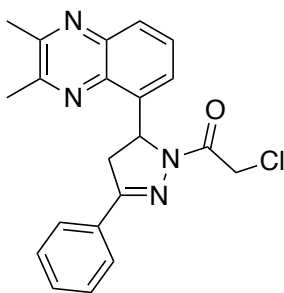
stirred at reflux for 1h and cooled to rt. The mixture was then concentrated to dryness and the crude was redissolved in MeOH 95 mL. Water (90 mL) was then added and the precipitate filtered to provide the quinoxaline (1.17 g, 4.94 mmol, 92%) as a tan solid. **LC/MS calc.** 237.0, found 236.9. **<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.01 (t, *J* = 8.7 Hz, 2H), 7.57 (td, *J* = 7.9, 1.8 Hz, 1H), 2.85 (s, 3H), 2.81 (s, 3H).



**2,3-dimethylquinoxaline-5-carbaldehyde:** 5-bromo-2,3-dimethylquinoxaline (474 mg, 2.0 mmol) was dissolved in THF (10 mL) and cooled to -78 °C. *n*-BuLi in hexanes (960 mL, 2.4 mmol) was then added dropwise, and the mixture stirred at -78 °C for 30 min. DMF (311 mL, 4.0 mmol) was then added at -78 °C and the reaction stirred at rt for 1 hour. The reaction was then quenched with sat. aq. NH<sub>4</sub>Cl at -78 °C, warmed to rt, diluted with water and extracted with EtOAc (3x10mL). Extracts were combined, washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated and purified by silica gel chromatography (0-50% EtOAc/Hex) to obtain the aldehyde (169 mg, 0.91 mmol, 45%). **LC/MS calc.** 187.1, found 187.1. **<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  11.41 (s, 1H), 8.04 (dd, *J* = 6.3, 3.5 Hz, 1H), 7.83 (t, *J* = 7.7 Hz, 1H), 7.72 (dd, *J* = 6.4, 3.4 Hz, 1H), 2.84 (s, 3H), 2.82 (s, 3H).



**(E)-3-(2,3-dimethylquinoxalin-5-yl)-1-phenylprop-2-en-1-one:** General Procedure A was followed starting from 2,3-dimethylquinoxaline-5-carbaldehyde (93 mg, 0.50 mmol). After stirring at room temperature for 5h, the mixture was diluted with water and extracted with EtOAc (3x5mL). Extracts were combined, washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated and purified by silica gel chromatography (0-25% EtOAc/Hex) to obtain the chalcone (59 mg, 0.21 mmol, 41%). **LC/MS calc.** 289.1, found 289.1. **<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.96 (d, *J* = 16.0 Hz, 1H), 8.17 – 8.11 (m, 2H), 8.09 (d, *J* = 7.4 Hz, 2H), 8.04 (s, 1H), 7.76 (t, *J* = 7.9 Hz, 1H), 7.68 – 7.63 (m, 1H), 7.61 – 7.54 (m, 2H), 2.84 (s, 3H), 2.80 (s, 3H).

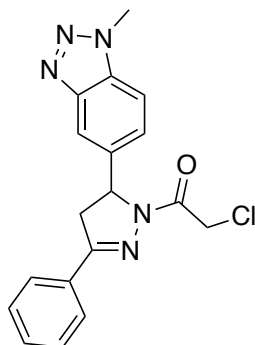


NJH-01-161



**2-chloro-1-(5-(2,3-dimethylquinoxalin-5-yl)-3-phenyl-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (NJH-01-161):** (*E*)-3-(2,3-dimethylquinoxalin-5-yl)-1-phenylprop-2-en-1-one (30 mg, 0.10 mmol) was converted via General Procedure F and purified by silica gel chromatography (0-45% EtOAc/Hex) to give **NJH-01-161** (30 mg, 0.080 mmol, 80%) as a white solid. HRMS calc. 379.1320, found 379.1342. <sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ 7.85 (d, *J* = 8.3 Hz, 1H), 7.71 (s, 2H), 7.54 (t, *J* = 7.7 Hz, 1H), 7.39 (s, 4H), 6.55 (d, *J* = 11.8 Hz, 1H), 4.62 (dd, *J* = 42.9, 13.4 Hz, 2H), 3.92 (dd, *J* = 17.9, 12.7 Hz, 1H), 3.18 (dd, *J* = 17.4, 4.2 Hz, 1H), 2.67 (s, 3H), 2.61 (s, 3H). <sup>13</sup>C NMR (151 MHz, DMSO) δ 163.7, 156.8, 154.5, 153.6, 141.2, 138.5, 137.8, 131.4, 131.1, 129.2, 129.0, 127.9, 127.4, 57.0, 42.9, 42.3, 23.6, 23.2.

**NJH-01-166:**



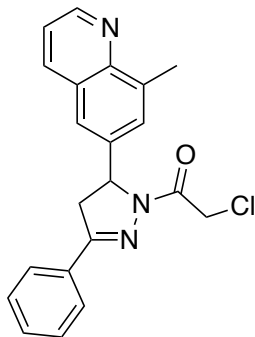
**(1-methyl-1H-benzo[d][1,2,3]triazol-5-yl)methanol:** LiAlH<sub>4</sub> (5.64 mmol, 1M in THF) was added dropwise to a solution of 1-methylbenzotriazole-5-carboxylic acid (500 mg, 2.52 mmol) in THF (10 mL) at 0 °C. The mixture was allowed to warm to rt and stirred for 5 h. The mixture was then diluted with Et<sub>2</sub>O (40mL), and 0.2 mL water was added at 0 °C, followed by 0.6 mL 1M NaOH aq., 0.2 mL water, and finally Na<sub>2</sub>SO<sub>4</sub> after 15 min at rt. The mixture was filtered, concentrated and purified by silica gel chromatography (0-80% EtOAc/Hex) to obtain the alcohol (308 mg, 1.89 mmol, 67%). **LC/MS** calc. 164.1, found 164.1.

**1-methyl-1H-benzo[d][1,2,3]triazole-5-carbaldehyde:** (1-methyl-1H-benzo[d][1,2,3]triazol-5-yl)methanol (200 mg, 1.22 mmol) and 4-methylmorpholine N-oxide (500 mg, 4.27 mmol) were combined in DCM (5 mL) with 4Å MS. TPAP (42 mg, 0.12 mmol) was then added and the mixture stirred as 2 hours before the alcohol was consumed. The mixture was diluted with Et<sub>2</sub>O and filtered to remove NMO, filtrate concentrated and purified by silica gel chromatography (0-30% EtOAc/Hex) to obtain the aldehyde (138 mg, 0.86 mmol, 70%). **LC/MS** calc. 162.1, found 162.0.

**2-chloro-1-(5-(1-methyl-1H-benzo[d][1,2,3]triazol-5-yl)-3-phenyl-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (NJH-01-166):** General Procedure A was followed starting from 1-methyl-1H-benzo[d][1,2,3]triazole-5-carbaldehyde (80 mg, 0.50 mmol). After 16 h., water was added and the precipitate filtered to obtain the chalcone (117 mg, 0.44 mmol, 88%) as a tan powder. This chalcone (*E*)-3-(1-methyl-1H-benzo[d][1,2,3]triazol-5-yl)-1-phenylprop-2-en-1-one (40 mg, 0.15 mmol) was converted via General Procedure F and purified by silica gel chromatography (0-60% EtOAc/Hex) to give **NJH-01-166** (16 mg, 0.046 mmol, 31%) as a light yellow foam. **HRMS** calc. 354.1116, found 354.1093. <sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ 7.99 (s, 1H), 7.82 (d, *J* = 7.8 Hz, 2H), 7.59 – 7.43 (m, 5H), 5.85 – 5.76 (m, 1H), 4.64 (d, *J* = 4.8 Hz, 2H), 4.32 (s, 3H), 3.93 (dd, *J* = 18.0, 11.2 Hz, 1H), 3.34 (dd, *J* = 17.9, 5.0 Hz, 1H). <sup>13</sup>C NMR (151 MHz, DMSO) δ 163.8, 156.2,

145.7, 138.0, 133.3, 131.2, 131.1, 129.3, 127.4, 125.8, 116.2, 111.7, 60.5, 43.0, 42.6, 34.7.

**NJH-01-181:**

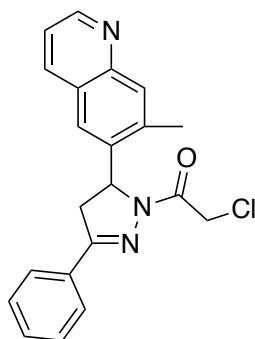


NJH-01-181

**2-chloro-1-(5-(8-methylquinolin-6-yl)-3-phenyl-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (NJH-01-181):** 6-bromo-8-methylquinoline (500 mg, 2.25 mmol) was converted to 8-methylquinoline-6-carbaldehyde via General Procedure K and purified by silica gel chromatography (0-40% EtOAc/Hex) to provide the aldehyde (184 mg, 1.08 mmol, 48%) as a pale orange solid.

Then, General Procedure A was followed starting from 8-methylquinoline-6-carbaldehyde (100mg, 0.58 mmol). After stirring overnight water (15 mL) was added, and the solids filtered to provide the chalcone (134 mg, 0.49 mmol, 85%) as a tan solid. (E)-3-(8-methylquinolin-6-yl)-1-phenylprop-2-en-1-one (75 mg, 0.27 mmol) was converted via General Procedure F and purified by silica gel chromatography (0-45% EtOAc/Hex) to give **NJH-01-181** (49 mg, 0.13 mmol, 49%) as a white powder. **HRMS** calc. 364.1211, found 364.1219. **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ 8.94 – 8.85 (m, 1H), 8.33 (d, *J* = 8.4 Hz, 1H), 7.90 – 7.81 (m, 2H), 7.67 (s, 1H), 7.56 – 7.47 (m, 5H), 5.74 (dd, *J* = 11.9, 5.0 Hz, 1H), 4.81 (dd, *J* = 32.2, 13.6 Hz, 2H), 3.99 (dd, *J* = 17.9, 11.4 Hz, 1H), 3.34 (dd, *J* = 17.9, 5.4 Hz, 1H), 2.71 (s, 3H). **<sup>13</sup>C NMR** (151 MHz, DMSO) δ 163.9, 156.2, 150.0, 146.6, 139.6, 137.8, 136.8, 131.2, 131.1, 129.3, 128.2, 127.7, 127.5, 122.8, 122.1, 60.5, 43.0, 42.5, 18.3.

**NJH-01-182:**

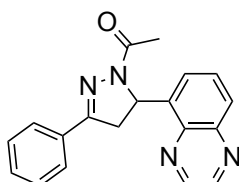


NJH-01-182

**2-chloro-1-(5-(7-methylquinolin-6-yl)-3-phenyl-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (NJH-01-182):** 7-methylquinoline-6-carbaldehyde:6-bromo-8-methylquinoline (500 mg, 2.25 mmol) was converted to 7-methylquinoline-6-carbaldehyde via general procedure K and purified by silica gel chromatography (0-50%

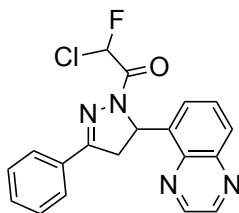
EtOAc/Hex) to provide the aldehyde (275 mg, 1.61 mmol, 71%) as a pale orange solid. LC/MS calc. 172.1, found 172.1.

General Procedure A was followed starting from 7-methylquinoline-6-carbaldehyde (100 mg, 0.58 mmol). The reaction mixture was stirred for 16 hours at rt and an immiscible orange oil formed. Water was added and the mixture extracted with EtOAc (3x5mL). Extracts were combined, washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated, and purified by silica gel chromatography (0 to 50% EtOAc/Hex.) to obtain the chalcone (151 mg, 0.55 mmol, 95%) as a yellow gummy solid. This chalcone (E)-3-(7-methylquinolin-6-yl)-1-phenylprop-2-en-1-one (75 mg, 0.27 mmol) was converted via General Procedure F and purified by silica gel chromatography (0-65% EtOAc/Hex) to give **NJH-01-182** (25 mg, 0.069 mmol, 25%) as a white solid. **HRMS** calc. 364.1211, found 364.1222. **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ 8.83 (dd, *J* = 4.2, 1.7 Hz, 1H), 8.26 (dd, *J* = 8.2, 1.8 Hz, 1H), 7.91 (s, 1H), 7.87 – 7.82 (m, 2H), 7.57 (s, 1H), 7.51 – 7.41 (m, 4H), 5.87 (dd, *J* = 11.7, 4.9 Hz, 1H), 4.87 (dd, *J* = 37.1, 13.8 Hz, 2H), 4.05 (dd, *J* = 18.1, 11.9 Hz, 1H), 3.24 (dd, *J* = 18.1, 4.9 Hz, 1H), 2.62 (s, 3H). **<sup>13</sup>C NMR** (151 MHz, DMSO) δ 163.8, 156.3, 150.9, 147.5, 139.3, 137.4, 136.3, 131.2, 131.1, 130.0, 129.3, 127.4, 126.8, 123.2, 121.4, 58.0, 43.0, 41.8, 20.1.



LEB-02-150

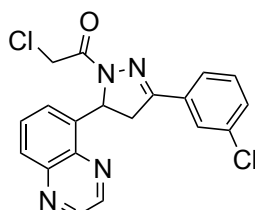
**1-(3-phenyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (LEB-02-150)**. General Procedure A was followed starting from acetophenone (0.40 mmol) and quinoxaline-5-carbaldehyde (0.40 mmol). The corresponding chalcone precipitated out during the reaction and was filtered by gravity filtration to yield product (0.20 mmol) as a light yellow solid. General Procedure F was followed starting from the above product (1.0 equiv, 0.20 mmol). Purification by flash column chromatography (EtOAc/hexane 50:50) yielded the chloroacetamide **LEB-02-150** (0.08 mmol, 22.1% yield across all steps). **<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*) δ 8.97 – 8.89 (m, 2H), 8.07 (d, *J* = 8.4 Hz, 1H), 7.82 – 7.71 (m, 3H), 7.56 – 7.50 (m, 1H), 7.48 – 7.40 (m, 3H), 7.30 (s, 3H), 6.73 (dd, *J* = 11.9, 4.8 Hz, 1H), 4.05 (dd, *J* = 17.9, 12.1 Hz, 1H), 3.18 (dd, *J* = 17.5, 4.9 Hz, 1H). **<sup>13</sup>C NMR** (151 MHz, DMSO) δ 168.0, 155.2, 146.2, 145.2, 143.1, 140.3, 139.7, 131.6, 130.8, 130.6, 129.2, 128.7, 127.1, 125.5, 56.0, 42.4, 22.2. **HRMS** (ESI): exact mass calculated for C<sub>19</sub>H<sub>16</sub>N<sub>4</sub>O [(M+H)<sup>+</sup>] 317.1397, found 317.1405.



LEB-02-157

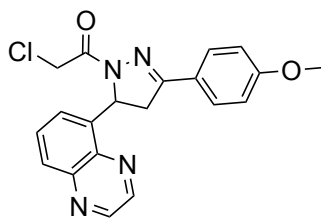
**2-chloro-2-fluoro-1-(3-phenyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (LEB-02-157)**. General Procedure A was followed starting from

acetophenone (0.40 mmol) and quinoxaline-5-carbaldehyde (0.40 mmol). The corresponding chalcone precipitated out during the reaction and was filtered by gravity filtration to yield product (0.38 mmol) as a yellow solid. Hydrazine monohydrate (2.0 equiv) was added to a suspension of the above product (1.0 equiv, 0.38 mmol) in EtOH (0.3 M). The resulting reaction mixture was stirred at reflux temperature for 4 h before it was concentrated under reduced pressure. The crude pyrazoline (1.0 equiv, 0.36 mmol) was added to a screw-cap oven-dried vial charged with a stir-bar, chloroacetate (1.2 equiv, 0.44 mmol), EDC hydrochloride (1.2 equiv, 0.44 mmol), HOBT (1.2 equiv, 0.44 mmol), DIEA (2.5 equiv, 0.92 mmol), and 2 mL DMF. Reaction proceeded overnight, and was extracted in EtOAc, washed 3X with brine, and purified via flash column chromatography (EtOAc/hexane 50:50), yielding the corresponding chloroacetamide **LEB-02-157** (0.026 mmol, 6.4% yield across all steps). **<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.97 – 8.84 (m, 2H), 8.13 – 8.08 (m, 1H), 8.06 – 7.95 (m, 1H), 7.83 – 7.75 (m, 3H), 7.63 – 7.57 (m, 1H), 7.54 – 7.44 (m, 3H), 6.70 (dd, *J* = 11.7, 5.1 Hz, 1H), 4.16 – 4.04 (m, 1H), 3.36 – 3.22 (m, 1H). **HRMS** (ESI): exact mass calculated for C<sub>19</sub>H<sub>14</sub>ClFN<sub>4</sub>O [(M+H)<sup>+</sup>] 369.0913, found 369.0914.



LEB-02-172

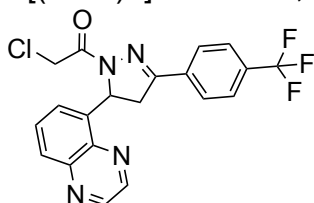
**2-chloro-1-(3-(3-chlorophenyl)-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (LEB-02-172)**. General Procedure A was followed starting from 3-chloroacetophenone (1.2 equiv, 0.76 mmol) and quinoxaline-5-carbaldehyde (1.0 equiv, 0.63 mmol). The corresponding chalcone precipitated out during the reaction and was filtered by gravity filtration to yield product (0.50 mmol) as a yellow solid. General Procedure F was followed starting from the above product (1.0 equiv, 0.50 mmol). Purification by flash column chromatography (EtOAc/hexane 50:50) yielded the chloroacetamide **LEB-02-172** (0.058 mmol, 9.2% yield across all steps). **<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.92 (dt, *J* = 21.3, 1.8 Hz, 2H), 8.10 (dd, *J* = 8.5, 1.7 Hz, 1H), 7.81 – 7.74 (m, 2H), 7.68 – 7.62 (m, 1H), 7.58 (d, *J* = 7.3 Hz, 1H), 7.49 – 7.43 (m, 1H), 7.43 – 7.37 (m, 1H), 6.74 – 6.66 (m, 1H), 4.79 (dd, *J* = 13.5, 1.6 Hz, 1H), 4.63 (dd, *J* = 13.7, 1.4 Hz, 1H), 4.10 – 3.99 (m, 1H), 3.28 – 3.19 (m, 1H). **<sup>13</sup>C NMR** (151 MHz, DMSO)  $\delta$  164.0, 155.6, 146.3, 145.3, 143.1, 139.7, 139.3, 134.1, 133.3, 131.2, 130.8, 130.5, 129.0, 126.9, 126.1, 56.9, 43.0, 42.2. **HRMS** (ESI): exact mass calculated for C<sub>19</sub>H<sub>14</sub>Cl<sub>2</sub>N<sub>4</sub>O [(M+H)<sup>+</sup>] 385.0617, found 385.063.



LEB-02-173

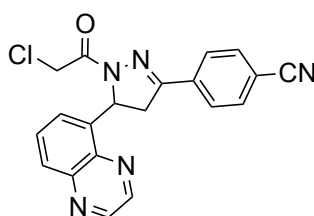
**2-chloro-1-(3-(4-methoxyphenyl)-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (LEB-02-173)**. General Procedure A was followed starting from 4-

methoxyacetophenone (1.2 equiv, 0.76 mmol) and quinoxaline-5-carbaldehyde (1.0 equiv, 0.63 mmol). The corresponding chalcone precipitated out during the reaction and was filtered by gravity filtration to yield product (0.47 mmol) as a light yellow solid. General Procedure F was followed starting from the above product (1.0 equiv, 0.47 mmol). Purification by flash column chromatography (EtOAc/hexane 50:50) yielded the chloroacetamide **LEB-02-173** (0.025 mmol, 4.0% yield across all steps). **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ 9.04 (q, *J* = 1.8 Hz, 2H), 8.05 (dd, *J* = 8.3, 1.4 Hz, 1H), 7.82 (dd, *J* = 8.5, 7.3 Hz, 1H), 7.77 – 7.70 (m, 2H), 7.53 (d, *J* = 7.2 Hz, 1H), 7.04 – 6.98 (m, 2H), 6.54 (dd, *J* = 11.9, 4.7 Hz, 1H), 4.90 (d, *J* = 13.8 Hz, 1H), 4.76 (d, *J* = 13.8 Hz, 1H), 4.04 (dd, *J* = 18.0, 11.9 Hz, 1H), 3.80 (s, 3H), 3.21 (dd, *J* = 18.0, 4.8 Hz, 1H). **HRMS** (ESI): exact mass calculated for C<sub>20</sub>H<sub>17</sub>ClN<sub>4</sub>O<sub>2</sub> [(M+H)<sup>+</sup>] 381.1113, found 381.1088.



LEB-02-174

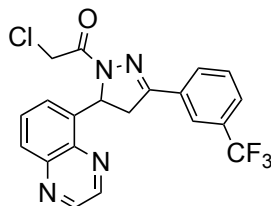
**2-chloro-1-(5-(quinoxalin-5-yl)-3-(4-(trifluoromethyl)phenyl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (LEB-02-174)**. General Procedure A was followed starting from 4-(trifluoromethyl)acetophenone (1.2 equiv, 0.76 mmol) and quinoxaline-5-carbaldehyde (1.0 equiv, 0.63 mmol). The corresponding chalcone precipitated out during the reaction and was filtered by gravity filtration to yield product (0.55 mmol) as a yellow solid. General Procedure F was followed starting from the above product (1.0 equiv, 0.55 mmol). Purification by flash column chromatography (EtOAc/hexane 50:50) yielded the chloroacetamide **LEB-02-174** (0.11 mmol, 18.9% yield across all steps). **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ 9.05 – 9.00 (m, 2H), 8.00 (d, *J* = 8.2 Hz, 2H), 7.85 – 7.79 (m, 3H), 7.49 – 7.43 (m, 2H), 4.97 – 4.77 (m, 1H), 4.17 – 3.99 (m, 1H), 3.40 (s, 2H), 3.31 (dd, *J* = 18.2, 5.1 Hz, 1H). **HRMS** (ESI): exact mass calculated for C<sub>20</sub>H<sub>14</sub>ClF<sub>3</sub>N<sub>4</sub>O [(M+H)<sup>+</sup>] 419.0881, found 419.0865.



LEB-02-175

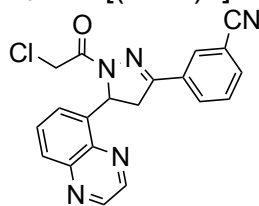
**4-(1-(2-chloroacetyl)-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-3-yl)benzotrile (LEB-02-175)**. General Procedure A was followed starting from 4-acetylbenzotrile (1.2 equiv, 0.76 mmol) and quinoxaline-5-carbaldehyde (1.0 equiv, 0.63 mmol). The corresponding chalcone precipitated out during the reaction and was filtered by gravity filtration to yield product (0.57 mmol) as a yellow solid. General Procedure F was followed starting from the above product (1.0 equiv, 0.57 mmol). Purification by flash column chromatography (EtOAc/hexane 50:50) yielded the chloroacetamide **LEB-02-175** (0.050 mmol, 7.9% yield across all steps). **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ 9.06 – 9.00 (m, 2H), 8.06 (dd, *J* = 8.4, 1.4 Hz, 1H), 8.00 – 7.91 (m, 4H), 7.83 (dd, *J* = 8.5, 7.3 Hz, 1H), 7.58 (dd, *J* = 7.3, 1.4 Hz, 1H), 4.15 – 3.99 (m, 2H), 3.39 (s, 2H), 3.31 (dd, *J* = 18.2, 5.2 Hz,

1H). **HRMS** (ESI): exact mass calculated for C<sub>20</sub>H<sub>14</sub>ClN<sub>5</sub>O [(M+H)<sup>+</sup>] 376.096, found 376.0977.



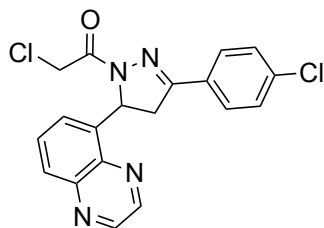
LEB-02-176

**2-chloro-1-(5-(quinoxalin-5-yl)-3-(3-(trifluoromethyl)phenyl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (LEB-02-176)**. General Procedure A was followed starting from 3-(trifluoromethyl)acetophenone (1.2 equiv, 0.76 mmol) and quinoxaline-5-carbaldehyde (1.0 equiv, 0.63 mmol). The corresponding chalcone precipitated out during the reaction and was filtered by gravity filtration to yield product (0.57 mmol) as a yellow solid. General Procedure F was followed starting from the above product (1.0 equiv, 0.57 mmol). Purification by flash column chromatography (EtOAc/hexane 50:50) yielded the chloroacetamide **LEB-02-176** (0.22 mmol, 35.3% yield across all steps). **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ 9.06 – 9.03 (m, 2H), 8.12 – 8.04 (m, 3H), 7.84 (dt, *J* = 8.5, 6.7 Hz, 2H), 7.71 (t, *J* = 7.8 Hz, 1H), 7.57 (dd, *J* = 7.2, 1.4 Hz, 1H), 6.58 (dd, *J* = 12.0, 5.0 Hz, 1H), 4.97 (d, *J* = 14.1 Hz, 1H), 4.84 (d, *J* = 14.1 Hz, 1H), 4.12 (dd, *J* = 18.3, 12.1 Hz, 1H), 3.37 – 3.34 (m, 1H). **<sup>13</sup>C NMR** (151 MHz, DMSO) δ 164.1, 155.7, 146.3, 145.3, 143.1, 139.7, 139.3, 132.3, 131.3, 130.5, 130.5, 130.4, 130.2, 130.0, 129.8, 129.0, 127.5, 127.4, 127.4, 127.1, 125.9, 125.3, 123.8, 123.8, 123.7, 123.5, 121.7, 57.0, 43.0, 42.1. **HRMS** (ESI): exact mass calculated for C<sub>20</sub>H<sub>14</sub>ClF<sub>3</sub>N<sub>4</sub>O [(M+H)<sup>+</sup>] 419.0881, found 419.0857.



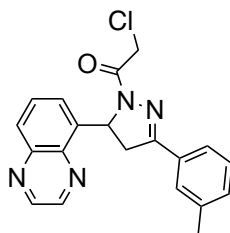
LEB-02-182

**3-(1-(2-chloroacetyl)-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-3-yl)benzonitrile (LEB-02-182)**. General Procedure A was followed starting from 3-acetylbenzonitrile (1.2 equiv, 0.76 mmol) and quinoxaline-5-carbaldehyde (1.0 equiv, 0.63 mmol). The corresponding chalcone precipitated out during the reaction and was filtered by gravity filtration to yield product (0.46 mmol) as a light yellow solid. General Procedure F was followed starting from the above product (1.0 equiv, 0.46 mmol). Purification by flash column chromatography (EtOAc/hexane 50:50) yielded the chloroacetamide **LEB-02-182** (0.025 mmol, 4.0% yield across all steps). **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ 9.09 – 9.00 (m, 2H), 8.28 – 8.23 (m, 1H), 8.14 (dt, *J* = 8.0, 1.4 Hz, 1H), 8.07 (dd, *J* = 8.4, 1.4 Hz, 1H), 7.96 (dt, *J* = 7.8, 1.4 Hz, 1H), 7.83 (dd, *J* = 8.4, 7.2 Hz, 1H), 7.68 (t, *J* = 7.9 Hz, 1H), 7.58 – 7.55 (m, 1H), 6.57 (dd, *J* = 11.9, 5.0 Hz, 1H), 4.96 (d, *J* = 14.1 Hz, 1H), 4.83 (d, *J* = 14.1 Hz, 1H), 4.14 – 4.01 (m, 1H), 3.34 – 3.27 (m, 1H). **HRMS** (ESI): exact mass calculated for C<sub>20</sub>H<sub>14</sub>ClN<sub>5</sub>O [(M+H)<sup>+</sup>] 376.096, found 376.097.



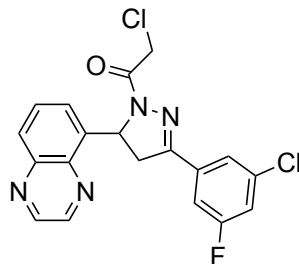
LEB-02-187

**2-chloro-1-(3-(4-chlorophenyl)-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (LEB-02-187).** General Procedure A was followed starting from 4-chloroacetophenone (1.2 equiv, 0.76 mmol) and quinoxaline-5-carbaldehyde (1.0 equiv, 0.63 mmol). The corresponding chalcone precipitated out during the reaction and was filtered by gravity filtration to yield product (0.63 mmol) as a light yellow solid. General Procedure F was followed starting from the above crude product (1.0 equiv, 0.63 mmol). Purification by flash column chromatography (EtOAc/hexane 50:50) yielded the chloroacetamide **LEB-02-187** (0.12 mmol, 19.2% yield across all steps). **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ 9.08 – 8.98 (m, 2H), 8.06 (dd, *J* = 8.4, 1.4 Hz, 1H), 7.87 – 7.77 (m, 3H), 7.58 – 7.49 (m, 3H), 6.55 (dd, *J* = 11.9, 5.0 Hz, 1H), 4.91 (d, *J* = 13.9 Hz, 1H), 4.78 (d, *J* = 13.9 Hz, 1H), 4.07 (dd, *J* = 18.2, 12.0 Hz, 1H), 3.26 (dd, *J* = 18.1, 5.0 Hz, 1H). **HRMS** (ESI): exact mass calculated for C<sub>19</sub>H<sub>14</sub>Cl<sub>2</sub>N<sub>4</sub>O [(M+H)<sup>+</sup>] 385.0617, found 385.0628.



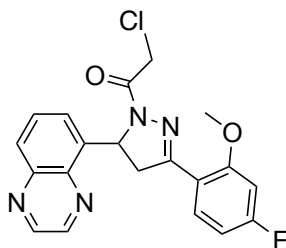
LEB-03-001c

**2-chloro-1-(5-(quinoxalin-5-yl)-3-(*m*-tolyl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (LEB-03-001c).** General Procedure A was followed starting from 3-methylacetophenone (1.2 equiv, 0.76 mmol) and quinoxaline-5-carbaldehyde (1.0 equiv, 0.63 mmol). The corresponding chalcone precipitated out during the reaction and was filtered by gravity filtration to yield product (0.53 mmol) as a yellow solid. General Procedure F was followed starting from the above crude product (1.0 equiv, 0.53 mmol). Purification by flash column chromatography (EtOAc/hexane 50:50) yielded the chloroacetamide **LEB-03-001c** (0.29 mmol, 46.7% yield across all steps). **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ 9.28 – 8.93 (m, 2H), 8.17 – 7.87 (m, 2H), 7.85 – 7.79 (m, 1H), 7.69 – 7.49 (m, 1H), 7.40 – 7.13 (m, 2H), 5.01 – 4.71 (m, 2H), 3.27 – 3.04 (m, 1H), 2.51 (s, 3H), 2.35 (d, *J* = 14.1 Hz, 3H). **HRMS** (ESI): exact mass calculated for C<sub>20</sub>H<sub>17</sub>ClN<sub>4</sub>O [(M+H)<sup>+</sup>] 365.1164, found 365.1185.



LEB-03-004c

**2-chloro-1-(3-(4-fluoro-2-methoxyphenyl)-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (LEB-03-004c).** General Procedure A was followed starting from 3-chloro-5-fluoroacetophenone (1.2 equiv, 0.76 mmol) and quinoxaline-5-carbaldehyde (1.0 equiv, 0.63 mmol). The corresponding chalcone precipitated out during the reaction and was filtered by gravity filtration to yield product (0.51 mmol) as a light yellow solid. General Procedure F was followed starting from the above crude product (1.0 equiv, 0.51 mmol). Purification by flash column chromatography (EtOAc/hexane 50:50) yielded the chloroacetamide **LEB-03-004c** (0.23 mmol, 37.2% yield across all steps). **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ 9.04 (dd, *J* = 4.3, 1.8 Hz, 2H), 8.08 – 7.91 (m, 2H), 7.85 (dt, *J* = 15.5, 7.9 Hz, 1H), 7.54 (d, *J* = 7.2 Hz, 1H), 6.89 (qd, *J* = 8.7, 2.4 Hz, 1H), 6.50 (dd, *J* = 11.9, 4.8 Hz, 1H), 4.91 – 4.71 (m, 1H), 4.15 – 4.00 (m, 1H), 3.72 (s, 2H), 3.18 (dd, *J* = 18.6, 4.9 Hz, 1H), 2.51 (dt, *J* = 3.7, 1.9 Hz, 3H). **<sup>13</sup>C NMR** (151 MHz, DMSO) δ 164.2, 163.5, 161.8, 154.8, 154.8, 146.3, 145.3, 143.1, 139.7, 139.2, 135.1, 135.0, 134.9, 134.8, 130.5, 129.1, 126.0, 123.5, 123.5, 118.4, 118.2, 113.3, 113.1, 57.2, 51.8, 43.0, 42.1, 25.0, 16.7. **HRMS** (ESI): exact mass calculated for C<sub>20</sub>H<sub>16</sub>ClFN<sub>4</sub>O [(M+H)<sup>+</sup>] 403.0523, found 403.0546.

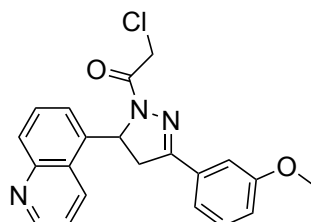


LEB-03-005c

**2-chloro-1-(3-(4-fluoro-2-methoxyphenyl)-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (LEB-03-005c).** General Procedure A was followed starting from 4-fluoro-2-methoxyacetophenone (1.2 equiv, 0.76 mmol) and quinoxaline-5-carbaldehyde (1.0 equiv, 0.63 mmol). The corresponding chalcone precipitated out during the reaction and was filtered by gravity filtration to yield product (0.51 mmol) as a white solid. General Procedure F was followed starting from the above crude product (1.0 equiv, 0.51 mmol). Purification by flash column chromatography (EtOAc/hexane 50:50) yielded the chloroacetamide **LEB-03-005c** (0.13 mmol, 20.0% yield across all steps). **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ 9.04 (dd, *J* = 4.3, 1.8 Hz, 2H), 8.08 – 7.91 (m, 2H), 7.85 (dt, *J* = 15.5, 7.9 Hz, 1H), 7.54 (d, *J* = 7.2 Hz, 1H), 6.89 (qd, *J* = 8.7, 2.4 Hz, 1H), 6.50 (dd, *J* = 11.9, 4.8 Hz, 1H), 4.91 – 4.71 (m, 1H), 4.15 – 4.00 (m, 1H), 3.72 (s, 2H), 3.18 (dd, *J* = 18.6, 4.9 Hz, 1H), 2.51 (dt, *J* = 3.7, 1.9 Hz, 3H). **<sup>13</sup>C NMR** (151 MHz, DMSO) δ 165.8, 164.1, 163.7, 160.3, 160.2, 155.1, 146.3, 145.3, 143.1, 139.6, 131.2, 131.1, 130.6, 128.9,



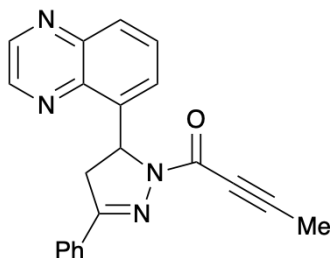
125.7, 116.7, 116.7, 108.1, 108.0, 101.1, 101.0, 56.8, 56.3, 45.3, 43.0. **HRMS** (ESI): exact mass calculated for C<sub>20</sub>H<sub>16</sub>ClFN<sub>4</sub>O [(M+H)<sup>+</sup>] 399.1019, found 399.1042



LEB-03-008c

**2-chloro-1-(3-(3-methoxyphenyl)-5-(quinolin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (LEB-03-008c)**. General Procedure A was followed starting from 3-methoxyacetophenone (1.2 equiv, 1.14 mmol) and quinoline-5-carbaldehyde (1.0 equiv, 0.95 mmol). The reaction was run for 15 minutes at 0°C. The corresponding chalcone was purified via flash chromatography (EtOAc/Hexane 50:50) to yield a light yellow solid (0.69 mmol). General Procedure F was followed starting from the above product (1.0 equiv, 0.69 mmol). Purification by flash column chromatography (EtOAc/hexane 50:50) yielded the chloroacetamide **LEB-03-008c** (0.29 mmol, 30.3% yield across all steps). **<sup>1</sup>H NMR** (400 MHz, DMSO-d<sub>6</sub>) δ 8.99 (dd, J = 4.2, 1.6 Hz, 1H), 8.67 (d, J = 8.6 Hz, 1H), 7.98 (d, J = 8.5 Hz, 1H), 7.75 – 7.61 (m, 2H), 7.44 – 7.27 (m, 4H), 4.89 (d, J = 30.0 Hz, 1H), 4.11 (dd, J = 18.2, 11.9 Hz, 1H), 3.79 (s, 3H), 3.36 (s, 2H), 3.26 (d, J = 4.9 Hz, 1H). **<sup>13</sup>C NMR** (151 MHz, DMSO) δ 167.9, 164.3, 164.0, 162.4, 160.0, 159.9, 157.8, 156.5, 156.4, 151.5, 151.2, 150.8, 148.6, 148.6, 148.0, 144.8, 143.9, 140.8, 140.7, 139.7, 139.6, 137.9, 137.6, 137.5, 135.9, 133.3, 132.5, 132.4, 132.2, 132.1, 132.0, 131.7, 130.6, 130.4, 129.7, 129.6, 129.3, 129.1, 128.7, 127.0, 125.8, 125.7, 125.1, 122.8, 122.6, 122.2, 122.0, 120.1, 119.9, 117.4, 117.2, 112.8, 112.6, 112.4, 59.9, 57.4, 56.9, 55.9, 55.8, 55.5, 43.0, 42.7, 42.6, 42.5, 42.3, 24.8. **HRMS** (ESI): exact mass calculated for C<sub>21</sub>H<sub>18</sub>ClN<sub>3</sub>O<sub>2</sub> [(M+H)<sup>+</sup>] 380.116, found 380.1185.

### 1-(3-Phenyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)but-2-yn-1-one (CMZ 07)

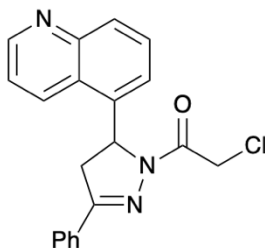


**Step 1:** Following General Procedure A, the chalcone was prepared from quinoxaline-5-carbaldehyde (500 mg, 3.16 mmol), acetophenone (443 μL, 3.79 mmol) and 5% (w/w) aq NaOH (3.8 mL, 4.74 mmol) in EtOH (10 mL). Filtration under reduced pressure gave the chalcone as a pale yellow solid (823 mg, quant) which was used without purification.

*Step 2:* Hydrazine hydrate (136  $\mu$ L, 1.54 mmol, 50–60% wt solution) was added to a solution of the chalcone (200 mg, 0.77 mmol) in EtOH (2 mL) at rt and the resultant mixture was heated at reflux for 4 h before being cooled to rt. One half of the reaction mixture was then concentrated *in vacuo* and the crude pyrazoline was used immediately without purification.

*Step 3:* A solution of tetrolic acid (32 mg, 0.38 mmol) in DMF (0.5 mL) was added to a solution of the crude pyrazoline (0.38 mmol) in DMF (0.5 mL) at 0 °C. The reaction mixture was then treated sequentially with NMM (422  $\mu$ L, 3.84 mmol) and T3P (642  $\mu$ L, 1.08 mmol, 50 wt% in EtOAc) and allowed to warm to rt and stirred at rt for 16 h. Water (2 mL) and EtOAc (2 mL) were then added and the aqueous layer was extracted with EtOAc (3  $\times$  2 mL). The combined organic extracts were washed sequentially with satd aq NaHCO<sub>3</sub> (6 mL) and brine (6 mL), then dried and concentrated *in vacuo*. Purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 55%) gave **CM 7** as a white solid (47 mg, 36% over 2 steps from the respective chalcone); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$ <sub>H</sub> 2.08 (s, 3H), 3.11 (dd, *J* = 17.8, 4.8 Hz, 1H), 3.98 (dd, *J* = 17.8, 11.8 Hz, 1H), 6.62 (dd, *J* = 11.8, 4.8 Hz, 1H), 7.25 – 7.40 (m, 3H), 7.46 (d, *J* = 7.2 Hz, 1H), 7.59 – 7.75 (m, 3H), 7.92 – 7.99 (m, 1H), 8.77 – 8.83 (m, 2H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>)  $\delta$ <sub>C</sub> 4.5, 42.5, 56.7, 73.9, 90.3, 125.6, 126.9, 128.7, 129.0, 130.1, 130.7, 131.1, 138.3, 140.2, 143.2, 143.9, 144.7, 151.3, 156.8. HRMS (ESI): exact mass calculated for C<sub>21</sub>H<sub>16</sub>N<sub>4</sub>O [(M+H)<sup>+</sup>] 341.1397, found 341.1376.

## 2-Chloro-1-(3-phenyl-5-(quinolin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (CMZ 13)

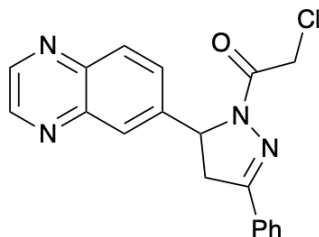


*Step 1:* Following General Procedure A, the chalcone was prepared from quinoline-5-carbaldehyde (55 mg, 0.35 mmol), acetophenone (49  $\mu$ L, 0.42 mmol) and 5% (w/w) aq NaOH (0.4 mL, 0.52 mmol) in EtOH (1 mL). Filtration under reduced pressure gave the chalcone as a brown solid (58 mg, 65%).

*Step 2:* Following General Procedure F, the pure title compound was prepared from chalcone (58 mg, 0.22 mmol) and hydrazine hydrate (40  $\mu$ L, 0.45 mmol, 50–60% wt solution) in EtOH (0.9 mL); then chloroacetyl chloride (27  $\mu$ L, 0.34 mmol) and triethylamine (94  $\mu$ L, 0.67 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (0.9 mL). Purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 61%) gave **CMZ 13** as a white solid (35 mg, 44% over 2 steps from the respective chalcone). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$ <sub>H</sub> 3.13 (dd, *J* = 17.6, 4.9 Hz, 1H), 3.90 (dd, *J* = 17.6, 11.9 Hz, 1H), 4.55 (d, *J* = 13.4 Hz, 1H), 4.65 (d, *J* = 13.4 Hz, 1H), 6.21 (dd, *J* = 11.9, 4.9 Hz, 1H), 7.27 (d, *J* = 7.2 Hz, 1H), 7.31 – 7.44 (m, 4H), 7.57 (dd, *J* = 8.6, 7.2 Hz, 1H), 7.63 – 7.71 (m, 2H), 7.98 (d, *J* = 8.6 Hz, 1H), 8.29 (d, *J* = 8.6 Hz, 1H), 8.89 (dd, *J* = 4.3,

1.6 Hz, 1H).  $^{13}\text{C}$  NMR (151 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  42.1, 42.4, 57.1, 121.4, 123.3, 125.2, 127.0, 129.0, 129.2, 130.0, 130.6, 131.2, 132.3, 136.6, 147.9, 149.6, 155.9, 164.5. **HRMS (ESI)**: exact mass calculated for  $\text{C}_{20}\text{H}_{16}\text{ClN}_3\text{O}$   $[(\text{M}+\text{H})^+]$  350.1055, found 350.1029.

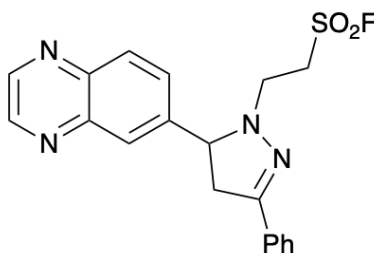
**2-Chloro-1-(3-phenyl-5-(quinoxalin-6-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (CMZ 15)**



*Step 1*: Following General Procedure A, the chalcone was prepared from quinoxaline-6-carbaldehyde (301 mg, 1.90 mmol), acetophenone (267  $\mu\text{L}$ , 2.28 mmol) and 5% (w/w) aq NaOH (2.28 mL, 2.85 mmol) in EtOH (9 mL). Filtration under reduced pressure gave the chalcone as a pale brown solid (445 mg, 90%).

*Step 2*: Following General Procedure F, the pure title compound was prepared from chalcone (100 mg, 0.38 mmol) and hydrazine hydrate (68  $\mu\text{L}$ , 0.77 mmol, 50–60% wt solution) in EtOH (1.6 mL); then chloroacetyl chloride (46  $\mu\text{L}$ , 0.58 mmol) and triethylamine (161  $\mu\text{L}$ , 1.15 mmol) in  $\text{CH}_2\text{Cl}_2$  (1.5 mL). Purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 59%) gave **CMZ 15** as a white solid (50 mg, 37% over 2 steps from the respective chalcone);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  3.23 (dd,  $J = 17.9, 4.9$  Hz, 1H), 3.84 (dd,  $J = 17.9, 11.8$  Hz, 1H), 4.49 – 4.61 (m, 2H), 5.74 (dd,  $J = 11.8, 4.9$  Hz, 1H), 7.32 – 7.44 (m, 3H), 7.60 (dd,  $J = 8.7, 2.1$  Hz, 1H), 7.61 – 7.72 (m, 2H), 7.91 (d,  $J = 2.1$  Hz, 1H), 8.02 (d,  $J = 8.7$  Hz, 1H), 8.73 (app s, 2H).  $^{13}\text{C}$  NMR (151 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  42.1, 42.1, 60.2, 126.0, 126.8, 128.0, 128.8, 130.4, 130.5, 131.0, 142.4, 142.8, 142.9, 145.0, 145.2, 155.3, 164.2. **HRMS (ESI)**: exact mass calculated for  $\text{C}_{19}\text{H}_{15}\text{ClN}_4\text{O}$   $[(\text{M}+\text{MeCN}+\text{H})^+]$  392.1272, found 392.1243.

**2-(3-Phenyl-5-(193quinoxaline-6-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethane-1-sulfonyl fluoride (CMZ 17)**



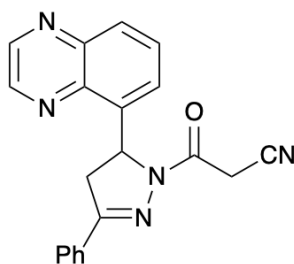
*Step 1*: Following General Procedure A, the chalcone was prepared from quinoxaline-6-carbaldehyde (301 mg, 1.90 mmol), acetophenone (267  $\mu\text{L}$ , 2.28 mmol) and 5% (w/w) aq NaOH (2.3 mL, 2.85 mmol) in EtOH (9 mL). Filtration under reduced pressure gave the chalcone as a pale brown solid (495 mg, 90%) which was used without purification.

*Step 2*: Hydrazine hydrate (48  $\mu\text{L}$ , 0.54 mmol, 50–60% wt solution) was added to a solution of the chalcone (71 mg, 0.27 mmol) in EtOH (1 mL) at rt and the resultant mixture

was heated at reflux for 1.5 h before being cooled to rt. One third of the reaction mixture was then concentrated *in vacuo* and the crude pyrazoline was used immediately without purification.

**Step 3:** Ethene sulfonyl fluoride (25  $\mu$ L, 0.30 mmol) was added to a solution of the crude pyrazoline (0.27 mmol) in THF (0.6 mL) at rt and stirred at 80°C for 4 h, then concentrated *in vacuo*. Purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 43%) gave **CMZ 17** as a pale yellow oil (63 mg, 60% over 2 steps from the respective chalcone); **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta_{\text{H}}$  3.06 (dd,  $J = 16.3, 13.8$  Hz, 1H), 3.22 – 3.33 (m, 1H), 3.39 (app dt,  $J = 12.6, 7.7$  Hz, 1H), 3.54 (dd,  $J = 16.3, 10.0$  Hz, 1H), 3.84 – 3.99 (m, 2H), 4.45 (dd,  $J = 13.8, 10.0$  Hz, 1H), 7.26 – 7.42 (m, 3H), 7.55 – 7.63 (m, 2H), 7.89 (dd,  $J = 8.7, 2.0$  Hz, 1H), 8.05 – 8.13 (m, 2H), 8.80 (app s, 2H). **<sup>13</sup>C NMR** (151 MHz, CDCl<sub>3</sub>)  $\delta_{\text{C}}$  43.2, 47.7, 49.5, 49.6, 71.2, 126.1, 128.2, 128.6, 129.0, 129.5, 130.3, 131.8, 141.6, 142.7, 142.9, 145.1, 145.2, 151.6. **HRMS (ESI):** exact mass calculated for C<sub>19</sub>H<sub>17</sub>FN<sub>4</sub>O<sub>2</sub>S [(M+H)<sup>+</sup>] 385.1129, found 385.1111.

### 3-Oxo-3-(3-phenyl-5-(quinoxaline-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)propanenitrile (CMZ 22)



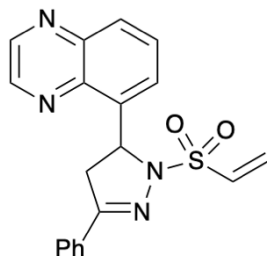
**Step 1:** Following General Procedure A, the chalcone was prepared from quinoxaline-5-carbaldehyde (500 mg, 3.16 mmol), acetophenone (443  $\mu$ L, 3.79 mmol) and 5% (w/w) aq NaOH (3.8 mL, 4.74 mmol) in EtOH (10 mL). Filtration under reduced pressure gave the chalcone as a pale yellow solid (823 mg, quant) which was used without purification.

**Step 2:** Hydrazine hydrate (68  $\mu$ L, 0.77 mmol, 50–60% wt solution) was added to a solution of the chalcone (100 mg, 0.38 mmol) in EtOH (2 mL) at rt and the resultant mixture was heated at reflux for 2 h before being cooled to rt. The reaction mixture was then concentrated *in vacuo* and the crude pyrazoline was used immediately without purification.

**Step 3:** A solution of cyanoacetic acid (33 mg, 0.38 mmol) in DMF (0.5 mL) was added to a solution of the crude pyrazoline (0.38 mmol) in DMF (0.5 mL) at 0 °C. The reaction mixture was then treated sequentially with NMM (422  $\mu$ L, 3.84 mmol) and T3P (643  $\mu$ L, 1.08 mmol, 50 wt% in EtOAc) and allowed to warm to rt and stirred at rt for 16 h. Water (2 mL) and EtOAc (2 mL) were then added and the aqueous layer was extracted with EtOAc (3  $\times$  2 mL). The combined organic extracts were washed sequentially with satd aq NaHCO<sub>3</sub> (6 mL) and brine (6 mL), then dried and concentrated *in vacuo*. Purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 48%) gave **CMZ 22** as a white solid (11 mg, 8% over 2 steps from the respective chalcone); **<sup>1</sup>H NMR** (600 MHz, DMSO-*d*<sub>6</sub>)  $\delta_{\text{H}}$  3.27 (dd,  $J = 18.2, 5.0$  Hz, 1H), 4.10 (dd,  $J = 18.2, 11.9$  Hz, 1H), 4.36 (d,  $J = 19.2$  Hz, 1H), 4.48 (d,  $J = 19.1$  Hz, 1H), 6.55 (dd,  $J = 11.9, 5.0$  Hz, 1H), 7.43 – 7.53 (m, 3H), 7.59 (dd,  $J = 7.4, 1.3$  Hz, 1H), 7.78 – 7.82

(m, 2H), 7.82 – 7.86 (m, 1H), 8.06 (dd,  $J = 8.4, 1.3$  Hz, 1H), 9.04 (dd,  $J = 10.5, 1.8$  Hz, 2H).  $^{13}\text{C}$  NMR (151 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  25.5, 40.1, 42.3, 56.3, 116.0, 125.5, 127.0, 128.7, 128.9, 130.1, 130.7, 130.9, 138.9, 139.3, 142.7, 144.9, 145.9, 156.7, 160.5. HRMS (ESI): exact mass calculated for  $\text{C}_{20}\text{H}_{15}\text{N}_5\text{O}$   $[(\text{M}+\text{H})^+]$  342.135, found 342.1336.

### 5-(3-Phenyl-1-(vinylsulfonyl)-4,5-dihydro-1H-pyrazol-5-yl)quinoxaline (CMZ 23)

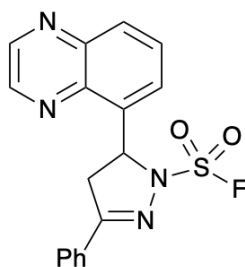


**Step 1:** Following General Procedure A, the chalcone was prepared from quinoxaline-5-carbaldehyde (500 mg, 3.16 mmol), acetophenone (443  $\mu\text{L}$ , 3.79 mmol) and 5% (w/w) aq NaOH (3.8 mL, 4.74 mmol) in EtOH (10 mL). Filtration under reduced pressure gave the chalcone as a pale yellow solid (823 mg, quant) which was used without purification.

**Step 2:** Hydrazine hydrate (204  $\mu\text{L}$ , 2.31 mmol, 50–60% wt solution) was added to a solution of the chalcone (300 mg, 1.15 mmol) in EtOH (4.5 mL) at rt and the resultant mixture was heated at reflux for 2 h before being cooled to rt. The reaction mixture was then concentrated *in vacuo* and the crude pyrazoline was used immediately without purification.

**Step 3:** Following General Procedure G, the pure title compound was prepared from crude pyrazoline (1.15 mmol), triphenylphosphine oxide (712 mg, 2.56 mmol), trifluoromethanesulfonic anhydride (194  $\mu\text{L}$ , 1.15 mmol), vinyl sulfonic acid (125 mg, 1.15 mmol), pyridine (93  $\mu\text{L}$ , 1.15 mmol) and triethylamine (321  $\mu\text{L}$ , 2.31 mmol) in  $\text{CH}_2\text{Cl}_2$  (17 mL). Purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 53%; then SNAP Ultra C18, 20% grading to 100% MeCN in water, product eluted at 43%) gave **CMZ 23** as a white solid (33 mg, 24% over 2 steps from the respective chalcone);  $^1\text{H}$  NMR (400 MHz,  $\text{DMSO}-d_6$ )  $\delta_{\text{H}}$  3.29 (dd,  $J = 17.9, 9.0$  Hz, 1H), 4.14 (dd,  $J = 17.9, 11.9$  Hz, 1H), 6.18 (dd,  $J = 11.9, 9.0$  Hz, 1H), 6.26 – 6.40 (m, 2H), 7.02 (dd,  $J = 16.5, 10.0$  Hz, 1H), 7.42 – 7.54 (m, 3H), 7.73 – 7.79 (m, 2H), 7.90 – 7.97 (m, 1H), 7.98 – 8.02 (m, 1H), 8.10 (dd,  $J = 8.4, 1.7$  Hz, 1H), 8.95 – 9.08 (m, 2H). HRMS (ESI): exact mass calculated for  $\text{C}_{19}\text{H}_{16}\text{N}_4\text{O}_2\text{S}$   $[(\text{M}+\text{H})^+]$  365.1067, found 365.1061.

### 3-Phenyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazole-1-sulfonyl fluoride (CMZ 24)

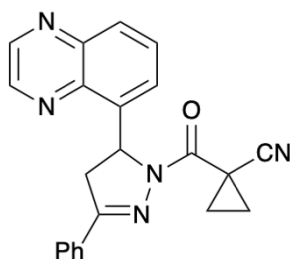


**Step 1:** Following General Procedure A, the chalcone was prepared from quinoxaline-5-carbaldehyde (500 mg, 3.16 mmol), acetophenone (443  $\mu$ L, 3.79 mmol) and 5% (w/w) aq NaOH (3.8 mL, 4.74 mmol) in EtOH (10 mL). Filtration under reduced pressure gave the chalcone as a pale yellow solid (823 mg, quant) which was used without purification.

**Step 2:** Hydrazine hydrate (204  $\mu$ L, 2.31 mmol, 50–60% wt solution) was added to a solution of the chalcone (300 mg, 1.15 mmol) in EtOH (4.5 mL) at rt and the resultant mixture was heated at reflux for 4 h before being cooled to rt. One third of the reaction mixture was then concentrated *in vacuo* and the crude pyrazoline was used immediately without purification.

**Step 3:** 1-(Fluorosulfonyl)-2,3-dimethyl-1H-imidazol-3-ium trifluoromethanesulfonate (126 mg, 0.38 mmol) was added to a solution of the crude pyrazoline (0.38 mmol) in MeCN (0.7 mL) at rt and stirred at this temperature for 16 h. Water (2 mL) was then added and the aqueous layer was extracted with EtOAc (3  $\times$  2 mL). The combined organic extracts were washed sequentially with brine (2  $\times$  2 mL), then dried and concentrated *in vacuo*. Purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 25%) gave **CMZ 24** as a white solid (34 mg, 25% over 2 steps from the respective chalcone); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$ <sub>H</sub> 3.24 (dd, *J* = 17.7, 8.2 Hz, 1H), 4.18 (dd, *J* = 17.7, 11.4 Hz, 1H), 6.50 (dd, *J* = 11.4, 8.2 Hz, 1H), 7.30 – 7.46 (m, 3H), 7.68 – 7.74 (m, 2H), 7.74 – 7.78 (m, 1H), 7.84 – 7.92 (m, 1H), 8.04 (dd, *J* = 8.5, 1.6 Hz, 1H), 8.73 – 8.88 (m, 2H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>)  $\delta$ <sub>C</sub> 44.3, 61.4, 127.2, 127.6, 129.0, 129.8, 130.0, 130.3, 131.7, 138.0, 140.3, 143.2, 144.3, 145.4, 160.2. HRMS (ESI): exact mass calculated for C<sub>17</sub>H<sub>13</sub>FN<sub>4</sub>O<sub>2</sub>S [(M+MeCN+H)<sup>+</sup>] 398.1081, found 398.1097.

### 1-(3-Phenyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazole-1-carbonyl)cyclopropane-1-carbonitrile (CMZ 29)

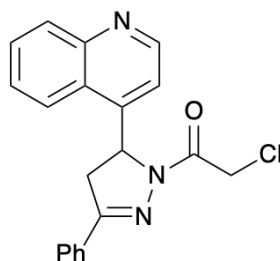


**Step 1:** Following General Procedure A, the chalcone was prepared from quinoxaline-5-carbaldehyde (500 mg, 3.16 mmol), acetophenone (443  $\mu$ L, 3.79 mmol) and 5% (w/w) aq NaOH (3.8 mL, 4.74 mmol) in EtOH (10 mL). Filtration under reduced pressure gave the chalcone as a pale yellow solid (823 mg, quant) which was used without purification.

*Step 2:* Hydrazine hydrate (204  $\mu\text{L}$ , 2.31 mmol, 50–60% wt solution) was added to a solution of the chalcone (300 mg, 1.15 mmol) in EtOH (4.5 mL) at rt and the resultant mixture was heated at reflux for 4 h before being cooled to rt. One third of the reaction mixture was then concentrated *in vacuo* and the crude pyrazoline was used immediately without purification.

*Step 3:* A solution of 1-cyano-1-cyclopropane carboxylic acid (85 mg, 0.77 mmol) in DMF (1 mL) was added to a solution of the crude pyrazoline (0.77 mmol) in DMF (1 mL) at 0  $^{\circ}\text{C}$ . The reaction mixture was then treated sequentially with NMM (845  $\mu\text{L}$ , 7.68 mmol) and T3P (1.3 mL, 2.15 mmol, 50 wt% in EtOAc) and allowed to warm to rt and stirred at rt for 16 h. Water (5 mL) and EtOAc (5 mL) were then added and the aqueous layer was extracted with EtOAc (3  $\times$  5 mL). The combined organic extracts were washed sequentially with satd aq  $\text{NaHCO}_3$  (15 mL) and brine (15 mL), then dried and concentrated *in vacuo*. Purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 39%) gave **CMZ 29** as a white solid (41 mg, 15% over 2 steps from the respective chalcone);  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  1.57 – 1.66 (m, 3H), 1.72 (ddd,  $J = 9.2, 4.7, 3.2$  Hz, 1H), 3.13 (dd,  $J = 17.9, 5.4$  Hz, 1H), 3.95 (dd,  $J = 17.9, 11.8$  Hz, 1H), 6.63 (dd,  $J = 11.8, 5.4$  Hz, 1H), 7.27 – 7.42 (m, 3H), 7.48 (d,  $J = 7.2$  Hz, 1H), 7.67 (app t,  $J = 8.0$  Hz, 1H), 7.74 – 7.87 (m, 2H), 7.93 – 8.04 (m, 1H), 8.60 – 8.95 (m, 2H).  $^{13}\text{C NMR}$  (151 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  14.2, 17.9, 18.6, 42.1, 57.4, 120.9, 125.7, 127.3, 129.0, 129.4, 130.3, 131.0, 131.1, 138.9, 140.2, 143.2, 144.3, 145.1, 156.4, 162.5. **HRMS (ESI):** exact mass calculated for  $\text{C}_{22}\text{H}_{17}\text{N}_5\text{O}$  [(M+H) $^+$ ] 368.1506, found 368.1524.

### 2-Chloro-1-(3-phenyl-5-(quinolin-4-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (CMZ 35)

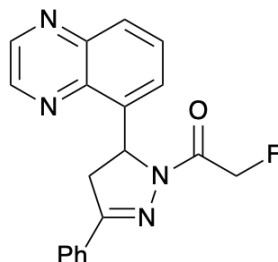


*Step 1:* Following General Procedure A, the chalcone was prepared from quinoline-4-carbaldehyde (137 mg, 0.87 mmol), acetophenone (204  $\mu\text{L}$ , 1.74 mmol) and 5% (w/w) aq NaOH (0.3 mL, 0.38 mmol) in EtOH (5 mL). Extraction with EtOAc and purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 42%) gave the chalcone as a yellow solid (39 mg, 17%).

*Step 2:* Following General Procedure F, the pure title compound was prepared from chalcone (39 mg, 0.15 mmol) and hydrazine hydrate (27  $\mu\text{L}$ , 0.30 mmol, 50–60% wt solution) in EtOH (1.5 mL); then chloroacetyl chloride (18  $\mu\text{L}$ , 0.23 mmol) and triethylamine (63  $\mu\text{L}$ , 0.45 mmol) in  $\text{CH}_2\text{Cl}_2$  (1.5 mL). Purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 46%) gave **CMZ 35** as a white solid (31 mg, 58% over 2 steps from the respective chalcone);  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  3.12 (dd,  $J = 17.6, 5.1$  Hz, 1H), 3.95 (dd,  $J = 17.6, 12.0$  Hz, 1H), 4.53 (d,  $J = 13.2$  Hz, 1H), 4.69 (d,  $J = 13.2$  Hz, 1H), 6.23 (dd,  $J = 12.0, 5.1$  Hz, 1H), 7.09 (d,  $J = 4.5$  Hz, 1H), 7.28 – 7.41 (m, 4H), 7.52 – 7.59 (m, 1H), 7.61 –

7.67 (m, 2H), 7.67 – 7.71 (m, 1H), 7.88 – 7.93 (m, 1H), 8.07 – 8.12 (m, 1H), 8.77 (d,  $J = 4.5$  Hz, 1H).  $^{13}\text{C}$  NMR (151 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  41.9, 42.0, 57.1, 116.5, 122.9, 125.1, 127.0, 127.8, 129.1, 129.9, 130.3, 130.3, 131.4, 146.6, 147.5, 149.5, 156.0, 164.7. HRMS (ESI): exact mass calculated for  $\text{C}_{20}\text{H}_{16}\text{ClN}_3\text{O}$   $[(\text{M}+\text{H})^+]$  350.1055, found 350.1066.

**2-Fluoro-1-(3-phenyl-5-(quinoxalin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (CMZ 38)**



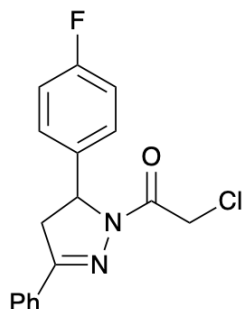
*Step 1:* Following General Procedure A, the chalcone was prepared from quinoxaline-5-carbaldehyde (500 mg, 3.16 mmol), acetophenone (443  $\mu\text{L}$ , 3.79 mmol) and 5% (w/w) aq NaOH (3.8 mL, 4.74 mmol) in EtOH (10 mL). Filtration under reduced pressure gave the chalcone as a pale yellow solid (823 mg, quant) which was used without purification.

*Step 2:* Hydrazine hydrate (94  $\mu\text{L}$ , 1.06 mmol, 50–60% wt solution) was added to a solution of the chalcone (138 mg, 0.53 mmol) in EtOH (2.2 mL) at rt and the resultant mixture was heated at reflux for 2 h before being cooled to rt. The reaction mixture was then concentrated *in vacuo* and the crude pyrazoline was used immediately without purification.

*Step 3:* A solution of fluoroacetic acid (41 mg, 0.53 mmol) in DMF (0.7 mL) was added to a solution of the crude pyrazoline (0.53 mmol) in DMF (0.7 mL) at 0 °C. The reaction mixture was then treated sequentially with NMM (583  $\mu\text{L}$ , 5.30 mmol) and T3P (887  $\mu\text{L}$ , 1.48 mmol, 50 wt% in EtOAc) and allowed to warm to rt and stirred at rt for 16 h. Water (2 mL) and EtOAc (2 mL) were then added and the aqueous layer was extracted with EtOAc (3  $\times$  2 mL). The combined organic extracts were washed sequentially with satd aq  $\text{NaHCO}_3$  (6 mL) and brine (6 mL), then dried and concentrated *in vacuo*. Purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 46%) gave **CMZ 38** as a white solid (20 mg, 11% over 2 steps from the respective chalcone);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  3.12 (dd,  $J = 17.9, 5.3$  Hz, 1H), 3.94 (dd,  $J = 17.9, 11.9$  Hz, 1H), 5.21 – 5.63 (m, 2H), 6.59 (dd,  $J = 11.9, 5.3$  Hz, 1H), 7.28 – 7.41 (m, 3H), 7.42 – 7.50 (m, 1H), 7.58 – 7.69 (m, 3H), 7.98 (dd,  $J = 8.4, 1.4$  Hz, 1H), 8.69 – 8.89 (m, 2H).  $^{13}\text{C}$  NMR (151 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  42.1, 56.9, 78.7, 79.8, 126.1, 126.9, 129.0, 129.2, 130.3, 130.9, 131.0, 138.6, 140.4, 143.1, 144.1, 144.7, 156.7, 165.2, 165.3. HRMS (ESI): exact mass calculated for  $\text{C}_{19}\text{H}_{15}\text{FN}_4\text{O}$   $[(\text{M}+\text{H})^+]$  335.1303, found 335.1304.

**2-Chloro-1-(5-(4-fluorophenyl)-3-phenyl-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (CMZ 47)**

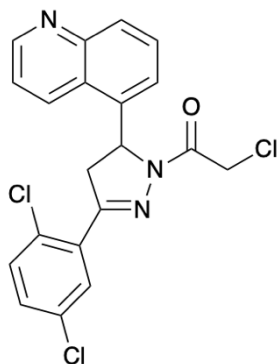




*Step 1:* Following General Procedure A, the chalcone was prepared from 4-fluorobenzaldehyde (259  $\mu$ L, 2.42 mmol), acetophenone (339  $\mu$ L, 2.90 mmol) and 5% (w/w) aq NaOH (2.9 mL, 3.63 mmol) in EtOH (3 mL). Filtration under reduced pressure gave the chalcone as a pale yellow solid (547 mg, quant).

*Step 2:* Following General Procedure F, the pure title compound was prepared from chalcone (100 mg, 0.44 mmol) and hydrazine hydrate (78  $\mu$ L, 0.88 mmol, 50–60% wt solution) in EtOH (3 mL); then chloroacetyl chloride (53  $\mu$ L, 0.66 mmol) and triethylamine (185  $\mu$ L, 1.33 mmol) in  $\text{CH}_2\text{Cl}_2$  (1.5 mL). Purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 16%) gave **CMZ 47** as a white solid (44 mg, 31% over 2 steps from the respective chalcone);  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  3.15 (dd,  $J = 17.9, 4.7$  Hz, 1H), 3.73 (dd,  $J = 17.9, 11.8$  Hz, 1H), 4.44 – 4.57 (m, 2H), 5.51 (dd,  $J = 11.8, 4.7$  Hz, 1H), 6.90 – 7.00 (m, 2H), 7.11 – 7.23 (m, 2H), 7.33 – 7.45 (m, 3H), 7.65 – 7.72 (m, 2H).  $^{13}\text{C NMR}$  (151 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  42.3, 42.3, 60.0, 115.9, 116.1, 127.0, 127.0, 127.8, 129.0, 129.0, 130.8, 131.1, 136.8, 136.8, 155.5, 161.6, 163.3, 164.1. **HRMS (ESI):** exact mass calculated for  $\text{C}_{17}\text{H}_{14}\text{ClFN}_2\text{O}$  [(M+H) $^+$ ] 317.0852, found 317.0845.

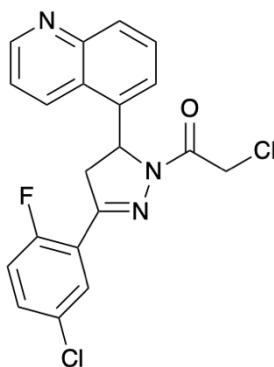
### 2-Chloro-1-(3-(2,5-dichlorophenyl)-5-(quinolin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (CMZ 51)



*Step 1:* Following General Procedure A, the chalcone was prepared from quinoline-5-carbaldehyde (200 mg, 1.27 mmol), 2',5'-dichloroacetophenone (367  $\mu$ L, 2.55 mmol) and 5% (w/w) aq NaOH (438  $\mu$ L, 0.55 mmol) in EtOH (4 mL). Filtration under reduced pressure and purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 32%) gave the chalcone as a yellow oil (168 mg, 40%).

*Step 2:* Following General Procedure F, the pure title compound was prepared from chalcone (168 mg, 0.51 mmol) and hydrazine hydrate (91  $\mu$ L, 1.03 mmol, 50–60% wt solution) in EtOH (4 mL); then chloroacetyl chloride (61  $\mu$ L, 0.77 mmol) and triethylamine (214  $\mu$ L, 1.54 mmol) in  $\text{CH}_2\text{Cl}_2$  (1.5 mL). Purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 45%) gave **CMZ 51** as a white solid (53 mg, 25% over 2 steps from the respective chalcone);  **$^1\text{H}$  NMR** (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  3.28 (dd,  $J = 18.1, 4.9$  Hz, 1H), 4.10 (dd,  $J = 18.1, 12.0$  Hz, 1H), 4.51 (d,  $J = 13.5$  Hz, 1H), 4.62 (d,  $J = 13.5$  Hz, 1H), 6.23 (dd,  $J = 12.0, 4.9$  Hz, 1H), 7.23 – 7.34 (m, 3H), 7.43 (dd,  $J = 8.6, 4.2$  Hz, 1H), 7.61 (dd,  $J = 8.6, 7.3$  Hz, 1H), 7.69 (d,  $J = 2.3$  Hz, 1H), 8.01 (d,  $J = 8.6$  Hz, 1H), 8.29 (d,  $J = 8.6$  Hz, 1H), 8.90 (dd,  $J = 4.2, 1.6$  Hz, 1H).  **$^{13}\text{C}$  NMR** (151 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  42.1, 44.9, 57.6, 121.5, 122.8, 125.0, 129.6, 129.8, 130.2, 131.2, 131.4, 131.5, 131.6, 132.4, 133.3, 136.0, 148.6, 150.1, 154.1, 164.7. **HRMS (ESI):** exact mass calculated for  $\text{C}_{20}\text{H}_{14}\text{Cl}_3\text{N}_3\text{O}$  [(M+H) $^+$ ] 418.0275, found 418.029.

**2-Chloro-1-(3-(5-chloro-2-fluorophenyl)-5-(quinolin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (CMZ 53)**

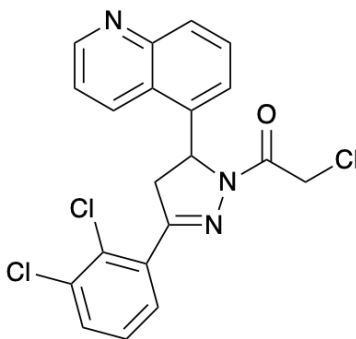


*Step 1:* Following General Procedure A, the chalcone was prepared from quinoline-5-carbaldehyde (207 mg, 1.31 mmol), 5'-chloro-2'-fluoroacetophenone (352  $\mu$ L, 2.63 mmol) and 5% (w/w) aq NaOH (452  $\mu$ L, 0.57 mmol) in EtOH (4 mL). Filtration under reduced pressure and purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 31%) gave the chalcone as a yellow solid (56 mg, 14%).

*Step 2:* Following General Procedure F, the pure title compound was prepared from chalcone (54 mg, 0.17 mmol) and hydrazine hydrate (31  $\mu$ L, 0.35 mmol, 50–60% wt solution) in EtOH (4 mL); then chloroacetyl chloride (21  $\mu$ L, 0.26 mmol) and triethylamine (72  $\mu$ L, 0.52 mmol) in  $\text{CH}_2\text{Cl}_2$  (1.5 mL). Purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 52%) gave **CMZ 53** as a white solid (31 mg, 44% over 2 steps from the respective chalcone);  **$^1\text{H}$  NMR** (400 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{H}}$  3.09 – 3.31 (m, 1H), 3.88 – 4.04 (m, 1H), 4.52 (app dd,  $J = 13.4, 1.8$  Hz, 1H), 4.63 (app dd,  $J = 13.4, 1.8$  Hz, 1H), 6.22 (dd,  $J = 12.2, 5.0$  Hz, 1H), 6.99 (ddd,  $J = 10.7, 8.7, 1.9$  Hz, 1H), 7.24 (d,  $J = 7.3$  Hz, 1H), 7.28 – 7.36 (m, 1H), 7.43 (ddd,  $J = 8.7, 4.1, 1.9$  Hz, 1H), 7.52 – 7.66 (m, 1H), 7.86 – 7.95 (m, 1H), 8.00 (d,  $J = 8.7$  Hz, 1H), 8.28 (d,  $J = 8.7$  Hz, 1H), 8.90 (dt,  $J = 4.1, 1.9$  Hz, 1H).  **$^{13}\text{C}$  NMR** (151 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  42.0,

44.4, 44.5, 57.3, 118.2, 118.4, 120.2, 120.3, 121.5, 122.9, 125.1, 128.5, 128.5, 129.6, 129.8, 130.2, 130.2, 131.9, 132.4, 132.5, 136.2, 148.3, 149.9, 151.5, 151.5, 159.0, 160.7, 164.7. **HRMS (ESI)**: exact mass calculated for C<sub>20</sub>H<sub>14</sub>Cl<sub>2</sub>FN<sub>3</sub>O [(M+H)<sup>+</sup>] 402.0571, found 402.0560.

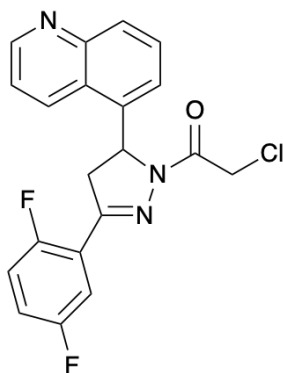
**2-Chloro-1-(3-(2,3-dichlorophenyl)-5-(quinolin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (CMZ 56)**



*Step 1*: Following General Procedure A, the chalcone was prepared from quinoline-5-carbaldehyde (213 mg, 1.36 mmol), 2',3'-dichloroacetophenone (396  $\mu$ L, 2.71 mmol) and 5% (w/w) aq NaOH (466  $\mu$ L, 0.58 mmol) in EtOH (4 mL). Filtration under reduced pressure and purification gave the chalcone as a pale brown solid (445 mg, quant).

*Step 2*: Following General Procedure F, the pure title compound was prepared from chalcone (306 mg, 0.93 mmol) and hydrazine hydrate (165  $\mu$ L, 1.87 mmol, 50–60% wt solution) in EtOH (4 mL); then chloroacetyl chloride (111  $\mu$ L, 1.40 mmol) and triethylamine (390  $\mu$ L, 2.80 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (4.7 mL). Purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 40%; then SNAP Ultra C18, 20% grading to 100% MeCN in water, product eluted at 63%) gave **CMZ 56** as a white solid (9 mg, 2.4% over 2 steps from the respective chalcone); **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$ <sub>H</sub> 3.26 (dd, *J* = 17.7, 5.1 Hz, 1H), 4.09 (dd, *J* = 17.7, 11.9 Hz, 1H), 4.49 (d, *J* = 12.8 Hz, 1H), 4.65 (d, *J* = 12.8 Hz, 1H), 5.89 (dd, *J* = 11.9, 5.1 Hz, 1H), 7.03 (dd, *J* = 8.0, 1.6 Hz, 1H), 7.14 (t, *J* = 8.0 Hz, 1H), 7.36 (dd, *J* = 8.0, 1.6 Hz, 1H), 7.53 (dd, *J* = 8.7, 4.2 Hz, 1H), 7.58 (dd, *J* = 7.3, 1.5 Hz, 1H), 7.65 (dd, *J* = 8.7, 7.3 Hz, 1H), 8.14 (app dt, *J* = 8.4, 1.1 Hz, 1H), 8.95 (dd, *J* = 4.2, 1.5 Hz, 1H), 9.49 - 9.59 (m, 1H). **<sup>13</sup>C NMR** (151 MHz, CDCl<sub>3</sub>)  $\delta$ <sub>C</sub> 41.9, 43.5, 57.6, 122.6, 124.0, 126.2, 127.4, 128.0, 129.5, 129.6, 130.1, 130.2, 131.6, 134.1, 137.2, 129.5, 146.8, 149.2, 155.1, 164.3. **HRMS (ESI)**: exact mass calculated for C<sub>20</sub>H<sub>14</sub>Cl<sub>3</sub>N<sub>3</sub>O [(M+H)<sup>+</sup>] 418.0275, found 418.0287.

**2-Chloro-1-(3-(2,5-difluorophenyl)-5-(quinolin-5-yl)-4,5-dihydro-1H-pyrazol-1-yl)ethan-1-one (CMZ 63)**



*Step 1:* Following General Procedure A, the chalcone was prepared from quinoline-5-carbaldehyde (200 mg, 1.27 mmol), 2'5'-difluoroacetophenone (161  $\mu$ L, 1.27 mmol) and 5% (w/w) aq NaOH (438  $\mu$ L, 0.55 mmol) in EtOH (4 mL). Filtration under reduced pressure and purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 44%) gave the chalcone as a yellow solid (59 mg, 16%).

*Step 2:* Following General Procedure F, the pure title compound was prepared from chalcone (59 mg, 0.20 mmol), hydrazine hydrate (35  $\mu$ L, 0.40 mmol, 50–60% wt solution) in EtOH (4 mL); then chloroacetyl chloride (24  $\mu$ L, 0.30 mmol) and triethylamine (84  $\mu$ L, 0.60 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (1.5 mL). Purification *via* flash column chromatography (Sfär Silica HC D, 0% grading to 100% EtOAc in hexane, product eluted at 56%) gave **CMZ 63** as a white solid (30 mg, 38% over 2 steps from the respective chalcone); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta_{\text{H}}$  3.23 (ddd,  $J = 18.5, 5.1, 3.0$  Hz, 1H), 4.00 (ddd,  $J = 18.5, 12.0, 3.0$  Hz, 1H), 4.51 (d,  $J = 13.4$  Hz, 1H), 4.62 (d,  $J = 13.4$  Hz, 1H), 6.23 (dd,  $J = 12.0, 5.1$  Hz, 1H), 6.94 – 7.12 (m, 2H), 7.26 (d,  $J = 7.2$  Hz, 1H), 7.45 (dd,  $J = 8.6, 4.3$  Hz, 1H), 7.58 – 7.69 (m, 2H), 8.03 (d,  $J = 8.6$  Hz, 1H), 8.31 (d,  $J = 8.6$  Hz, 1H), 8.91 (dd,  $J = 4.3, 1.6$  Hz, 1H). **HRMS (ESI):** exact mass calculated for C<sub>20</sub>H<sub>14</sub>ClF<sub>2</sub>N<sub>3</sub>O [(M+H)<sup>+</sup>] 386.0866, found 386.0881.

## APPENDIX 3

### CHAPTER 4 METHODS

#### Materials

Cysteine-reactive covalent ligand libraries were either previously synthesized and described or for the compounds starting with “EN” were purchased from Enamine, including EN523. Lumacaftor was purchased from Medchemexpress LLC.

#### Cell Culture

CFBE41o-4.7  $\Delta$ F508-CFTR Human CF Bronchial Epithelial cells were purchased from Millipore Sigma (SCC159). CFBE41o-4.7  $\Delta$ F508-CFTR Human CF Bronchial Epithelial cells were cultured in MEM (Gibco) containing 10% (v/v) fetal bovine serum (FBS) and maintained at 37 °C with 5% CO<sub>2</sub>.

#### Gel-Based ABPP

Recombinant OTUB1 (0.1  $\mu$ g/sample) was pre-treated with either DMSO vehicle or covalent ligand or DUBTACs at 37 °C for 30 min in 25  $\mu$ L of PBS, and subsequently treated with IA-Rhodamine (concentrations designated in figure legends) (Setareh Biotech) at room temperature for 1 h. The reaction was stopped by addition of 4 $\times$ reducing Laemmli SDS sample loading buffer (Alfa Aesar). After boiling at 95 °C for 5 min, the samples were separated on precast 4–20% Criterion TGX gels (Bio-Rad). Probe-labeled proteins were analyzed by in-gel fluorescence using a ChemiDoc MP (Bio-Rad).

#### NJH-2-057 Probe Labeling of Recombinant OTUB1

Recombinant and pure OTUB1 protein (0.5  $\mu$ g) per sample per replicate was suspended in 50  $\mu$ L total PBS. 1  $\mu$ L of either DMSO or NJH-2-075 (to give final concentrations of 50, 10, 1, and 0.1  $\mu$ M) was added, followed by a 1.5 h incubation at 37 °C. Next, 7.8  $\mu$ L of a solution composed of 9.4  $\mu$ L of 5mM Azide-Fluor 545 (in DMSO), 112  $\mu$ L of TBTA ligand (Stock 1.7 mM in 4 parts t-butanol + 1 part DMSO), 37.5  $\mu$ L of 50 mM TCEP (in water), and 37.5  $\mu$ L of 50 mM Copper (II) sulfate was added to each sample and the samples were incubated for 1 hour at room temperature. Following CuAAC, 30  $\mu$ L of Laemmli Sample Buffer (4 x) was added to each sample, vortexed and boiled for 6 min at 95 °C. Samples were loaded on an SDS/PAGE gel and analyzed for in-gel fluorescence.

#### Deubiquitinase Activity Assay

Previously described methods were used to assess EN523 effects on OTUB1 activity<sup>22</sup>. Recombinant OTUB1 (500 nM) was pre-incubated with DMSO or EN523 (50  $\mu$ M) for 1 hr. To initiate assay pre-treated OTUB1 enzyme was mixed 1:1 with di-Ub reaction mix for final concentrations of 250 nM OTUB1, 1.5  $\mu$ M di-Ub, 12.5  $\mu$ M UBE2D1 and 5 mM DTT. The appearance of mono-Ub was monitored by Western blotting over time by removing a portion of the reaction mix and adding Laemmli's buffer to terminate reaction. Blot shown is a representative gel from n=3 biologically independent experiments/group.

### **Bio-NMR Analysis of EN523-OTUB1 Interactions**

We recorded all NMR spectra on a Bruker 600 MHz spectrometer, equipped with a 5 mm QCI-F cryo probe with z-gradient, and kept the temperature constant at 298K during all experiments. To probe compound and E2 ligase binding to OTUB1, we recorded  $^1\text{H}$ -1D and  $^{13}\text{C}$ -SOFAST-HMQC experiments. We used 3 mm NMR tubes filled with 160  $\mu\text{L}$  of 50  $\mu\text{M}$   $\{\text{U}\}$ - $^2\text{H}$ ,  $^1\text{H}/^{13}\text{C}$ -methyl-Ile/Leu/Val/Ala(ILVA),  $\{\text{U}\}$ - $^{15}\text{N}$  labeled OTUB1, 25 mM d-Tris, pH 7.5, 150 mM NaCl, 5%  $\text{D}_2\text{O}$  (to lock), 100  $\mu\text{M}$  DSS (internal standard), 75  $\mu\text{M}$  EN-523 (dissolved in 100%  $\text{d}_6$ -DMSO; for compound binding study) and/or 100  $\mu\text{M}$  E2 D2 / Ub-E2 D2 (for ligase binding studies). To allow for complete binding of the compound to OTUB1, we chose an incubation period of  $\sim$ 40 hours. We also recorded reference spectra with the adequate volumes of pure  $\text{d}_6$ -DMSO and/or E2 buffer to compensate for solvent induced effects, and repeated experiments after 40 hours to make sure that any spectral changes were not related to protein oxidation.

### **Labeling of Endogenous OTUB1 in HEK293T Cells with NJH-2-075 Probe**

One plate of 70% confluent HEK293T cells per condition per replicate were treated with either DMSO vehicle or NJH-02-075 (50  $\mu\text{M}$ ) for 2 hours. Cells were harvested by scraping, suspended in 600  $\mu\text{L}$  of PBS, lysed by probe sonication, and centrifuged for 10 min at 5000 rpm to remove debris. Lysate was normalized to 3.1 mg/mL and 85  $\mu\text{L}$  removed for Western blotting analysis of input. 500  $\mu\text{L}$  of lysate was then incubated for 1 hour at room temperature with 10  $\mu\text{L}$  of 5 mM biotin picolyl azide (in water), 10  $\mu\text{L}$  of 50mM TCEP (in water), 30  $\mu\text{L}$  TBTA ligand (Stock 1.7 mM in 4 parts t-butanol + 1 part DMSO), and 10  $\mu\text{L}$  of 50 mM Copper (II) sulfate. Following CuAAC, precipitated proteins were washed 3 x with cold methanol and resolubilized in 200  $\mu\text{L}$  1.2% SDS/PBS. To ensure solubility, proteins were heated to 90  $^\circ\text{C}$  for 5 min following resuspension. 1 mL of PBS was then added to each sample, followed by 50  $\mu\text{L}$  of high-capacity streptavidin beads. Samples were then incubated overnight on a rocker at 4  $^\circ\text{C}$ . The following morning the samples were warmed to room temperature, and non-specific binding proteins were washed away with 3 x PBS washes followed by 3 x water washes. Beads were then resuspended in 100  $\mu\text{L}$  PBS and 30  $\mu\text{L}$  Laemmli Sample Buffer (4 x) and boiled for 13 min at 95  $^\circ\text{C}$ . Samples were vortexed and loaded onto an SDS/PAGE gel along with saved input samples for Western blotting analysis.

### **Western Blotting**

Proteins were resolved by SDS/PAGE and transferred to nitrocellulose membranes using the Trans-Blot Turbo transfer system (Bio-Rad). Membranes were blocked with 5% BSA in Tris-buffered saline containing Tween 20 (TBS-T) solution for 30 min at RT, washed in TBS-T, and probed with primary antibody diluted in recommended diluent per manufacturer overnight at 4 $^\circ\text{C}$ . After 3 washes with TBS-T, the membranes were incubated in the dark with IR680- or IR800-conjugated secondary antibodies at 1:10,000 dilution in 5 % BSA in TBS-T at RT for 1 h. After 3 additional washes with TBST, blots were visualized using an Odyssey Li-Cor fluorescent scanner. The membranes were stripped using ReBlot Plus Strong Antibody Stripping Solution (EMD Millipore) when additional primary antibody incubations were performed. Antibodies used in this study were CFTR (Cell Signaling Technologies, Rb mAb

#78335, **Figures 3 and 4**), CFTR (R&D Systems, Ms mAb, #MAB25031, **Extended Data Figure 3**), CFTR (Millipore, Ms mAb, #MAB3484, **Extended Data Figure 3**), CFTR (Prestige, Rb pAb, #HPA021939, **Extended Data Figure 3**), GAPDH (Proteintech, Ms mAb, #60004-1-Ig), OTUB1 (Abcam, Rb mAb, #ab175200, [EPR13028(B)]), and CTNNB1 (Cell Signaling Technologies, Rb mAb, #8480).

### **IsoTOP-ABPP Chemoproteomic Experiments**

IsoTOP-ABPP studies were done as previously reported<sup>14,29,46</sup>. Our aggregate chemoproteomic data analysis of DUBs were obtained from 455 distinct isoTOP-ABPP experiments performed in the Nomura Research Group. These data are aggregated from various human cell lines, including 231MFP, A549, HeLa, HEK293T, HEK293A, UM-Chor1, PaCa2, PC3, HUH7, NCI-H460, THP1, SKOV3, U2OS, and K562 cells. Some of the chemoproteomic data have been previously reported as part of other studies<sup>32,46–55</sup>. All of the isoTOP-ABPP datasets were prepared as described below using the IA-alkyne probe. Cells were lysed by probe sonication in PBS and protein concentrations were measured by BCA assay. Cells were treated for 4 h with either DMSO vehicle or EN4 (from 1,000x DMSO stock) before cell collection and lysis. Proteomes were subsequently labeled with IA-alkyne labeling (100  $\mu$ M for DUB ligandability analysis and 200  $\mu$ M for profiling cysteine-reactivity of NJH-2-057) for 1 h at room temperature. CuAAC was used by sequential addition of tris(2-carboxyethyl)phosphine (1 mM, Strem, 15-7400), tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (34  $\mu$ M, Sigma, 678937), copper(II) sulfate (1 mM, Sigma, 451657) and biotin-linker-azide—the linker functionalized with a tobacco etch virus (TEV) protease recognition sequence as well as an isotopically light or heavy valine for treatment of control or treated proteome, respectively. After CuAAC, proteomes were precipitated by centrifugation at 6,500g, washed in ice-cold methanol, combined in a 1:1 control:treated ratio, washed again, then denatured and resolubilized by heating in 1.2% SDS–PBS to 80 °C for 5 min. Insoluble components were precipitated by centrifugation at 6,500g and soluble proteome was diluted in 5 ml 0.2% SDS–PBS. Labeled proteins were bound to streptavidin-agarose beads (170  $\mu$ l resuspended beads per sample, Thermo Fisher, 20349) while rotating overnight at 4 °C. Bead-linked proteins were enriched by washing three times each in PBS and water, then resuspended in 6 M urea/PBS, and reduced in TCEP (1 mM, Strem, 15-7400), alkylated with iodoacetamide (18 mM, Sigma), before being washed and resuspended in 2 M urea/PBS and trypsinized overnight with 0.5  $\mu$ g/ $\mu$ L sequencing grade trypsin (Promega, V5111). Tryptic peptides were eluted off. Beads were washed three times each in PBS and water, washed in TEV buffer solution (water, TEV buffer, 100  $\mu$ M dithiothreitol) and resuspended in buffer with Ac-TEV protease (Invitrogen, 12575-015) and incubated overnight. Peptides were diluted in water and acidified with formic acid (1.2 M, Fisher, A117-50) and prepared for analysis.

### **IsoTOP-ABPP Mass Spectrometry Analysis**

Peptides from all chemoproteomic experiments were pressure-loaded onto a 250  $\mu$ m inner diameter fused silica capillary tubing packed with 4 cm of Aqua C18 reverse-phase resin (Phenomenex, 04A-4299), which was previously equilibrated on an Agilent 600 series high-performance liquid chromatograph using the gradient from 100% buffer A to 100% buffer B over 10 min, followed by a 5 min wash with 100% buffer B and a 5 min

wash with 100% buffer A. The samples were then attached using a MicroTee PEEK 360  $\mu\text{m}$  fitting (Thermo Fisher Scientific p-888) to a 13 cm laser pulled column packed with 10 cm Aqua C18 reverse-phase resin and 3 cm of strong-cation exchange resin for isoTOP-ABPP studies. Samples were analyzed using an Q Exactive Plus mass spectrometer (Thermo Fisher Scientific) using a five-step Multidimensional Protein Identification Technology (MudPIT) program, using 0, 25, 50, 80 and 100% salt bumps of 500 mM aqueous ammonium acetate and using a gradient of 5–55% buffer B in buffer A (buffer A: 95:5 water:acetonitrile, 0.1% formic acid; buffer B 80:20 acetonitrile:water, 0.1% formic acid). Data were collected in data-dependent acquisition mode with dynamic exclusion enabled (60 s). One full mass spectrometry (MS1) scan (400–1,800 mass-to-charge ratio ( $m/z$ )) was followed by 15 MS2 scans of the  $n$ th most abundant ions. Heated capillary temperature was set to 200 °C and the nanospray voltage was set to 2.75 kV.

Data were extracted in the form of MS1 and MS2 files using Raw Extractor v.1.9.9.2 (Scripps Research Institute) and searched against the Uniprot human database using ProLuCID search methodology in IP2 v.3-v.5 (Integrated Proteomics Applications, Inc.)<sup>56</sup>. Cysteine residues were searched with a static modification for carboxyamino-methylation (+57.02146) and up to two differential modifications for methionine oxidation and either the light or heavy TEV tags (+464.28596 or +470.29977, respectively). Peptides were required to be fully tryptic peptides and to contain the TEV modification. ProLUCID data were filtered through DTASelect to achieve a peptide false-positive rate below 5%. Only those probe-modified peptides that were evident across two out of three biological replicates were interpreted for their isotopic light to heavy ratios. For those probe-modified peptides that showed ratios greater than two, we only interpreted those targets that were present across all three biological replicates, were statistically significant and showed good quality MS1 peak shapes across all biological replicates. Light versus heavy isotopic probe-modified peptide ratios are calculated by taking the mean of the ratios of each replicate paired light versus heavy precursor abundance for all peptide-spectral matches associated with a peptide. The paired abundances were also used to calculate a paired sample  $t$ -test  $P$  value in an effort to estimate constancy in paired abundances and significance in change between treatment and control.  $P$  values were corrected using the Benjamini–Hochberg method.

### **Knockdown studies**

RNA interference was performed using siRNA purchased from Dharmacon. CFBE41o-4.7 cells were seeded at 400,000 cells per 6 cm plate and allowed to adhere overnight. Cells were transfected with 33 nM of either nontargeting (ON-TARGETplus Non-targeting Control Pool, Dharmacon #D-001810-10-20) or anti-CFTR siRNA (Dharmacon, custom) using 8  $\mu\text{L}$  of transfection reagent: either DharmaFECT 1 (Dharmacon #T-2001-02), DharmaFECT 4 (Dharmacon, T-2004-02) or Lipofectamine 2000 (ThermoFisher #11668027). Transfection reagent was added to OPTIMEM (ThermoFisher #31985070) media, allowed to incubate for 5 min at room temperature. Meanwhile siRNA was added to an equal amount of OPTIMEM. Solutions of transfection reagent and siRNA in OPTIMEM were then combined and allowed to incubate for 30 minutes at room temperature. These combined solutions were diluted



with complete MEM to provide 33nM siRNA and 8  $\mu$ L of transfection reagent per 4 mL MEM, and the media exchanged. Cells were incubated with transfection reagents for 24h, at which point the media replaced with media containing DMSO or 10  $\mu$ M NJH-2-057 and incubated for another 24h. Cells were then harvested, and protein abundance analyzed by Western blotting.

### **Quantitative TMT Proteomics Analysis**

Quantitative TMT-based proteomic analysis was performed as previously described<sup>29</sup>. Acquired MS data was processed using Proteome Discoverer v. 2.4.0.305 software (Thermo) utilizing Mascot v 2.5.1 search engine (Matrix Science, London, UK) together with Percolator validation node for peptide-spectral match filtering<sup>57</sup>. Data was searched against Uniprot protein database (canonical human sequences, EBI, Cambridge, UK) supplemented with sequences of common contaminants. Peptide search tolerances were set to 10 ppm for precursors, and 0.8 Da for fragments. Trypsin cleavage specificity (cleavage at K, R except if followed by P) allowed for up to 2 missed cleavages. Carbamidomethylation of cysteine was set as a fixed modification, methionine oxidation, and TMT-modification of N-termini and lysine residues were set as variable modifications. Data validation of peptide and protein identifications was done at the level of the complete dataset consisting of combined Mascot search results for all individual samples per experiment via the Percolator validation node in Proteome Discoverer. Reporter ion ratio calculations were performed using summed abundances with most confident centroid selected from 20 ppm window. Only peptide-to-spectrum matches that are unique assignments to a given identified protein within the total dataset are considered for protein quantitation. High confidence protein identifications were reported using a Percolator estimated <1% false discovery rate (FDR) cut-off. Differential abundance significance was estimated using ANOVA with Benjamini-Hochberg correction to determine adjusted p-values.

### **Data Availability Statement**

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

### **Code Availability Statement**

Data processing and statistical analysis algorithms from our lab can be found on our lab's Github site: <https://github.com/NomuraRG>, and we can make any further code from this study available at reasonable request.

### **Chemical Synthesis and Characterization**

Starting materials, reagents and solvents were purchased from commercial suppliers and were used without further purification unless otherwise noted. All reactions were monitored by TLC (TLC Silica gel 60 F<sub>254</sub>, Sepulco Millipore Sigma). Reaction products were purified by flash column chromatography using a Biotage Isolera with Biotage Sfar® or Silicycle normal-phase silica flash columns (5 g, 10 g, 25 g, or 40 g). <sup>1</sup>H NMR and <sup>13</sup>C NMR spectra were recorded on a 400 MHz Bruker Avance I spectrometer or a 600 MHz Bruker Avance III spectrometer equipped with a 5 mm <sup>1</sup>H/BB Prodigy cryo-

probe. Chemical shifts are reported in parts per million (ppm,  $\delta$ ) downfield from tetramethylsilane (TMS). Coupling constants (J) are reported in Hz. Spin multiplicities are described as br (broad), s (singlet), d (doublet), t (triplet), q (quartet) and m (multiplet).

#### **General Procedure A:**

Carboxylic acid (1.0 Eq.) was dissolved in DCM (0.1 M). An amine (1.25 Eq.) was added, followed by DIEA (4.0 Eq.), HOBt (0.2 Eq.) and EDCI (2.0 Eq.). The reaction mixture was stirred overnight at rt, water was added, and the mixture extracted three times with DCM. Combined organic extracts were washed with 1M HCl, washed with brine, dried over sodium sulfate, and concentrated. The crude product was purified by silica gel chromatography to provide the amide.

#### **General Procedure B:**

Boc-protected amine was dissolved in DCM (0.1 M), and TFA was added to give a 1:2 TFA:DCM ratio. The solution was allowed to stir for 1h. The volatiles were then evaporated, and the resulting oil redissolved in DCM and treated with aqueous saturated  $\text{NaHCO}_3$ . The resulting mixture was then extracted with DCM three times, the combined organic extracts dried over  $\text{Na}_2\text{SO}_4$ , and concentrated to provide the amine without further purification.

#### **General Procedure C:**

Tert-butyl ester such as Intermediate 3 (30 mg, 0.086 mmol, 3.0 Eq) was dissolved in DCM (600 mL). TFA (300 mL) was added and the solution stirred for 1h. Volatiles were evaporated under vacuum, and DCM (1 mL) was added and evaporated to give the carboxylic acid intermediate, though some excess TFA remained. This intermediate was dissolved in DMF (500 mL) and DIEA (150 mL, 30 Eq.) and the appropriate amine (0.029 mmol, 1.0 Eq) were added, followed by HATU (30 mg, 0.079 mmol, 2.7 Eq.). The reaction mixture was allowed to stir for 1h at rt. Water was added, and the mixture extracted three times with EtOAc for CFTR DUBTACS, or 4:1  $\text{CHCl}_3$ :IPA for WEE1 DUBTACS. Organic extracts were combined, washed with brine, dried over sodium sulfate, and concentrated. Crude residues were purified by silica gel chromatography to provide the final compounds.

#### **General Procedure D:**

To a solution of the appropriate bromide dissolved in dioxane, N,N'-dimethylethylenediamine (0.25 eq),  $\text{K}_2\text{CO}_3$  (3.0 eq), CuI (0.1 eq), and the appropriate amide coupling partner (1.0 eq) were added. The reaction mixture was degassed, the atmosphere exchanged for nitrogen, and stirred at 100 °C overnight. Saturated  $\text{NH}_4\text{Cl}$  was added to the completed reaction mixture once cooled, which was stirred for 20 minutes, then filtered through celite and the celite pad was washed with EtOAc. The mixture was extracted with EtOAc three times, washed with brine twice, and dried by  $\text{NaSO}_4$ , before concentration *in vacuo*. Resulting crude mixtures were purified via silica gel column chromatography.

#### **General Procedure E:**

The appropriate amine was dissolved in THF and water (2:1 THF:H<sub>2</sub>O) with potassium carbonate (3.0 eq). Benzyl chloroformate (1-2 eq) was added dropwise to the reaction mixture, which was then stirred vigorously overnight at room temperature. Water was added and the mixture extracted with EtOAc three times. Organic extracts were combined, washed with brine twice, concentrated and the resulting crude purified using flash column chromatography.

**General Procedure F:**

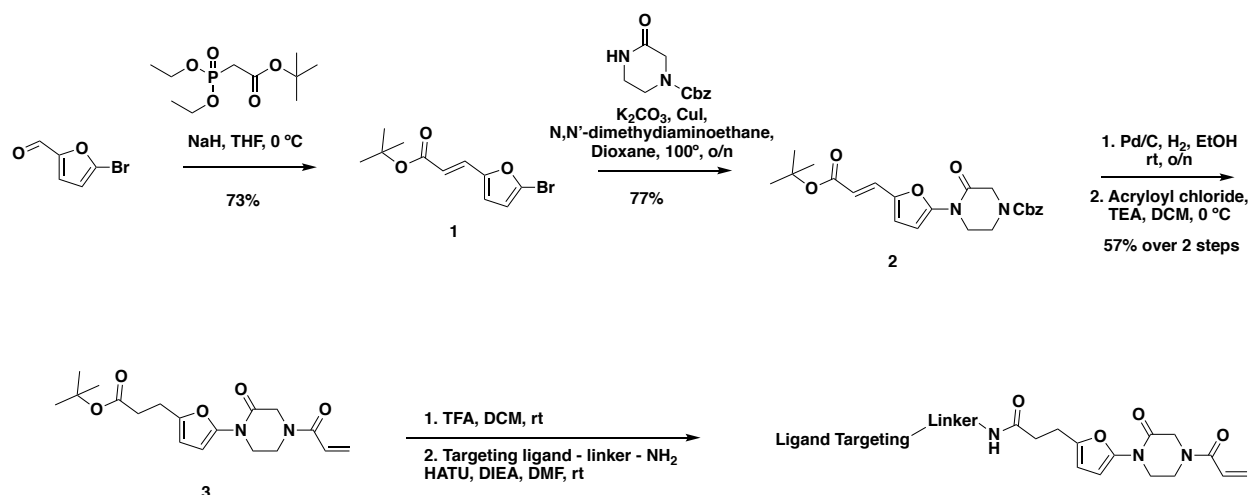
The coupled product was dissolved in DCM, followed by a dropwise addition of trifluoroacetic acid (1:2 TFA:DCM) until consumption of starting material was observed via TLC (15-30 min). The mixture was then washed with DCM twice and immediately used without further purification.

**General Procedure G:**

Pd/C (10% wt.) was added to a mixture of the Cbz-protected compound in EtOH (0.2 M), and the atmosphere was exchanged for H<sub>2</sub> (balloon). The reaction mixture was stirred vigorously overnight, before being diluted with DCM, filtered through a syringe filter (0.45  $\mu$ m), concentrated, and purified using silica gel column chromatography.

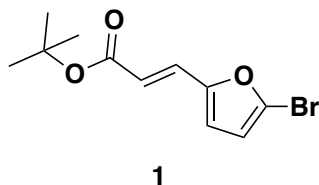
**General Procedure H:**

The amine starting material was dissolved in DCM on ice. TEA (3.0 eq) and acryloyl chloride (1.5 eq) were then added to the reaction mixture until consumption of the starting material was observed by TLC (0.5 – 2 hrs). Water was added, and the mixture reaction was extracted with DCM three times. Organic extracts were combined, washed with H<sub>2</sub>O then brine, concentrated, and purified via silica gel column chromatography.



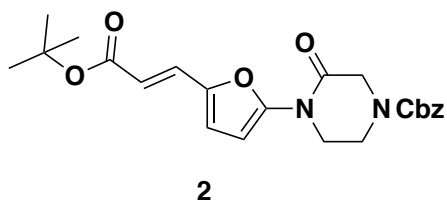
Scheme 1. General scheme describing synthetic route to bifunctional DUBTACs containing EN523 as an OTUB1 recruiter.

## SYNTHESIS OF CFTR DUBTACs



**tert-butyl (*E*)-3-(5-bromofuran-2-yl)acrylate (1):** tert-butyl diethylphosphonoacetate (971 mg, 0.908 mL, 3.85 mmol) was dissolved in THF (22 mL) and the solution cooled to 0 °C. Then, 5-bromofuran-2-carbaldehyde (613 mg, 3.50 mmol) was added portion-wise over 5 minutes. The reaction was stirred for 20 minutes at 0 °C as a gummy solid precipitated. Water was added and the resulting mixture was extracted with EtOAc three times. Combined organic extracts were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated. The crude residue was purified by silica gel chromatography (0-15% EtOAc/Hex) to provide the title compound as a light yellow oil (782 mg, 2.86 mmol, 82%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.26 (d, *J* = 15.7 Hz, 1H), 6.55 (d, *J* = 3.5 Hz, 1H), 6.42 (d, *J* = 3.4 Hz, 1H), 6.29 (d, *J* = 15.7 Hz, 1H), 1.55 (s, 9H).

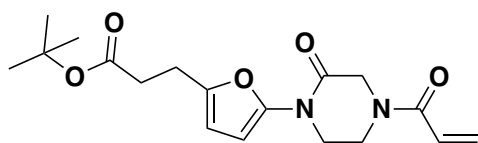


**benzyl (*E*)-4-(5-(3-(tert-butoxy)-3-oxoprop-1-en-1-yl)furan-2-yl)-3-oxopiperazine-1-carboxylate (2):** tert-butyl (*E*)-3-(5-bromofuran-2-yl)acrylate (1.62 g, 5.94 mmol) was dissolved in dioxane (30 mL) and benzyl 3-oxopiperazine-1-carboxylate (1.4 g, 5.94 mmol), K<sub>2</sub>CO<sub>3</sub> (2.46 g, 17.8 mmol), *N,N'*-dimethyldiaminoethane (0.167 mL, 1.49 mmol),

and CuI (114 mg, 0.59 mmol) were added. The mixture was stirred under nitrogen at reflux for 40 h, then cooled to rt. 5 mL saturated aq. NH<sub>4</sub>Cl was added and the mixture stirred for 30 min. Then the mixture was diluted in EtOAc, filtered through celite, water was added, the mixture partitioned, and the aqueous layer extracted with EtOAc. The extracts were combined, washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated, and purified by silica gel chromatography (0-35% EtOAc/Hex) to provide the title compound as an orange oil (1.95 g, 4.59 mmol, 77%).

**LC/MS:** [M+2H-tBu]<sup>+</sup> m/z calc. 371.18, found 373.1.

**<sup>1</sup>H NMR** (400 MHz, DMSO) δ 7.45 – 7.24 (m, 6H), 6.98 (s, 1H), 6.57 (s, 1H), 6.08 (dd, J = 15.7, 3.4 Hz, 1H), 5.14 (dd, J = 4.4, 2.3 Hz, 2H), 4.22 (s, 2H), 4.01 (s, 2H), 3.77 (s, 2H), 1.47 (s, 9H).



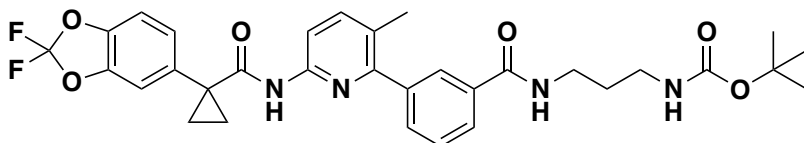
3

**tert-butyl 3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)propanoate (Intermediate (3)):** benzyl (*E*)-4-(5-(3-(tert-butoxy)-3-oxoprop-1-en-1-yl)furan-2-yl)-3-oxopiperazine-1-carboxylate (1.95 g, 4.59 mmol) was dissolved in EtOH (25 mL) and Pd/C (200 mg, 10% wt. Pd) was added. The reaction was placed under an atmosphere of H<sub>2</sub> and stirred vigorously overnight, before being filtered through celite twice and then concentrated. The crude product was then redissolved in DCM (25 mL), cooled to 0 °C and treated with TEA (1.28 mL, 9.18 mmol) before a solution of acryloyl chloride (445 μL, 5.51 mmol) in DCM (5 mL) was added over 2 minutes. After stirring for 20 min, water was added, and the mixture extracted with DCM three times. Combined organic extracts were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated, and the resulting crude oil was purified by silica gel chromatography (0-75% EtOAc/Hex) to obtain the title compound as a light yellow oil (846 mg, 2.43 mmol, 53% over two steps).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 6.64 – 6.46 (m, 1H), 6.41 (dd, J = 16.7, 2.0 Hz, 1H), 6.29 (d, J = 3.2 Hz, 1H), 6.04 (d, J = 3.3 Hz, 1H), 5.82 (dd, J = 10.2, 2.0 Hz, 1H), 4.42 (d, J = 24.9 Hz, 2H), 4.06 – 3.82 (m, 4H), 2.88 (t, J = 7.8 Hz, 2H), 2.54 (d, J = 7.6 Hz, 2H), 1.44 (s, 9H).

**LC/MS:** [M+2H-tBu]<sup>+</sup> m/z calc. 293.1, found 293.1.

*Note: Intermediate 3 is prone to decomposition, likely through polymerization. Care should be taken to store at -20 °C. Attempts to dry thoroughly (e.g. leaving on vacuum at rt overnight) occasionally led to decomposition of ~50% of the material.*

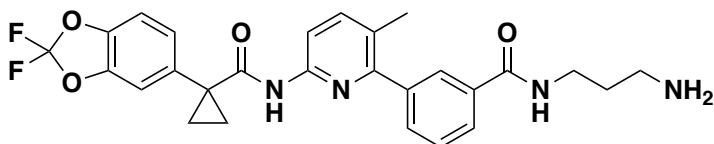


4a

**tert-butyl (3-(3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamido)propyl)carbamate (4a):** Lumacaftor (3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzoic acid) (18 mg, 0.04 mmol), tert-butyl (3-aminopropyl)carbamate (14 mg, 0.08 mmol), DIEA (35  $\mu$ L, 0.20 mmol), and HOBt (5.4 mg, 0.04 mmol) were dissolved in DCM (1 mL), followed by the addition of EDCI HCl (15 mg, 0.05 mmol). The reaction was stirred at rt for 2 days before water was added, the mixture partitioned, and the aqueous layer extracted with DCM twice. The combined organic extracts were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated, and the resulting crude oil was purified by silica gel chromatography (0-60% EtOAc/Hex) to obtain the **4a** as a clear oil (23 mg, 0.038 mmol, 94%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$  8.13 (d, J = 8.4 Hz, 1H), 7.95 (s, 1H), 7.88 (d, J = 7.6 Hz, 1H), 7.74 (s, 1H), 7.62 (d, J = 8.5 Hz, 1H), 7.60 – 7.49 (m, 2H), 7.34 (s, 1H), 7.30 – 7.18 (m, 2H), 7.11 (d, J = 8.2 Hz, 1H), 4.96 (s, 1H), 3.54 (q, J = 6.2 Hz, 2H), 3.27 (q, J = 6.3 Hz, 2H), 2.31 (s, 3H), 1.78 (q, J = 3.9 Hz, 2H), 1.76 – 1.70 (m, 2H), 1.47 (s, 9H), 1.19 (q, J = 3.9 Hz, 2H).

**LC/MS:** [M+H]<sup>+</sup> m/z calc. 609.24, found 609.3.

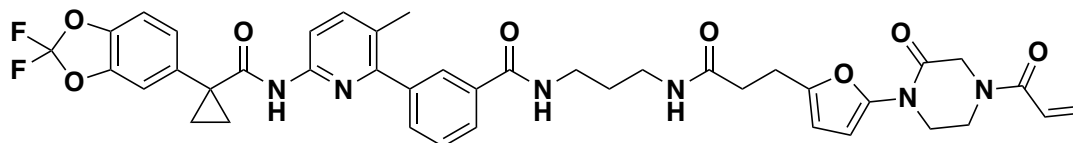


5a

**N-(3-aminopropyl)-3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide (5a):** Intermediate **4a** (23 mg, 0.038 mmol) was dissolved in DCM (1 mL) and TFA (1 mL) was added and the solution stirred for 2 hours. The volatiles were then evaporated and the resulting oil redissolved in DCM and treated with aqueous saturated NaHCO<sub>3</sub>. The resulting mixture was then extracted with DCM three times, combined organic extracts dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated to provide **5a** (15 mg, 0.029 mmol, 78%) as a colorless oil which was used in the next step without further purification.

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$  10.73 (s, 1H), 8.96 (s, 1H), 8.66 (t, J = 5.7 Hz, 1H), 7.95 – 7.85 (m, 3H), 7.79 – 7.66 (m, 2H), 7.60 (d, J = 7.6 Hz, 1H), 7.56 – 7.49 (m, 2H), 7.41 – 7.30 (m, 2H), 3.33 (q, J = 6.4 Hz, 2H), 2.88 – 2.77 (m, 2H), 2.21 (s, 3H), 1.79 (p, J = 6.9 Hz, 2H), 1.52 (dd, J = 4.9, 2.5 Hz, 2H), 1.19 – 1.15 (m, 2H).

**LC/MS:** [M+H]<sup>+</sup> m/z calc. 509.19, found 509.2.



NJH-2-056

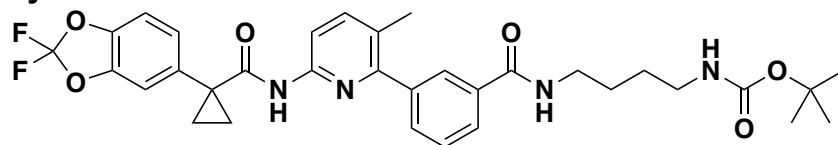
***N*-(3-(3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)propanamido)propyl)-3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide (NJH-2-056)**: Intermediate **3** (tert-butyl 3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)propanoate) (14 mg, 0.04 mmol) was dissolved in DCM (0.6 mL) and TFA (0.3 mL) was added and the solution was stirred for 1 h at rt until starting material was consumed as monitored by TLC. Volatiles were evaporated, DCM was added and evaporated again. The residue was dissolved in DCM (1.5 mL) and DIEA (140  $\mu$ L, 0.80 mmol) was added followed by *N*-(3-aminopropyl)-3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide (5.4 mg, 0.1 mmol). EDCI HCl (15 mg, 0.08 mmol) was then added, and the mixture stirred for 16h. Water was added and the resulting suspension was extracted with DCM three times. Combined organic extracts were washed with brine and dried over Na<sub>2</sub>SO<sub>4</sub> before being concentrated. The crude residue was purified by silica gel chromatography (0-5% MeOH/DCM) to obtain **NJH-2-056** (9.5 mg, 0.012 mmol, 30%) as a white powder following lyophilization from 1:1 water:acetonitrile (2 mL).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$  8.09 (d, *J* = 8.4 Hz, 1H), 7.93 – 7.87 (m, 1H), 7.83 (dt, *J* = 7.5, 1.6 Hz, 1H), 7.72 (s, 1H), 7.59 (d, *J* = 8.5 Hz, 1H), 7.57 – 7.45 (m, 2H), 7.29 (s, 1H), 7.23 (dd, *J* = 8.2, 1.8 Hz, 1H), 7.19 (d, *J* = 1.7 Hz, 1H), 7.07 (d, *J* = 8.1 Hz, 1H), 6.50 (s, 1H), 6.43 – 6.33 (m, 2H), 6.19 (d, *J* = 3.2 Hz, 1H), 6.07 (d, *J* = 3.3 Hz, 1H), 5.81 (d, *J* = 10.1 Hz, 1H), 4.47 – 4.31 (m, 2H), 4.04 – 3.78 (m, 4H), 3.36 (q, *J* = 6.2 Hz, 2H), 3.32 – 3.23 (m, 2H), 2.96 (t, *J* = 7.2 Hz, 2H), 2.55 (t, *J* = 7.2 Hz, 2H), 2.26 (s, 3H), 1.74 (q, *J* = 3.9 Hz, 2H), 1.69 – 1.58 (m, 2H), 1.16 (q, *J* = 3.9 Hz, 2H).

**<sup>13</sup>C NMR** (151 MHz, CDCl<sub>3</sub>)  $\delta$  172.5, 171.8, 167.4, 165.0, 155.5, 149.8, 148.9, 145.0, 144.1, 143.6, 141.0, 140.2, 134.9, 134.6, 131.8, 131.7, 130.0, 128.5, 127.8, 127.0, 126.6, 126.5, 126.3, 112.9, 112.4, 110.2, 107.6, 101.3, 36.0, 35.9, 35.2, 31.2, 29.5, 24.4, 19.2, 17.2.

**HRMS**: [M+H]<sup>+</sup> *m/z* calc. 783.2949, found 783.2954.

#### Synthesis of NJH-2-132



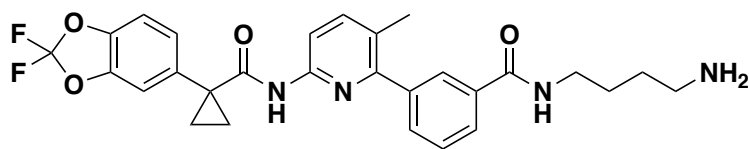
4b

**tert-butyl (4-(3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamido)butyl)carbamate (4b)**: Lumacaftor (100 mg, 0.22 mmol), tert-butyl (4-aminobutyl)carbamate were reacted according to

General Procedure A and purified by silica gel chromatography (0-60% EtOAc/Hex) to obtain **4b** as a clear colorless oil (128 mg, 0.20 mmol, 93%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 8.14 (d, J = 8.4 Hz, 1H), 7.87 (s, 1H), 7.83 (d, J = 7.7 Hz, 1H), 7.74 (s, 1H), 7.63 (d, J = 8.5 Hz, 1H), 7.57 (dt, J = 7.7, 1.5 Hz, 1H), 7.51 (t, J = 7.6 Hz, 1H), 7.27 (dd, J = 8.1, 1.8 Hz, 1H), 7.23 (d, J = 1.7 Hz, 1H), 7.12 (d, J = 8.2 Hz, 1H), 6.58 (s, 1H), 4.68 (s, 1H), 3.52 (q, J = 6.4 Hz, 2H), 3.20 (q, J = 6.6 Hz, 2H), 2.29 (s, 3H), 1.79 (q, J = 3.9 Hz, 2H), 1.71 – 1.67 (m, 2H), 1.65 (s, 9H), 1.65 – 1.57 (m, 2H), 1.20 (q, J = 3.9 Hz, 2H).

**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 623.3, found 623.3.

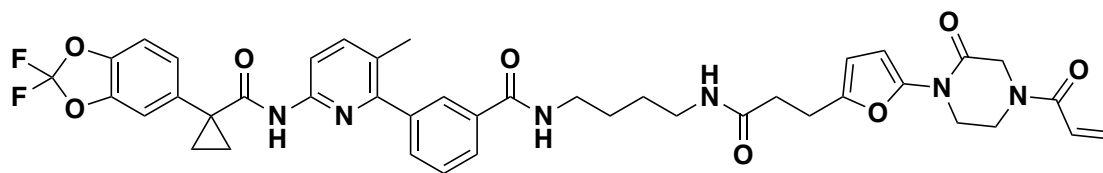


**5b**

**N-(4-aminobutyl)-3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide (**5b**):** **4b** (128 mg, 0.20 mmol) was deprotected according to General Procedure B to provide the amine **5b** (104 mg, 0.20 mmol, quant.) as a colorless oil.

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 8.13 (dd, J = 8.4, 1.7 Hz, 1H), 7.85 (tt, J = 8.5, 1.8 Hz, 1H), 7.81 (dt, J = 7.6, 1.6 Hz, 1H), 7.73 (s, 1H), 7.62 (dd, J = 8.5, 2.1 Hz, 1H), 7.56 (ddt, J = 7.7, 2.9, 1.5 Hz, 1H), 7.50 (td, J = 7.6, 3.0 Hz, 1H), 7.27 (dd, J = 8.2, 1.8 Hz, 1H), 7.22 (t, J = 1.8 Hz, 1H), 7.11 (d, J = 8.1 Hz, 1H), 7.03 (d, J = 5.3 Hz, 1H), 3.57 – 3.46 (m, 2H), 3.27 (d, J = 6.7 Hz, 1H), 2.80 (t, J = 6.7 Hz, 1H), 2.28 (d, J = 2.5 Hz, 3H), 1.98 (d, J = 1.4 Hz, 1H), 1.86 (s, 1H), 1.79 (q, J = 3.9 Hz, 2H), 1.72 (dd, J = 8.1, 6.3 Hz, 1H), 1.63 – 1.53 (m, 1H), 1.20 (qd, J = 4.0, 1.1 Hz, 2H).

**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 523.2, found 523.2.



**NJH-2-132**

**N-(4-(3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)propanamido)butyl)-3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide (**NJH-2-132**):** Intermediate **3** (tert-butyl 3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)propanoate) (30 mg, 0.086 mmol) was dissolved in DCM (0.6 mL) and TFA (0.3 mL) was added and the solution stirred for 1 h until



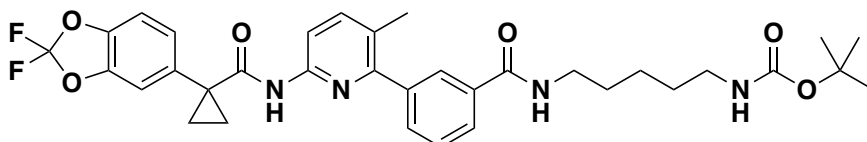
starting material was consumed. Volatiles were evaporated, DCM was added and evaporated again. The residue was dissolved in DCM (1.5 mL) and DIEA (150 mL, 0.86 mmol) was added followed by N-(4-aminobutyl)-3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide (**5b**) (15 mg, 0.029 mmol). HATU (30mg, 0.079 mmol) was then added and the mixture stirred for 16h. Water was added and the resulting suspension was extracted with DCM three times. Combined organic extracts were washed brine and dried over sodium sulfate before being concentrated. The crude residue was purified by silica gel chromatography (0-5% MeOH/DCM) to obtain the title compound (9.5 mg, 0.012 mmol, 30%) as a white solid.

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 8.12 (d, J = 8.4 Hz, 1H), 7.88 (t, J = 1.8 Hz, 1H), 7.84 (dt, J = 7.6, 1.6 Hz, 1H), 7.74 (s, 1H), 7.62 (d, J = 8.5 Hz, 1H), 7.56 (dt, J = 7.7, 1.5 Hz, 1H), 7.51 (t, J = 7.6 Hz, 1H), 7.26 (dd, J = 8.2, 1.7 Hz, 1H), 7.22 (d, J = 1.7 Hz, 1H), 7.11 (d, J = 8.2 Hz, 1H), 6.80 (s, 1H), 6.53 (d, J = 24.7 Hz, 1H), 6.41 (dd, J = 16.7, 2.0 Hz, 1H), 6.20 (d, J = 3.2 Hz, 1H), 6.07 (d, J = 3.3 Hz, 2H), 5.83 (dd, J = 10.2, 2.0 Hz, 1H), 4.38 (d, J = 28.2 Hz, 2H), 4.07 – 3.79 (m, 4H), 3.73 (tt, J = 9.8, 4.9 Hz, 1H), 3.45 (q, J = 6.4 Hz, 2H), 3.27 (q, J = 6.2 Hz, 2H), 3.20 (qd, J = 7.4, 3.4 Hz, 1H), 2.94 (q, J = 6.1, 5.0 Hz, 2H), 2.52 (t, J = 7.2 Hz, 2H), 2.28 (s, 3H), 1.77 (q, J = 3.9 Hz, 2H), 1.63 – 1.51 (m, 2H), 1.19 (q, J = 3.9 Hz, 2H).

**<sup>13</sup>C NMR** (151 MHz, CDCl<sub>3</sub>) δ 171.8, 167.4, 165.0, 155.5, 149.9, 148.9, 144.7, 144.1, 143.6, 141.0, 140.2, 134.9, 134.7, 133.4, 131.8, 131.7, 128.5, 127.6, 127.0, 126.6, 126.4, 112.9, 112.4, 110.2, 107.4, 101.2, 55.5, 43.5, 39.6, 39.0, 34.9, 31.2, 26.8, 26.7, 24.3, 19.2, 18.6, 17.2, 17.2, 12.5.

**HRMS (ESI):** *m/z* calc. 797.3032, found 797.3109.

#### Synthesis of NJH-2-057

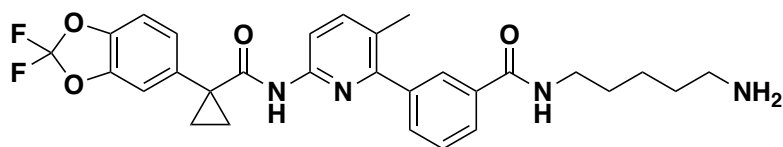


**4c**

**tert-butyl (5-(3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide)pentyl)carbamate (**4c**):** Lumacaftor (3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzoic acid) (181 mg, 0.40 mmol), tert-butyl (5-aminopentyl)carbamate (121 mg, 0.60 mmol), DIEA (350  $\mu$ L, 2.00 mmol), and HOBt (54 mg, 0.4mmol) were dissolved in DCM (6 mL), followed by addition of EDCI HCl (153 mg, 0.50 mmol). The reaction was stirred at rt for 16 hours before water was added, the mixture partitioned, and the aqueous layer extracted with DCM twice. The combined organic extracts were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated, and the resulting crude oil was purified by silica gel chromatography (0-50% EtOAc/Hex) to obtain Intermediate **4c** as a clear oil (240 mg, 0.38 mmol, 95%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 8.14 (d, J = 8.4 Hz, 1H), 7.84 (s, 1H), 7.80 (dt, J = 7.6, 1.6 Hz, 1H), 7.73 (s, 1H), 7.63 (d, J = 8.5 Hz, 1H), 7.57 (dt, J = 7.7, 1.5 Hz, 1H), 7.51 (t, J = 7.6 Hz, 1H), 7.27 (dd, J = 8.1, 1.8 Hz, 1H), 7.23 (d, J = 1.7 Hz, 1H), 7.12 (d, J = 8.2 Hz, 1H), 6.25 (s, 1H), 3.17 (d, J = 6.8 Hz, 2H), 4.61 (s, 1H), 3.49 (q, J = 7.0, 6.8, 6.3 Hz, 2H), 2.29 (s, 3H), 1.79 (q, J = 3.9 Hz, 2H), 1.56 (q, J = 7.2 Hz, 2H), 1.46 (s, 11H), 1.36 – 1.27 (m, 2H), 1.20 (q, J = 3.9 Hz, 2H), 0.97 – 0.89 (m, 2H).

**LC/MS** [M+H]<sup>+</sup> m/z calc. 637.28, found 637.3.

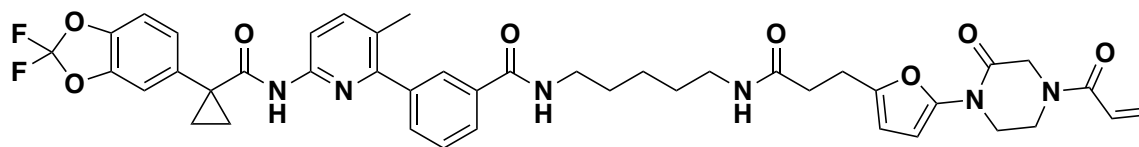


**5c**

***N*-(5-aminopentyl)-3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide (5c)**: Intermediate **4c** (240 mg, 0.038 mmol) was dissolved in DCM (2 mL) and TFA (2 mL) was added and the solution stirred for 2 hours. The volatiles were then evaporated and the resulting oil redissolved in DCM and treated with aqueous saturated NaHCO<sub>3</sub>. The layers were separated and the aqueous layer was then extracted with DCM three times. The combined organic extracts were dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated to provide the title compound (184 mg, 0.34 mmol, 85% over two steps) as a colorless oil which was used in the next step without further purification.

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 8.09 (d, J = 8.4 Hz, 1H), 7.80 (t, J = 1.8 Hz, 1H), 7.76 (dd, J = 7.7, 1.5 Hz, 1H), 7.69 (s, 1H), 7.59 (d, J = 8.5 Hz, 1H), 7.57 – 7.50 (m, 1H), 7.47 (t, J = 7.6 Hz, 1H), 7.23 (dd, J = 8.2, 1.7 Hz, 1H), 7.19 (d, J = 1.8 Hz, 1H), 7.08 (d, J = 8.2 Hz, 1H), 6.30 (s, 1H), 3.45 (q, J = 6.7 Hz, 2H), 2.74 (t, J = 6.8 Hz, 2H), 2.25 (s, 3H), 1.65 – 1.59 (m, 2H), 1.57 – 1.47 (m, 2H), 1.48 – 1.40 (m, 2H), 1.33 – 1.23 (m, 2H), 1.20 – 1.12 (m, 2H), 0.91 – 0.85 (m, 2H).

**LC/MS** [M+H]<sup>+</sup> m/z calc. 537.22, found 537.2.



**NJH-2-057**

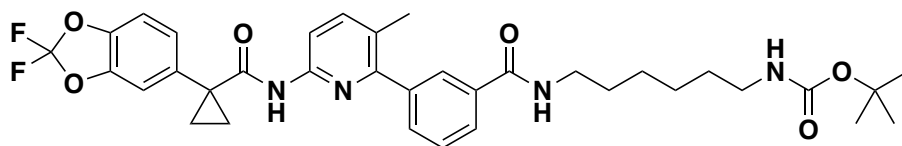
***N*-(5-(3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)propanamido)pentyl)-3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide (NJH-2-057)**: Intermediate **3** (tert-butyl 3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)propanoate) (70 mg, 0.20 mmol) was dissolved in DCM (1.0 mL) and TFA (0.8 mL) was added and the solution stirred for 1 h until starting material was consumed as monitored by TLC. The volatiles were evaporated, DCM was added and evaporated again. The residue was dissolved in DMF (1.5 mL) and DIEA (150 μL, 0.86 mmol) was added followed by intermediate **5c** (*N*-(5-aminopentyl)-3-(6-(1-

(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide) (54 mg, 0.1 mmol). HATU (152 mg, 0.4 mmol) was then added and the mixture stirred for 1 h. Water was added, and the resulting suspension was extracted with DCM three times. Combined organic extracts were washed twice with 1M HCl twice, saturated NaHCO<sub>3</sub>, twice with 5% LiCl, brine, and dried over Na<sub>2</sub>SO<sub>4</sub> before being concentrated. The crude residue was purified by silica gel chromatography (0-4% MeOH/DCM) to obtain the title compound (35 mg, 0.043 mmol, 43%) as a white powder following lyophilization from 1:1 water:acetonitrile (2 mL).

**<sup>1</sup>H NMR** (600 MHz, CDCl<sub>3</sub>) δ 8.11 (d, J = 8.4 Hz, 1H), 7.85 (t, J = 1.8 Hz, 1H), 7.81 (dt, J = 7.8, 1.5 Hz, 1H), 7.71 (s, 1H), 7.61 (d, J = 8.5 Hz, 1H), 7.55 (dt, J = 7.7, 1.4 Hz, 1H), 7.49 (t, J = 7.6 Hz, 1H), 7.25 (dd, J = 8.2, 1.8 Hz, 1H), 7.21 (d, J = 1.8 Hz, 1H), 7.10 (d, J = 8.2 Hz, 1H), 6.53 (s, 1H), 6.41 (dd, J = 16.7, 1.8 Hz, 2H), 6.22 (d, J = 3.3 Hz, 1H), 6.03 (d, J = 3.3 Hz, 1H), 5.82 (dd, J = 10.4, 1.8 Hz, 2H), 4.54 – 4.32 (m, 2H), 4.07 – 3.79 (m, 4H), 3.45 (q, J = 6.6 Hz, 2H), 3.24 (q, J = 6.6 Hz, 2H), 2.91 (t, J = 7.3 Hz, 2H), 2.46 (t, J = 7.3 Hz, 2H), 2.27 (s, 3H), 1.77 (q, J = 3.9 Hz, 2H), 1.65 – 1.59 (m, 2H), 1.52 (p, J = 7.0 Hz, 2H), 1.40 – 1.32 (m, 2H), 1.18 (q, J = 3.9 Hz, 2H).

**<sup>13</sup>C NMR** (151 MHz, CDCl<sub>3</sub>) δ 171.7, 167.4, 165.0, 155.5, 148.9, 144.8, 144.1, 143.6, 141.0, 140.2, 134.9, 134.8, 131.8, 128.4, 127.5, 127.0, 126.6, 126.6, 126.3, 112.9, 112.4, 110.2, 107.4, 100.9, 39.7, 39.1, 31.2, 29.0, 24.2, 23.7, 19.2, 17.2.

**HRMS** [M+H]<sup>+</sup> m/z calc. 811.3262, found 811.3267.

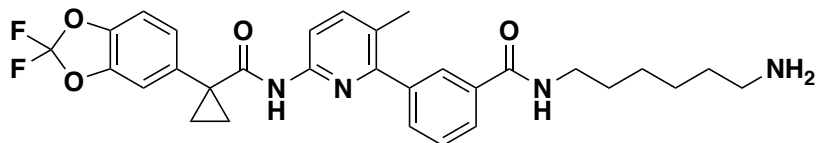


4d

**tert-butyl (6-(3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide)hexyl)carbamate (4d)**: Lumacaftor (100 mg, 0.22 mmol) and tert-butyl (6-aminohexyl)carbamate were reacted according to General Procedure A and purified by silica gel chromatography (0-60% EtOAc/Hex) to obtain intermediate **4d** (114 mg, 0.18 mmol, 80%) as a clear colorless oil.

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 8.14 (d, J = 8.4 Hz, 1H), 7.86 (s, 1H), 7.82 (d, J = 7.7 Hz, 1H), 7.72 (s, 1H), 7.63 (d, J = 8.5 Hz, 1H), 7.57 (dt, J = 7.6, 1.5 Hz, 1H), 7.51 (t, J = 7.6 Hz, 1H), 7.27 (dd, J = 8.2, 1.8 Hz, 1H), 7.23 (d, J = 1.7 Hz, 1H), 7.12 (d, J = 8.1 Hz, 1H), 6.37 (s, 1H), 4.58 (s, 1H), 3.48 (q, J = 6.7 Hz, 2H), 3.17 (q, J = 6.7 Hz, 2H), 2.29 (s, 3H), 1.79 (q, J = 3.9 Hz, 2H), 1.69 – 1.64 (m, 1H), 1.58 – 1.49 (m, 1H), 1.46 (s, 9H), 1.45 – 1.38 (m, 6H), 1.20 (q, J = 3.9 Hz, 2H).

**LC/MS** [M+H]<sup>+</sup> m/z calc. 651.3, found 651.2.

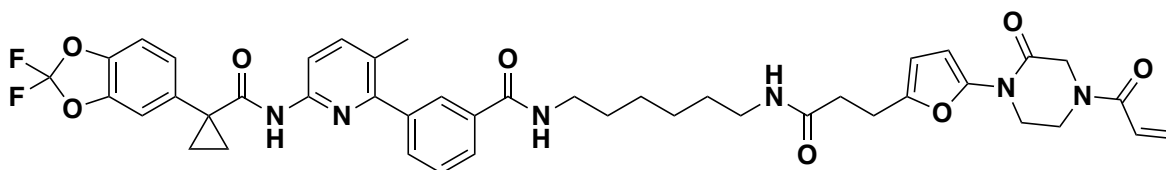


5d

**N-(6-aminohexyl)-3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide (5d):** **4d** (114 mg, 0.18 mmol) was deprotected according to General Procedure B to provide the amine **5d** (99 mg, 0.18 mmol, quant.) as a colorless oil.

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 8.10 (d, J = 8.4 Hz, 1H), 7.78 (s, 1H), 7.74 (dt, J = 7.5, 1.6 Hz, 1H), 7.69 (s, 1H), 7.59 (d, J = 8.4 Hz, 1H), 7.53 (dt, J = 7.7, 1.5 Hz, 1H), 7.47 (t, J = 7.6 Hz, 1H), 7.22 (dd, J = 8.2, 1.8 Hz, 1H), 7.18 (d, J = 1.6 Hz, 1H), 7.07 (d, J = 8.2 Hz, 1H), 6.17 (s, 1H), 3.44 (td, J = 7.2, 5.8 Hz, 2H), 2.68 (t, J = 6.8 Hz, 2H), 2.24 (s, 3H), 1.99 (s, 1H), 1.81 (s, 1H), 1.74 (q, J = 3.9 Hz, 2H), 1.67 – 1.55 (m, 3H), 1.51 – 1.33 (m, 5H), 1.16 (q, J = 3.9 Hz, 2H).

**LC/MS** [M+H]<sup>+</sup> *m/z* calc. 551.2, found 551.2.



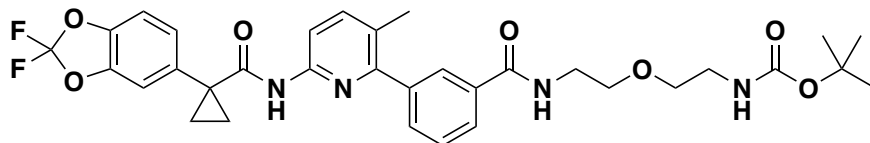
NJH-2-133

**N-(6-(3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)propanamido)hexyl)-3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide (NJH-2-133):** Intermediate **3** (30 mg, 0.086 mmol) was deprotected and coupled to intermediate **5d** (16 mg, 0.029 mmol) following General Procedure C to provide **NJH-2-133** (17.4 mg, 0.021 mmol, 73%) as a clear colorless oil.

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 8.12 (d, J = 8.4 Hz, 1H), 7.89 – 7.79 (m, 2H), 7.73 (s, 1H), 7.62 (d, J = 8.5 Hz, 1H), 7.56 (dt, J = 7.7, 1.5 Hz, 1H), 7.51 (t, J = 7.6 Hz, 1H), 7.27 (dd, J = 8.2, 1.8 Hz, 1H), 7.22 (d, J = 1.7 Hz, 1H), 7.11 (d, J = 8.2 Hz, 1H), 6.54 (d, J = 31.0 Hz, 2H), 6.41 (dd, J = 16.8, 1.9 Hz, 1H), 6.25 (d, J = 3.3 Hz, 1H), 6.07 (d, J = 3.3 Hz, 1H), 5.98 (d, J = 39.7 Hz, 1H), 5.83 (dd, J = 10.3, 2.0 Hz, 1H), 4.42 (d, J = 21.6 Hz, 2H), 4.05 – 3.81 (m, 4H), 3.74 (p, J = 6.7 Hz, 2H), 3.45 (q, J = 6.7 Hz, 2H), 3.22 (dq, J = 13.2, 6.9 Hz, 3H), 2.94 (q, J = 6.4, 5.5 Hz, 2H), 2.52 (t, J = 7.4 Hz, 2H), 2.28 (s, 3H), 1.78 (q, J = 3.9 Hz, 2H), 1.61 (p, J = 6.9 Hz, 2H), 1.42 – 1.30 (m, 3H), 1.20 (q, J = 3.9 Hz, 2H).

**<sup>13</sup>C NMR** (151 MHz, CDCl<sub>3</sub>) δ 171.8, 167.3, 155.5, 149.7, 148.9, 144.7, 144.1, 143.6, 141.0, 140.2, 134.9, 134.9, 133.4, 131.7, 131.7, 130.0, 128.5, 127.5, 127.0, 126.6, 126.6, 126.4, 112.9, 112.4, 110.2, 107.3, 100.8, 55.6, 43.6, 39.6, 39.1, 34.8, 31.2, 29.4, 29.3, 26.0, 25.9, 24.2, 19.1, 18.6, 17.2, 12.5.

**HRMS (ESI):** [M+H]<sup>+</sup> *m/z* calc. 825.3345, found 825.3425.



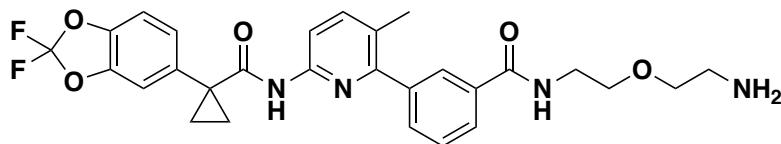
4e

**tert-butyl (2-(2-(3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamido)ethoxy)ethyl)carbamate (4e):**

Lumacaftor (100 mg, 0.22 mmol) and tert-butyl (2-(2-aminoethoxy)ethyl)carbamate (57 mg, 0.28 mmol) were reacted according to General Procedure A and purified by silica gel chromatography (0-60% EtOAc/Hex) to obtain intermediate **4e** (122 mg, 0.19 mmol, 87%) as a clear colorless oil.

**<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.14 (d, *J* = 8.4 Hz, 1H), 7.88 (t, *J* = 1.8 Hz, 1H), 7.81 (dt, *J* = 7.5, 1.6 Hz, 1H), 7.72 (s, 1H), 7.63 (d, *J* = 8.5 Hz, 1H), 7.59 (dt, *J* = 7.7, 1.5 Hz, 1H), 7.52 (t, *J* = 7.6 Hz, 1H), 7.27 (dd, *J* = 8.2, 1.8 Hz, 1H), 7.23 (d, *J* = 1.7 Hz, 1H), 7.12 (d, *J* = 8.2 Hz, 1H), 6.60 (s, 1H), 4.87 (s, 1H), 3.74 – 3.62 (m, 4H), 3.58 (t, *J* = 5.2 Hz, 2H), 3.41 – 3.31 (m, 2H), 2.29 (s, 3H), 1.79 (q, *J* = 3.9 Hz, 2H), 1.46 (s, 9H), 1.20 (q, *J* = 3.9 Hz, 2H).

**LC/MS** [M+H]<sup>+</sup> *m/z* calc. 639.3, found 639.2.



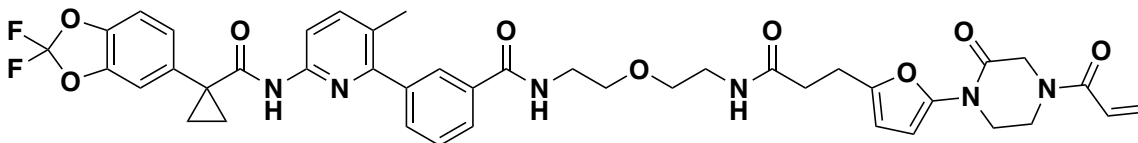
5e

**N-(2-(2-aminoethoxy)ethyl)-3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide (5e):**

Intermediate **4e** (122 mg, 0.19 mmol) was deprotected according to General Procedure B to provide the amine **5e** (102 mg, 0.19 mmol, quant.) as a colorless oil.

**<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.14 (d, *J* = 8.4 Hz, 1H), 7.88 (t, *J* = 1.7 Hz, 1H), 7.85 (s, 1H), 7.77 (s, 1H), 7.63 (d, *J* = 8.5 Hz, 1H), 7.57 (dt, *J* = 7.7, 1.5 Hz, 1H), 7.51 (t, *J* = 7.6 Hz, 1H), 7.27 (dd, *J* = 8.1, 1.8 Hz, 1H), 7.23 (d, *J* = 1.7 Hz, 1H), 7.12 (d, *J* = 8.2 Hz, 1H), 6.91 (s, 0H), 3.70 (tdd, *J* = 7.9, 4.0, 1.2 Hz, 4H), 3.55 (t, *J* = 5.2 Hz, 2H), 2.91 (t, *J* = 5.2 Hz, 2H), 2.29 (s, 3H), 1.79 (q, *J* = 3.9 Hz, 2H), 1.20 (q, *J* = 3.9 Hz, 2H).

**LC/MS**: [M+H]<sup>+</sup> *m/z* calc. 539.2 found 639.2.



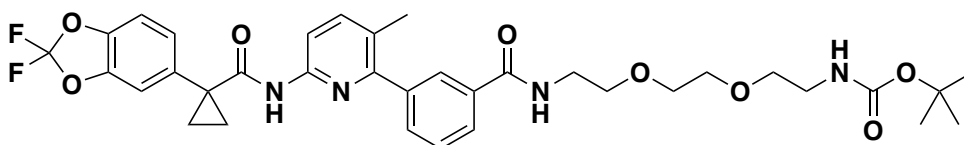
LEB-3-149

**N-(2-(2-(3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)propanamido)ethoxy)ethyl)-3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide (LEB-03-149).** Intermediate **3** (30 mg, 0.086 mmol) was deprotected and coupled to intermediate **5e** (23 mg, 0.043 mmol) following General Procedure C to provide **LEB-03-149** (10.9 mg, 0.0134 mmol, 31% yield) as a white foam.

**<sup>1</sup>H NMR** (600 MHz, Chloroform-*d*)  $\delta$  8.10 (d, *J* = 8.4 Hz, 1H), 7.88 (t, *J* = 1.8 Hz, 1H), 7.82 (dt, *J* = 7.7, 1.5 Hz, 1H), 7.72 (s, 1H), 7.60 (d, *J* = 8.5 Hz, 1H), 7.55 (dt, *J* = 7.6, 1.4 Hz, 1H), 7.48 (t, *J* = 7.7 Hz, 1H), 7.25 (dd, *J* = 8.2, 1.8 Hz, 1H), 7.21 (d, *J* = 1.7 Hz, 1H), 7.09 (d, *J* = 8.2 Hz, 1H), 6.86 (s, 1H), 6.39 (dd, *J* = 16.7, 1.8 Hz, 1H), 6.20 (d, *J* = 3.3 Hz, 1H), 6.02 (d, *J* = 3.2 Hz, 1H), 5.81 (dd, *J* = 10.4, 1.8 Hz, 1H), 3.82 (s, 2H), 3.73 (hept, *J* = 6.6 Hz, 2H), 3.63 (d, *J* = 4.1 Hz, 4H), 3.53 (t, *J* = 5.1 Hz, 2H), 3.41 (q, *J* = 5.3 Hz, 2H), 3.19 (q, *J* = 7.4 Hz, 2H), 2.89 (t, *J* = 7.5 Hz, 2H), 2.47 (t, *J* = 7.3 Hz, 2H), 2.26 (s, 3H), 1.48 (t, *J* = 7.4 Hz, 3H), 1.18 (q, *J* = 3.9 Hz, 2H), 0.12 – 0.06 (m, 1H).

**<sup>13</sup>C NMR** (151 MHz, CDCl<sub>3</sub>)  $\delta$  171.77, 167.53, 165.03, 155.44, 149.75, 148.91, 144.71, 144.11, 143.59, 140.95, 140.22, 134.94, 134.55, 131.91, 131.68, 128.46, 127.72, 126.98, 126.64, 126.36, 112.96, 112.39, 110.21, 107.27, 100.91, 69.63, 69.50, 55.72, 53.43, 43.65, 39.83, 39.18, 34.69, 31.20, 24.08, 19.14, 17.18, 12.52.

**HRMS (ESI):** [M+H]<sup>+</sup> *m/z* calc. 813.31, found 813.3055.



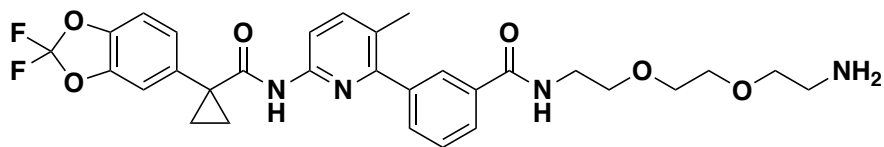
4f

**tert-butyl (2-(2-(2-(3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamido)ethoxy)ethoxy)ethyl)carbamate (4f):**

Lumacaftor (100 mg, 0.22 mmol) and tert-butyl (2-(2-(2-aminoethoxy)ethoxy)ethyl)carbamate (70 mg, 0.28 mmol) were reacted according to General Procedure A and purified by silica gel chromatography (0-80% EtOAc/Hex) to obtain intermediate **4f** (127 mg, 0.19 mmol, 85%) as a clear colorless oil.

**<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.14 (d, *J* = 8.4 Hz, 1H), 7.88 (s, 1H), 7.80 (d, *J* = 7.5 Hz, 1H), 7.77 – 7.72 (m, 1H), 7.63 (d, *J* = 8.4 Hz, 1H), 7.57 (d, *J* = 7.5 Hz, 1H), 7.52 (t, *J* = 7.6 Hz, 1H), 7.27 (dd, *J* = 8.2, 1.8 Hz, 1H), 7.23 (d, *J* = 1.7 Hz, 1H), 7.11 (d, *J* = 8.2 Hz, 1H), 6.74 (s, 1H), 5.02 (s, 1H), 3.75 – 3.61 (m, 8H), 3.56 (t, *J* = 5.4 Hz, 2H), 3.31 (d, *J* = 5.8 Hz, 2H), 2.28 (s, 3H), 1.79 (q, *J* = 3.9 Hz, 2H), 1.45 (s, 9H), 1.20 (q, *J* = 3.9 Hz, 2H).

**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 683.3, found 683.3.

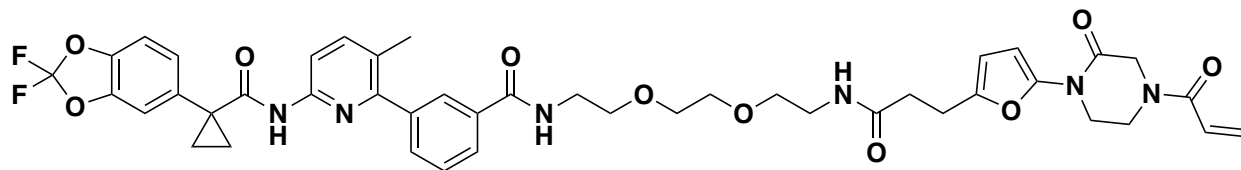


5f

**N-(2-(2-(2-aminoethoxy)ethoxy)ethyl)-3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide (5f):** Intermediate 4f (127 mg, 0.19 mmol) was deprotected according to General Procedure B to provide the amine 5f (111 mg, 0.19 mmol, quant.) as a colorless oil.

**<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.13 (d, *J* = 8.4 Hz, 1H), 7.89 (t, *J* = 1.7 Hz, 1H), 7.83 (dt, *J* = 7.7, 1.5 Hz, 1H), 7.78 (s, 1H), 7.63 (d, *J* = 8.5 Hz, 1H), 7.56 (dt, *J* = 7.7, 1.5 Hz, 1H), 7.51 (t, *J* = 7.6 Hz, 1H), 7.27 (dd, *J* = 8.2, 1.7 Hz, 1H), 7.23 (d, *J* = 1.7 Hz, 1H), 7.12 (d, *J* = 8.2 Hz, 1H), 7.08 (s, 1H), 3.73 – 3.62 (m, 9H), 3.51 (t, *J* = 5.2 Hz, 2H), 2.82 (t, *J* = 5.1 Hz, 2H), 2.28 (s, 3H), 1.79 (q, *J* = 3.9 Hz, 2H), 1.20 (q, *J* = 3.9 Hz, 2H).

**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 583.2, found 583.3.



LEB-3-150

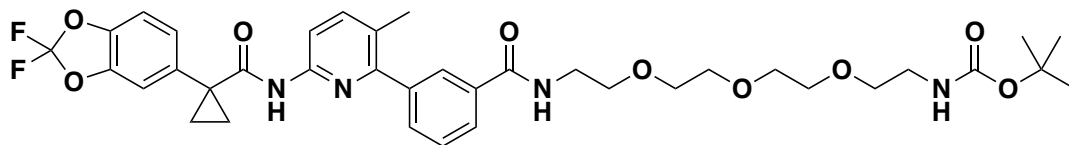
**N-(2-(2-(2-(3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)propanamido)ethoxy)ethoxy)ethyl)-3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide (LEB-03-150).**

Intermediate 3 (30 mg, 0.086 mmol) was deprotected and coupled to intermediate 5f (25 mg, 0.043 mmol) following General Procedure C to provide **LEB-03-150** (11.6 mg, 0.0134 mmol, 31% yield) as a clear colorless oil.

**<sup>1</sup>H NMR** (600 MHz, Chloroform-*d*)  $\delta$  8.11 (d, *J* = 8.4 Hz, 1H), 7.86 (tt, *J* = 1.8, 1.2 Hz, 1H), 7.79 (ddd, *J* = 7.7, 1.8, 1.2 Hz, 1H), 7.72 (s, 1H), 7.62 – 7.58 (m, 1H), 7.55 (ddd, *J* = 7.6, 1.7, 1.2 Hz, 1H), 7.48 (td, *J* = 7.7, 0.6 Hz, 1H), 7.25 (dd, *J* = 8.2, 1.8 Hz, 1H), 7.21 (d, *J* = 1.7 Hz, 1H), 7.09 (d, *J* = 8.2 Hz, 1H), 6.83 (d, *J* = 5.8 Hz, 1H), 6.41 (dd, *J* = 16.7, 1.8 Hz, 1H), 6.24 (d, *J* = 3.2 Hz, 1H), 6.18 (s, 1H), 6.05 (dd, *J* = 3.3, 1.0 Hz, 1H), 5.82 (dd, *J* = 10.5, 1.8 Hz, 1H), 5.32 (s, 1H), 4.40 (d, *J* = 39.8 Hz, 2H), 3.94 (d, *J* = 47.9 Hz, 1H), 3.85 (s, 2H), 3.70 – 3.58 (m, 7H), 3.50 (dd, *J* = 5.6, 4.8 Hz, 2H), 3.39 (q, *J* = 5.4 Hz, 2H), 2.93 (t, *J* = 7.5 Hz, 2H), 2.47 (t, *J* = 7.5 Hz, 2H), 2.25 (s, 3H), 2.19 (s, 1H), 1.76 (q, *J* = 3.8 Hz, 2H), 1.47 (d, *J* = 12.2 Hz, 1H), 1.18 (p, *J* = 3.8 Hz, 2H).

**<sup>13</sup>C NMR** 151 MHz, CDCl<sub>3</sub>)  $\delta$  171.78, 171.54, 167.31, 164.98, 155.46, 148.91, 144.68, 144.12, 143.60, 140.94, 140.25, 134.93, 134.64, 131.88, 131.68, 128.46, 127.72, 126.98, 126.63, 126.51, 126.34, 113.00, 112.39, 110.19, 107.18, 100.77, 70.23, 70.18, 69.80, 55.62, 53.43, 43.58, 39.81, 39.16, 34.67, 31.20, 30.92, 23.97, 19.13, 17.19, 12.47, 1.02.

**HRMS (ESI):**  $[M+H]^+$   $m/z$  calc. 857.33, found 857.3319.

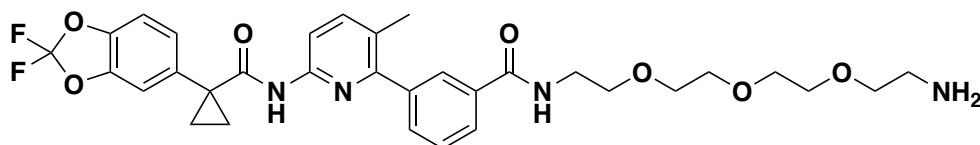


**4g**

**tert-butyl (1-(3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)phenyl)-1-oxo-5,8,11-trioxa-2-azatridecan-13-yl)carbamate (4g):** Lumacaftor (100 mg, 0.22 mmol) and tert-butyl (2-(2-(2-(2-aminoethoxy)ethoxy)ethoxy)ethyl)carbamate (82 mg, 0.28 mmol) were reacted according to General Procedure A and purified by silica gel chromatography (0-100% EtOAc/Hex) to obtain intermediate **4g** (139 mg, 0.19 mmol, 87%) as a colorless oil.

**<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.14 (d,  $J$  = 8.4 Hz, 1H), 7.89 (s, 1H), 7.82 (d,  $J$  = 7.5 Hz, 1H), 7.73 (s, 1H), 7.63 (d,  $J$  = 8.5 Hz, 1H), 7.57 (d,  $J$  = 7.5 Hz, 1H), 7.51 (t,  $J$  = 7.6 Hz, 1H), 7.27 (dd,  $J$  = 8.2, 1.8 Hz, 1H), 7.23 (d,  $J$  = 1.7 Hz, 1H), 7.12 (d,  $J$  = 8.2 Hz, 1H), 6.80 (s, 1H), 3.73 – 3.66 (m, 9H), 3.64 (dd,  $J$  = 6.1, 3.2 Hz, 2H), 3.59 (dd,  $J$  = 6.1, 3.2 Hz, 2H), 3.50 (t,  $J$  = 5.1 Hz, 2H), 3.30 (d,  $J$  = 5.7 Hz, 2H), 2.28 (s, 3H), 1.79 (q,  $J$  = 3.9 Hz, 2H), 1.46 (s, 9H), 1.20 (q,  $J$  = 3.9 Hz, 2H).

**LC/MS:**  $[M+H]^+$   $m/z$  calc. 727.3, found 727.2.



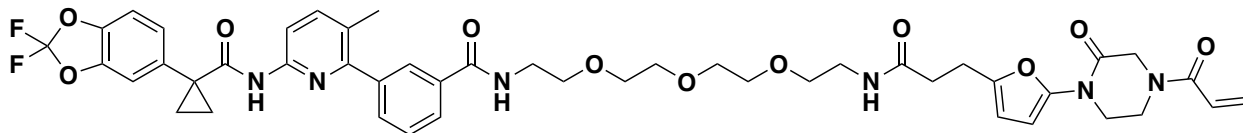
**5g**

**N-(2-(2-(2-(2-aminoethoxy)ethoxy)ethoxy)ethyl)-3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide (5g):** Intermediate **4g** (139 mg, 0.19 mmol) was deprotected according to General Procedure B to provide the amine **5g** (119 mg, 0.19 mmol, quant.) as a colorless oil.

**<sup>1</sup>H NMR** (400 MHz, Chloroform-*d*)  $\delta$  8.13 (d,  $J$  = 8.4 Hz, 1H), 7.93 (t,  $J$  = 1.8 Hz, 1H), 7.87 (dt,  $J$  = 7.6, 1.6 Hz, 1H), 7.76 (s, 1H), 7.62 (d,  $J$  = 8.5 Hz, 1H), 7.60 (s, 1H), 7.55 (dt,  $J$  = 7.7, 1.5 Hz, 1H), 7.50 (t,  $J$  = 7.6 Hz, 1H), 7.27 (dd,  $J$  = 8.2, 1.7 Hz, 1H), 7.23 (d,  $J$  = 1.7 Hz, 1H), 7.12 (d,  $J$  = 8.2 Hz, 1H), 3.73 – 3.63 (m, 9H), 3.61 (dt,  $J$  = 6.0, 1.8 Hz, 4H), 3.48 – 3.43 (m, 2H), 2.82 – 2.75 (m, 2H), 2.29 (s, 3H), 1.79 (q,  $J$  = 3.9 Hz, 2H), 1.20 (q,  $J$  = 3.9 Hz, 2H).

**LC/MS:**  $[M+H]^+$   $m/z$  calc. 627.3, found 627.3.





LEB-3-151

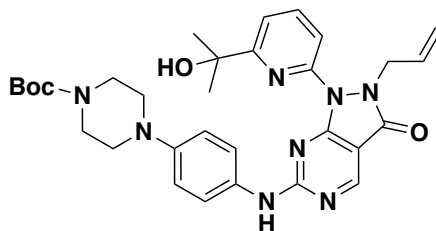
**N-(15-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)-13-oxo-3,6,9-trioxa-12-azapentadecyl)-3-(6-(1-(2,2-difluorobenzo[d][1,3]dioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)benzamide (LEB-03-151)**. Intermediate **3** (30 mg, 0.086 mmol) was deprotected and coupled to intermediate **5g** (27 mg, 0.043 mmol) following General Procedure C to provide **LEB-03-151** (13.7 mg, 0.0152 mmol, 35% yield) as a colorless oil.

**<sup>1</sup>H NMR** (600 MHz, Chloroform-*d*)  $\delta$  8.10 (d, *J* = 8.5 Hz, 1H), 7.87 (t, *J* = 1.8 Hz, 1H), 7.80 (dt, *J* = 7.8, 1.5 Hz, 1H), 7.73 (s, 1H), 7.60 (d, *J* = 8.5 Hz, 1H), 7.54 (dt, *J* = 7.7, 1.4 Hz, 1H), 7.48 (t, *J* = 7.7 Hz, 1H), 7.25 (dd, *J* = 8.2, 1.8 Hz, 1H), 7.21 (d, *J* = 1.7 Hz, 1H), 7.09 (d, *J* = 8.2 Hz, 1H), 6.93 (d, *J* = 6.0 Hz, 1H), 6.41 (dd, *J* = 16.7, 1.8 Hz, 1H), 6.25 (d, *J* = 3.2 Hz, 1H), 6.05 (d, *J* = 3.3 Hz, 1H), 5.82 (dd, *J* = 10.4, 1.8 Hz, 1H), 4.41 (d, *J* = 35.7 Hz, 2H), 3.95 (d, *J* = 50.4 Hz, 3H), 3.85 (s, 2H), 3.70 – 3.62 (m, 8H), 3.62 – 3.57 (m, 2H), 3.57 – 3.52 (m, 2H), 3.47 (dd, *J* = 5.6, 4.6 Hz, 2H), 3.39 (q, *J* = 5.3 Hz, 2H), 2.93 (t, *J* = 7.6 Hz, 2H), 2.47 (t, *J* = 7.6 Hz, 2H), 2.25 (s, 3H), 2.19 (s, 1H), 1.76 (q, *J* = 3.9 Hz, 2H), 1.18 (q, *J* = 3.9 Hz, 2H).

**<sup>13</sup>C NMR** (151 MHz, CDCl<sub>3</sub>)  $\delta$  171.78, 171.50, 167.25, 164.97, 155.52, 148.90, 144.64, 144.13, 143.60, 140.92, 140.21, 134.93, 134.65, 133.37, 131.81, 131.68, 129.98, 128.40, 127.78, 126.98, 126.62, 126.59, 126.35, 112.96, 112.38, 110.20, 107.12, 100.70, 70.43, 70.38, 70.18, 70.07, 69.85, 69.82, 53.43, 39.81, 39.19, 34.59, 31.20, 30.92, 23.93, 19.13, 17.18.

**HRMS (ESI)**: [M+H]<sup>+</sup> *m/z* calc. 901.36, found 901.3584.

### Synthesis of WEE1 DUBTACs:

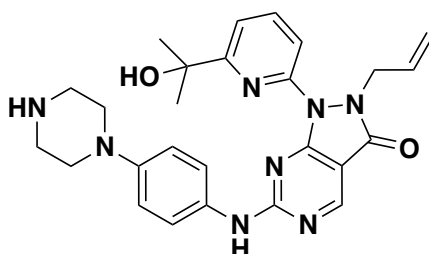


Intermediate 7

**tert-butyl 4-(4-((2-allyl-1-(6-(2-hydroxypropan-2-yl)pyridin-2-yl)-3-oxo-2,3-dihydro-1H-pyrazolo[3,4-d]pyrimidin-6-yl)amino)phenyl)piperazine-1-carboxylate (Intermediate 7)**. Commercially available **Intermediate 6** 2-allyl-1-(6-(2-hydroxypropan-2-yl)pyridin-2-yl)-6-(methylthio)-1,2-dihydro-3H-pyrazolo[3,4-d]pyrimidin-3-one (250 mg, 0.7 mmol) was dissolved in 7 mL of toluene and cooled to 0°C. meta-Chloroperoxybenzoic acid (190 mg, 0.77 mmol) was added to the reaction mixture on ice, and the reaction mixture was warmed to room temperature and stirred for 1 hour. N,N-Diisopropylethylamine (365 mL, 2.1 mmol) and 1-Piperazinecarboxylic acid, 4-(4-

aminophenyl)-, 1,1-dimethylethyl ester (232 mg, 0.84 mmol) were then added slowly and the reaction mixture was stirred overnight. The reaction mixture was extracted in EtOAc, washed 3X with brine, and dried on silica. Purification by flash column chromatography (DCM/Hexane 5:95) yielded **Intermediate 7** (0.445 mmol, 64% yield).

<sup>1</sup>H NMR (400 MHz, Chloroform-d) δ 8.99 (s, 1H), 7.95 (t, J = 7.9 Hz, 1H), 7.80 (dd, J = 8.1, 0.8 Hz, 1H), 7.44 (dd, J = 7.7, 0.8 Hz, 1H), 5.83 – 5.65 (m, 1H), 5.13 – 5.04 (m, 1H), 4.97 (dq, J = 17.1, 1.4 Hz, 1H), 4.85 (dt, J = 6.2, 1.4 Hz, 2H), 3.80 (s, 1H), 2.63 (s, 3H), 1.63 (s, 6H).

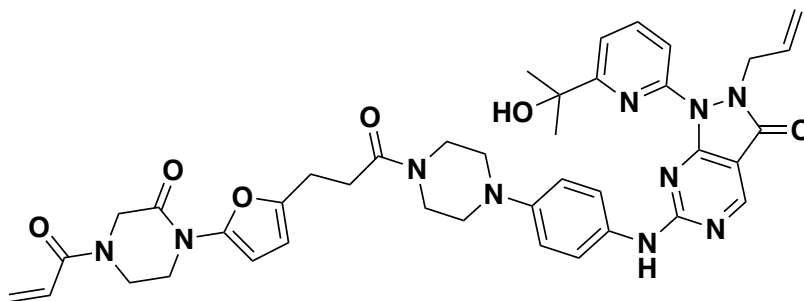


**Intermediate 8**

**2-allyl-1-(6-(2-hydroxypropan-2-yl)pyridin-2-yl)-6-((4-(piperazin-1-yl)phenyl)amino)-1,2-dihydro-3H-pyrazolo[3,4-d]pyrimidin-3-one (Intermediate 8).** Intermediate 7 (261 mg, 0.445 mmol) was dissolved in 4mL of DCM and cooled to 0°C. 1mL of trifluoroacetic acid was added dropwise on ice. The reaction mixture was stirred at room temperature for 1 hour, then extracted in DCM, washed 3X with brine, and dried on silica. Purification by flash column chromatography (DCM/Hexane 5:95) yielded **Intermediate 8** (0.398 mmol, 89% yield).

<sup>1</sup>H NMR (400 MHz, Chloroform-d) δ 8.84 (s, 1H), 7.86 (t, J = 7.9 Hz, 1H), 7.75 (d, J = 8.1 Hz, 1H), 7.48 (d, J = 8.5 Hz, 2H), 7.34 (d, J = 7.6 Hz, 1H), 6.93 (d, J = 9.0 Hz, 2H), 5.78 – 5.59 (m, 1H), 5.04 (d, J = 10.2 Hz, 1H), 4.94 (d, J = 17.0 Hz, 1H), 4.74 (d, J = 6.2 Hz, 2H), 3.94 (s, 1H), 3.60 (t, J = 5.1 Hz, 4H), 3.11 (t, J = 5.1 Hz, 4H), 2.05 (s, 1H), 1.59 (s, 6H), 1.49 (s, 9H).

### Synthesis of LEB-03-153



**LEB-03-153**

**6-((4-(4-(3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)propanoyl)piperazin-1-yl)phenyl)amino)-2-allyl-1-(6-(2-hydroxypropan-2-yl)pyridin-2-yl)-1,2-dihydro-3H-**

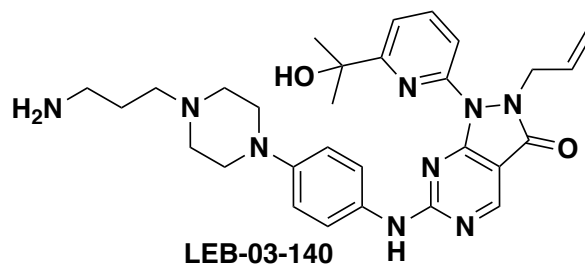
**pyrazolo[3,4-d]pyrimidin-3-one (LEB-03-153).** LEB-03-139 (0.0449 mmol) was dissolved in 3 mL of DCM and the reaction mixture was cooled on ice. 1 mL of trifluoroacetic acid was added dropwise and the solution was warmed to room temperature and stirred for 1 hour. The deprotected amine salt was washed twice with DCM and dried under vacuum. Immediately following deprotection, the crude product was dissolved in 0.5 mL DMF and **deprotected intermediate 3** (.0898 mmol) was added to the mixture, followed by DIPEA (0.449 mmol) and HATU (0.0898 mmol). The reaction was stirred for 30 minutes before water was added. The mixture was extracted three times with EtOAc, and combined organic extracts were washed with brine, dried over sodium sulfate, and concentrated. Purification by flash column chromatography (MeOH:DCM 8:92) yielded **LEB-03-153** as a light-yellow solid (12.9 mg, 0.0169 mmol, 38% yield).

**<sup>1</sup>H NMR** (600 MHz, Chloroform-*d*)  $\delta$  8.76 (s, 1H), 7.80 (t,  $J = 7.9$  Hz, 1H), 7.66 (d,  $J = 8.0$  Hz, 1H), 7.43 (d,  $J = 8.4$  Hz, 2H), 7.29 (d,  $J = 7.6$  Hz, 1H), 7.19 (s, 1H), 6.87 – 6.82 (m, 2H), 6.45 (s, 1H), 6.34 (dd,  $J = 16.7, 1.8$  Hz, 1H), 6.20 (d,  $J = 3.2$  Hz, 1H), 6.01 (d,  $J = 3.2$  Hz, 1H), 5.74 (t,  $J = 11.1, 10.6$  Hz, 1H), 5.67 – 5.59 (m, 1H), 5.23 (s, 1H), 4.97 (dd,  $J = 9.8, 0.8$  Hz, 1H), 4.87 (dd,  $J = 17.4, 0.8$  Hz, 1H), 4.67 (d,  $J = 6.2$  Hz, 2H), 4.38 (s, 1H), 4.31 (s, 1H), 3.73 (t,  $J = 5.2$  Hz, 2H), 3.55 (t,  $J = 5.1$  Hz, 2H), 3.06 (t,  $J = 5.2$  Hz, 4H), 2.92 (t,  $J = 7.8$  Hz, 2H), 2.62 (d,  $J = 8.4$  Hz, 2H), 1.59 (s, 4H), 1.52 (s, 6H).

**<sup>13</sup>C NMR** (151 MHz, Chloroform-*d*)  $\delta$  169.96, 165.90, 165.00, 162.18, 161.36, 161.26, 156.36, 150.16, 147.68, 147.51, 144.67, 138.85, 131.56, 131.29, 126.30, 119.07, 117.20, 116.21, 116.12, 107.23, 101.12, 72.46, 50.15, 49.85, 49.49, 47.67, 45.40, 41.64, 31.55, 30.56, 23.71.

**HRMS (ESI):** [M+H]<sup>+</sup>  $m/z$  calc. 761.35, found 761.3522.

### Synthesis of LEB-03-144

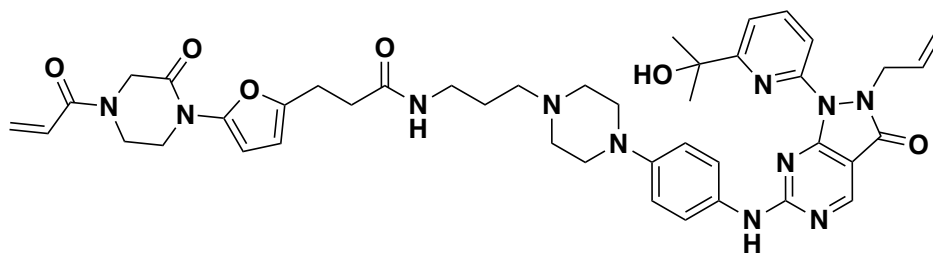


**2-allyl-6-((4-(4-(3-aminopropyl)piperazin-1-yl)phenyl)amino)-1-(6-(2-hydroxypropan-2-yl)pyridin-2-yl)-1,2-dihydro-3H-pyrazolo[3,4-d]pyrimidin-3-one (LEB-03-140).** Intermediate 8 (40 mg, 0.0823 mmol) was dissolved in 0.5 mL of DMF. tert-butyl (3-bromopropyl)carbamate (24 mg, 1.2 eq, 0.0987 mmol) and potassium carbonate (34 mg, 3.0 eq, 0.247 mmol) were added to the mixture, and the reaction was warmed to 50°C and stirred overnight. Water was added, the mixture extracted three times with EtOAc, combined organic extracts were washed with brine, and dried over sodium sulfate, and concentrated. Purification by flash column chromatography (EtOAc:Hexanes 50:50) yielded the boc-protected intermediate. This was immediately dissolved in 3 mL of DCM and the reaction mixture was cooled on ice. 1 mL of trifluoroacetic acid was added dropwise and the solution was warmed to room temperature and stirred for 1 hour. The deprotected amine TFA salt was washed twice

with DCM and dried under vacuum to yield **LEB-03-140** (33 mg, 0.0497 mmol, 60% yield over two steps) as a yellow oil.

**<sup>1</sup>H NMR** (300 MHz, Chloroform-d)  $\delta$  8.80 (s, 1H), 7.92 (t, J = 7.9 Hz, 1H), 7.63 (d, J = 8.0 Hz, 1H), 7.54 (d, J = 8.4 Hz, 3H), 6.90 (d, J = 8.9 Hz, 2H), 5.75 – 5.54 (m, 1H), 5.05 (d, J = 10.2 Hz, 1H), 4.89 (d, J = 17.1 Hz, 1H), 4.75 (d, J = 6.2 Hz, 2H), 3.66 (s, 1H), 3.43 (s, 9H), 3.28 (q, J = 9.4, 8.5 Hz, 2H), 3.19 (s, 1H), 3.06 (t, J = 7.1 Hz, 2H), 2.23 (d, J = 8.2 Hz, 3H), 1.59 (s, 6H).

**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 544.3, found 544.3.



**LEB-03-144**

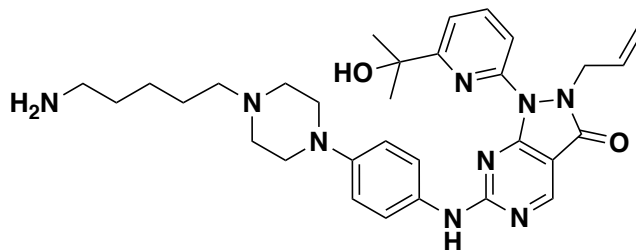
**3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)-N-(3-(4-(4-((2-allyl-1-(6-(2-hydroxypropan-2-yl)pyridin-2-yl)-3-oxo-2,3-dihydro-1H-pyrazolo[3,4-d]pyrimidin-6-yl)amino)phenyl)piperazin-1-yl)propyl)propenamide (LEB-03-144).** Intermediate **3** (19 mg, 0.0558 mmol) and **LEB-3-140** (0.0497 mmol) were reacted according to general procedure C. After hydrolysis, deprotected **3** and **LEB-3-140** were dissolved in DMF (0.5 mL), followed by DIPEA (43 mL, 0.249 mmol) and HATU (23 mg, 0.0596 mmol). The reaction was stirred for 30 minutes. Water was added and the mixture extracted three times with 4:1 CHCl<sub>3</sub>:IPA. Combined organic extracts were washed with brine, and dried over sodium sulfate, and concentrated. Purification by prep TLC (10% MeOH in DCM) yielded **LEB-03-144** as a light-yellow solid (8.1 mg, 0.0099 mmol, 20% yield).

**<sup>1</sup>H NMR** (600 MHz, DMSO)  $\delta$  10.07 (s, 1H), 8.75 (s, 1H), 7.97 (s, 1H), 7.83 (t, J = 5.6 Hz, 1H), 7.68 (d, J = 8.1 Hz, 1H), 7.54 (d, J = 7.7 Hz, 1H), 7.51 (s, 2H), 6.85 (d, J = 8.6 Hz, 2H), 6.80 – 6.72 (m, 1H), 6.16 – 6.08 (m, 2H), 6.04 (d, J = 3.2 Hz, 1H), 5.71 – 5.66 (m, 1H), 5.64 – 5.55 (m, 1H), 5.24 (s, 1H), 4.92 (d, J = 10.2 Hz, 1H), 4.76 (d, J = 17.0 Hz, 1H), 4.61 (d, J = 6.0 Hz, 2H), 4.27 (d, J = 93.6 Hz, 2H), 3.95 – 3.63 (m, 4H), 3.02 (q, J = 6.4 Hz, 6H), 2.73 (t, J = 7.5 Hz, 2H), 2.31 (t, J = 7.5 Hz, 2H), 2.24 (t, J = 7.2 Hz, 2H), 1.55 – 1.47 (m, 2H), 1.39 (s, 2H), 1.17 (s, 6H), 0.80 – 0.74 (m, 2H).

**<sup>13</sup>C NMR** (151 MHz, DMSO)  $\delta$  171.0, 168.0, 164.6, 161.6, 156.5, 150.1, 139.3, 132.7, 131.3, 128.8, 118.7, 116.8, 115.9, 106.9, 100.5, 72.8, 55.9, 53.2, 49.2, 47.6, 47.1, 46.9, 42.5, 37.4, 34.7, 33.8, 31.4, 30.9, 29.5, 26.9, 25.3, 24.0, 22.6, 22.5, 14.4.

**HRMS (ESI):** [M+H]<sup>+</sup> *m/z* calc. 818.41, found 818.4101.

### **Synthesis of LEB-3-145**

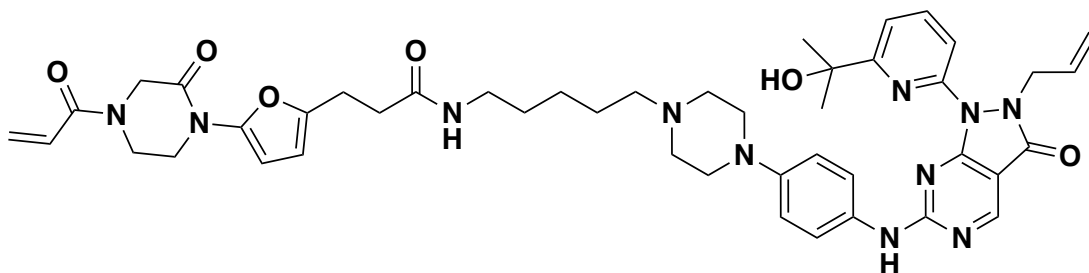


LEB-03-141

**2-allyl-6-((4-(4-(5-aminopentyl)piperazin-1-yl)phenyl)amino)-1-(6-(2-hydroxypropan-2-yl)pyridin-2-yl)-1,2-dihydro-3H-pyrazolo[3,4-d]pyrimidin-3-one (LEB-03-141).** Intermediate 8 (40 mg, 0.0823 mmol) was dissolved in 0.5 mL of DMF. tert-butyl (5-bromopentyl)carbamate (26 mg, 1.2 eq, 0.0987 mmol) and potassium carbonate (34 mg, 3.0 eq, 0.247 mmol) were added to the mixture, and the reaction was warmed to 50°C and stirred overnight. Water was added, the mixture extracted three times with EtOAc, combined organic extracts were washed with brine, and dried over sodium sulfate, and concentrated. Purification by flash column chromatography (EtOAc:Hexanes 50:50) yielded boc-protected intermediate. This was immediately dissolved in 3mL of DCM and the reaction mixture was cooled on ice. 1mL of trifluoroacetic acid was added dropwise and the solution was warmed to room temperature and stirred for 1 hour. The deprotected amine TFA salt was washed twice with DCM and dried under vacuum to yield **LEB-03-141** (21 mg, 0.0307 mmol, 37% yield over two steps) as a yellow oil.

<sup>1</sup>H NMR (300 MHz, Chloroform-d) δ 8.80 (s, 1H), 8.12 (s, 1H), 7.94 (t, J = 7.9 Hz, 1H), 7.64 (d, J = 8.0 Hz, 1H), 7.56 (d, J = 8.2 Hz, 3H), 6.92 (d, J = 8.8 Hz, 2H), 5.67 (dd, J = 16.8, 10.4 Hz, 1H), 5.07 (d, J = 10.2 Hz, 1H), 4.90 (d, J = 17.1 Hz, 1H), 4.76 (d, J = 6.2 Hz, 2H), 3.69 (s, 1H), 3.55 – 3.47 (m, 8H), 3.22 (s, 1H), 3.19 – 2.89 (m, 4H), 1.91 – 1.66 (m, 4H), 1.61 (s, 6H), 1.51 (s, 2H), 1.27 (s, 1H).

LC/MS: [M+H]<sup>+</sup> *m/z* calc. 572.3, found 572.3.



LEB-03-145

**3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)-N-(5-(4-(4-((2-allyl-1-(6-(2-hydroxypropan-2-yl)pyridin-2-yl)-3-oxo-2,3-dihydro-1H-pyrazolo[3,4-d]pyrimidin-6-yl)amino)phenyl)piperazin-1-yl)pentyl)propanamide (LEB-03-145).** Intermediate 3 (19 mg, 0.0558 mmol) and **LEB-2-141** (21 mg, 0.0307 mmol) were coupled according to general procedure C. After hydrolysis, deprotected 3 and **LEB-2-141** were dissolved in DMF (0.5 mL), followed by DIPEA (27 mL, 0.153 mmol) and HATU (14 mg, 0.0368 mmol). Water was added and the mixture extracted three times with 4:1 CHCl<sub>3</sub>:IPA. Combined organic extracts were washed with brine, and dried over sodium sulfate, and

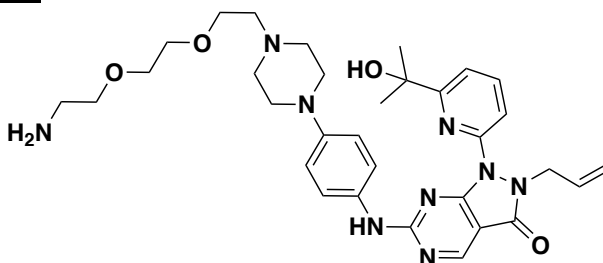
concentrated. Purification by prep TLC (8% MeOH in DCM) yielded **LEB-03-145** as a light-yellow solid (10.1 mg, 0.0119 mmol, 39% yield).

**<sup>1</sup>H NMR** (600 MHz, DMSO-d<sub>6</sub>) δ 10.15 (s, 1H), 8.83 (s, 1H), 8.05 (s, 1H), 7.86 (t, J = 5.6 Hz, 1H), 7.78 – 7.72 (m, 1H), 7.61 (d, J = 7.7 Hz, 2H), 7.58 (s, 2H), 6.92 (d, J = 8.7 Hz, 2H), 6.88 – 6.76 (m, 1H), 6.24 – 6.19 (m, 1H), 6.10 (d, J = 3.2 Hz, 1H), 5.76 (q, J = 9.8, 8.3 Hz, 1H), 5.72 – 5.61 (m, 1H), 5.36 – 5.26 (m, 1H), 5.00 (dq, J = 10.3, 1.3 Hz, 1H), 4.84 (dq, J = 17.2, 1.5 Hz, 1H), 4.69 (d, J = 6.0 Hz, 2H), 4.43 (s, 1H), 4.27 (s, 1H), 3.95 (d, J = 5.8 Hz, 1H), 3.86 (s, 1H), 3.82 – 3.73 (m, 2H), 3.13 – 3.08 (m, 4H), 3.08 – 3.01 (m, 2H), 2.80 (t, J = 7.5 Hz, 2H), 2.38 (t, J = 7.5 Hz, 2H), 2.30 (t, J = 7.4 Hz, 2H), 1.47 (s, 6H), 1.45 – 1.37 (m, 2H), 1.26 – 1.21 (m, 6H), 0.89 – 0.81 (m, 2H).

**<sup>13</sup>C NMR** (151 MHz, DMSO) δ 170.97, 168.04, 161.64, 156.46, 150.07, 139.28, 132.67, 128.77, 118.72, 115.93, 106.92, 100.44, 72.78, 58.33, 53.28, 49.17, 47.57, 47.06, 46.88, 42.46, 38.88, 33.80, 30.92, 29.54, 29.48, 29.16, 26.48, 24.81, 24.00, 22.56, 14.42.

**HRMS (ESI):** [M+H]<sup>+</sup> *m/z* calc. 846.44, found 846.4395.

### **Synthesis of LEB-3-146**

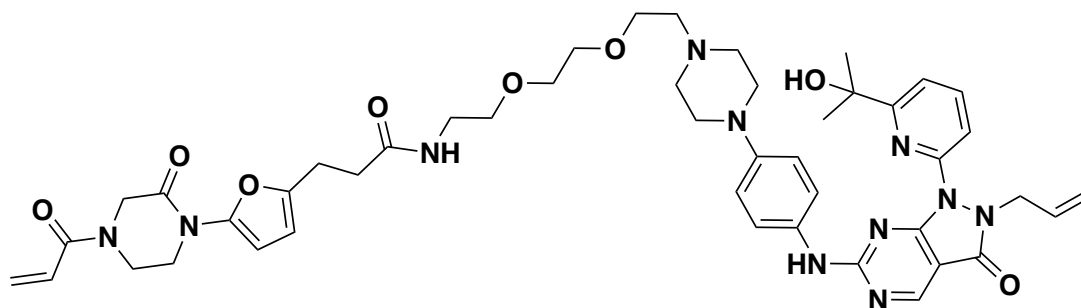


**LEB-03-142**

**2-allyl-6-((4-(4-(2-(2-(2-aminoethoxy)ethoxy)ethyl)piperazin-1-yl)phenyl)amino)-1-(6-(2-hydroxypropan-2-yl)pyridin-2-yl)-1,2-dihydro-3H-pyrazolo[3,4-d]pyrimidin-3-one (LEB-03-142).** Intermediate 8 (40 mg, 0.0823 mmol) was dissolved in 0.5 mL of DMF. tert-butyl (2-(2-(bromomethoxy)ethoxy)ethyl)carbamate (31 mg, 1.2 eq, 0.0987 mmol) and potassium carbonate (34 mg, 3.0 eq, 0.247 mmol) were added to the mixture, and the reaction was warmed to 50°C and stirred overnight. Water was added, the mixture extracted three times with EtOAc, combined organic extracts were washed with brine, and dried over sodium sulfate, and concentrated. Purification by flash column chromatography (EtOAc:Hexanes 50:50) yielded boc-protected intermediate. This was immediately dissolved in 3mL of DCM and the reaction mixture was cooled on ice. 1mL of trifluoroacetic acid was added dropwise and the solution was warmed to room temperature and stirred for 1 hour. The deprotected amine TFA salt was washed twice with DCM and dried under vacuum to yield **LEB-03-142** (28 mg, 0.0389 mmol, 47% yield).

**<sup>1</sup>H NMR** (300 MHz, Chloroform-d) δ 10.99 (s, 1H), 8.74 (s, 1H), 8.25 (s, 1H), 7.97 (t, J = 7.9 Hz, 1H), 7.60 (t, J = 8.9 Hz, 2H), 7.50 (d, J = 8.5 Hz, 2H), 6.87 (d, J = 8.7 Hz, 2H), 5.66 (ddd, J = 16.5, 10.3, 5.6 Hz, 1H), 5.07 (d, J = 10.2 Hz, 1H), 4.90 (d, J = 17.1 Hz, 1H), 4.75 (d, J = 6.3 Hz, 4H), 3.87 (d, J = 4.6 Hz, 4H), 3.76 – 3.69 (m, 4H), 3.65 (s, 4H), 3.39 – 3.10 (m, 8H), 1.61 (s, 6H).

**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 618.3, found 618.3.



LEB-03-146

**3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)-N-(2-(2-(2-(4-(4-((2-allyl-1-(6-(2-hydroxypropan-2-yl)pyridin-2-yl)-3-oxo-2,3-dihydro-1H-pyrazolo[3,4-d]pyrimidin-6-yl)amino)phenyl)piperazin-1-yl)ethoxy)ethoxy)ethyl)propanamide (LEB-03-146).**

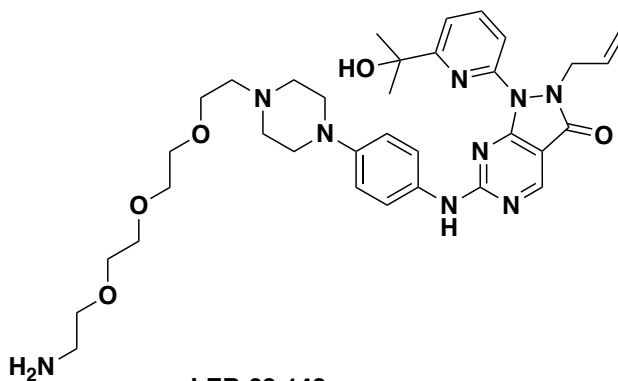
**Intermediate 3** (19 mg, 0.0558 mmol) and **LEB-3-142** (28 mg, 0.0389 mmol) were coupled according to general procedure C. After hydrolysis, deprotected **3** and **LEB-3-142** were dissolved in DMF (0.5 mL), followed by DIPEA (34 mL, 0.195 mmol) and HATU (18 mg, 0.0466 mmol). The reaction was stirred for 30 minutes. Water was added and the mixture extracted three times with 4:1 CHCl<sub>3</sub>:IPA. Combined organic extracts were washed with brine, and dried over sodium sulfate, and concentrated. Purification by prep TLC (8% MeOH in DCM) yielded **LEB-03-146** as a light-yellow solid (8.3 mg, 0.0093 mmol, 17% yield).

**<sup>1</sup>H NMR** (600 MHz, DMSO-d<sub>6</sub>) δ 8.83 (s, 1H), 8.05 (s, 1H), 7.97 (t, J = 5.8 Hz, 1H), 7.76 (s, 1H), 7.61 (d, J = 7.8 Hz, 1H), 7.58 (s, 3H), 6.92 (d, J = 8.5 Hz, 2H), 6.81 (d, J = 12.8 Hz, 1H), 6.23 – 6.15 (m, 2H), 6.10 (d, J = 3.2 Hz, 1H), 5.76 (d, J = 7.0 Hz, 2H), 5.67 (ddt, J = 16.5, 10.8, 6.0 Hz, 1H), 5.31 (s, 1H), 5.04 – 4.97 (m, 1H), 4.87 – 4.80 (m, 1H), 4.69 (s, 2H), 4.42 (s, 1H), 4.26 (s, 1H), 3.94 (s, 1H), 3.85 (s, 1H), 3.77 (d, J = 24.8 Hz, 2H), 3.56 (t, J = 5.8 Hz, 2H), 3.54 – 3.49 (m, 6H), 3.42 (t, J = 5.9 Hz, 2H), 3.22 (q, J = 5.8 Hz, 2H), 3.09 (d, J = 5.8 Hz, 4H), 2.79 (t, J = 7.6 Hz, 2H), 2.58 (t, J = 4.8 Hz, 4H), 2.44 – 2.36 (m, 4H), 1.47 (s, 6H), 0.86 (d, J = 7.4 Hz, 1H).

**<sup>13</sup>C NMR** (151 MHz, DMSO) δ 171.31, 168.04, 161.64, 156.46, 150.02, 139.28, 132.68, 128.76, 118.72, 115.93, 106.93, 100.44, 72.78, 70.12, 70.04, 69.58, 68.89, 57.72, 53.63, 49.14, 47.07, 46.88, 42.46, 39.07, 33.65, 30.92, 29.49, 23.89, 14.42.

**HRMS (ESI):** [M+H]<sup>+</sup> *m/z* calc. 892.45, found 892.4454.

### Synthesis of LEB-3-147

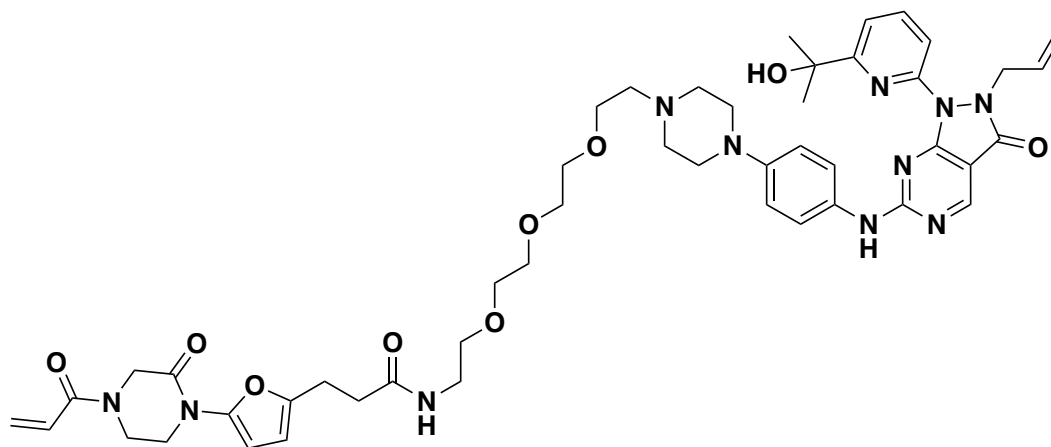


LEB-03-143

**2-allyl-6-((4-(4-(2-(2-(2-(2-aminoethoxy)ethoxy)ethoxy)ethyl)piperazin-1-yl)phenyl)amino)-1-(6-(2-hydroxypropan-2-yl)pyridin-2-yl)-1,2-dihydro-3H-pyrazolo[3,4-d]pyrimidin-3-one (LEB-03-143).** Intermediate 8 (40 mg, 0.0823 mmol) was dissolved in 0.5 mL of DMF. tert-butyl (2-(2-(2-(bromomethoxy)ethoxy)ethoxy)ethyl)carbamate (35 mg, 1.2 eq, 0.0987 mmol) and potassium carbonate (34 mg, 3.0 eq, 0.247 mmol) were added to the mixture, and the reaction was warmed to 50°C and stirred overnight. Water was added, the mixture extracted three times with EtOAc, combined organic extracts were washed with brine, and dried over sodium sulfate, and concentrated. Purification by flash column chromatography (EtOAc:Hexanes 50:50) yielded boc-protected intermediate. This was immediately dissolved in 3mL of DCM and the reaction mixture was cooled on ice. 1mL of trifluoroacetic acid was added dropwise and the solution was warmed to room temperature and stirred for 1 hour. The deprotected amine TFA salt was washed twice with DCM and dried under vacuum to yield **LEB-03-143** (22 mg, 0.0279 mmol, 34% yield) as a yellow oil.

<sup>1</sup>H NMR (300 MHz, Chloroform-d) δ 11.44 (s, 1H), 8.71 (s, 1H), 8.10 (s, 4H), 8.00 (t, J = 7.9 Hz, 1H), 7.65 (d, J = 8.0 Hz, 1H), 7.57 (t, J = 7.4 Hz, 3H), 6.91 (d, J = 8.6 Hz, 2H), 5.69 (ddt, J = 16.5, 10.1, 6.2 Hz, 1H), 5.11 (d, J = 10.1 Hz, 1H), 4.92 (d, J = 17.1 Hz, 1H), 4.78 (d, J = 6.3 Hz, 2H), 3.98 – 3.77 (m, 5H), 3.77 – 3.63 (m, 9H), 3.33 (d, J = 59.7 Hz, 8H), 1.64 (s, 6H), 1.29 (s, 1H).

LC/MS: [M+H]<sup>+</sup> m/z calc. 662.3, found 662.4



LEB-03-147

**3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)-N-(2-(2-(2-(2-(4-(4-((2-allyl-1-(6-(2-hydroxypropan-2-yl)pyridin-2-yl)-3-oxo-2,3-dihydro-1H-pyrazolo[3,4-d]pyrimidin-6-yl)amino)phenyl)piperazin-1-yl)ethoxy)ethoxy)ethoxy)ethyl)propanamide (LEB-03-147).** Intermediate 3 (19 mg, 0.0558 mmol) and **LEB-3-143** (22 mg, 0.0279 mmol) were coupled according to general procedure C. After hydrolysis, deprotected **3** and **LEB-3-142** were dissolved in DMF (0.5 mL), followed by DIPEA (49 mL, 0.279 mmol) and HATU (21 mg, 0.0558 mmol). The reaction was stirred for 30 minutes. Water was added and the mixture extracted three times with 4:1 CHCl<sub>3</sub>:IPA. Combined organic extracts were washed with brine, and dried over sodium sulfate, and concentrated. Purification by prep TLC (8% MeOH in DCM) yielded **LEB-03-147** as a light-yellow solid (10.0 mg, 0.0107 mmol, 19% yield).

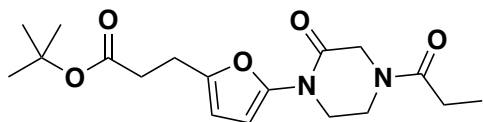


**<sup>1</sup>H NMR** (600 MHz, DMSO-*d*<sub>6</sub>) δ 10.14 (s, 1H), 8.83 (s, 1H), 8.05 (s, 1H), 7.97 (t, J = 5.6 Hz, 1H), 7.76 (d, J = 8.1 Hz, 1H), 7.61 (d, J = 7.4 Hz, 1H), 6.92 (d, J = 8.8 Hz, 2H), 6.87 – 6.75 (m, 1H), 6.21 (d, J = 3.2 Hz, 1H), 6.11 – 6.06 (m, 1H), 5.76 (s, 2H), 5.67 (ddt, J = 16.3, 10.2, 6.0 Hz, 1H), 5.32 (s, 1H), 5.00 (dq, J = 10.2, 1.4 Hz, 1H), 4.84 (dq, J = 17.1, 1.5 Hz, 1H), 4.69 (d, J = 6.0 Hz, 2H), 4.26 (s, 1H), 3.95 (s, 1H), 3.77 (d, J = 24.9 Hz, 2H), 3.59 – 3.48 (m, 9H), 3.41 (t, J = 5.9 Hz, 2H), 3.21 (q, J = 5.8 Hz, 2H), 3.09 (d, J = 5.3 Hz, 4H), 2.83 – 2.76 (m, 2H), 2.57 (t, J = 5.0 Hz, 4H), 2.51 (p, J = 1.9 Hz, 9H), 1.47 (s, 6H).

**<sup>13</sup>C NMR** (151 MHz, DMSO-*d*<sub>6</sub>) δ 171.30, 164.62, 156.47, 150.02, 147.70, 139.28, 132.68, 128.76, 118.72, 116.75, 115.93, 106.92, 100.44, 72.78, 70.26, 70.23, 70.17, 70.08, 69.59, 68.87, 57.72, 55.38, 53.62, 49.15, 47.54, 47.07, 46.87, 42.45, 39.05, 33.64, 30.92, 23.88.

**HRMS (ESI):** [M+H]<sup>+</sup> *m/z* calc. 936.47, found 936.4723.

### Synthesis of NJH-2-106



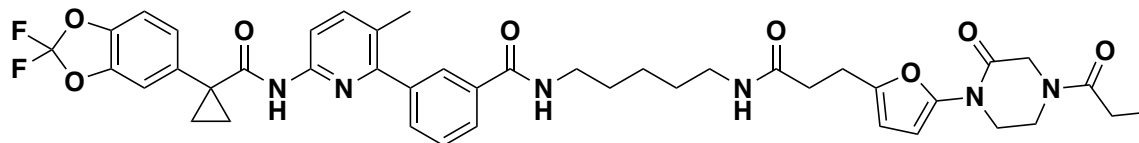
10

#### **tert-butyl 3-(5-(2-oxo-4-propionylpiperazin-1-yl)furan-2-yl)propanoate (10):**

Intermediate **2** (benzyl (E)-4-(5-(3-(tert-butoxy)-3-oxoprop-1-en-1-yl)furan-2-yl)-3-oxopiperazine-1-carboxylate) (85 mg, 0.20 mmol) was dissolved in EtOH (5 mL) and Pd/C (10 mg, 10% wt.) was added. The atmosphere was exchanged for hydrogen (balloon) and the mixture was stirred vigorously overnight. After 16h, the suspension was diluted with DCM and filtered through Celite to remove Pd/C, then concentrated. The crude residue was redissolved in DCM (2 mL), and TEA (83 mL, 0.60 mmol) was added. The solution was then cooled to 0 °C and propionyl chloride (25 mL, 0.31 mmol) was added and the mixture stirred for 30 min at 0 °C. Water was added and the mixture was extracted with DCM three times. Organic extracts were combined, washed with brine, dried over sodium sulfate, and concentrated. The crude residue was purified by silica gel chromatography to provide **Intermediate 10** (48 mg, 0.14 mmol, 69% yield over two steps) as a white solid.

**<sup>1</sup>H NMR** (600 MHz, CDCl<sub>3</sub>) δ 6.28 (d, J = 3.2 Hz, 1H), 6.04 (d, J = 3.2 Hz, 1H), 4.40 (s, 1H), 4.29 (s, 1H), 3.91 (dt, J = 30.8, 5.3 Hz, 2H), 3.85 – 3.78 (m, 2H), 2.88 (t, J = 7.6 Hz, 2H), 2.54 (t, J = 7.5 Hz, 2H), 2.43 – 2.34 (m, 2H), 1.44 (s, 9H), 1.19 (q, J = 6.9 Hz, 3H).

**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 351.2, found 351.2.



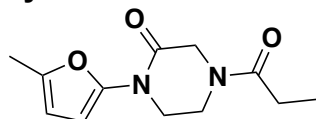
NJH-2-106

**3-(6-(1-(2,2-difluorobenzodioxol-5-yl)cyclopropane-1-carboxamido)-3-methylpyridin-2-yl)-N-(5-(3-(5-(2-oxo-4-propionylpiperazin-1-yl)furan-2-yl)propanamido)pentyl)benzamide (NJH-2-106):** Intermediate **10** (15 mg, 0.043 mmol) and Intermediate **5c** (15 mg, 0.029 mmol) were reacted according to General Procedure C to provide **NJH-2-106** (18 mg, 0.022 mmol, 76%) as a clear colorless oil. **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 8.10 (d, J = 8.4 Hz, 1H), 7.82 (s, 1H), 7.79 (d, J = 7.6 Hz, 1H), 7.72 (s, 1H), 7.59 (d, J = 8.5 Hz, 1H), 7.52 (d, J = 7.7 Hz, 1H), 7.46 (t, J = 7.6 Hz, 1H), 7.22 (dd, J = 8.1, 1.8 Hz, 1H), 7.19 (d, J = 1.7 Hz, 1H), 7.07 (d, J = 8.1 Hz, 1H), 6.51 – 6.39 (m, 1H), 6.22 – 6.15 (m, 1H), 6.02 – 5.97 (m, 1H), 5.88 – 5.76 (m, 1H), 4.31 (d, J = 50.6 Hz, 2H), 3.96 – 3.71 (m, 4H), 3.42 (q, J = 6.6 Hz, 2H), 3.21 (q, J = 6.5 Hz, 2H), 2.92 – 2.82 (m, 2H), 2.44 (t, J = 7.4 Hz, 2H), 2.41 – 2.29 (m, 2H), 2.24 (s, 3H), 1.74 (q, J = 3.9 Hz, 2H), 1.64 – 1.56 (m, 2H), 1.53 – 1.43 (m, 2H), 1.38 – 1.26 (m, 2H), 1.21 – 1.09 (m, 5H).

**<sup>13</sup>C NMR** (151 MHz, CDCl<sub>3</sub>) δ 206.9, 172.2, 171.8, 167.3, 155.4, 149.8, 148.9, 144.9, 144.1, 143.6, 141.0, 140.1, 134.9, 134.8, 131.7, 131.7, 128.4, 127.5, 127.0, 126.6, 113.0, 112.4, 110.2, 107.3, 100.9, 53.4, 49.3, 47.2, 42.4, 39.7, 39.1, 38.7, 34.9, 31.2, 29.0, 26.5, 24.2, 23.8, 19.1, 17.2, 9.0.

**HRMS (ESI):** [M+H]<sup>+</sup> *m/z* calc. 813.3345, found 813.3422.

#### Synthesis of NJH-2-080:



NJH-2-080

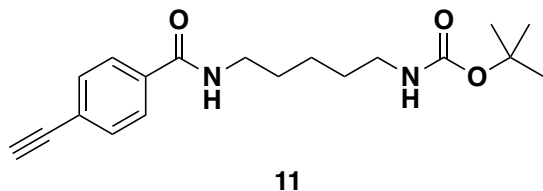
**1-(5-methylfuran-2-yl)-4-propionylpiperazin-2-one (NJH-2-080):** 1-(5-methylfuran-2-yl)piperazin-2-one (30 mg, 0.17 mmol) was dissolved in DCM (2 mL). The solution was cooled to 0 °C and TEA (69 mL, 0.50 mmol) and propionyl chloride (21 mL, 0.25 mmol) were added. After stirring at 0 °C for 30 min, water was added, and the reaction extracted three times with DCM. Organic extracts were combined, washed with brine, dried over sodium sulfate, and concentrated. The crude residue was purified by silica gel chromatography (0-100% EtOAc/Hex) to provide **NJH-2-080** (17.3 mg, 0.073 mmol, 43%) as a white solid.

**<sup>1</sup>H NMR** (600 MHz, CDCl<sub>3</sub>) δ 6.25 (d, J = 3.2 Hz, 1H), 6.00 (d, J = 2.2 Hz, 1H), 4.41 (s, 1H), 4.29 (s, 1H), 3.97 – 3.86 (m, 2H), 3.82 (t, J = 5.5 Hz, 2H), 2.45 – 2.34 (m, 2H), 2.27 (s, 3H), 1.23 – 1.16 (m, 3H). **<sup>13</sup>C NMR** (151 MHz, CDCl<sub>3</sub>) δ 172.2, 163.4, 147.5, 144.4, 107.2, 101.0, 49.3, 46.9, 38.8, 26.5, 13.4, 9.0.

**<sup>13</sup>C NMR** (151 MHz, CDCl<sub>3</sub>) δ 172.2, 163.4, 147.5, 144.4, 107.2, 101.0, 49.3, 46.9, 38.8, 26.5, 13.4, 9.0.

**HRMS (ESI):**  $[M+H]^+$   $m/z$  calc. 259.1160, found 259.1053.

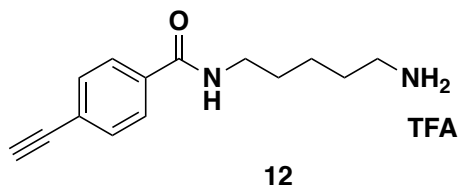
### **SYNTHESIS OF NJH-2-075**



**tert-butyl (5-(4-ethynylbenzamido)pentyl)carbamate (11):** 4-ethynylbenzoic acid (27 mg, 0.19 mmol), N-Boc-1,5-diaminopentane (47 mg, 0.23 mmol), HOBt (26 mg, 0.19 mmol), and DIEA (165  $\mu$ L, 0.95 mmol) were dissolved in DCM (1.5 mL) and EDCI $\cdot$ HCl (73 mg, 0.38 mmol) was added. After stirring the mixture for 16h at rt, water was added, the mixture partitioned, and the aqueous phase extracted with DCM. Combined organic extracts were washed with brine and dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated, and the crude residue was purified by silica gel chromatography (0-50% EtOAc/Hex) to obtain the Boc-protected amine **11** (27 mg, 0.082 mmol, 43%) as a white solid.

**<sup>1</sup>H NMR** (300 MHz, CDCl<sub>3</sub>)  $\delta$  7.78 (d, J = 8.3 Hz, 2H), 7.59 (d, J = 8.7 Hz, 2H), 6.32 (s, 1H), 4.63 (s, 1H), 3.50 (td, J = 7.0, 5.7 Hz, 2H), 3.23 (s, 1H), 3.18 (q, J = 6.5 Hz, 2H), 1.70 (d, J = 7.5 Hz, 2H), 1.62 – 1.52 (m, 2H), 1.46 (s, 11H).

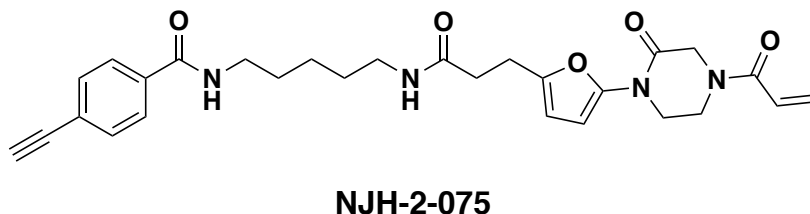
**LC/MS**  $[M+H]^+$   $m/z$  calc. 331.19, found 331.1.



**N-(5-aminopentyl)-4-ethynylbenzamide (12):** tert-butyl (5-(4-ethynylbenzamido)pentyl)carbamate **11** (27 mg, 0.082 mmol) was dissolved in DCM (1 mL) and TFA (0.5 mL) was added. After stirring at rt for 2h, the mixture was diluted in DCM and evaporated repeatedly to remove volatiles and provide the amine as a TFA salt and an orange oil (32 mg, 0.096 mmol, 117%), which was used without further purification.

**<sup>1</sup>H NMR** (400 MHz, DMSO)  $\delta$  8.55 (t, J = 5.7 Hz, 1H), 7.84 (d, J = 8.2 Hz, 2H), 7.63 (s, 2H), 7.57 (d, J = 8.1 Hz, 2H), 4.39 (s, 1H), 3.26 (q, J = 6.6 Hz, 2H), 2.83 – 2.74 (m, 2H), 1.62 – 1.48 (m, 4H), 1.40 – 1.32 (m, 2H).

**LC/MS**  $[M+H]^+$   $m/z$  calc. 231.14, found 231.1.



***N*-(5-(3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)propanamido)pentyl)-4-ethynylbenzamide (NJH-2-075):** Intermediate 3, tert-butyl 3-(5-(4-acryloyl-2-oxopiperazin-1-yl)furan-2-yl)propanoate, (20 mg, 0.057 mmol) was dissolved in DCM (0.5 mL) and treated with TFA (0.25 mL). The mixture was stirred at rt for 45 minutes until the starting material was consumed, followed by dilution with DCM and evaporation to remove volatiles. The carboxylic acid was then dissolved in DMF, and intermediate **12** *N*-(5-aminopentyl)-4-ethynylbenzamide TFA (22 mg, 0.062 mmol), DIEA (50  $\mu$ L, 0.29 mmol), and HATU (43 mg, 0.11 mmol) were added. After stirring the mixture at rt for 1 h, water was added. The resulting suspension was extracted three times with DCM. Combined organic extracts were washed brine and dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated, and the crude residue was purified by silica gel chromatography (0-4% MeOH/DCM) to obtain **NJH-2-075** (7.6 mg, 0.016 mmol, 27%) as a pale yellow oil.

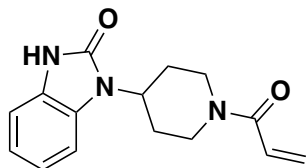
**<sup>1</sup>H NMR** (300 MHz, CDCl<sub>3</sub>)  $\delta$  7.82 (d, *J* = 8.3 Hz, 2H), 7.58 (d, *J* = 8.3 Hz, 2H), 6.77 – 6.50 (m, 2H), 6.43 (dd, *J* = 16.7, 2.1 Hz, 1H), 6.24 (d, *J* = 3.2 Hz, 1H), 6.06 (d, *J* = 3.3 Hz, 1H), 5.93 (s, 1H), 5.86 (dd, *J* = 10.1, 2.1 Hz, 1H), 4.44 (d, *J* = 17.4 Hz, 2H), 4.01 (s, 2H), 3.91 – 3.84 (m, 2H), 3.46 (q, *J* = 6.6 Hz, 2H), 3.32 – 3.19 (m, 3H), 2.93 (t, *J* = 7.2 Hz, 2H), 2.50 (t, *J* = 7.3 Hz, 2H), 1.72 – 1.61 (m, 2H), 1.60 – 1.46 (m, 2H), 1.44 – 1.35 (m, 2H).

**<sup>13</sup>C NMR** (151 MHz, DMSO)  $\delta$  171.0, 165.7, 164.6, 150.1, 135.2, 132.1, 128.8, 127.9, 124.7, 106.9, 100.5, 83.4, 83.1, 38.9, 33.8, 29.3, 29.2, 24.3, 24.0.

**HRMS** [M+H]<sup>+</sup> *m/z* calc. 380.1586, found 380.1581.

## Synthesis of EN523 Analogs:

### Synthesis of EZ-1-032



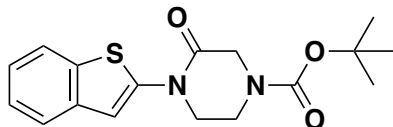
**1-(1-acryloylpiperidin-4-yl)-1,3-dihydro-2H-benzo[d]imidazol-2-one (EZ-1-032):** 1-(piperidin-4-yl)-1,3-dihydro-2H-benzo[d]imidazol-2-one (50 mg, 0.23 mmol) was acylated via general procedure H and the crude residue was purified by silica gel chromatography (0 to 20% MeOH/DCM) to afford the title compound as a clear yellow oil (11.8 mg, 0.043 mmol, 19%).

**<sup>1</sup>H NMR** (400 MHz, DMSO)  $\delta$  10.87 (s, 1H), 7.29 – 7.17 (m, 1H), 7.05 – 6.95 (m, 3H), 6.88 (ddd, *J* = 16.1, 10.5, 3.3 Hz, 1H), 6.16 (d, *J* = 2.4 Hz, 1H), 5.70 (dd, *J* = 10.4, 2.4 Hz, 1H), 4.61 (d, *J* = 13.1 Hz, 1H), 4.44 (tt, *J* = 12.0, 3.9 Hz, 1H), 4.21 (d, *J* = 13.8 Hz, 1H), 3.21 (t, *J* = 13.3 Hz, 1H), 2.76 (t, *J* = 12.9 Hz, 1H), 2.34 – 2.07 (m, 2H), 1.75 (d, *J* = 12.4 Hz, 2H).

**<sup>13</sup>C NMR** (151 MHz, DMSO)  $\delta$  164.8, 154.2, 129.7, 129.0, 129.0, 127.7, 121.1, 120.9, 109.3, 109.0, 50.3, 45.1, 41.6, 29.9, 29.0.

**HRMS (ESI):** [M+H]<sup>+</sup> *m/z* calc. 272.14, found 272.1394.

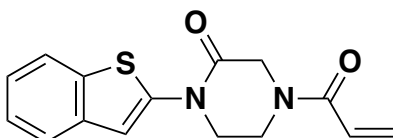
### Synthesis of EZ-1-038



**tert-butyl 4-(benzo[*b*]thiophen-2-yl)-3-oxopiperazine-1-carboxylate (EZ-1-035):** 2-bromobenzo[*b*]thiophene (100 mg, 0.47 mmol) was coupled to *tert*-butyl 3-oxopiperazine-1-carboxylate (93.5 mg, 0.47 mmol) via general procedure D and the crude residue was purified by silica gel chromatography (0 to 100% EtOAc/hexane) to yield a pale yellow solid (22.3 mg, 0.116 mmol, 14%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.81 (d, *J* = 7.8 Hz, 1H), 7.72 (d, *J* = 7.7 Hz, 1H), 7.30 (s, 2H), 6.92 (s, 1H), 4.40 (s, 2H), 4.01 (t, *J* = 5.4 Hz, 2H), 3.92 (t, *J* = 5.4 Hz, 2H), 1.54 (s, 9H).

**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 333.1, found 333.1



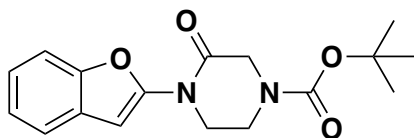
**4-acryloyl-1-(benzo[*b*]thiophen-2-yl)piperazin-2-one (EZ-1-038):** *tert*-butyl 4-(benzo[*b*]thiophen-2-yl)-3-oxopiperazine-1-carboxylate (EZ-1-035) (18 mg, 0.05 mmol) was deprotected and acylated via general procedures F and H respectively. The crude residue was purified by silica gel chromatography (0 to 100% EtOAc/Hex) to afford the title compound as a pale yellow solid (6.6 mg, 0.023 mmol, 46%).

**<sup>1</sup>H NMR** (400 MHz, DMSO) δ 7.86 (d, *J* = 7.9 Hz, 1H), 7.74 (t, *J* = 7.2 Hz, 1H), 7.45 – 7.32 (m, 1H), 7.28 (q, *J* = 6.8 Hz, 1H), 7.11 (s, 1H), 6.98 – 6.77 (m, 1H), 6.21 (d, *J* = 16.7 Hz, 1H), 5.83 – 5.74 (m, 1H), 4.50 (d, *J* = 68.5 Hz, 2H), 4.18 – 3.91 (m, 4H).

**<sup>13</sup>C NMR** (151 MHz, DMSO) δ 164.7, 142.0, 136.7, 136.2, 128.9, 128.0, 124.9, 123.9, 122.8, 122.1, 108.0, 49.2, 48.4, 47.6, 46.8.

**HRMS (ESI):** [M+Na]<sup>+</sup> *m/z* calc. 309.0674, found 309.0667.

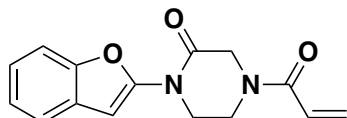
### Synthesis of EZ-1-045



**tert-butyl 4-(benzofuran-2-yl)-3-oxopiperazine-1-carboxylate (EZ-1-044):** 2-bromobenzofuran (200 mg, 1.02 mmol) was coupled with *tert*-butyl 3-oxopiperazine-1-carboxylate (204.24 mg, 1.02 mmol) via general procedure D and purified by silica gel chromatography (0 to 50% EtOAc/hexane) to yield a yellow solid (44.3 mg, 0.14 mmol, 14%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.65 – 7.52 (m, 1H), 7.48 – 7.39 (m, 1H), 7.26 (dd, J = 6.0, 3.3 Hz, 2H), 6.96 (d, J = 1.2 Hz, 1H), 4.35 (s, 2H), 4.19 – 4.05 (m, 2H), 3.86 (d, J = 5.6 Hz, 2H), 1.53 (d, J = 1.6 Hz, 9H).

**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 316.1, found 316.2



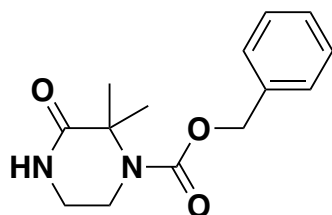
**4-acryloyl-1-(benzofuran-2-yl)piperazin-2-one (EZ-1-045):** *tert*-butyl 4-(benzofuran-2-yl)-3-oxopiperazine-1-carboxylate (EZ-1-044) (44.3 mg, 0.14 mmol) was deprotected and acylated via general procedures F and H and purified by silica gel chromatography (0 to 50% EtOAc/hexane) to afford the title compound as a pale yellow solid (9.3 mg, 0.034 mmol, 25%).

**<sup>1</sup>H NMR** (300 MHz, CDCl<sub>3</sub>) δ 7.63 – 7.51 (m, 1H), 7.43 (dt, J = 7.1, 3.8 Hz, 1H), 7.29 (td, J = 6.3, 2.8 Hz, 2H), 6.98 (d, J = 1.0 Hz, 1H), 6.57 (d, J = 9.8 Hz, 1H), 6.47 (dd, J = 16.7, 2.2 Hz, 1H), 5.88 (dd, J = 10.1, 2.2 Hz, 1H), 4.52 (s, 2H), 4.24 – 3.92 (m, 4H).

**<sup>13</sup>C NMR** (151 MHz, DMSO) δ 165.0, 150.1, 149.5, 129.0, 128.8, 128.1, 123.9, 123.9, 121.2, 111.1, 94.6, 49.5, 47.1, 46.6, 42.4.

**HRMS (ESI):** [M+H]<sup>+</sup> *m/z* calc. 271.1004, found 271.1078.

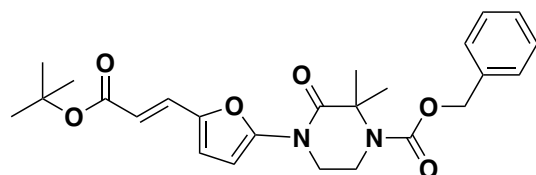
### Synthesis of EZ-1-055



**benzyl 2,2-dimethyl-3-oxopiperazine-1-carboxylate (EZ-1-050):** 3,3-dimethylpiperazin-2-one (400 mg, 3.12 mmol) was protected with benzyl chloroformate via general procedure E and purified by silica gel chromatography (0 to 10% MeOH/DCM) to yield a fluffy white powder (492.1 mg, 1.88 mmol, 60%).

**<sup>1</sup>H NMR** (300 MHz, CDCl<sub>3</sub>) δ 7.41 (s, 5H), 6.02 (s, 1H), 5.19 (s, 2H), 3.87 – 3.74 (m, 2H), 3.49 – 3.35 (m, 2H), 1.75 (s, 6H).

**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 263.1, found 263.1.

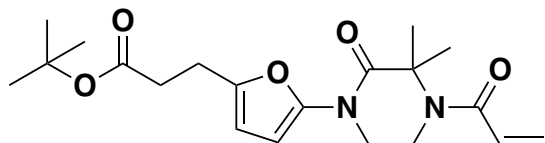


**benzyl (E)-4-(5-(3-(*tert*-butoxy)-3-oxoprop-1-en-1-yl)furan-2-yl)-2,2-dimethyl-3-oxopiperazine-1-carboxylate (EZ-1-052):** *tert*-butyl (E)-3-(5-bromofuran-2-yl)acrylate (Intermediate 2) (104 mg, 0.38 mmol) and benzyl 2,2-dimethyl-3-oxopiperazine-1-carboxylate (EZ-1-050) (100 mg, 0.38 mmol) were coupled via general procedure D and

purified by silica gel chromatography (0 to 50% EtOAc/hexane) to yield a clear yellow oil that solidified upon standing (133.7 mg, 0.29 mmol, 77%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.43 (d, J = 5.1 Hz, 6H), 6.66 (q, J = 3.6 Hz, 2H), 6.12 (d, J = 15.6 Hz, 1H), 5.22 (s, 2H), 4.04 – 3.98 (m, 2H), 3.91 (d, J = 5.2 Hz, 2H), 1.80 (s, 6H), 1.56 (d, J = 4.0 Hz, 9H).

**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 455.2, found 455.2



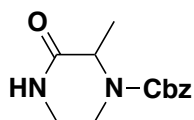
**tert-butyl 3-(5-(4-acryloyl-3,3-dimethyl-2-oxopiperazin-1-yl)furan-2-yl)propanoate (EZ-1-055):** benzyl (*E*)-4-(5-(3-(*tert*-butoxy)-3-oxoprop-1-en-1-yl)furan-2-yl)-2,2-dimethyl-3-oxopiperazine-1-carboxylate (EZ-1-052) (30 mg, 0.066 mmol) was deprotected and acylated via general procedures G and H and purified by silica gel chromatography (0-70% EtOAc/hexane) to afford the title compound as a clear colorless oil (7.2 mg, 0.019 mmol, 29% over two steps).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 6.51 (ddd, J = 16.8, 10.6, 2.3 Hz, 1H), 6.29 (t, J = 2.9 Hz, 1H), 6.23 (dt, J = 16.8, 2.1 Hz, 1H), 6.03 (d, J = 3.2 Hz, 1H), 5.70 (dt, J = 10.5, 2.1 Hz, 1H), 3.88 (dd, J = 6.4, 3.4 Hz, 2H), 3.78 (dd, J = 6.1, 3.6 Hz, 2H), 2.87 (t, J = 7.6 Hz, 2H), 2.54 (td, J = 7.9, 2.3 Hz, 2H), 1.83 (d, J = 2.3 Hz, 6H), 1.44 (d, J = 2.3 Hz, 9H).

**<sup>13</sup>C NMR** (151 MHz, DMSO) δ 171.6, 171.1, 166.3, 149.1, 146.2, 131.5, 127.2, 107.2, 99.7, 80.4, 63.6, 47.5, 42.7, 28.2, 23.8, 23.5.

**HRMS (ESI):** [M+Na]<sup>+</sup> *m/z* calc. 399.1896, found 399.1883.

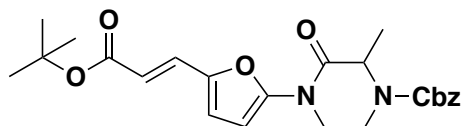
### Synthesis of EZ-1-058



**Benzyl 2-methyl-3-oxopiperazine-1-carboxylate (EZ-1-049):** 3-methylpiperazin-2-one (400 mg, 3.5 mmol) was protected with benzyl chloroformate via general procedure E and purified by silica gel chromatography (0 to 10% MeOH/DCM) to yield a pale yellow solid (123.9 mg, 0.5 mmol, 14%).

**<sup>1</sup>H NMR** (300 MHz, CDCl<sub>3</sub>) δ 7.36 (s, 5H), 5.96 (s, 1H), 5.16 (s, 2H), 4.69 (s, 1H), 4.18 (s, 1H), 3.47 (d, J = 12.1 Hz, 1H), 3.27 (d, J = 12.2 Hz, 2H), 1.46 (d, J = 7.1 Hz, 3H).

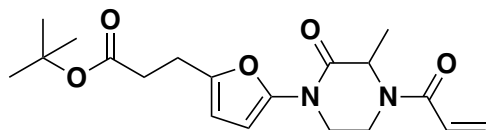
**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 249.1, found 249.1.



**benzyl (E)-4-(5-(3-(tert-butoxy)-3-oxoprop-1-en-1-yl)furan-2-yl)-2-methyl-3-oxopiperazine-1-carboxylate (EZ-1-051):** Benzyl 2-methyl-3-oxopiperazine-1-carboxylate (EZ-1-049) (60 mg, 0.24 mmol) and *tert*-butyl (E)-3-(5-bromofuran-2-yl)acrylate (EZ-1-048) (66 mg, 0.24 mmol) were coupled via general procedure D and purified by silica gel chromatography (0 to 50% EtOAc/hexane) to yield a yellow solid (69.3 mg, 0.16 mmol, 66%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.42 (s, 5H), 7.32 – 7.24 (m, 1H), 6.70 – 6.62 (m, 2H), 6.12 (d, J = 15.4 Hz, 1H), 5.23 (d, J = 2.5 Hz, 2H), 4.89 (s, 1H), 4.35 (s, 1H), 4.00 (d, J = 13.9 Hz, 2H), 3.50 (s, 1H), 1.72 – 1.49 (m, 12H).

**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 441.2, found 441.2.



**tert-butyl 3-(5-(4-acryloyl-3-methyl-2-oxopiperazin-1-yl)furan-2-yl)propanoate (EZ-1-058):**

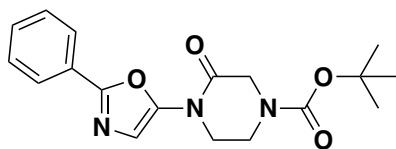
benzyl (E)-4-(5-(3-(tert-butoxy)-3-oxoprop-1-en-1-yl)furan-2-yl)-2-methyl-3-oxopiperazine-1-carboxylate (EZ-1-051) (52.3 mg, 0.12 mmol) was deprotected and acylated via general procedures G and H and purified by silica gel chromatography (0-100% EtOAc/hexane) to yield the title compound as a clear colorless oil (17.9 mg, 0.05 mmol, 42% over two steps).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 6.66 – 6.51 (m, 1H), 6.46 (d, J = 16.7 Hz, 1H), 6.32 (d, J = 3.2 Hz, 1H), 6.07 (dd, J = 3.2, 1.0 Hz, 1H), 5.84 (d, J = 9.8 Hz, 1H), 4.74 (s, 1H), 4.23 – 3.23 (m, 4H), 2.91 (t, J = 7.6 Hz, 2H), 2.57 (dd, J = 8.2, 6.9 Hz, 3H), 1.63 (s, 3H), 1.47 (s, 9H).

**<sup>13</sup>C NMR** (151 MHz, DMSO) δ 171.6, 167.6, 164.2, 149.5, 145.9, 128.7, 128.2, 107.2, 100.7, 80.4, 60.2, 54.5, 52.0, 48.2, 33.4, 28.2, 23.5, 17.0.

**HRMS (ESI):** [M+Na]<sup>+</sup> *m/z* calc. 385.1739, found 385.1728.

### **Synthesis of EZ-1-067**

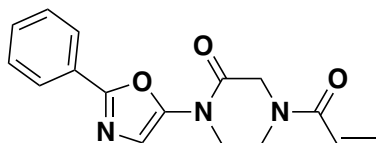


**tert-butyl 3-oxo-4-(2-phenyloxazol-5-yl)piperazine-1-carboxylate (EZ-1-066):** 5-bromo-2-phenyloxazole (50 mg, 0.22 mmol) was coupled with *tert*-butyl 3-oxopiperazine-1-carboxylate (44.7 mg, 0.22 mmol) via general procedure D and purified by silica gel chromatography (0 to 60% EtOAc/hexane) to yield a white solid (40.4 mg, 0.117 mmol, 54%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 8.05 – 7.98 (m, 2H), 7.49 (dd, J = 5.7, 1.8 Hz, 3H), 7.38 (s, 1H), 4.36 (s, 2H), 4.04 (t, J = 5.4 Hz, 2H), 3.89 (t, J = 5.3 Hz, 2H), 1.55 (s, 9H).

**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 344.2, found 344.1.





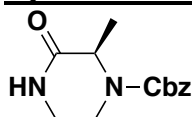
**4-acryloyl-1-(2-phenyloxazol-5-yl)piperazin-2-one (EZ-1-067):** *tert*-butyl 3-oxo-4-(2-phenyloxazol-5-yl)piperazine-1-carboxylate (EZ-1-066) (40.4 mg, 0.117 mmol) was deprotected and acylated via general procedures F and H and purified by silica gel chromatography (0 to 80% EtOAc/hexane) to afford the title compound as a yellow solid (34.6 mg, 0.116 mmol, 45% over two steps)

**<sup>1</sup>H NMR** (300 MHz, CDCl<sub>3</sub>) δ 8.01 (dd, *J* = 6.8, 3.0 Hz, 2H), 7.54 – 7.46 (m, 3H), 7.39 (s, 1H), 6.59 (s, 1H), 6.54 – 6.42 (m, 1H), 5.90 (d, *J* = 11.6 Hz, 1H), 4.53 (s, 2H), 4.10 (s, 4H).

**<sup>13</sup>C NMR** (151 MHz, DMSO) δ 164.6, 155.1, 146.6, 130.8, 129.6, 128.9, 128.3, 128.1, 127.1, 125.9, 116.2, 49.4, 47.2, 46.9.

**HRMS (ESI):** [M+H]<sup>+</sup> *m/z* calc. 298.1113, found 298.1187.

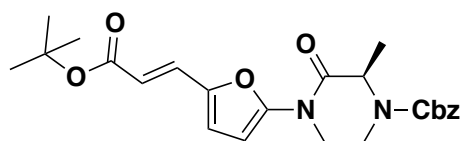
### Synthesis of EZ-1-068



**phenyl (*R*)-2-methyl-3-oxopiperazine-1-carboxylate (EZ-1-062):** (*R*)-3-methylpiperazin-2-one (100 mg, 0.88 mmol) was protected with benzyl chloroformate (186 mL, 0.876 mmol) via general procedure E and purified by silica gel chromatography (0 to 100% EtOAc/hexane) to yield a white solid (47.2 mg, 0.25 mmol, 22%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 6.15 (s, 1H), 5.21 (s, 2H), 4.73 (s, 1H), 4.24 (s, 1H), 3.51 (d, *J* = 12.5 Hz, 1H), 3.31 (d, *J* = 12.6 Hz, 2H), 1.50 (d, *J* = 7.0 Hz, 3H).

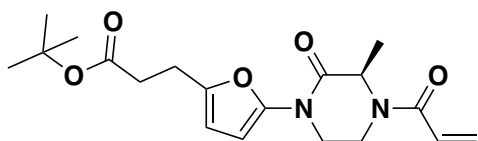
**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 248.1, found 248.1.



**benzyl (*R,E*)-4-(5-(3-(*tert*-butoxy)-3-oxoprop-1-en-1-yl)furan-2-yl)-2-methyl-3-oxopiperazine-1-carboxylate (EZ-1-064):** phenyl (*R*)-2-methyl-3-oxopiperazine-1-carboxylate (EZ-1-062) (44.6 mg, 0.18 mmol) was coupled to *tert*-butyl (*E*)-3-(5-bromofuran-2-yl)acrylate (EZ-1-048) (49.1 mg, 0.18 mmol) via general procedure D and purified by silica gel chromatography (0 to 35% EtOAc/hexane) to yield a clear yellow oil (56.7 mg, 0.13 mmol, 72%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.42 (d, *J* = 5.3 Hz, 5H), 7.30 (s, 1H), 6.74 – 6.62 (m, 2H), 6.12 (d, *J* = 15.6 Hz, 1H), 5.23 (d, *J* = 2.3 Hz, 2H), 4.89 (s, 1H), 4.34 (s, 1H), 4.02 (s, 2H), 3.49 (s, 1H), 1.61 (s, 3H), 1.56 (d, *J* = 5.5 Hz, 9H).

**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 441.2, found 441.2.



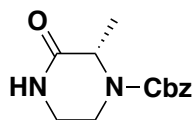
**tert-butyl (R)-3-(5-(4-acryloyl-3-methyl-2-oxopiperazin-1-yl)furan-2-yl)propanoate (EZ-1-068):** benzyl (R,E)-4-(5-(3-(tert-butoxy)-3-oxoprop-1-en-1-yl)furan-2-yl)-2-methyl-3-oxopiperazine-1-carboxylate (EZ-1-064) (31.2 mg, 0.07 mmol) was deprotected and acylated via general procedures F and H and purified by silica gel chromatography (0 to 100% EtOAc/hexane) to afford the title compound as a yellow solid (18.9 mg, 0.052 mmol, 68% over two steps).

**<sup>1</sup>H NMR** (300 MHz, CDCl<sub>3</sub>) δ 6.65 – 6.40 (m, 2H), 6.33 (d, J = 3.4 Hz, 1H), 6.08 (d, J = 3.3 Hz, 1H), 5.90 – 5.81 (m, 1H), 4.76 (s, 1H), 3.93 – 3.34 (m, 4H), 2.92 (t, J = 7.5 Hz, 2H), 2.58 (dd, J = 8.3, 6.8 Hz, 2H), 2.22 (s, 3H), 1.48 (s, 9H).

**<sup>13</sup>C NMR** (151 MHz, DMSO) δ 171.6, 167.6, 164.2, 149.4, 145.9, 128.7, 128.2, 107.2, 100.6, 80.4, 52.0, 48.2, 47.2, 33.4, 28.2, 23.5.

**HRMS (ESI):** [M+Na]<sup>+</sup> *m/z* calc. 385.1739, found 385.1730.

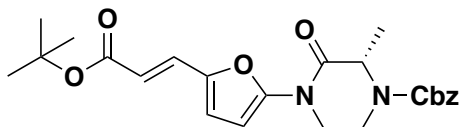
### Synthesis of EZ-1-071



**benzyl (S)-2-methyl-3-oxopiperazine-1-carboxylate (EZ-1-063):** (S)-3-methylpiperazin-2-one (100mg, 0.88mmol) was protected with benzyl chloroformate (149.4 mg, 0.88 mmol) via general procedure E and purified by silica gel chromatography (0 to 100% EtOAc/hexane) to yield a white solid (89.4 mg, 0.36 mmol, 41%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.40 (d, J = 4.6 Hz, 5H), 6.13 (s, 1H), 5.21 (s, 2H), 4.72 (s, 1H), 4.24 (s, 1H), 3.53 (s, 1H), 3.31 (d, J = 12.5 Hz, 2H).

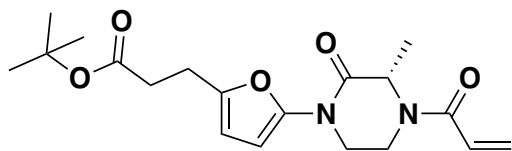
**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 248.1, found 248.1.



**benzyl (S,E)-4-(5-(3-(tert-butoxy)-3-oxoprop-1-en-1-yl)furan-2-yl)-2-methyl-3-oxopiperazine-1-carboxylate (EZ-1-069):** benzyl (S)-2-methyl-3-oxopiperazine-1-carboxylate (EZ-1-063) (41.6 mg, 0.17 mmol) was coupled to *tert*-butyl (E)-3-(5-bromofuran-2-yl)acrylate (EZ-1-048) (46.8 mg, 0.17 mmol) via general procedure D and purified by silica gel chromatography (0 to 50% EtOAc/hexane) to yield a clear yellow oil (41.3 mg, 0.09 mmol, 56%).

**<sup>1</sup>H NMR** (300 MHz, CDCl<sub>3</sub>) δ 7.45 – 7.37 (m, 5H), 7.31 (d, J = 1.3 Hz, 1H), 6.72 – 6.61 (m, 2H), 6.12 (d, J = 15.6 Hz, 1H), 5.23 (d, J = 1.3 Hz, 2H), 4.90 (s, 1H), 4.34 (s, 1H), 4.05 – 3.92 (m, 2H), 3.49 (s, 1H), 1.62 (s, 3H), 1.56 (d, J = 3.1 Hz, 9H).

**LC/MS:**  $[M+H]^+$   $m/z$  calc. 441.2, found 441.2.



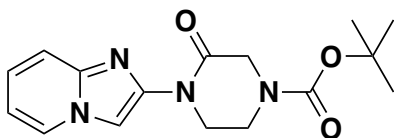
**tert-butyl (S)-3-(5-(4-acryloyl-3-methyl-2-oxopiperazin-1-yl)furan-2-yl)propanoate (EZ-1-071):** benzyl (S,E)-4-(5-(3-(tert-butoxy)-3-oxoprop-1-en-1-yl)furan-2-yl)-2-methyl-3-oxopiperazine-1-carboxylate (EZ-1-069) (35.4 mg, 0.08 mmol) was deprotected and acylated via general procedures F and H and purified by silica gel chromatography (0 to 100% EtOAc/hexane) to afford the title compound as a clear colorless oil (16.9 mg, 0.047 mmol, 58% over two steps).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$  6.63 – 6.41 (m, 2H), 6.32 (d,  $J$  = 3.4 Hz, 1H), 6.07 (d,  $J$  = 3.5 Hz, 1H), 5.85 (d,  $J$  = 10.3 Hz, 1H), 4.77 (s, 2H), 3.88 (s, 2H), 3.34 (s, 1H), 2.91 (t,  $J$  = 7.5 Hz, 2H), 2.58 (dt,  $J$  = 8.8, 5.2 Hz, 2H), 1.74 (s, 3H), 1.48 (d,  $J$  = 4.0 Hz, 9H).

**<sup>13</sup>C NMR** (151 MHz, DMSO)  $\delta$  171.6, 164.2, 149.5, 145.9, 128.7, 128.1, 107.2, 100.7, 80.4, 54.4, 52.0, 48.2, 33.4, 28.2, 23.5.

**HRMS (ESI):**  $[M+Na]^+$   $m/z$  calc. 385.1739, found 385.1726.

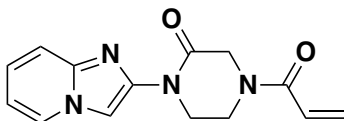
### **Synthesis of EZ-1-073**



**tert-butyl 4-(imidazo[1,2-a]pyridin-2-yl)-3-oxopiperazine-1-carboxylate (EZ-1-072):** 2-bromoimidazo[1,2-a]pyridine (50 mg, 0.25 mmol) was coupled to *tert*-butyl 3-oxopiperazine-1-carboxylate (50.8 mg, 0.25 mmol) via general procedure D and purified by silica gel chromatography (0 to 80% EtOAc/hexane) to yield a clear colorless oil (35.7 mg, 0.11 mmol, 45%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$  8.33 (s, 1H), 8.15 (d,  $J$  = 6.7 Hz, 1H), 7.54 (d,  $J$  = 9.1 Hz, 1H), 7.22 (ddd,  $J$  = 8.7, 6.9, 1.4 Hz, 1H), 6.85 (td,  $J$  = 6.8, 1.3 Hz, 1H), 4.34 (s, 2H), 4.31 (t,  $J$  = 5.5 Hz, 2H), 3.83 (t,  $J$  = 5.4 Hz, 2H), 1.53 (s, 9H).

**LC/MS:**  $[M+H]^+$   $m/z$  calc. 317.2, found 317.2.



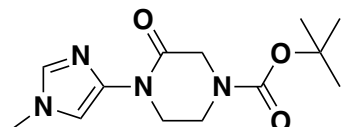
**4-acryloyl-1-(imidazo[1,2-a]pyridin-2-yl)piperazin-2-one (EZ-1-073):** *tert*-butyl 4-(imidazo[1,2-a]pyridin-2-yl)-3-oxopiperazine-1-carboxylate (EZ-1-072) (23.4 mg, 0.074 mmol) was deprotected and acylated via general procedures F and H and purified by

silica gel chromatography (0 to 100% EtOAc/hexane) to afford the title compound as an off white solid (3.8 mg, 0.014 mmol, 19% over two steps).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 8.32 (s, 1H), 8.15 (d, J = 6.9 Hz, 1H), 7.55 (d, J = 9.0 Hz, 1H), 7.24 (t, J = 7.9 Hz, 1H), 6.87 (t, J = 6.8 Hz, 1H), 6.62 (s, 1H), 6.46 (d, J = 16.7 Hz, 1H), 5.86 (d, J = 10.5 Hz, 1H), 4.53 (d, J = 23.7 Hz, 2H), 4.38 (s, 2H), 4.04 (d, J = 33.0 Hz, 2H).

**HRMS (ESI):** [M+H]<sup>+</sup> *m/z* calc. 271.1117, found 271.1190.

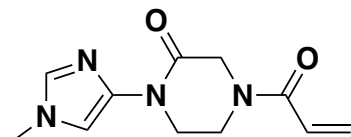
### **Synthesis of EZ-1-085**



***tert*-butyl 4-(1-methyl-1*H*-imidazol-4-yl)-3-oxopiperazine-1-carboxylate (EZ-1-084):** 4-bromo-1-methyl-1*H*-imidazole (155 mL, 1.55 mmol) was coupled to *tert*-butyl 3-oxopiperazine-1-carboxylate (311 mg, 1.55 mmol) via general procedure D and the crude residue was purified by silica gel chromatography (0-100% EtOAc/Hex) to yield a pale yellow solid (412 mg, 1.47 mmol, 95%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.58 – 7.50 (m, 1H), 7.39 – 7.26 (m, 1H), 4.27 (d, J = 9.5 Hz, 2H), 4.18 – 4.06 (m, 3H), 3.80 – 3.61 (m, 4H), 1.51 (d, J = 4.1 Hz, 9H).

**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 281.2, found 281.2.



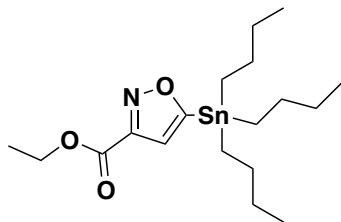
**4-acryloyl-1-(1-methyl-1*H*-imidazol-4-yl)piperazin-2-one (EZ-1-085):** *tert*-butyl 4-(1-methyl-1*H*-imidazol-4-yl)-3-oxopiperazine-1-carboxylate (EZ-1-084) (100 mg, 0.36 mmol) was deprotected and acylated via general procedures F and H and the crude residue was purified by silica gel chromatography (0 to 10% MeOH/DCM) to afford the title compound as a white solid (27.7 mg, 0.12 mmol, 33%).

**<sup>1</sup>H NMR** (300 MHz, CDCl<sub>3</sub>) δ 7.54 (s, 1H), 7.27 (s, 1H), 6.58 (s, 1H), 6.44 (dd, J = 16.7, 2.0 Hz, 1H), 5.88 – 5.81 (m, 1H), 4.46 (d, J = 16.0 Hz, 2H), 4.18 (s, 2H), 3.99 (d, J = 23.0 Hz, 2H), 3.73 (s, 3H).

**<sup>13</sup>C NMR** (151 MHz, DMSO) δ 163.4, 162.9, 138.9, 133.9, 128.6, 128.5, 128.2, 46.9, 44.9, 42.6, 33.7.

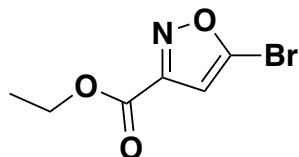
**HRMS (ESI):** [M+H]<sup>+</sup> *m/z* calc. 235.1117, found 235.1190.

### **Synthesis of EZ-1-099**



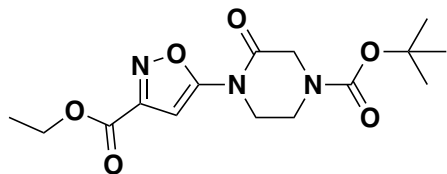
**ethyl 5-(tributylstannyl)isoxazole-3-carboxylate (EZ-1-090):** To a solution of ethyl-2-chloro-2-(hydroxyiminoacetate) (481 mg, 3.17 mmol) dissolved in anhydrous DCM (15 mL), potassium carbonate (482.5mg, 3.5mmol) and tributyl(ethynyl)stannane (872 mL, 3.17 mmol) were added and stirred at room temperature overnight. The reaction was then quenched with water, extracted with DCM and dried over anhydrous sodium sulfate. The organic layer was purified via silica gel column chromatography (0 to 10% EtOAc/hexane) to give the product as a clear colorless watery oil (753 mg, 1.75 mmol, 55%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 6.84 (s, 1H), 4.48 (q, J = 7.1 Hz, 2H), 1.70 – 1.10 (m, 27H), 0.94 (s, 3H).



**Ethyl 5-bromoisoxazole-3-carboxylate (EZ-1-091):** Br<sub>2</sub> (134 mL, 2.62 mmol) was added to a solution of ethyl 5-(tributylstannyl)isoxazole-3-carboxylate (EZ-1-090) (753 mg, 1.74 mmol) and sodium carbonate (203 mg, 1.91 mmol) dissolved in DCM (10 mL), and stirred at room temperature overnight. The reaction mixture was then quenched with saturated sodium thiosulfate (8 mL) before extracting with DCM and washing with brine. The organic layer was dried over anhydrous sodium sulfate and purified via silica gel column chromatography (0 to 15% EtOAc/hexane) to produce a clear colorless oil (241.8 mg, 1.1 mmol, 63%) that crystallized upon standing.

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 6.76 (s, 1H), 4.49 (q, J = 7.1 Hz, 2H), 1.47 (dt, J = 9.6, 6.9 Hz, 3H).



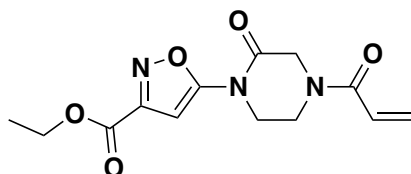
**Ethyl 5-(4-(tert-butoxycarbonyl)-2-oxopiperazin-1-yl)isoxazole-3-carboxylate (EZ-1-097):**

Anhydrous dioxane (3 mL) was added to a vial flushed with N<sub>2</sub> containing ethyl 5-bromoisoxazole-3-carboxylate (EZ-1-091) (94.6 mg, 0.43 mmol), *tert*-butyl 3-oxopiperazine-1-carboxylate (0.43mmol, 86.1mg), cesium carbonate (280.2 mg, 0.86 mmol), Xantphos (19 mg, 0.032 mmol), Pd(dba)<sub>3</sub> (10 mg, 0.011 mmol) and the suspension was degassed. The reaction mixture was stirred at 90 °C overnight. The

product was extracted with EtOAc, washed with brine, and purified via silica gel column chromatography (0 to 75% EtOAc/hexane) to afford a clear yellow oil (14 mg, 0.04 mmol, 9.6%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 4.48 (q, J = 7.1 Hz, 2H), 4.38 (s, 2H), 4.13 (q, J = 5.5 Hz, 2H), 3.91 – 3.84 (m, 2H), 1.54 (d, J = 2.8 Hz, 9H), 1.46 (t, J = 7.1 Hz, 3H).

**LC/MS:** [M+H]<sup>+</sup> *m/z* calc. 340.1, found 340.2



**Ethyl 5-(4-acryloyl-2-oxopiperazin-1-yl)isoxazole-3-carboxylate (EZ-1-099):**

Ethyl 5-(4-(tert-butoxycarbonyl)-2-oxopiperazin-1-yl)isoxazole-3-carboxylate (EZ-1-097) (14 mg, 0.04 mmol) was deprotected and acylated via general procedures F and H respectively and the crude residue was purified by silica gel chromatography (0 to 100% EtOAc/hexane) to afford the title compound as a clear colorless oil (5.0 mg, 0.017 mmol, 42%).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.01 (s, 1H), 6.57 (s, 1H), 6.47 (dd, J = 16.8, 2.0 Hz, 1H), 5.90 (dd, J = 10.1, 2.0 Hz, 1H), 4.56 (s, 2H), 4.48 (q, J = 7.1 Hz, 2H), 4.18 (d, J = 5.3 Hz, 2H), 4.07 (s, 2H), 1.46 (t, J = 7.1 Hz, 3H).

**<sup>13</sup>C NMR** (151 MHz, DMSO) δ 173.4, 144.1, 143.7, 135.1, 121.6, 119.5, 118.2, 117.4, 64.5, 44.7, 27.3, 16.5, 9.9.

**HRMS (ESI):** [M+Na]<sup>+</sup> *m/z* calc. 316.0909, found 316.0907.



