Chemical probe development for the KLHL12 E3 ligase



Evmorfia Vasiliki Dalietou

Prof. Alex N. Bullock, Prof. Paul E. Brennan

Nuffield Department of Medicine & St Anne's College

University of Oxford

Trinity Term 2025

Thesis submitted for the degree of Doctor of Philosophy in Interdisciplinary Bioscience

ABSTRACT

E3 ligases in the ubiquitin-proteasome system regulate protein homeostasis through the selective degradation of substrate proteins. However, their diverse functions and assembly mechanisms remain incompletely understood. I aimed to develop structures, assays, and chemical probes to better understand and modulate two atypical E3 assemblies, $CUL3^{KLHL12}$ and $CUL2/CUL4^{NRBP1}$, implicated in developmental disorders and cancer. A crystal structure of the KLHL12 Kelch domain bound to antagonist PLEKHA4 revealed a conserved binding mode for substrates, co-factors, and inhibitors supporting spatial sequestration as a control mechanism. Substrate displacement assays enabled identification of small molecule inhibitors, including fragment **EVD-1** that selectively inhibited KLHL12 at sub-micromolar concentrations over other family members. Covalent binding to Cys511 was observed by crystallography and key determinants of compound reactivity established by structure-activity relationships. A tractable synthetic route for future analogues was identified as well as exit vectors for elaboration and potential PROTAC development. Proteins, binding assays and preliminary cryo-EM studies were also advanced for characterisation of putative CUL2/ $CUL4^{NRBP1}$ complexes and a molecular glue-induced KLHL12-USP21 complex. Together, these studies expand our understanding of E3 ligase assemblies. **EVD-1** further establishes a powerful tool to investigate KLHL12 biology, or for functionalisation into KLHL12-based PROTACs. In summary, this work expands the repertoire of chemically tractable E3 ligases, positioning KLHL12 as a promising target for small-molecule inhibitors or in the context of targeted protein degradation strategies.

DECLARATION

I declare that there are no parts of this thesis or its research herein have been reproduced or accepted for another award or degree or diploma at any other university or learning institution. This thesis contains no other person's work except where stated in text.

LAB MOVES AND PANDEMIC IMPACT

My DPhil studies started in Michaelmas 2020, at the height of the COVID-19 pandemic. I began my research under lockdown and completed a three-month rotation at the Rosalind Franklin Institute in the laboratories of Prof. Jim Naismith and Dr. Ray Owens, focusing on the engineering of a nanobody with improved affinity for the SARS-CoV-2 spike protein. This work culminated in a publication, where I contributed as a joint first author (Mikolajek et al., 2022)[1]. Following this rotation, I joined Prof. Alex Bullock and Dr. Paul Brennan to begin work on the contents of this thesis. The initial three months of this period were impacted by pandemic-related restrictions, including shift patterns and reduced working hours. Once the effects of the pandemic began to subside, our department Centre for Medicines Discovery moved from the Old Road Campus Research Building (ORCRB) to Biochemistry Phase II. After one year, we relocated again, this time from Biochemistry to the NDM Research Building (NDMRB) at the Old Road Campus in Headington. Both of these relocations caused significant disruptions to my DPhil work. Given the lab-based nature of my research, progress was severely hindered when labs were either completely inaccessible or non-functional. The packing, unpacking, and recommissioning of equipment resulted in an estimated overall disruption of six months. Despite these initial setbacks, later in my studies, during the summer of 2023, I had the opportunity to complete an internship at AstraZeneca, Cambridge, where I received training in cryo-electron microscopy (cryo-EM) for studying E3 ligase complexes, work that is not included in this thesis. This experience greatly broadened my scientific perspective and deepened my appreciation for the real-world applications of my research.

[1] Mikolajek H, Weckener M, Brotzakis ZF, Huo J, Dalietou EV, Le Bas A, Sormanni P, Harrison PJ, Ward PN, Truong S, Moynie L, Clare DK, Dumoux M, Dormon J, Norman C, Hussain N, Vogirala V, Owens RJ, Vendruscolo M, Naismith JH. Correlation between the binding affinity and the conformational entropy of nanobody SARS-CoV-2 spike protein complexes. Proc Natl Acad Sci U S A. 2022 Aug 2;119(31):e2205412119.

ACKNOWLEDGMENTS

I would first like to express my gratitude to all the colleagues I have had the privilege of working with over the past four years. First and foremost, I would like to thank Dr. Zhuoyao Chen, whose incredible support has been instrumental in the work presented in this thesis. She not only shared her expertise and generously offered her time to guide me through the early stages of my studies but also provided unwavering encouragement, serving as both a mentor and a friend. I would also like to extend my sincere thanks to Dr. Xiangrong Chen and Dr. William Richardson, who shared their vast knowledge of small-molecule discovery and crystallography. Their hard work, team spirit, and expertise have been a constant source of inspiration. Special thanks are also due to Ms. Charlotte Manning, Dr. Jong Fu, and Dr. Ellie Williams, who guided me through my first steps in the lab during the turbulent pandemic period. Their mentorship instilled in me fundamental principles that continue to serve me well. I am also deeply grateful to all past and present members of the GFSU lab, Dr. Melissa Sweeney, Dr. Roxanne Lau, Beatriz Montes Tan, Dr. Julien Cros, Dr. Jesús Baltanás Copado, as well as my fellow DPhil students, Dr. Ashley Ramdass, William Seaton-Burn, and Dr. David Cruz Walma, for their support throughout this journey.

I have been fortunate to be part of not just one but two laboratories, bridging both biology and chemistry. I would like to express my deepest appreciation to Dr. Tryfon Zarganis-Tzitzikas, who played a pivotal role in coordinating the small-molecule work described in this thesis. I am also sincerely grateful to Dr. Laura Ortega Varga for her unwavering patience and the countless hours she spent meticulously working with me on small-molecule docking. Of course, none of the chemistry in this work would have been possible without the laboratory expertise of Dr. Lukas Scheibelberger and

Mr. Thomas Grimes, whose expertise is matched only by their generosity and who - fortunately - kept any cyanide-related incidents at bay. I am also thankful to Dr. Gonzalo Vera Namuncura, Ethan Chidlow, and Mr. Renan Vinicius de Araújo for their invaluable advice and experimental support. My heartfelt thanks go out to all members of the Brennan lab; Dr. Matthew Holland, Ms. Leonard Lee, Mr. Matthew Fry, Ms. Alice Page, Dr. Mia Callens, Dr. Katherine England, Ms. Amber Truepenny, Dr. Sheenagh Aiken, Dr. Lisa Dollhopf, Ms. Hadeel Abozenadah, and Ms. Amelia Henley, who have cultivated a welcoming and supportive environment, one that I am truly grateful to have been part of.

I consider myself fortunate to have been part of the Centre for Medicines Discovery (CMD), where knowledge, techniques, and resources are shared generously among colleagues. Special thanks to Karly Buchan for her guidance, her experimental support, and her expert insights into small-molecule screening and biophysical assays, which played a pivotal role in this thesis. Her deep understanding of both the challenges and opportunities inherent in this field proved invaluable. I am also grateful to Dr. Rod Chalk and Kavya Clement for their support with mass spectrometry experiments and for the engaging discussions. I am particularly grateful to Dr. Michael Fairhead, whose encyclopedic knowledge and generosity know no bounds. I would also like to thank the entire Biotech Team, and especially Ms. Leela Shrestha, Ms. Dong Wang, and Dr. Souvika Bakshi, for their expertise and time in supporting my cloning and protein expression experiments. Finally, I would like to deeply thank Dr. Ashley Pike, Dr. Gamma Chi, and Dr. William Bradshaw for their enduring patience in attempting to impart their expertise in all things structural. Only when I attain even a fraction of their understanding, I will consider myself a structural biologist.

Although this work is not presented in this thesis, it would be remiss not to thank Prof. Jim Naismith, Prof. Ray Owens, and the colleagues in their groups - especially Dr. Audrey Le Bas, Dr. Jiangdong Huo, Dr. Lucile Moynie, and Dr. Philip Ward - for being the first lab I worked in at the beginning of my studies. They welcomed a new student with open arms, and I remain deeply grateful for their kindness. I will also never forget the invaluable training in cryo-EM and, most importantly, the

professional and personal support I received from Dr. Fiona Shilliday, Dr. Taiana Maia de Oliveira, Dr. Shintaro Aibara, and Dr. Christopher Phillips at AstraZeneca Cambridge. My time in Cambridge was one of the highlights of my studies, thanks to the entire Mechanistic and Structural Biology team, including Dr. Kasim Sader, Dr. Juntai Liu, Dr. Catarina Rodrigues, Dr. Marianne Schimpl, Dr. Jon Read, Dr. Oliver Acton, and Dr. Judit Debreczeni.

Finally, I cannot find the words to express how fortunate I was to have Prof. Alex Bullock and Prof. Paul Brennan as my supervisors. I sincerely thank them both for their insights, enthusiasm, and continuous hands-on guidance and support throughout this project.

Beyond the science, I will always remember Alex's encouragement in all my endeavors - scientific and beyond - from exploring AI to teaching and even taking on high-risk chemistry projects. I will never forget our group's trip to Bletchley Park in the cold English winter, our visit to the boathouse in the cold English summer, the experience of seeing Cambridge colleges, and - almost miraculously - seeing Scotland under a brilliant sun. More than just a thesis supervisor, Alex helped me to think critically about the future, and shape my ideas about what is possible. He is not only a brilliant scientist but also an exceptional mentor.

Paul pushed me to embrace everything Oxford had to offer, reminded me how little I truly knew about chemistry, and, most importantly, believed in me from the very beginning. His constant support has been invaluable in shaping not only this thesis but also my growth as a scientist. Paul has the uncanny ability to identify every little thorny point in any experiment or project, face it head-on, and always encourage, never discourage. He is a brilliant person with a remarkable talent for leading and uplifting everyone around him and a great inspiration as a scientist and person.

Most importantly, I gratefully acknowledge the Biotechnology and Biological Sciences Research Council and the Oxford DTP for funding my studies over the past four years, as well as the Bodosaki Foundation for additional financial support. First and foremost, I warmly thank my parents for their unwavering encouragement and for being a constant source of inspiration.

LIST OF ABBREVIATIONS

ABPP ..... Activity-Based Protein Profiling

ALG2 ..... Alpha-1,3/1,6-mannosyltransferase

ALP ..... Autophagy-Lysosome Pathway

ALS ..... Amyotrophic Lateral Sclerosis

AP-1 ..... Activator Protein 1

APP ..... Amyloid Precursor Protein

ASU ..... Asymmetric Unit

AT7 ..... Acyl Transferase 7

ATE1 ..... Arginyl-tRNA-Protein Transferase

ATP ..... Adenosine Triphosphate

AUC ..... Analytical Ultracentrifugation

AUC ..... Area Under the Curve

AUX ..... Auxin

Aβ ..... Amyloid-β

B ..... Beads

BACK ..... BTB and C-terminal Kelch

BB ..... Binding Buffer

BCR BTB-Cul3-Rbx1

BEVS ..... Baculovirus Expression Vector System

BLI ..... Biolayer Interferometry

BRD4 ..... Bromodomain-Containing Protein 4

BSA ..... Bovin Serum Albumin

BTB ..... Broad-Complex, Tramtrack and Bric à Brac

BTK ..... Bruton's Tyrosine Kinase

BV ..... Baculovirus

CCT ..... Conserved Carboxy-Terminal

CHAPS ..... 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate

CMV ..... Cytomegalovirus

co-IP ..... co-Immonoprecipitation

COPII ..... Coat Protein Complex II

CRBN ..... Cereblon

CRISPR ..... Clustered Regularly Interspaced Short Palindromic Repeats

CRL ..... Cullin Ring Ubiquitin Ligase

cryo-EM ..... Cryogenic Electron Microscopy

CTD ..... Carboxy-Terminal Domain

CUL ..... Cullin

CV ..... Column Volume

DCM ..... Dichloromethane

DDB1 ..... DNA Damage-Binding Protein 1

DEL ..... DNA-Encoded Library

DMA ..... Dimethylacetamide

DMF ..... Dimethyl Fumarate

DMSO ..... Dimethyl Sulfoxide

DNA ..... DeoxyriboNucleic Acid

DTT ..... Dithiothreitol

DUB ..... Deubiquitinase

DVL ..... Segment Polarity Protein Dishevelled Homolog

EB ..... Elution Buffer

ER ..... Endoplasmic-Reticulum

ESI-TOF ..... Electrospray Ionization Time of Flight

ESI ..... Electrospray Ionization

EWG ..... Electron Withdrawing Group

FBD ..... Familial British Dementia

FDD ..... Familial Danish Dementia

FT ..... Flow-Through

GAP ..... GTPase-Activating Protein

GCI ..... Grating-Coupled Interferometry

GF ..... Gel Filtration

GFB ..... Gel Filtration Buffer

GPCR ..... G Protein-Coupled Receptor

GSH ..... Glutathione

GST ..... Glutathione S-Transferase

GTP ..... Guanosine-5'-Triphosphate

HECT ..... Homologous to E6-AP C Terminus

HEPES ..... 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HIF1α ..... Hypoxia-Inducible Factor-1α

HPLC ..... High Pressure Liquid Chromatography

hr ..... Hour

HRMS ..... High Resolution Mass Spectrometry

HRP ..... Horseradish Peroxidase

HSP70 ..... Heat Shock 70 kDa Protein

HTRF ..... Homogeneous Time Resolved Fluorescence

IAC ..... ImmunoAffinity Chromatography

IAM ..... Iodoacetamide

IC_{50} ..... Half-Maximal Inhibitory Concentration

IDR ..... Intrinsically Disordered Region

IEX ..... Ion-Exchange Chromatography

IL-1 ..... Interleukin-1

IMAC ..... Immobilized Metal Affinity Chromatography

IMiDs ..... Immunomodulatory Drugs

IPTG ..... Isopropyl β-d-1-thiogalactopyranoside

ITC ..... Isothermal Titration Calorimetry

ITM2B ..... Integral Membrane Protein 2B

K_a ..... Association constant

K_d ..... Dissociation Constant

KEAP1 ..... Kelch-like ECH-Associated Protein 1

KLHDC2 ..... Kelch Domain-Containing Protein 2

KLHL12 ..... Kelch Like Family Member 12

LB ..... Lysogeny Broth

LC-MS/MS ..... Liquid Chromatography with Tandem Mass Spectrometry

LC ..... Liquid Chromatography

LIC ..... Ligation-Independent Cloning

LiSH ..... Lissencephaly Homology Domain

MES ..... 2-(N-morpholino)ethanesulfonic acid

mM ..... millimolar

MQ ..... Milli-Q

MS ..... Mass Spectrometry

MST ..... Microscale Thermophoresis

MWCO ..... Molecular Weight Cut-Off

Ni-NTA ..... Nickel-Nitrilotriacetic Acid

NMP N-Methyl-2-Pyrrolidone

NMR ..... Nuclear Magnetic Resonance

NRBP1 ..... Nuclear Receptor Binding Protein 1

NTD ..... N-Terminal Domain

OD ..... Optical Density

PBS ..... Phosphate-Buffered Saline

PCR ..... Polymerase Chain Reaction

PDB ..... Protein Data Bank

PEF1 ..... Penta-EF-Hand Domain Containing 1

PEG ..... Polyethylene Glycol

PEI ..... Polyethylenimine

pLDDT ..... Predicted Local Distance Difference Test

PLEKHA4 ..... Pleckstrin Homology Domain-Containing Family A

POI ..... Protein of Interest

PROTAC ..... Proteolysis-Targeting Chimera

PVDF ..... Polyvinylidene Difluoride

Q-TOF ..... Quadrupole Time of Flight

RBR ..... RING-Between-RING

RBX1 ..... RING Box Protein 1

RCR ..... RING Cys Relay

RING ..... Really Interesting New Gene

RMSD ..... Root Mean Square Deviation

RNA ..... Ribodeoxyribonucleic Acid

RNAi ..... RNA Interference

RPM ..... Revolutions per Minute

RU ..... Response Unit

SALL4 ..... Spalt-Like Transcription Factor 4

SCF ..... SKP, Cullin, F-Box Containing Complex

scFvs ..... Single-Chain Variable Fragment

SD ..... Standard Deviation

SDS-PAGE ..... Sodium Dodecyl Sulphate PolyAcrylamide Gel Electrophoresis

SEC-MALS ..... Size-Exclusion Chromatography - Multi Angle Light Scattering

SEC ..... Size-Exclusion Chromatography

SFM ..... Serum-Free Media

SKP1 ..... Substrate-Recognition Component of the SCF

SN ..... Supernatant

SOCS ..... Suppressor of Cytokines Signalling

SPAK ..... SPS1-Related Proline/Alanine-Rich Kinase

SPR ..... Surface Plasmon Resonance

TAE ..... Tris/Acetate/EDTA

TB ..... Terrific Broth

TCEP ..... Tris(2-CarboxyEthyl)Phosphine

TEM ..... Transmission Electron Microscope

TEV ..... Tobacco Etch Virus Nuclear-Inclusion-a Endopeptidase

THF ..... Tetrahydrofuran

TLC ..... Thin Layer Chromatography

Tm ..... Melting Temperature

TNF ..... Tumor Necrosis Factor

TR-TUBE ..... Trypsin-Resistant Tandem Ubiquitin-Binding Entity

TSA ..... Thermal Shift Assay

TSC22DF ..... TSC22 Domain Family Member

Ub ..... Ubiquitin

UBR ..... Ubiquitin-Protein Ligase E3 Component n-Recognin

UCHs ..... Ub C-Terminal Hydrolases

UPS ..... Ubiquitin-Proteasome System

USP ..... Ub-Specific Proteases

UV ..... Ultraviolet

VHL ..... von Hippel-Lindau

VLS ..... Virtual Ligand Screening

WB ..... Wash Buffer

WCL ..... Whole Cell Lysate

WNK ..... With no K (lysine) Protein Kinase

Wnt ..... Wingless and Int-1

YT ..... Yeast Extract Tryptone Medium

µM ..... micromolar

TABLE OF CONTENTS

1 Introduction **1**

1.1 The Ubiquitin-Proteasome System: A Key Regulator of Protein Homeostasis 1

1.1.1 Protein Homeostasis and the Role of the Ubiquitin-Proteasome System (UPS) 2

1.1.2 E3 Ubiquitin Ligases: Specificity in ubiquitylation 5

1.1.3 Exploring Non-Canonical E3 Ligase Assemblies 11

1.2 KLHL12: A substrate adaptor of a homotypic Cullin-RING E3 Ligase 15

1.2.1 KLHL12 in embryonic and neural development 15

1.2.2 KLHL12 domains mediate its role as a substrate-specific adaptor for CRL3 . 18

1.2.3 KLHL12 recognises a "PGXPP" motif 19

1.3 NRBP1: A substrate adaptor of a heterotypic Cullin-RING E3 ligase 22

1.3.1 NRBP1 is an atypical pseudokinase that associates with transcriptional machinery components 22

1.3.2 NRBP1 as substrate adaptor in CRL E3 ligases 22

1.3.3 NRBP1 in pathophysiology 24

1.4 Chemical Probes and Small-Molecule Strategies for Modulating E3 Ligases 28

1.4.1 Chemical Probes are selective protein modulators 28

1.4.2 Chemical Probes in Understanding E3 Ligase Function 29

1.4.3 Targeted Protein Degradation: from chemical probes to proximity inducers . 30

1.5 Motivation & Objectives of this Thesis 35

1.5.1 Motivation 35

1.5.2 Objectives 36

2 Materials & Methods **40**

2.1 Production of recombinant human proteins 40

2.1.1 Ligation Independent Cloning. 40

2.1.2 Screening and protein production in insect cells 42

2.1.3 Large-scale protein expression and purification 43

2.1.4 Protein Purification 45

2.2 Co-purification assays 47

2.3 *In silico* methods 49

2.4 Structure determination of protein macromolecules 51

2.4.1 X-ray crystallography 51

2.4.2 Cryo-electron microscopy 52

2.5 Methods to probe molecular interactions 54

2.5.1 Surface Plasmon Resonance 54

2.5.2 Thermal Shift Assay 55

2.5.3 Grating-coupled interferometry 56

2.5.4 Bio-layer Interferometry 57

2.5.5 Microscale Thermophoresis 58

2.5.6 AlphaScreen 58

2.6 Mass Spectrometry (MS) 61

2.6.1 Electrospray Ionization Time of Flight (ESI-TOF) 61

2.6.2 In-gel digestion MS/MS analysis 62

2.6.3 High pressure liquid chromatography (HPLC) MS 63

2.6.4 RapidFire MS 63

2.7 Chemical Synthesis 63

2.7.1 General Reagent Information 63

2.7.2 Chromatography 64
2.7.3 Compound Characterisation 64

3 Structural insights into KLHL12 interactor binding 66

3.1 Introduction 66
3.2 Results 67
3.2.1 Structure determination of KLHL12-interactor complexes 67
3.2.2 Structural observations on binding mode of KLHL12 interactors 70
3.3 Discussion 80

4 Targeting KLHL12: Identification of Small Molecule Binders 87

4.1 Introduction 87
4.2 Investigating DNA-encoding library hits from ZebiAI 88
4.2.1 *In silico* filtering for initial compound selection 88
4.2.2 Screening of selected DEL hits 91
4.2.3 Extending screening of DEL hits 99
4.2.4 Orthogonal validation of AlphaScreen hits 100
4.3 Identification of covalent fragment hits for KLHL12 104
4.3.1 Cysteine tractability assessment 104
4.3.2 Validation of covalent hits following AlphaScreen selection 105
4.4 Discussion 110

5 Targeting KLHL12: Chemical Probe Development 118

5.1 Introduction 118
5.2 Results 119
5.2.1 Structure determination of KLHL12-covalent hit complex 119
5.2.2 Covalent hit **EVD-1** selectively targets C511 on KLHL12 binding site 123
5.2.3 *In silico* screening for compound selection 126

5.2.4 Covalent adduct formation assessment of potential new scaffolds 129
5.2.5 Covalent adduct formation assessment of hit SAR analogues 129
5.2.5.1 Effect of γ-nitrogen and ethyl ester modifications 131
5.2.5.2 Effect of modifications on position of nitrile warhead and ester group 136
5.2.5.3 Scaffold inversion affords more reactive compounds 138
5.2.6 Structure determination of KLHL12 in complex with an alternative covalent compound scaffold 141
5.2.7 Structural observations on the binding mode of an alternative fragment scaffold 146
5.3 Selectivity and reactivity characterisation 152
5.3.1 Selectivity over a panel of Kelch domain E3 ligases 152
5.3.2 Glutathione reactivity testing 152
5.4 Synthesis of hit elaborations 156
5.4.0.1 Synthesis of key dichloro intermediate 158
5.4.0.2 Synthesis of fragment elaborations 159
5.4.1 Synthesis of hit analogues for PROTAC functionalisation 161
5.5 Discussion 166

6 Exploring molecular glue mechanism in KLHL12-USP21 assembly 174

6.1 Introduction 174
6.2 $CUL3^{KLHL12}$ E3 ligase associates with USP21 in the presence of BAY-805 177
6.2.1 Expression and purification trials of dimeric KLHL12 177
6.2.2 Indications of BAY-805 mediated USP21-HiBIT-LgBIT KLHL12 Kelch domain assembly *in vitro* 188
6.2.3 Structural characterisation attempts for the KLHL12-USP21 complex mediated by BAY-805 189
6.3 Discussion 195

7 Exploring structural assemblies of the NRBP1 E3 ligase **203**

7.1 Introduction 203

7.2 *In vitro* exploration of putative heterodimeric CUL2/CUL4A^{NRBP1} ligase 204

7.2.1 ElonginBC co-expression rescues NRBP1 *E.coli* constructs 204

7.2.2 NRBP1 associates with ElonginBC stoichiometrically *in vitro* 215

7.2.3 Structural characterisation attempts for the NRBP1-ElonginBC complex . . . 218

7.2.4 Validating NRBP1-ElonginBC complex interactions 224

7.2.4.1 ElonginBC is not sufficient for CUL2 recruitment 224

7.2.4.2 Indications of DDB1 recruitment in the absence of ElonginBC . . 224

7.2.4.3 TSC22D4 transcription factor binds to NRBP1 225

7.2.4.4 Modeling studies of complex assembly 233

7.3 Discussion 239

8 Conclusions and Future Directions **243**

9 Experimental Procedures **255**

9.0.1 Experimental Procedures 255

Appendix A Cloning &Construct Design **i**

LIST OF FIGURES

1.1 Schematic representation of (de)ubiquitylation mechanism 6

1.2 Ubiquitinylation mechanisms of non-lysine residues 7

1.3 Substrate recognition by E3 ligases 9

1.4 Architecture of the modular Cullin-RING E3 Ligases (CRLs) 10

1.5 Productive multimerisation of E3 ligases enhances substrate targeting 13

1.6 Non-productive multimerisation of E3 ligases prevent substrate recruitment 14

1.7 Models of $CUL3^{KLHL12}$ E3 ligase assemblies 17

1.8 Overview of $CRL3^{KLHL12}$ architecture 20

1.9 Two different binding modes for dishevelled substrates binding to KLHL12 Kelch domain have been described in the literature 21

1.10 Proposed model for CRL2/$CRL4A^{NRBP1}$ Ligase 26

1.11 Structural basis of Targeted Protein Degradation: Molecular Glue vs. PROTAC . . . 32

1.12 Molecular Glues and PROTACs recruit neosubstrates to E3 ligases with different mechanisms 33

1.13 Overview of Thesis Chapters 37

3.1 Crystal packing contacts of apo KLHL12 68

3.2 Optimisation of KLHL12-PLEKHA4 co-crystallization using microseeding 71

3.3 Asymmetric Units of KLHL12-Interacting Peptide Co-Crystals 72

3.4 1.95 Å structure of KLHL12 Kelch domain bound to PLEKHA4 peptide 73

3.5 1.88 Å structure of KLHL12 Kelch domain bound to PEF1 peptide 74

3.6 Overview of the main interactions that tether PEF1 and PLEKHA4 peptides in the KLHL12 Kelch domain binding site 75
3.7 Overview of hydrophobic interactions of PEF1 peptide nested in the KLHL12 Kelch domain binding site 76
3.8 Overview of hydrophobic interactions of PLEKHA4 peptide nested in the KLHL12 Kelch domain binding site 78
3.9 Comparison of KLHL12 Kelch domain between the bound and apo states 81
3.10 Binding mode of KLHL12 interacting partners 82
3.11 Many PGXPP-containing proteins in the human proteome may be potential KLHL12 interactors. 86

4.1 Average B-factor analysis of KLHL12-DVL1 complex 89
4.2 Representative docked poses of DEL hits 90
4.3 Selection of a subset of DEL hits representing chemical diversity of the collection . . 92
4.4 Overview of preliminary assay steps 93
4.5 AlphaScreen and TSA hits from initial DEL selection 95
4.5 AlphaScreen and TSA hits from initial DEL selection 96
4.6 Orthogonal validation of DEL hits; representative SPR results 97
4.7 Screening of extended DEL library 101
4.8 Orthogonal validation of DEL hits; representative GCI results 103
4.9 Assessment of KLHL12 cysteine reactivity 106
4.10 Chemical diversity of covalent warheads in the screened collection 107
4.11 Heatmap of covalent hits targeting KLHL12, KLHL20, and KLHL6 108
4.12 Covalent hit selectively targets one cysteine of KLHL12 109
4.13 Overview of compounds that competitively inhibit cognate peptide binding on Kelch binding pockets 112
4.14 Possible mechanism of metal chelation by false-positive AlphaScreen hits 113

5.1 Co-crystallisation trials of KLHL12 and covalent hit 120
5.2 Representative crystal hits of EVD-1 in complex with KLHL12 121
5.3 Asymmetric unit of the EVD-1 and KLHL12 co-crystal complex 124
5.4 2.6 Å structure of covalent hit **EVD-1** bound to KLHL12 Kelch domain 125
5.5 2.6 Å structure of covalent hit **EVD-1** bound to KLHL12 peptide binding site. 126
5.6 Overview of *in silico* screening workflow in search for hit analogues 128
5.7 Validated hits from *in silico* screening of LUMC compound libraries 130
5.8 Covalent adduct formation assessment of hit SAR analogues 132
5.9 Effect of γ-nitrogen and ethyl ester modifications 134
5.10 Further Structure-Activity Relationship (SAR) 135
5.11 Effect of the position of the nitrile covalent warhead within **EVD-1** pyrazolopyrimidine heterocyclic core scaffold 137
5.12 Effect of the relative positions between nitrile and ethyl ester on varying heterocyclic cores 139
5.13 Structure-Activity Relationships of nitriles without a proximal heteroatom 140
5.14 Structure Activity Relationships of nitriles on 5-membered rings 142
5.15 IC_{50} curves of 5-membered ring nitriles 143
5.16 Co-crystallisation trials of KLHL12 and fragments containing 5-membered ring nitriles 144
5.17 Representative crystal hits of fragment 23-KLHL12 complex 145
5.18 Asymmetric unit of the fragment 23 and KLHL12 co-crystal complex 147
5.19 2.6 Å structure of fragment 23 bound to the KLHL12 substrate binding site 149
5.20 KLHL12-compound interaction diagram 150
5.21 Selectivity assay for covalent EVD-1 and fragment 23 153
5.22 Glutathione reactivity assay of selected compounds 155
5.23 **EVD-1**-KLHL12 Kelch domain co-crystal reveals accessible elaboration vector . . . 157
5.24 Amide analogue EVD-15 shows greater labelling efficiency than ester hit EVD-1 . . 164

5.25 Potential reaction mechanism of covalent bond formation of EVD-1 168
5.26 Compound binding induces Y512 shift in otherwise rigid KLHL12 binding site . . . 169
5.27 KLHL12 interactor recognition mechanism is evolutionary conserved 170
5.28 Structure activity relationship overview for EVD-1 172

6.1 USP21 interacts with engineered HiBIT-LgBIT KLHL12 dimer even in the absence of BTB-BACK 176
6.2 Full length KLHL12 expression optimisation 179
6.3 Large scale protein expression of full-length (FL) KLHL12 180
6.4 Cloning HiBIT- and LgBIT-KLHL12 fusions into bacterial expression vectors 182
6.5 Large scale bacterial expression of HiBIT and LgBIT-KLHL12 fusions as monomers. 183
6.6 Bacterial co-expression of HiBIT and LgBIT-KLHL12 monomer fusions with folding chaperones 184
6.7 Bacterial co-expression of HiBIT-LgBIT KLHL12 dimer 186
6.8 High-yielding HiBit-LgBit KLHL12 Kelch domain dimer from Sf9 expression . . . 187
6.9 Preliminary MST tests of NHS-labelled HiBIT-LgBIT KLHL12 Kelch domain . . . 190
6.10 Preliminary MST tests of NHS-labelled USP21 catalytic domain 191
6.11 BAY-805 may potentiate KLHL12-HiBiT-LgBiT binding to USP21 192
6.12 cryo-EM grid preparation overview for the KLHL12-USP21 complexes 194
6.13 USP21 has multiple KLHL12 PGXPP-like recognition sequences 197
6.14 Model of CUL3-RBX1-KLHL12 in complex with USP21 199
6.15 Predicted interactions between USP21 and KLHL12 200

7.1 NRBP1 AFv1 modeling and domain architecture 205
7.2 Mass Spectrometry analysis of NRBP1 pseudokinase domain 206
7.3 NRBP1 construct Sf9 expression result summary 210
7.4 NRBP1 construct *E. coli* expression result summary 211

7.5 Summary of NRBP1-ElonginBC samples for crystallisation trials 212
7.6 Representative purification of NRBP1-ElonginBC complex 213
7.7 Mass Spectrometry analysis of NRBP1-ElonginBC purified complexes 216
7.8 NRBP1:ElonginB:ElonginC is a 1:1:1 complex 217
7.9 Cryo-EM grid screening overview for NRBP1-ElonginBC complex 222
7.10 Cryo-EM data refinement workflow of NRBP1-ElonginBC complex 223
7.11 NRBP1-ElonginBC CUL2-box is not sufficient for CUL2 recruitment 225
7.12 NRBP1-ElonginBC binding precludes DDB1 association 226
7.13 BLI experimental setup to probe NRBP1-TSC22D interaction 228
7.14 Immobilisation of RFXV-containing biotinylated peptides on super-streptavidin BLI Sensors 228
7.15 NRBP1 binds to the TSC22D4 RFTV motif but not to TSC22D2 230
7.16 RFXV-containing biotinylated peptides immobilisation on SSA BLI Sensors 231
7.17 NRBP1-TSC22DF pull-down validation experiment 232
7.18 NRBP1 may associate with ElonginBC by bipartite binding 234
7.19 Putative model of TSC22DF mediated NRBP1 dimerisation 235
7.20 TSC22D4 may act as a scaffold to facilitate CUL2 binding with NRBP1 237
7.21 Putative model of $CUL4^{NRBP1}$ 238

8.1 **EVD-1** meets many covalent chemical probe criteria 247
8.2 Proposed bipartite binding by targeting both KLHL12 binding cysteines 249
8.3 Design of EVD-1 as a handle for KLHL12-based degraders 254

A.1 Oligonucleotide design for NRBP1 interactors for ***E.coli*** expression ii
A.2 Oligonucleotide design for NRBP1 interactors for **Sf9** expression ii
A.3 Oligonucleotide design for NRBP1 interactors for *E.coli* co-expression iii
A.4 Thermocycling conditions and PCR reaction for PCR amplification iii

A.5 Thermocycling conditions and PCR reaction for PCR amplification using Q5 polymerase iv
A.6 Purified PCR amplicons summary for baculo constructs: 96 successful amplifications v
A.7 Vector digestion of vectors for Sf9 expression vi
A.8 Representative Thermocycling conditions and PCR reaction for PCR bacmid screening vii
A.9 Representative Thermocycling conditions and PCR reaction for PCR bacmid screening vii
A.10 Colony PCR and bacmid screening of Sf9 constructs viii

LIST OF TABLES

1.1 KLHL12 Interacting Partners and Binding Motifs 19

1.2 Overview of identified NRBP1 interacting partners 27

1.3 Landscape of Chemical Probes for Kelch domains 34

1.4 Landscape of PROTACs for Kelch Substrate Receptors 34

3.1 Summary of KLHL12-PLEKHA4 fine screen designs 70

3.2 Summary of crystallographic data for KLHL12 and PEF1, PLEKHA4 complexes . . 79

4.1 Summary of virtual DEL hits validation 104

4.2 Cysteine Reactivity Prediction in the KLHL12 Kelch domain 105

5.1 Anisotropy of KLHL12-covalent compounds co-crystals 146

5.2 Crystallographic data for KLHL12-**EVD-1** and KLHL12-**23** complexes 151

5.3 Cyclisation conditions used for the formation of **EVD-3** 158

5.4 Conditions used for the cyanation of arylchlorides 161

5.5 Conditions used for the hydrolysis of ester **EVD-1** 163

5.6 Amide coupling conditions used in the synthesis of **EVD-18** 165

7.1 Construct design for NRBP1 and interactors for *E.coli* expression 208

7.2 Construct design for NRBP1 and interactors for Sf9 expression 209

7.3 Construct design for NRBP1 interactors for *E.coli* co-expression 214

7.4 NRBP1 Kinase domain crystallisation trials 218

7.5 NRBP1-ElonginBC complex crystallisation trials 219

7.6 NRBP1-c064[59:535]-ElonginBC complex grid screening conditions 220

7.7 NRBP1-c062[41:535]-ElonginBC complex grids for collection 221

LIST OF SCHEMES

1 General synthetic route to access hit elaborations 158

2 Synthesis of key intermediate for hit elaborations 159

3 Attempted synthesis of thioether and sulfonyl hit elaborations 160

4 Attempted synthesis of ether hit elaborations 161

5 Attempted synthesis of acid and amide analogues 163

6 Synthesis of amide analogue 165

CHAPTER 1

Introduction

1.1 The Ubiquitin-Proteasome System: A Key Regulator of Protein Homeostasis

In 1975, Gideon Goldstein reported that he had isolated a *"polypeptide of 8500 molecular weight [...] that is probably represented universally in living cells"* [1]. The ubiquitous polypeptide that showed a high degree of evolutionary conservation across eukaryotic cells of mammals and higher plants, and yeast, alike, was named *ubiquitin* [1]. Two years later, in 1977, Alfred Goldberg *et al.* reported that protein degradation in lysed reticulocytes required ATP, challenging the prevailing view that proteolysis occurred primarily in membrane-bound organelles like lysosomes and was energy-independent [2].

This discovery led to a series of studies by Ciechanover, Hershko, Rose and colleagues throughout the late 1970s and early 1980s to identify the components of this ATP-dependent proteolytic system. They employed centrifugation, protein precipitation by ammonium sulfate, and column chromatography to fractionate reticulocyte lysate in an effort to locate the proteolytic activity. They found that a factor present only in the non-adsorbed column fraction (fraction I) restored ATP-dependent degradation of isotope-labelled proteins when added back to the adsorbed fraction (fraction II) [3].

In 1979, this factor, designated ATP-dependent proteolysis factor 1 (APF-1), was found to be heat-stable, with an estimated molecular weight of 9 kDa [4]. It formed covalent conjugates with substrate proteins, which were resistant to mild acid, base, and heat treatment, but sensitive to N-ethylmaleimide, suggesting covalent bond via a cysteine residue [4–7]. The conjugates disassem-

bled upon ATP removal, releasing free APF-1, indicating a reversible process [6]. Parallel work by Wilkinson *et al.*, showed that APF-1 peptide identified by Ciechanover, Hershko, and Rose, was tautologous to the small heat-stable polypeptide, ubiquitin previously and independently identified by Goldstein [8]. The subsequent studies laid the groundwork for understanding the ubiquitinylation process, by identifying and affinity-purifying the components of the ubiquitin cascade, E1, E2, and E3, that mediate the sequential transfer of ubiquitin to abnormal protein substrates, leading them to their degradation, with E3 ligases catalyzing the final conjugation step [9–11]. Later, Chau *et al.* made a seminal discovery demonstrating that Lys48-linked polyubiquitin chains serve as critical signals for proteasomal degradation [12].

30 years after the discovery of ubiquitin by Goldberg, Aaron Ciechanover, Avram Hershko, and Irwin Rose were awarded the Nobel Prize in Chemistry for their contributions in characterising ubiquitin-mediated intracellular protein degradation. Prior to their findings proteins were largely considered stable entities, with degradation assumed to occur passively in lysosomes [13]. Years of extensive research on the small eukaryotic protein ubiquitin, revealed its role in the highly complex, temporally controlled and tightly regulated process of protein intracellular proteolysis. The discovery of regulated protein degradation reshaped our understanding of cellular control, elevating it to a level of importance comparable to transcriptional and translational regulation. Given the vast array of substrates and processes governed by this system, its dysfunction has been linked to disease pathogenesis, including cancer, neurodegenerative disorders, and immune-related conditions. This makes the pathway a crucial subject of study for gaining deeper insights into human physiology and the underlying causes of disease.

1.1.1 Protein Homeostasis and the Role of the Ubiquitin-Proteasome System (UPS)

In living cells, protein levels are balanced between protein synthesis and degradation to achieve protein homeostasis also termed as *proteostasis*. Cells constantly survey their proteome in order to identify and direct misfolded, aggregated or aberrantly expressed proteins to degradation, while sparing

the timely-expressed and properly folded proteins. The main protein degradation pathway in mammalian cells is the ubiquitin-proteasome system (UPS) [14, 15], which primarily degrades short-lived, misfolded, and damaged proteins [2, 16–18], playing a key role in regulating cell signaling, transcription, cell cycle progression, cell proliferation, and apoptosis [19–21]. Larger cellular components, including protein aggregates, damaged organelles, and membrane structures on the other hand, are usually eliminated by the autophagy-lysosome pathway (ALP) [22–26]. Both systems exhibit extensive crosstalk, with ubiquitin serving as a central regulator of proteasomal and autophagosomal degradation through coordinated and compensatory mechanisms [27–35]. The UPS will be the focus of this study since it degrades the vast majority of soluble cytosolic and nuclear proteins [36], including proteins of the endoplasmic reticulum exported into the cytoplasm [37].

Ubiquitylation is a post-translational modification in which ubiquitin is covalently attached to target proteins, typically through an isopeptide bond between ubiquitin's C-terminal glycine (G76) and an ε-amino group of a lysine residue on the substrate or another ubiquitin molecule [12, 38–40]. The ubiquitin molecule covalently bound to the substrate proteins, can form additional isopeptide bonds with other ubiquitin molecules through its N-terminal methionine M1 or one of its seven lysine residues (K6, K11, K27, K29, K33, K48, and K63), forming ubiquitin polymeric chains. These chains can be homotypic or heterotypic, branched, or linear. Different linkage types, govern distinct cellular functions, in their own "ubiquitin code" as coined by Komander *et al.* [41]. K11 or K48-linked chains marking proteins for proteasomal degradation [42, 43], while K63-linked chains have non-proteolyic functions such as affecting protein localisation including directing aggregate and organelle elimination through the autophagy-lysosome system [44–51]. Additionally, M1-linked linear chains playing roles in immune responses by activation of NF-κB signaling pathway [52, 53]. In summary, ubiquitin modification marks the substrate for proteasomal degradation, or acts as a regulatory element of its function by mediating protein-protein interactions or by affecting its subcellular localisation.

Substrate ubiquitin is mediated by the concerted action of three enzymes in an enzymatic cascade, as

outlined in Fig. 1.1. Briefly, an E1 ubiquitin-activating enzyme tethers a ubiquitin molecule to a cysteine in its binding site creating a E1 $\sim$ Ub thioester intermediate by consuming one ATP molecule. Subsequently, trans-thiolation of the activated ubiquitin occurs by an E2 ubiquitin-conjugating enzyme that transfers the activated ubiquitin molecule to a cysteine on its own binding site. Finally, an E3 ubiquitin ligase recruits substrate to enable transfer of this ubiquitin to the target protein through an isopeptide bond between ubiquitin's C-terminal glycine (G76) and a lysine residue on the substrate protein (or another ubiquitin in polyubiquitylation) [15, 54]. Ubiquitylation by this E1-E2-E3 enzymatic cascade can be iterative and yields proteins that either bear a single ubiquitin, multiple mono-ubiquitins or polyubiquitin chains of variable length and topology. Long polyubiquitin chains usually act as a signal for degradation of the substrate by the 26S proteasome [43, 55] [Fig. 1.1].

More recent studies have shown that ubiquitylation can also occur on non-lysine residues that contain a lone pair of electrons (hydroxyl, thiol, or amine groups) capable of nucleophilic attack, although the functional consequences of these modifications remain to be fully elucidated [Fig. 1.2].

For example, in addition to canonical ubiquitin conjugation to the ε-amine group of lysine residues [Fig. 1.1], ubiquitin can also be attached to the α-amino group at the N-terminus of a substrate protein. In this non-canonical mechanism, ubiquitin is conjugated to the free α-amine of the polypeptide backbone, rather than to a lysine side chain [Fig. 1.2 (a)]. A well-characterised example is Myoblast Determination Protein 1 (MyoD), which undergoes N-terminal ubiquitylation mediated by Atrogin-1/MAFbx, a muscle-specific E3 ubiquitin ligase that is upregulated during muscle atrophy [56, 57]. In an analogous manner, cysteine can also act as a nucleophile leading to its ubiquitinylation, with a recently identified example the HECT E3 NEDD4L E3 ligase that catalyses ubiquitin transfer to a cysteine of the substrate TRAF3, positively regulating innate antiviral immunity [58] [Fig. 1.2 (b)]. Finally, the endoplasmic reticulum-associated E2 conjugating enzyme UBE2J2 and E3 ligases HRD1, MYCBP2, and HOIL-1 exhibit esterification activity, enabling ubiquitination of the hyroxyl groups of serine and threonine residues on ER-associated degradation (ERAD) substrates and components of

immune signaling pathways [59–65] [Fig. 1.2 (c)].

While ubiquitin ligases mediate the covalent conjugation of ubiquitin to target proteins, **deubiquitinases (DUBs)** perform the opposing function by cleaving ubiquitin chains from substrates [Fig. 1.1 (b)]. The human genome encodes over 100 different DUBs, which are classified into seven major families. Six of these families are **cysteine proteases**, including Ub C-terminal hydrolases (UCHs) and Ub-specific proteases (USPs), which utilize a catalytic cysteine residue for deubiquitination. In contrast, the Jab1/Mov34/MPN+ protease (JAMM) family members are **metalloproteases** that require metal ions for their catalytic activity. These enzymes target various types of ubiquitin linkages, ranging from the removal of polyubiquitin chains to the cleavage of monoubiquitin from target proteins [66, 67]. DUBs, therefore, play an important regulatory role by reversing ubiquitin signalling, not only by completely removing ubiquitin from a substrate, but also by trimming ubiquitin molecules at the end of the elongating chain (distal ubiquitin), thereby modulating both the extent and type of ubiquitylation. For example, USP21 is a promiscuous DUB that can hydrolyse all ubiquitin chain types and even ubiquitin-like modifications [68], while OTULIN, an ovarian tumour deubiquitinase, specifically recognises and removes ubiquitin linked through its first methionine residue (Met1-Ub) [69]. On the other hand, the metallophone RPN11 removes the entire polyubiquitin chain at once, by recognising and cleaving the ubiquitin molecule tethered to the substrate at the start of the elongating chain (proximal ubiquitin), effectively acting as an "off" switch of ubiquitin signalling [70].

1.1.2 E3 Ubiquitin Ligases: Specificity in ubiquitylation

The ubiquitylation mechanism needs to be (i) universal to accommodate the heterogeneity of the proteome, (ii) selective to the target-protein substrates and (iii) be able to exert precise spatiotemporal control on the half-life of each substrate. Universality is achieved by the concerted action of the few and functionally conserved E1 ubiquitin-activating and E2 ubiquitin-conjugating enzymes, while E3 ubiquitin ligases confer specificity by selectively recruiting their substrates to the E2/E3 complex to be ubiquitylated [71]. The hierarchical role of the cascade is reflected in the number of the respective

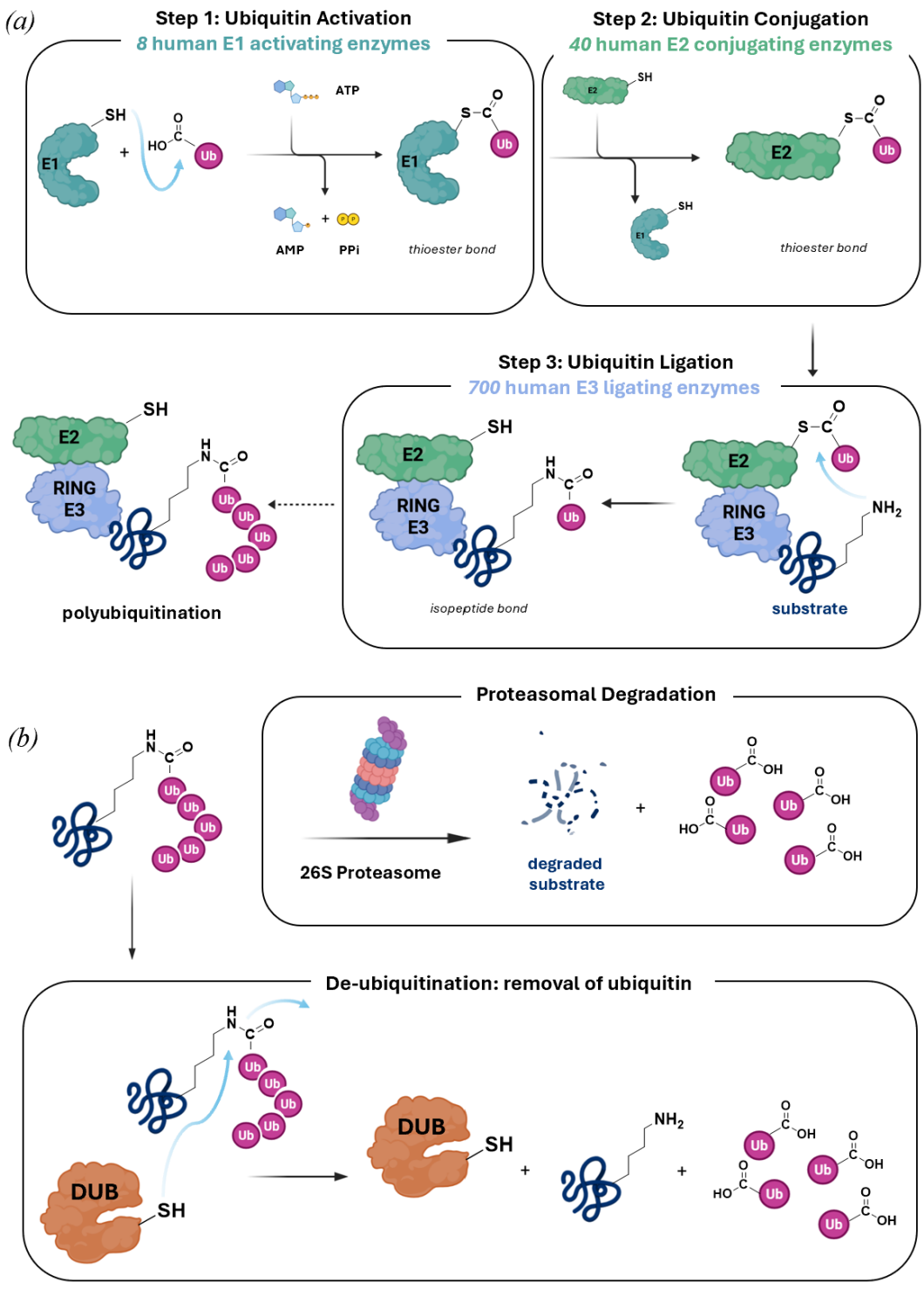


Figure 1.1: **Schematic representation of (de)ubiquitylation mechanism.** (a) Substrate-proteins get ubiquitylated by the concerted action the E1-E2-E3 enzymatic cascade: Initially, an E1-activating enzyme tethers a ubiquitin molecule to a cysteine on its binding site by consuming one ATP molecule. Subsequently, E2-conjugating enzyme transfers the activated ubiquitin to a cysteine on its own binding site in a trans-thiolation reaction. Finally, an E3-ligase transfers the ubiquitin from the E2 enzyme to a residue (typically lysine) of the substrate protein. The mono-ubiquitylation by this E1-E2-E3 enzymatic cascade can be iterative (designated by dashed arrow) to yield proteins modified with poly-ubiquitin chains. (b) Long poly-ubiquitin chains usually act as a signal for degradation of the substrate by the 26S proteasome. Proteasomal degradation releases free ubiquitin. The ubiquitin signal can be removed by deubiquitinating enzymes (DUBs), a family of cysteine proteases that hydrolyse the isopeptide bond between ubiquitin and substrate.

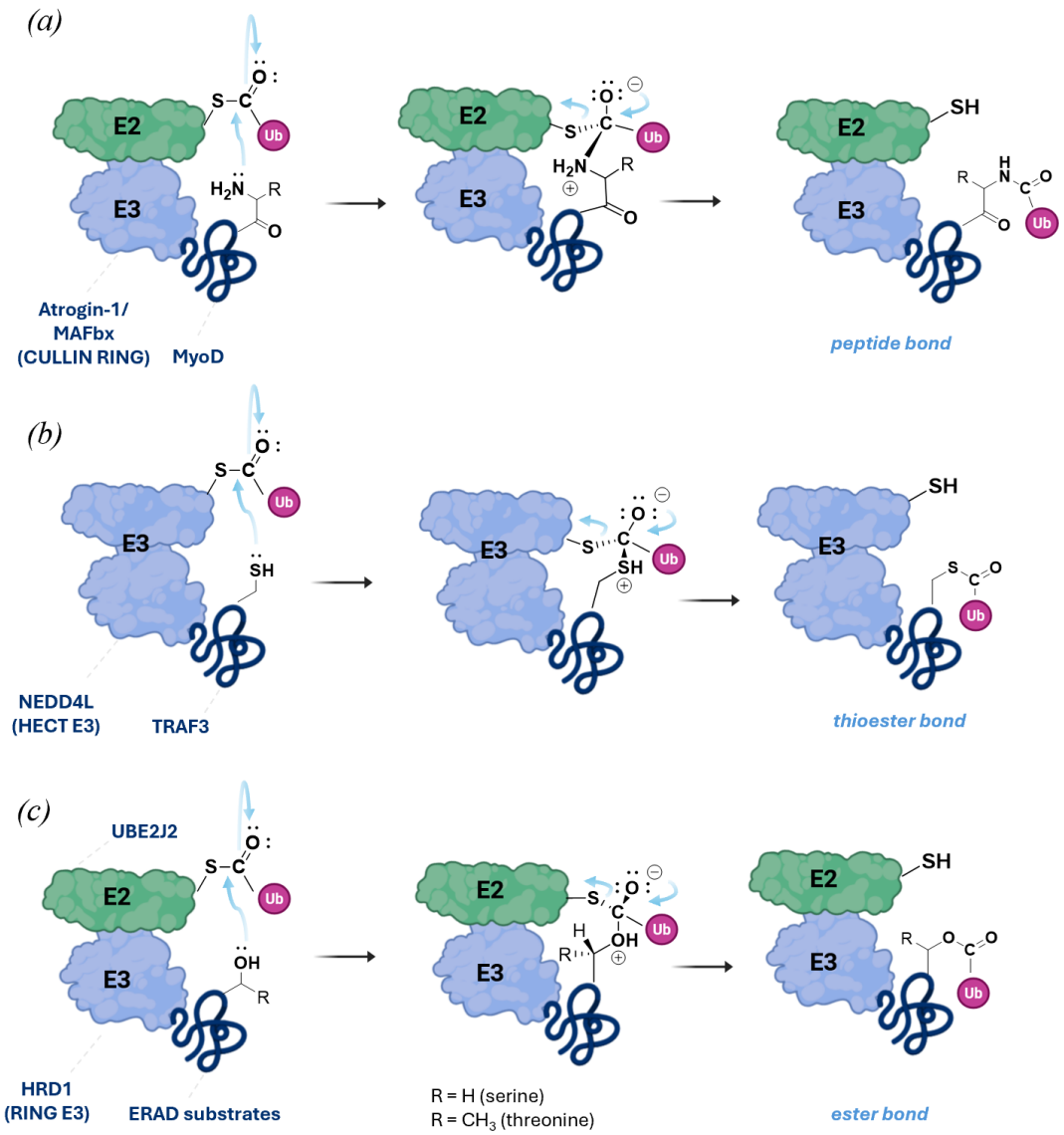


Figure 1.2: **Ubiquitinylation mechanisms of non-lysine residues.** The ubiquitylation machinery can also covalently tether ubiquitin on non-lysine protein residues with groups with a lone pair of electrons that are capable of nucleophilic attack. Representative examples follow; (a) The α-amino group at the N-terminus of MyoD serves as a nucleophile for ubiquitin conjugation, mediated by the CULLIN RING E3 ligase Atrogin-1/MAFbx, forming a peptide bond. (b) Hydroxyl groups on serine and threonine residues can also be ubiquitylated, forming oxyester linkages. This is often the case in the ubiquitinylation of ER-associated substrates by the concerted action of E2 UBE2J2 and the RING E3 ligase HRD1. (c) The thiol group of a cysteine can also serve as a nucleophile to accept ubiquitin via thioester bond formation, as observed in HECT E3 ligase NEDD4L that catalyses ubiquitin transfer by a transthiolation reaction from its own cysteine to a cysteine on the TRAF3 substrate.

enzymes, with two E1 enzymes, $\sim$ 40 E2 enzymes, and several hundred E3s in humans. An understanding of UPS requires structural insight of how the c. 700 human E3 ligases recognise and select their substrates [Fig. 1.1].

E3 ligases bind their physiological substrates through recognition of specific peptidic motifs, termed as degron motifs or degrons (as they promote degradation). Degrons can be short peptides in proteins such as the "PGXPP" motif recognised by Cullin RING (CRL) E3 ligase $CRL3^{KLHL12}$ [72], or the C-terminal di-Gly motif recognised by $CRL2^{KLHDC2}$ [73] [Fig. 1.3 (b)]. Degrons can also be "revealed" in a protein after being exposed by proteolytic cleavage as is the case with the ubiquitin-protein ligase E3 component n-recognin (UBR) ligases recognising N-degrons [74]. Degrons can also be formed upon post-translational modifications as exemplified by $Skp1^{\beta TrCP}$ recognising a phosphorylated β-catenin [75] [Fig. 1.3 (c)]. Entire folded domains can act as a signal for degradation, often also needing post-translational modifications as stimulus, as the Substrate-recognition component of the SCF (SKP1) in complex with F-box only protein 22 (FBXO22) $CRL1^{SKP1\text{-}FBXO22}$ ligase that recognises the BACH1 quaternary structure [76] [Fig. 1.3 (a)]. Finally, an interesting case is found in plants where two small-molecule phytohormones, jasmonate and auxin, mediate Aux/IAA recognition for degradation by being wedged between the substrate and SCF^{TIR1} E3 ligase [77, 78].

This complex landscape of ubiquitylation signals associated with an equally complex and diverse E3 ligase landscape. The **HECT** (**h**omologous to **E**6-AP **C** **t**erminus) class of E3 ligases form a covalent E3 $\sim$ Ub intermediate on a catalytic cysteine in their binding site, before transferring ubiquitin to the substrate. On the other hand, **RING** (**R**eally **I**nteresting **N**ew **G**ene) E3 ligases, adopt a scaffolding role, facilitating direct transfer of ubiquitin from E2 to the substrate without forming an intermediate E3 $\sim$Ub bond. The **RBR** (RING-between-RING) E3 ligases function through a hybrid mechanism, first acting like a RING E3 to bind E2, then transferring ubiquitin via a HECT-like intermediate catalytic step.

The largest class of E3 ubiquitin ligases are Cullin-RING ligases (CRLs) which are members of the

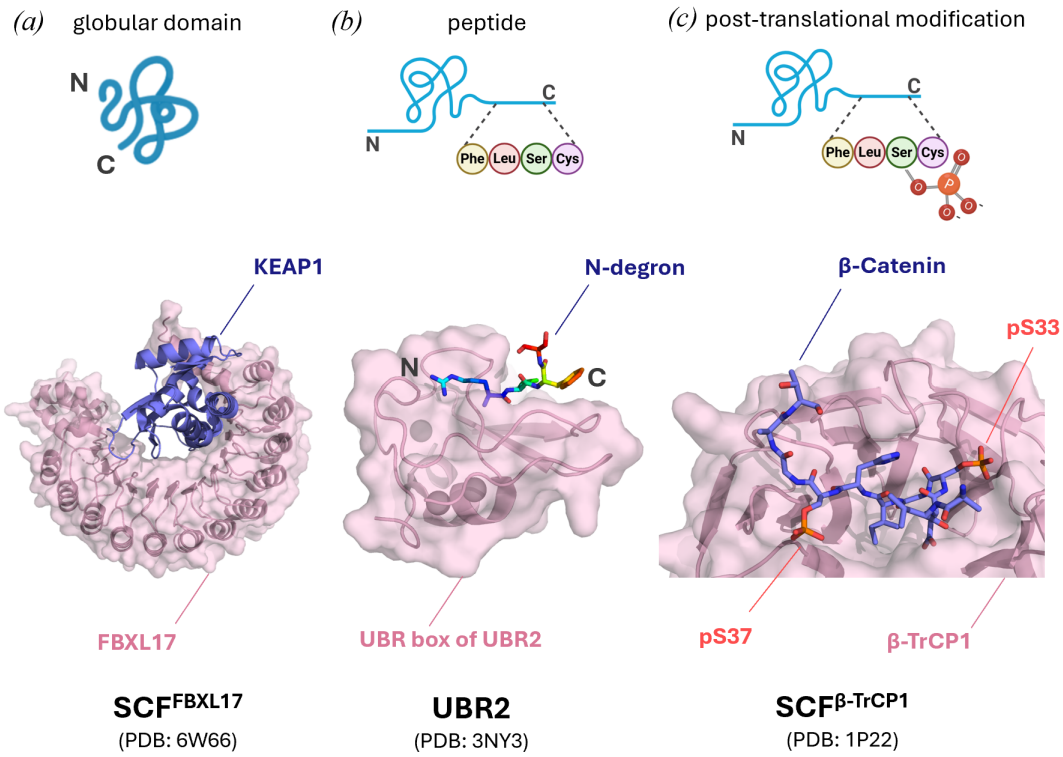
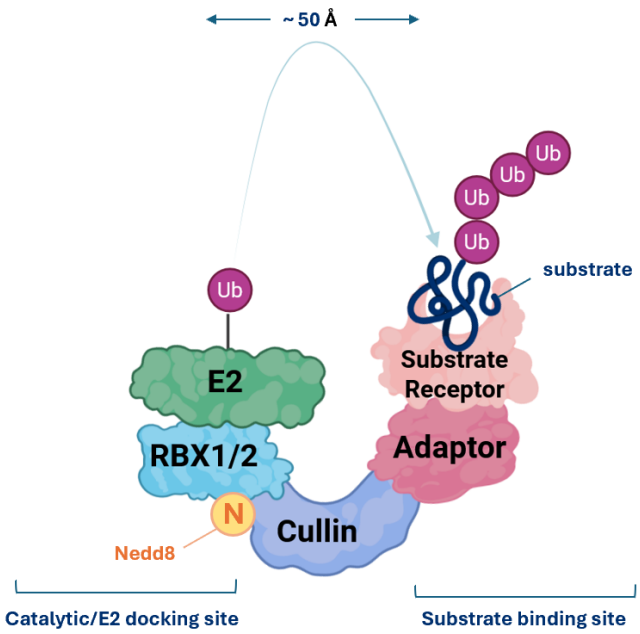


Figure 1.3: **Substrate recognition by E3 ligases**. E3 ligases recognise their target proteins for ubiquitination through specific degradation motifs known as degrons, which signal the protein for destruction by the proteasome. Degrons can be constitutive as in the cases of (a) **globular domains**, otherwise named structural degrons, such as SCF^{FBXL17} ligase that recognises a folded KEAP1 domain as a substrate for degradation [PDB: 6W66 [79]], or (b) **short linear aminoacid sequences**, as in the case of UBR2 ligase that recognises proteins with a positively charged N-terminus (bearing arginine or lysine residues) through its UBR-box [PDB: 3NY3 [80]]. (c) Degradation motifs can also be created by pros-translational modifications, such as $SCF^{\beta TrCP}$ ligase that recognises di-phosphorylated β-catenin (phosphodegron) [PDB: 1P22 [75]]. Structure figures were made with PyMOL [81].

(a)



(b)

Cullin-RING E3 ligase	Cullin	Adaptor	Substrate Receptor domain	# of human substrate receptors
CRL1	CUL1	SKP1	F-box	68
CRL2	CUL2	ELONGIN B/C	VHL-box	~20
CRL3	CUL3	-	BTB box	~70
CRL4/B	CUL4A/B	DDB1	DCAF	~25
CRL5	CUL5	ELONGIN B/C	SOCS	~30
CRL7	CUL7	SKP1	FBXW8	1
CRL9	CUL9	undefined	undefined	undefined

Figure 1.4: **Architecture of the modular Cullin-RING E3 Ligases (CRLs).** (a) Schematic representation of the general architecture of the Cullin-RING ligase family (CRL). CRLs are a large family of multi-subunit modular enzymes, that consist of a central structural scaffold Cullin protein that binds to a substrate-binding adaptor through its N-terminus, and a RING (really interesting new gene) domain recruiting an E2 enzyme through its C-terminus. Cullin has a scaffolding role bringing together the substrate and the ubiquitin-bearing E2 enzyme, allowing for ubiquitin transfer to the substrate directly from the charged E2 enzyme. Besides their scaffolding role, Cullins can also be activated by covalent attachment of Nedd8 protein (neddylation) that leads to conformational change promoting ubiquitin transfer to the substrate. (b) Overview of human CRL components and composition. CRLs can recruit a variety of E2 enzymes and substrate receptors through adaptor proteins. These interchangeable components can be swapped out to recognise different substrates and interact with different cellular signaling pathways. Despite human Cullin protein members being only 8 - seven canonical (CUL1, CUL2, CUL3, CUL4A, CUL4B, CUL5, CUL7) and one atypical (CUL9) - the combinatorial assembly of each Cullin with many varied substrate receptors leads to more than 500 possible CRLs. Skp1 = S-phase kinase-associated protein 1, VHL = von Hippel-Lindau, BTB = Bric-a-brac, Tramtrack, Broad-complex domain, DDB1 = DNA damage-binding protein 1, DCAF = DDB1-CUL4 associated factor, SOCS = Suppressors of cytokine signaling, FBXW8 = F-box and WD repeat domain containing protein 8.

RING E3 ligase family, and will be the topic of investigation of this thesis. CRLs are modular, multi-subunit ligase assemblies, where catalytic cores can bind diverse substrate receptors interchangeably [Fig. 1.4]. A handful of well-studied catalytic cores termed Cullins (CUL1, CUL2, CUL3, CUL4A, CUL4B, and CUL5) are the central scaffolding proteins of the assembly adopting an elongated core. The Cullins bind the substrate receptor directly or through dedicated substrate adaptor proteins via their N-terminal domains. The C-terminal domain of the Cullins bind the RING zinc finger protein RBX1 or RBX2, which recruits the E2 ∼ubiquitin intermediate. This architecture brings the E2-bound ubiquitin proximal to the substrate associated with the substrate receptor to facilitate ubiquitin transfer. Each scaffolding cullin associates with specific substrate receptors, which in turn recognise specific "degrons", affording ∼ more than 500 estimated distinct CRL complexes [Fig. 1.4]. This combinatorial assembly of this limited set of proteins enables the system to accommodate a vast array of substrates, ensuring broad specificity and adaptability.

1.1.3 Exploring Non-Canonical E3 Ligase Assemblies

In addition to the well-characterised architecture of RING and HECT ligases briefly described above, a growing number of E3 ligases exhibit non-canonical architectures and atypical assembly mechanisms. Another layer adding to the complexity of substrate recognition and E3 ligase activity regulation is E3 ligase multimerisation.

Some ligases function as obligate dimers or higher-order oligomers, where multimerisation is essential for their catalytic activity or substrate recognition. For example, RNF4 dimerises through its RING domain, to facilitate ubiquitylation of SUMOylated substrates [82]. $CRL3^{KLHL22}$ dimerises via KLHL22 BTB domain to target the same degron in a multimeric substrate (mitochondrial protein GDH1 hexamer) and promote its poly-ubiquitylation [83] [Fig. 1.5 (a)]. Another BTB-containing ligase, $CRL3^{KEAP1}$, also dimerises via KEAP1 BTB domain to target different degrons in a monomeric substrate NRF2 at the same time. The high affinity ETGE motif in the substrate binds first to one Kelch domain, whereas binding of the weaker DLG motif on the second Kelch domain follows in a

coined *"hinge and latch mechanism"* [84] [Fig. 1.5 (b)].

DDB1-associated substrate adaptor, DCAF1, had been proven to dimerise through its lissencephaly homology LisH domain [85] to form an active form of a $CRL4^{DCAF1}$ ligase that successful ubiquitinates substrate Uracil-DNA glycosylase 2 (UNG2). However, later studies revealed that the protomer ligase can form an inactive tetramerised state, with the WD40 domain of DCAF1 mediating binding to the cullin C-terminal domain (CTD) and the RBX1 subunit of a neighboring $CRL4^{DCAF1}$ protomer [86] [Fig. 1.6 (b)]. A similar mechanism of multimerisation leading to autoinhibition is observed in the dimerised $CRL2^{KLHLDC2}$, where KLHLDC2's C-terminal Gly-Ser motif mimics a degron and engages the substrate-binding domain of another protomer [87] [Fig. 1.6 (a)].

These unconventional assemblies expand the functional diversity of the ubiquitin-proteasome system, enabling specialised regulatory roles beyond classical paradigms. This thesis will investigate two of these CRLs, namely the homotypic $CRL3^{KLHL12}$ and the heterotypic $CRL2/CRL4^{NRBP1}$, whose architecture and relevance to physiology will be introduced below.

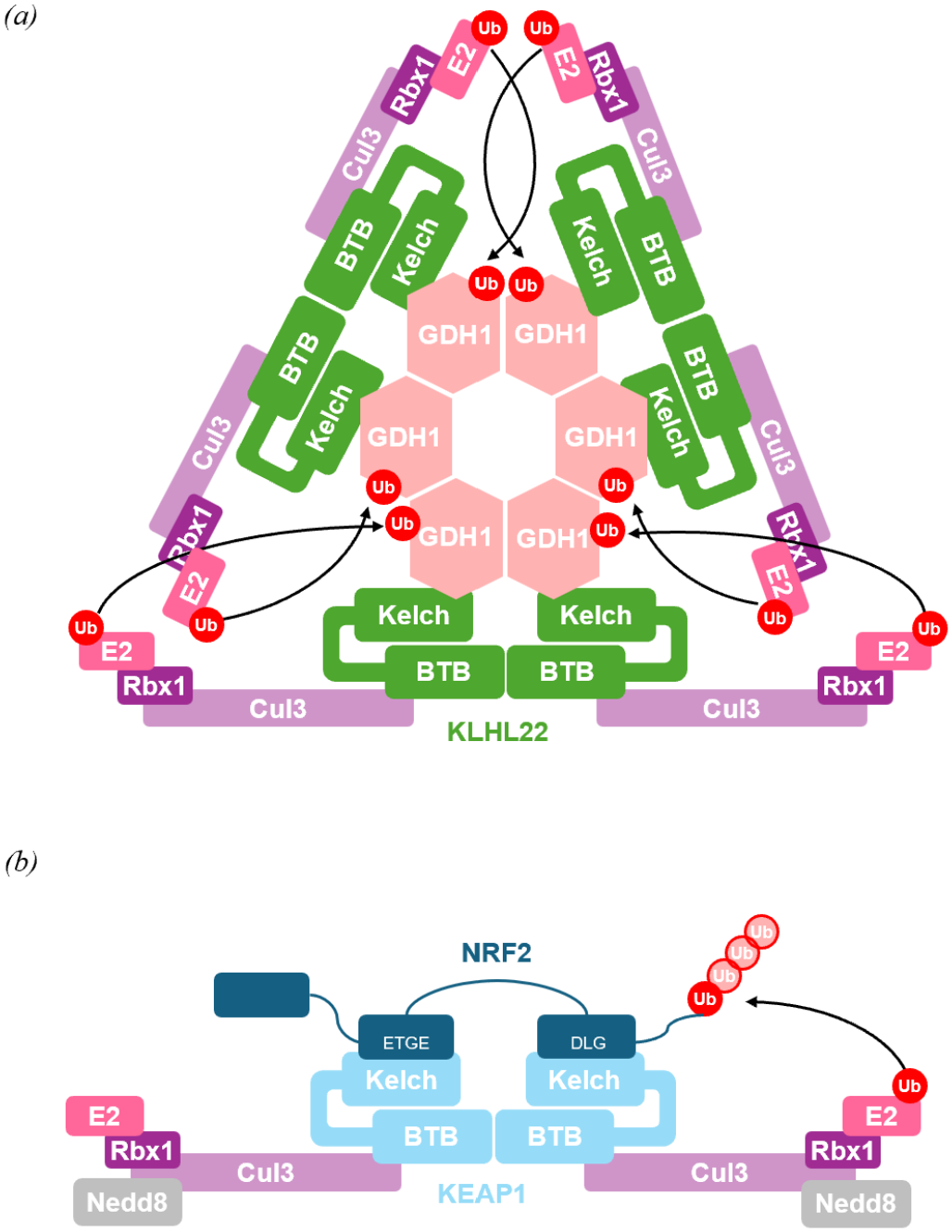


Figure 1.5: **Productive multimerisation of E3 ligases enhances substrate targeting.** Many E3 ligases multimerise to more effectively target their substrates for ubiquitination. This multimerisation can serve multiple purposes: it may enable recognition of multimeric substrates or allow simultaneous engagement of multiple degrons within a single substrate. Oligomerisation may also increase the local concentration of the E3 complex, thereby enhancing the efficiency and speed of ubiquitination. (a) A representative example of an E3 dimerising to target the same degron in a multimeric substrate is $CRL3^{KLHL22}$ ligase protomer that dimerises through its N-terminal BTB domain to target hexameric mitochondrial protein GDH1 [model based on PDB: 8KHP [83]]. (b) A representative example of an E3 ligase multimerising to target different substrate degrons is $CRL3^{KEAP1}$ protomer that dimerises through its N-terminal BTB domain to target different degrons (ETGE, and DLG motifs) in a monomeric substrate NRF2 simultaneously [model based on PDB: 2FLU, 3WN7 [84]].

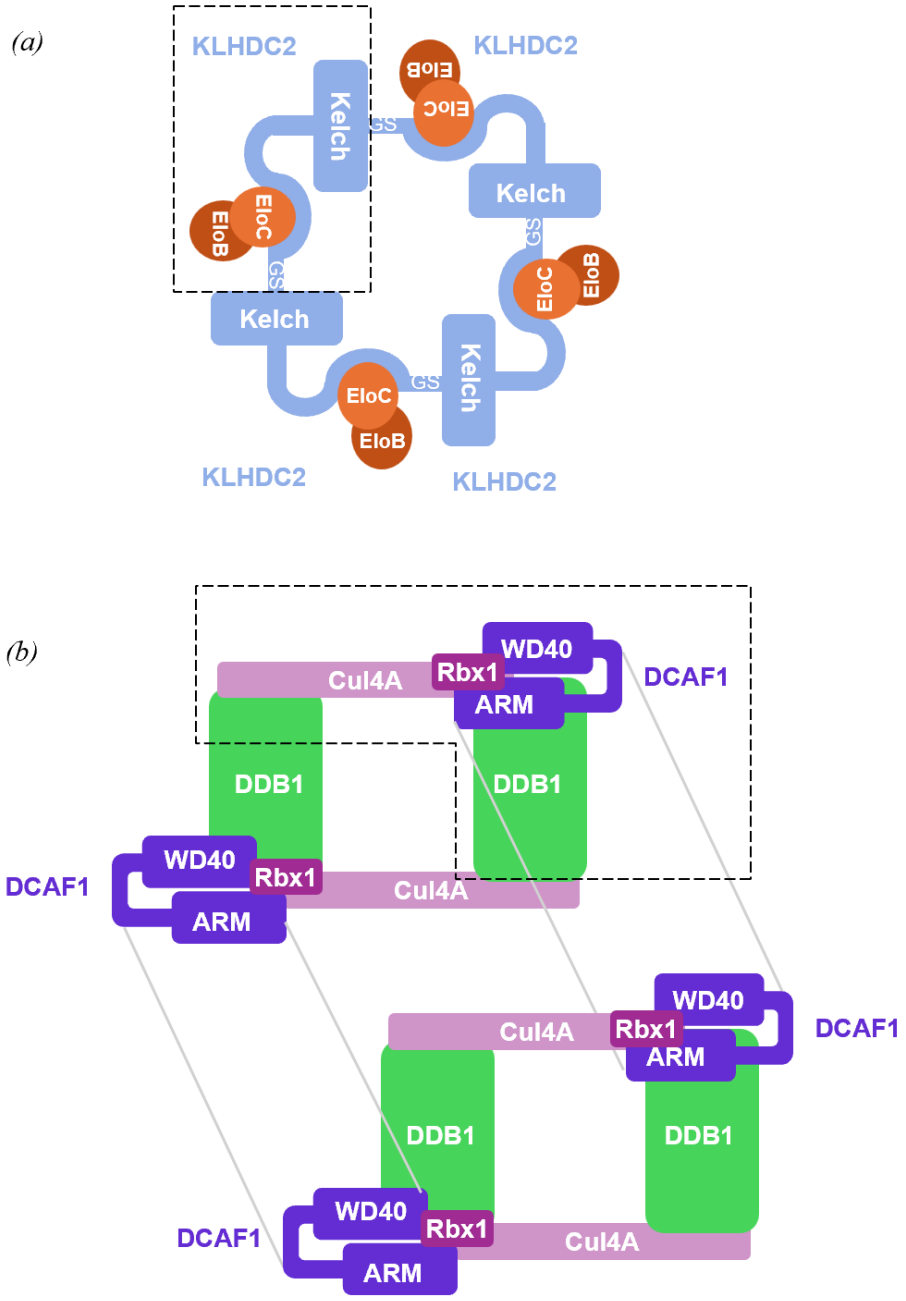


Figure 1.6: **Non-productive multimerisation of E3 ligases prevent substrate recruitment.** Some E3 ligases can self-associate into multimeric assemblies that are unable to engage substrates, forming auto-inhibited complexes. This self-assembly provides an additional layer of regulation, controlling the timing and specificity of ubiquitination activity. (a) A representative example of an E3 ligase rendered inactive through multimerisation is $CRL2^{KLHLDC2}$ ligase. KLHDC2, which normally recognises C-terminal di-Gly degrons on substrates, self-associates by engaging its own di-Gly motif. This interaction drives the formation of an auto-inhibited homotetramer that is unable to bind external substrates [87]. (b) Another example of auto-inhibition via self-association is the $CRL4^{DCAF1}$ that forms an inactive tetramer. Tetramerisation is mediated by two distinct intermolecular interfaces between the RBX1 and DCAF WD40 domain at one dimerisation interface, and the DCAF1 N-terminal LisH motifs at the other. Therefore, two dimers associate to form the inactive tetramer [model based on PDB: 7OKQ [86]].

1.2 KLHL12: A substrate adaptor of a homotypic Cullin-RING E3 Ligase

1.2.1 KLHL12 in embryonic and neural development

KLHL12 is widely expressed across various tissues, with higher expression levels in testis and the brain, and moderate expression in heart, skeletal muscle, liver, lung, and kidney [88], indicating its involvement in fundamental cellular processes. As a substrate-specific adapter of BCR (BTB-CUL3-RBX1) E3 ubiquitin ligase complexes, KLHL12 can engage target-specific co-adaptors to ubiquitylate different substrates, not only promoting their proteasomal degradation but also affecting their function and localisation.

The $CRL3^{KLHL12}$ ligase was first described as a negative regulator of the developmental Wnt signalling pathway by poly-ubiquitinating Dishevelled paralogs DVL1-3 to promote their degradation [89] [Fig. 1.7 (a)]. Pleckstrin homology domain-containing family A, member 4 (PLEKHA4) has been observed to act as an antagonist of the KLHL12-mediated dishevelled's poly-ubiquitylation , thereby acting as a positive regulator of the Wnt pathway [90]. PLEKHA4 oligomers localise into clusters at the PI(4,5)P2-rich (phosphatidylinositol 4,5-bisphosphate) membrane regions and employ their proline-rich domains to sequester $CRL3^{KLHL12}$ ligases away from the free pool of cytoplasmic DVL [Fig. 1.7 (c)]. Since DVL is an integrator of both β-catenin dependent and independent branches, KLHL12 has been posited as a modulator of cell differentiation and proliferation, as well as cell polarity and motility respectively, highlighting KLHL12's role in embryogenesis and tumorigenesis [91, 92]

KLHL12 has also been implicated in vesicle formation, size, and function by modulating export and Endoplasmic-Reticulum (ER)/Golgi localisation of Coat Protein Complex II (COPII) vesicles bearing collagen and apoB100-VLDL particles [93, 94]. In the case of collagen, $CRL3^{KLHL12}$ ligase engages calcium-dependent complex PEF1/ALG2 that acts as co-factor diverting substrate selectivity of the ligase away from DVLs to COPII vesicle protein SEC31 [95] [Fig. 1.7 (b)]. Contrary to dishevelled paralogs, substrate SEC31 is monoubiquitylated, while co-factor PEF1 is not modified

by the ligase. Neither co-factor PEF1, nor substrate SEC31 are degraded as a consequence of their binding to KLHL12 [94, 95].

Collagen is expressed in neurons [96, 97] and plays a crucial role in regulating axonal outgrowth and synaptic differentiation [98]. Its dysregulation has been implicated in cellular stress responses and Alzheimer's disease [99]. Furthermore, mutations in *SEC23A* that disrupt ER-to-Golgi trafficking result in the accumulation of large proteins, including collagen and lipid particles, leading to cranio-lenticulo-sutural dysplasia [100]. Impaired ER-Golgi transport, along with other trafficking defects, has also been described in amyotrophic lateral sclerosis (ALS) [101]. These findings highlight KLHL12 as a critical regulator of COPII vesicle formation and transport, linking its function to neurodevelopmental disorders.

In addition to the ubiquitylation of soluble cytosolic proteins by $CRL3^{KLHL12}$, a recent proteomic study identified a trans-membrane protein, Lunapark, as another substrate of the ligase [102]. Lunapark is an ER-shaping protein localized to ER three-way junctions, and its ubiquitylation by KLHL12 leads to the recruitment of mTORC1 to the lysosome [102]. Interestingly, unlike typical substrates targeted for degradation, Lunapark's ubiquitylation does not serve as a degradation signal but rather functions as a signal for mTORC1 lysosomal recruitment [102]. The study further demonstrated that inhibiting Lunapark ubiquitylation results in neurodevelopmental defects, suggesting that KLHL12-dependent ubiquitylation is crucial for normal neural development.

KLHL12 has also been found to play a non-canonical role in dopamine signaling. Specifically, it regulates dopamine responses by ubiquitinylating lysine and non-lysine residues in the polymorphic repeats of the human D4.2 receptor, without promoting receptor degradation [103–105]. This suggests an additional layer of KLHL12's involvement in neural function beyond its canonical role in protein degradation.

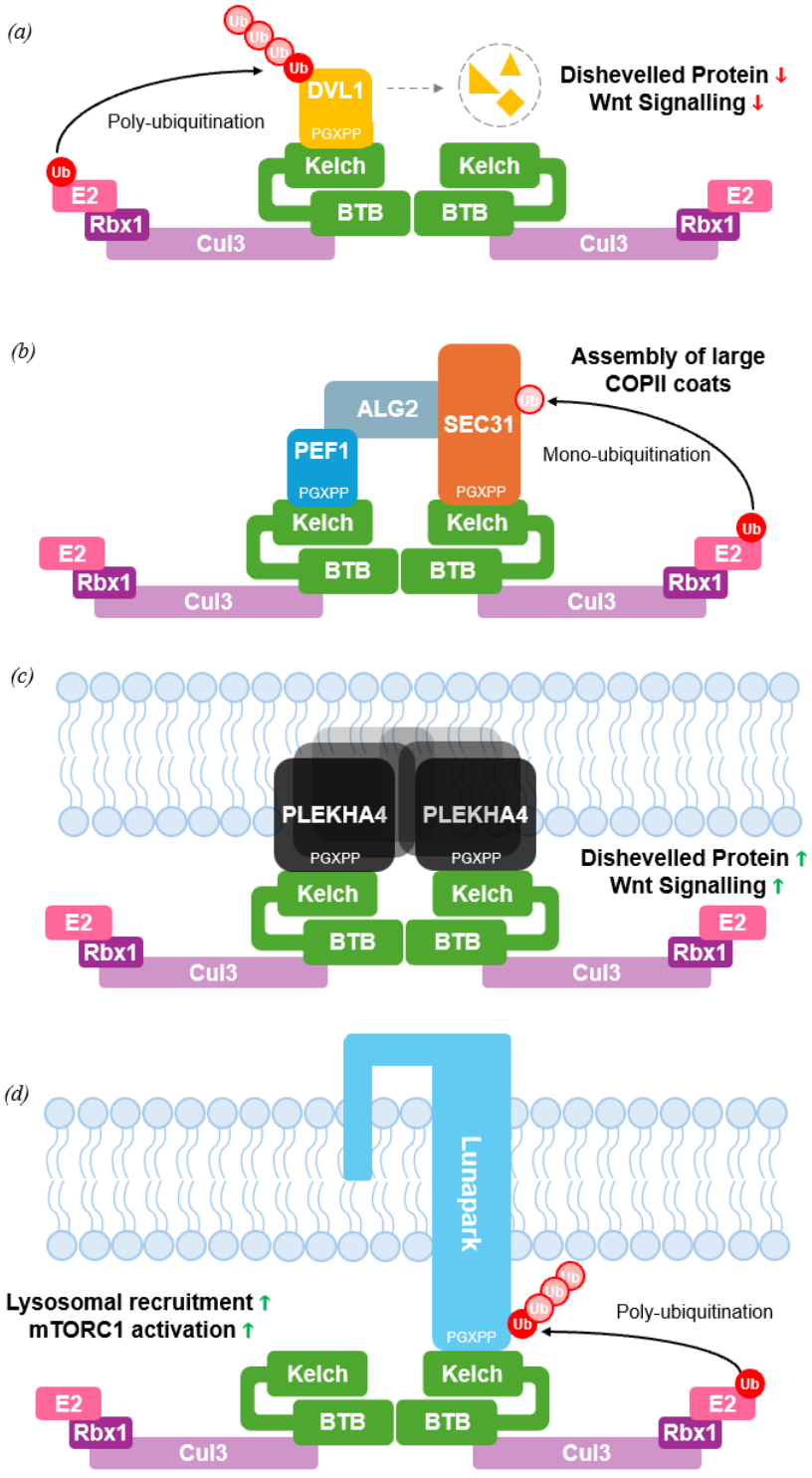


Figure 1.7: **Models of CUL3^{KLHL12} E3 ligase assemblies.** KLHL12 homodimerises through its N-terminal BTB domain and binds to interacting partners through a conserved "PGXPP" motif. (a) CUL3^{KLHL12} poly-ubiquitylates substrate DVL1 leading to its degradation (b) one CUL3^{KLHL12} protomer binds substrate-specific co-adaptor ALG2-PEF1, while the other binds substrate SEC31 to be monoubiquitylated (adapted from McGourty *et al.* [95]). (c) CUL3^{KLHL12} is sequestered to the membrane associating with PLEKHA4 oligomers. (d) CUL3^{KLHL12} recognises ER-membrane protein Lunapark for poly-ubiquitylation (adapted from Yuniati *et al.* [102]).

1.2.2 KLHL12 domains mediate its role as a substrate-specific adaptor for CRL3

Kelch-like protein 12 (KLHL12) is a member of the BTB-BACK-Kelch family of proteins, which function both as Cullin-RING ligase (CRL) adaptors and substrate recognition modules [89, 106]. KLHL12's N-terminal BTB (Bric-a-brac, Tramtrack, Broad complex) and downstream BACK (BTB and C-terminal Kelch) domains confer its scaffolding role by recruiting Cullin 3 (CUL3), while its C-terminal Kelch domain recognises substrates or target-specific co-adaptors [106–108] [Fig. 1.8 (a)].

The structural basis of CRL3 assemblies has been elucidated, revealing by homology that KLHL12's BTB domain mediates CUL3 binding together with a two-helix C-terminal BTB extension, which forms a hydrophobic groove known as the *3-box* [108] [Fig. 1.8 (e, f)]. The 3-box and downstream helices in KLHL12 form the BACK domain, which form a linker to the C-terminal Kelch domain in addition to CUL3 binding [109].BTB residues are also known to facilitate homodimerisation [110], a feature that is functionally significant for CRL3 (Cullin-RING Ligase 3) ligase activity [111] [Fig. 1.8 (b, c)]. Interestingly, in addition to the canonical homodimerisation between KLHL12 protomers, KLHL12 has been also found to heterodimerise with KLHL26 [112], hinting at an even greater modularity mediated by BTB domains.

$CRL3^{KLHL12}$ ligase recruits substrates for ubiquitylation via its C-terminal Kelch domain [Fig. 1.8 (d)]. Kelch domains, widespread across eukaryotic species, are characterised by their β-propeller structure, which primarily facilitates protein-protein interactions. These domains are the most prevalent substrate-recognition modules in CRL3 complexes [106]. A typical β-propeller fold consists of four to eight blades, each formed by antiparallel β-sheets arranged radially around a central axis. While the β-strands are crucial for maintaining the central propeller form, the flexible loops connecting each strand are highly variable and serve as the binding sites for interactors. The variability of these loops provides specificity in engaging diverse cellular targets. KLHL12's Kelch domain adopts this typical β-propeller fold, with six Kelch repeats (numbered I-VI), each consisting of four-stranded

antiparallel β-sheets (labeled a-d) [106] [Fig. 1.8 (d)].

1.2.3 KLHL12 recognises a "PGXPP" motif

Despite the diversity of interactors identified for KLHL12, recent studies have mapped the binding of direct interacting partners to a consensus "PGXPP" motif [Table 1.1], where X represents a small non-polar residue [72, 114]. This finding was in agreement with preliminary evidence that both PLEKHA4 and DVL bind KHLH12 through proline-rich regions [90, 95]. The structures of the Kelch domain of KLHL12 in complex with peptides of substrates DVL1 [72] and DVL3 [114] have been also resolved, showing both peptides bound to the shallow accessible KLHL12 Kelch domain. Even though both peptides adopt a U-shaped conformation that manages to contact all six blades of the β- propeller domain of the C-terminal KLHL12 Kelch motif, two divergent binding modes were revealed [Fig. 1.9 (a), (b)]. Specifically, KLHL12 distinct interactions with the DVL1-**PP**GG**PP** and DVL3-**P**GA**PP** peptides were visualized through buried surface analysis. These differences are summarized in Fig. 1.9 (c)-(f). Interestingly, Chen *et al.* proved that this proline rich motif is both a recognition and degron motif [72].

Table 1.1: KLHL12 Interacting Partners and Binding Motifs

KLHL12 Interacting Partner	**KLHL12 Binding Motif**	**Relation to KLHL12**	**Study**
Human DVL1	658-P**PG**G**PP**-663	Substrate	[72, 95, 114]
Human DVL2	698-P**PG**A**PP** -703	Substrate	[72, 95]
Human DVL3	678-P**PG**A**PP** -683	Substrate	[72, 89, 95, 114]
Human PLEKHA4	176-G**PG**G**PP**-181	Antagonist	[90]
Human PEF1	17-A**PG**A**PP**-22	Co-adaptor	[95]
Human SEC31	957-G**PG**A**PP**-962	Substrate	[94, 95]
Human Lunapark	203-A**PG**G**PP**-208	Substrate	[102]
Human DRD4	298-A**PG**L**PP**-303	Substrate	[103–105, 114]
Consensus	**PGXPP**		

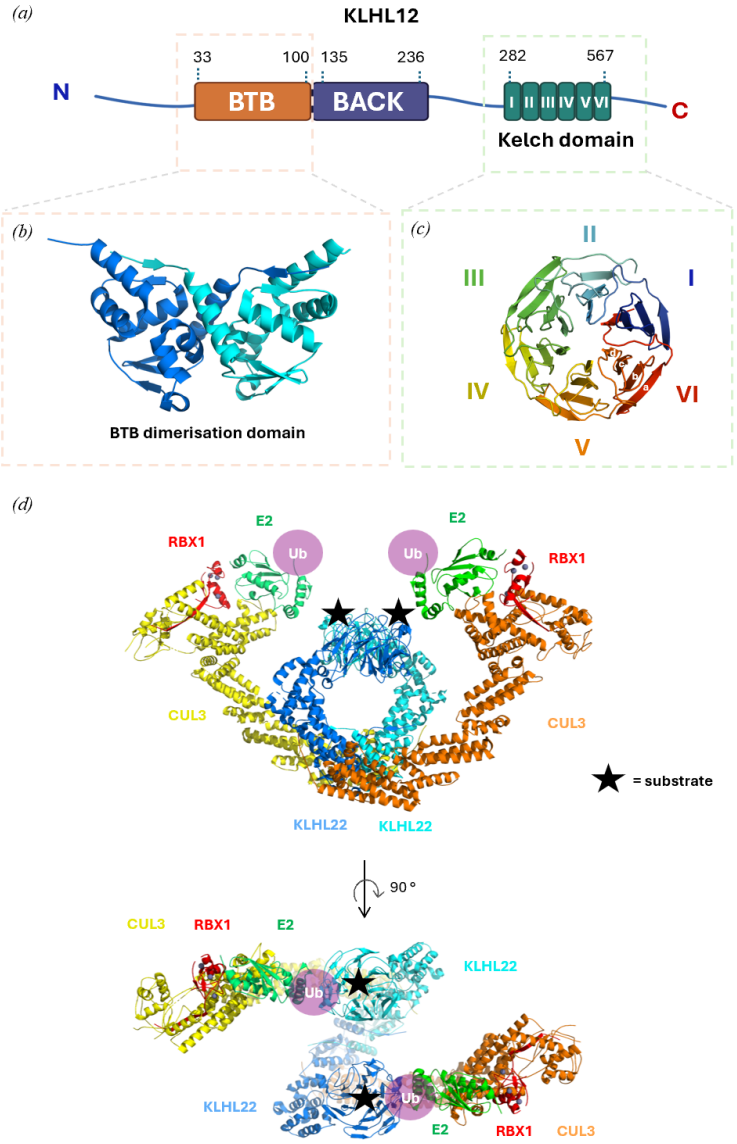


Figure 1.8: **Overview of CRL3^{KLHL12} architecture.**
(a) Overview of human KLHL12 domain architecture, with numbers referring to KLHL12 residues. The domain boundaries of the BTB, BACK, and Kelch domains of KLHL12 are indicated, as well as the six Kelch repeats that constitute the complete Kelch domain. (b) KLHL12 dimerises through its N-terminal BTB domain. Here is presented a model of the BTB-mediated KLHL22 homodimer based on PDB: 8KH [83]. One KLHL22-BTB domain is coloured in cyan while the other in blue. KLHL12 BTB-mediated homodimer should be analogous. (c) Kelch domain of KLHL12 [PDB: 2VPJ [106]]. The six Kelch repeats are numbered I-VI, and the four antiparallel βstrands that form each Kelch repeat are labelled a-d. (d) Model of CRL3^{KLHL22} ligase. KLHL22 dimerises through its N-terminal BTB domain and each monomer recruits a substrate (denoted by black star) via its C-terminal Kelch domain. Each KLHL22 monomer also recruits one CUL3 scaffold protein through its N-terminal BTB-BACK domain. On the other end of the assembly, CUL3 binds to RBX1 and to a ubiquitin-charged cognate E2 enzyme. This supramolecular assembly brings the ubiquitin molecule in proximity to the substrate. The KLHL12-CUL3 assembly should be analogous to the presented model. The model was generated by aligning the CUL3-KLHL22-RBX1 complex [PDB: 8KHP [83]] and the structure of an RBX1-UBC12 ~NEDD8-CUL1-DCN1 complex [4P5O [113]] through their common elements (RBX1-Cullin scaffold). Structure figures were made with PyMOL [81].

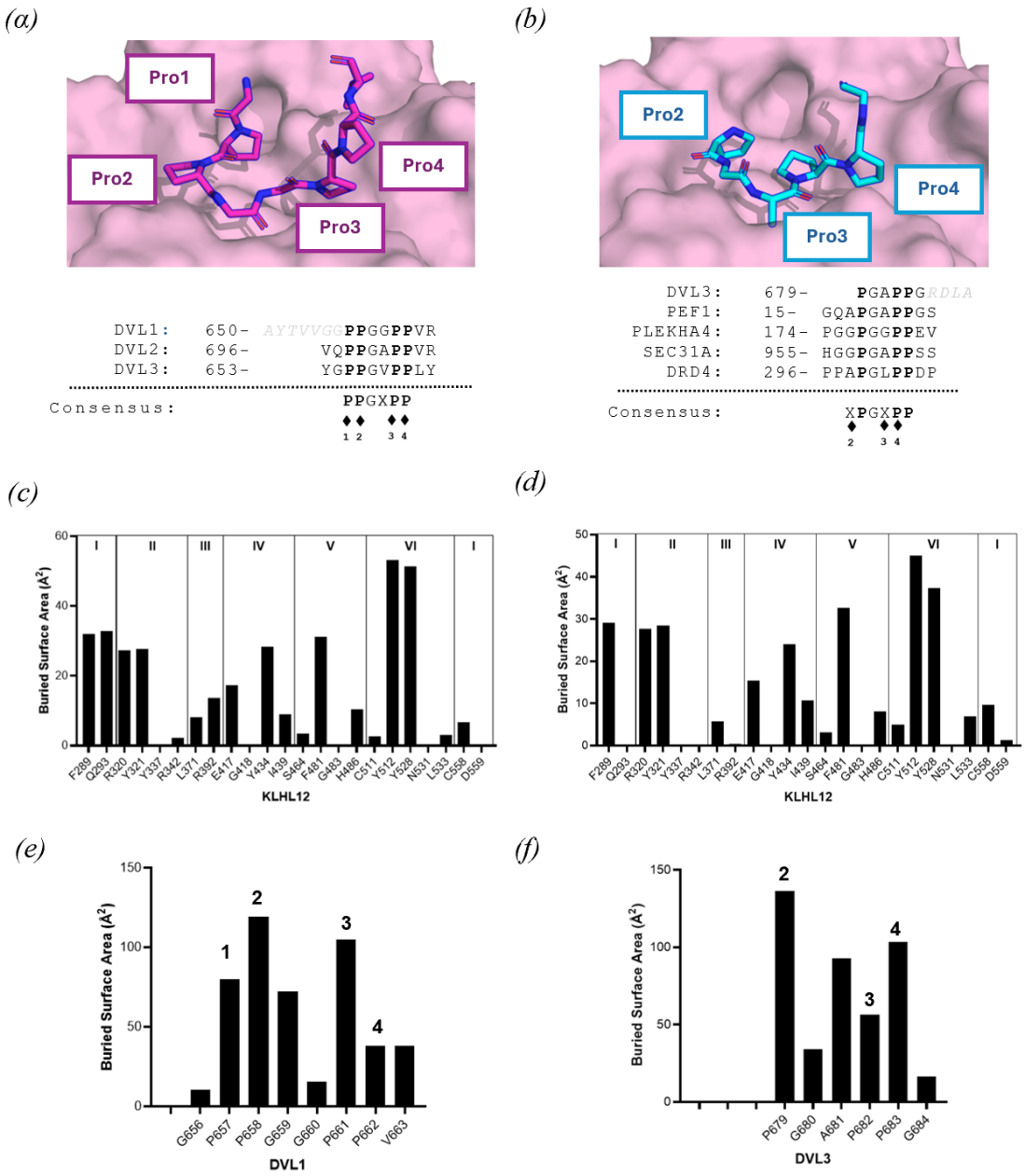


Figure 1.9: **Two different binding modes for dishevelled substrates binding to KLHL12 Kelch domain have been described in the literature.** (a) DVL1 peptide (magenta) bound to the surface of KLHL12 Kelch domain (pink). The sequence of the co-crystallized DVL1 peptide is provided below, with peptide residues that were not resolved in the co-crystal structure coloured in light gray. (b) **DVL3** peptide (cyan) bound to the surface of KLHL12 Kelch domain (pink). The sequence of the co-crystallized DVL3 peptide is provided below, with peptide residues that were not resolved in the co-crystal structure coloured in light gray. Sequence alignment of KLHL12 interactors that led the authors to identify the "PGXPP" consensus motif is provided below. Conserved proline residues are numbered 1-4 in the sequence alignment. In the DVL3 structure (right panel) Pro2 has shifted to occupy the position of DVL1 Pro1 (left panel), and consequently DVL3 Pro4 has shifted to occupy the position of DVL1 Pro3. It is unclear if that represents a divergent biding mode or is an artifact of DVL3 N-terminal truncation. (c), (e) Buried surface area of residues involved in the **DVL1-KLHL12** interaction (KLHL12 chain A, DVL1 chain E). (d), (f) Buried surface area of residues involved in the **DVL3-KLHL12** interaction (KLHL12 chain A, DVL3 chain C). The numbers above the bars in (e) and (f) correspond to the conserved proline numbering from (a) and (b) sequence alignments. The KLHL12-DVL1 complex is based on PDB: 6TTK, and the KLHL12-DVL3 complex is based on PDB: 6V7O. Buried surface area was calculated using the EMBI-PISA server ([115]).

1.3 NRBP1: A substrate adaptor of a heterotypic Cullin-RING E3 ligase

1.3.1 NRBP1 is an atypical pseudokinase that associates with transcriptional machinery components

Nuclear receptor-binding protein 1 (NRBP1) is a ubiquitously expressed and highly conserved adapter protein with a kinase-like domain. Sequence alignment revealed that kinase domain lacks all three critical motifs that confer kinase activity (DFG, HRD, VAIK), classifying it as class G pseudokinase [116, 117]. The lack of kinase activity has been also experientially validated in vitro as it lacked the ability to phosphorylate substrates [118, 119]. NRBP1 is thus named, as it carries two putative nuclear receptor-binding motifs, a putative binding domain for SH2 (Src homology 2) domain-containing proteins, nuclear export signals and a nuclear localization signal [117]. NRBP1 shuttles between the cytoplasm and the nucleus, acting as a regulator of protein localisation and transcription-factor activity [117–119].

NRBP1's interaction with members of the transcription factor family has been well documented. NRBP1 directly interacts with the mammalian proteins TSC22D1, TSC22D2, TSC22D3, and TSC22D4 - mammalian transcription factors with putative coiled coil helical regions that could from leucine zipper hetero or homodimers [120]. This finding was recapitulated in Drosophila [121], human and mouse cell-lines [117]. NRBP1 also directly binds another transcription factor, the zinc-finger Sall4 [117, 122], that has a role in embryonic stem cell pluripotency and development [123]. In addition to transcription factors, NRBP1 has also been reliably reported to bind to the ElonginBC complex [117, 120, 122], that was initially identified as a positive activator of RNA polymerase II (Pol II) elongation factor Elongin A [124].

1.3.2 NRBP1 as substrate adaptor in CRL E3 ligases

ElonginBC, in addition to its role in transcriptional elongation, is a known adaptor complex that recruits substrate receptor proteins containing a SOCS-box or ElonginBC-box motif to assemble with

cullin complexes CUL2-Rbx1 and CUL5-Rbx2 [125]. This led to the hypothesis that NRBP1 could function similarly to SOCS-box proteins as a substrate adaptor in Cullin-Ring Ligase (CRL) complexes. Initial studies on the characterisation of cullin-box sequences that are responsible for recruitment of CUL2/CUL5 to Elongin BC-based ligase complexes failed to show association of NRBP1 with either Cullin [120]. Based on sequence alignment analysis of suppressor of cytokines signalling (SOCS)-box containing proteins, the authors hypothesised that the variant residues (mainly 337A instead of 337C) in NRBP1's atypical SOCS-box is responsible for this inability to recruit cullin proteins [120].

However, more recent findings from immonoprecipitation experiments in HeLa Cells over-expressing NRBP1 and cullins have shown NRBP1 binding ElonginBC through its SOCS-box [331-341 a.a.] and subsequent CUL2 recruitment [126], only in the presence of TSC22D2/D3/D4. The same authors also observe CUL4 recruitment through DDB1 association, only in the presence of TSC22D2/D3/D4, through a proposed H-box overlapping with ElonginBC [126] [Fig. 1.10 (b)]. The H-box motif, first described in hepatitis B virus X protein (HBx), is a short peptide sequence, Φ-Φ-X_3-Φ, where Φ represents a hydrophobic amino acid (like leucine, isoleucine, or valine), and X represents any amino acid, that mediates DDB1 binding [127].

Based on these findings, the authors propose a model whereby homodimeric NRBP1 assembles into a heterotypic CUL2/CULl4A Cullin Ring Ligase (CRL) [126] [Fig. 1.10 (a)]. Leucine zipper mammalian transcription factor family members TSC22D3/4 which are known NRBP1 binding partners [120, 128] seem to act as molecular chaperones to promote NRBP1-Cullin assembly by multiple mechanisms; either by enhancing NRBP1 stability or by mediating direct interaction that has not been fully elucidated.

NRBP1 homodimerisation is likely mediated through a conserved lissencephaly homology (LisH) motif L-X_2-L-$X_{3\text{-}5}$-L-$X_{3\text{-}5}$-L in its C-terminal domain (461-480 a.a.) [118, 126], as observed for the dimerisation of DDB1/CUL4-associated factor 1 (DCAF1) [85]. LisH is a dimerisation do-

main, whose mutations result in lissencephaly (smooth brain), a debilitating developmental syndrome caused by the impaired ability of postmitotic neurons to migrate to their correct destination in the cerebral cortex [129]. Yasukawa *et al.* reported that the homodimerisation of NRBP1 through the LisH and/or TSCC2D association, was necessary for the ligase's ability to ubiquitylate its substrates but not necessary for substrate recruitment [126].

The same authors employed a trypsin-resistant tandem ubiquitin-binding entity (TR-TUBE) assay coupled with mass spectrometry-based proteomics to identify the amyloidogenic peptides BRI2 and BRI3 as substrates of the CRL4/CRL2^{NRBP1} ligase [126]. These new data confirming NRBP1's ligase activity provide context for previous studies reporting that the knockdown of NRBP1 in human colon cancer cells led to the accumulation of the transcription factor SALL4. In light of later studies implicating NRBP1 as an E3 ligase substrate receptor that directs its substrates to proteasomal degradation [126], these earlier findings raise the question of whether SALL4 itself is a substrate of the putative ligase [122].

1.3.3 NRBP1 in pathophysiology

NRBP1, through its association with multiple transcription factors and key pathway regulators, as summarised in Table 1.2, has been implicated in cell differentiation, proliferation, tissue homeostasis, and pathologies like cancer and Alzheimer's disease.

NRBP1 has shown tumour suppressing properties by (i) associating directly and inhibiting the oncogenic myeloid leukaemia factor 1 (MLF1) [118], and (ii) exerting a JAB-1 inhibitory effect on the activation of activator protein (AP-1) [119], a transcription factor exhibiting both tumorigenic and pro-apoptotic effects [130], (iii) a *C. elegans* RNAi screen revealed that NRBP1 is down-regulated in a range of intestinal tumors and correlated with poor prognosis [122]. This finding was later validated in adult mice by Tan *et al.*, who proposed that NRBP1 negatively regulates SALL4, possibly through an Elongin BC ubiquitin complex, suppressing Wnt signalling that is linked with aberrant proliferation, differentiation and localisation of cells along the crypt villus axis [131].

A study published at the time of writing of this thesis, also characterised a novel substrate of NRBP1 ligase, and shed light into its inhibitory role of interleukin-1 (IL-1)-mediated inflammatory pathway in squamous cell carcinoma (SCC) [132]. The study found that NRBP1 ligase can degrade tumor necrosis factor receptor-associated factor 6 (TRAF-6), while oncogenic TSC22D4 inhibits that degradation by directly binding and inactivating NRBP1 ligase, promoting IL-1 signaling. NRBP1 knockout mice develop intestinal tumors and hematologic malignancy [132].

Physiological inhibitors of amyloid precursor protein (APP) processing and Aβ oligomerisation BRI2 and BRI3 [133] were identified as substrates of the NRBP1-containing ubiquitin ligase [126], implicating NRBP1 in Alzheimer's pathogenesis. Amyloidogenic peptides arising from mutations in the BRI2 gene, FBD-BRI2 and FDD-BRI2 which are implicated in Familial British Dementia (FBD) and Danish Dementia (FDD) respectively [133], were also identified as substrates [126]. RNAi-mediated NRBP1 depletion in neuronal cells lead to an increase in abundance of BRI2/3 peptides and reduced Aβ production [126], positing NRBP1 inhibition as a viable strategy for pharmacological intervention in amyloidogenesis. Further genome and transcriptome wide association studies have also linked NRBP1 with a variety of neuropsychiatric pathologies other than amyloidogenic disease; reduced drinking behaviour through mTOR signalling [134], lower Parkinson's disease (PD) risk [135], a biomarker related to the treatment of post-stroke depression [136], and functional outcome of ischemic stroke [137].

A recent study investigating the close link between NRBP1 and the largely disordered transcription factors of the TSC22D family, suggested that NRBP1, TSC22D family members, and homologous to NRBP1 with-no-lysine WNK kinases have co-evolved in metazoans to co-regulate rapid cell-volume changes in response to osmolarity. The study proposed that the intrinsically disordered regions of the transcription factors allow these proteins to form biomolecular condensates, as a response to hyperosmotic stress [138]. These findings combined with genome-wide association studies posit NRBP1 as a possible therapeutic target in metabolic disorders, such as gout [139] and diabetes [140].

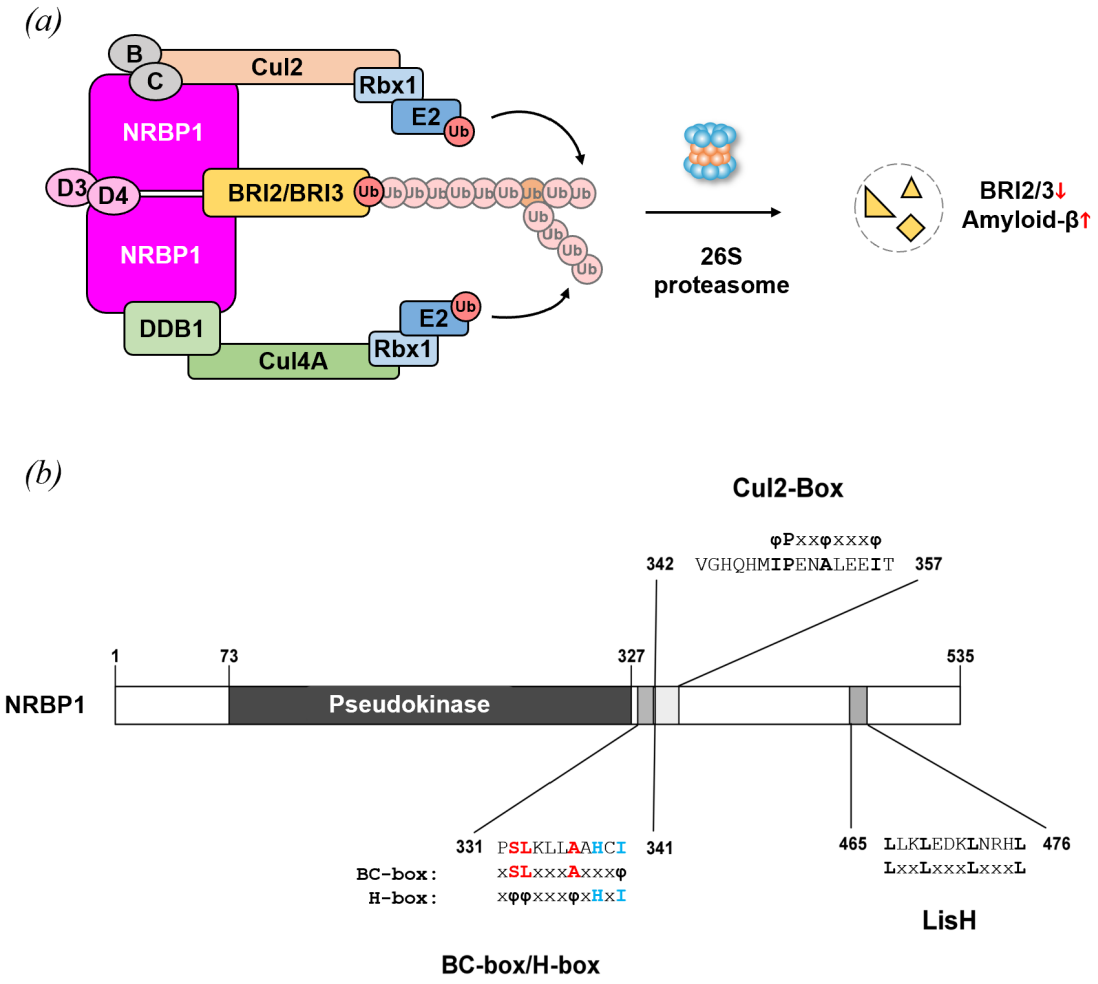


Figure 1.10: **Proposed model for CRL2/CRL4A^{NRBP1} Ligase** (a) Figure of the putative assembly of heterotypic NRBP1-containing CRL2/CRL4 E3 ligase, adapted from Yasukawa *et al.* [126]. The E3 ligase targets physiological Aβ inhibitor BRI2/BI3 for degradation, leading to higher levels of Aβ production and aggregation. (b) NRBP1 domain architecture, with conserved sequences mediating ElonginBC (BC-box), CUL2 (Cul-2 box), DDB1 (H-box) recruitment and NRBP1 dimerisation (LisH motif) as proposed by Yasukawa *et al.* [126].

Table 1.2: Overview of identified NRBP1 interacting partners

NRBP1 Interactor	Binding Observed	Method	Study
ElonginBC	yes	TR-TUBE, MS, co-IP	[120]
	yes	IAC-MS, co-IP	[122]
	yes	co-IP	[117]
	yes	co-IP	[126]
DDB1	yes	co-IP	[126]
CUL2	no	IAC-MS, co-IP	[120]
	+ TSC22D3/D4	co-IP	[126]
CUL5	no	IAC-MS, co-IP	[120]
	yes	co-IP	[117]
	yes	IAC-MS, co-IP	[122]
	inferred via EloBC	N/A	[126]
CUL4	+ TSC22D3/D4	co-IP	[126]
BunA[1]	yes	IAC-MS	[121]
TSC22D2	yes	co-IP	[117]
	yes	IAC-MS, co-IP	[122]
	yes	co-IP	[126]
TSC22D3	yes	co-IP	[126]
TSC22D4	yes	IAC-MS, co-IP	[120]
	yes	co-IP	[117]
	yes	IAC-MS, co-IP	[122]
	yes	co-IP	[126]
BRI2[2]	yes	TR-TUBE, MS, co-IP	[126]
BRI3	yes	TR-TUBE, MS, co-IP	[126]
Sall4	yes	IAC-MS	[117]
	yes	IAC-MS, co-IP	[122]
NRBP1	yes	co-IP	[118]
	+ TSC22D3/D4	co-IP	[126]

IAC: ImmunoAffinity Chromatography, co-IP: co-immonoprecipitation, MS: Mass Spectrometry, TR-TUBE: trypsin-resistant tandem ubiquitin-binding entity, +TSC22D: binding observed only in transcription factor's presence, [1]Drosophila homologue of TSC22D, [2] Also FBD-BRI2, FDD-BRI2 disease mutants.

1.4 Chemical Probes and Small-Molecule Strategies for Modulating E3 Ligases

Dysregulation of E3 ligases is often implicated in disease, leading to aberrant activation or inhibition of signaling pathways, improper assembly of protein complexes, accumulation of misfolded proteins, or mislocalisation of proteins from their functional cellular compartments [141, 142]. Despite the critical role of E3 ligases in cellular proteostasis, the diverse functions and assembly mechanisms of this extended family of ~ 700 members remain incompletely understood. Selective modulation of E3 ligases can shed light into their function and enable the interrogation of UPS components.

1.4.1 Chemical Probes are selective protein modulators

An attractive approach to achieving selective protein modulation is through the use of chemical probes; small-molecule tools capable of traversing cellular membranes to selectively bind with high affinity to a target protein. These probes modulate protein interactions by enhancing, inhibiting, or stabilizing them. They can be employed to investigate mechanistic and phenotypic aspects of protein function in *in vitro*, *in cellulo*, and in animal studies [143].

Chemical probe inhibition is complementary to gene silencing methods such as CRISPR and RNAi, providing a means to experimentally validate hypotheses generated by genetic associations or qualitative biological studies on protein function [144]. While gene knockout and silencing act at the transcriptional or translational level, chemical probes function in the post-translational landscape. Data generated from chemical probes can be integrated with chemical proteomics and biophysical methods as orthogonal assays to delineate the catalytic or scaffolding roles [145] of ubiquitin-proteasome system (UPS) components. Chemical probes also offer temporal control, allowing for precise modulation through controlled administration, and can be used in combination screens, such as synthetic lethality experiments, to explore pathway interconnections, including those between autophagy and proteasomal degradation.

Despite their potential, only a limited number of chemical probes meet the quality criteria necessary to

generate biologically reliable data [143, 146]. Some probes suffer from inadequate characterisation, while others exhibit poor properties such as low solubility, non-selectivity, or the presence of promiscuous aggregating and pan-assay interference structures [147, 148]. To address these challenges, the scientific community has advocated for standardised quality benchmarks and the development of open-access, community-led databases cataloging chemical probes, their targets, and properties[1].

A high-quality chemical probe should meet three key criteria: (i) high in vitro potency against the target protein, (ii) high selectivity relative to other proteins within the same family and other pharmacologically relevant off-targets, and (iii) demonstrated target engagement *in cellulo* [143, 146].

1.4.2 Chemical Probes in Understanding E3 Ligase Function

Small molecules can function as traditional occupancy-based inhibitors, mimicking a ligase knockout phenotype and shedding light on its function. These molecules can be identified through high-throughput screening of large small-molecule libraries using in vitro assays or, if the degron is known, by attempting degron mimicry. A prominent example in the E3 ligase field is the discovery of ligands for the $CRL2^{VHL}$ complex, which recognizes and ubiquitinates hydroxylated hypoxia-inducible factor-1α (HIF1α), targeting it for proteasomal degradation under normoxic conditions [149, 150]. The degron motif of HIF1α initially served as a peptide ligand for $CRL2^{VHL}$ [151], though with limited affinity. Computational approaches and medicinal chemistry optimization led to the development of small molecules that mimic the hydroxyproline core, which mediates critical interactions, ultimately yielding a high-affinity small-molecule probe [152, 153].

As an alternative to non-covalent inhibition, and since ubiquitin transfer relies on cysteine chemistry, the catalytic or other critical cysteine residues of E1, E2, and E3 enzymes can also be targeted with covalent modifiers. A recent example is the development of covalent modulators of $CRL3^{KEAP1}$, where a triterpene derivative, Omaveloxolone, acts as a reversible ligase covalent binder to KEAP1 BTB domain. Omaveloxolone was recently approved for the treatment of Friedreich's Ataxia and

[1] For example, see http://www.chemicalprobes.org/

relapsing multiple sclerosis [154].

Beyond E3 ligase inhibition, thiol-reactive covalent chemical probes can be employed to capture catalytic cysteines in E1, E2, or E3 enzymes, providing insights into ubiquitylation mechanisms and the structural organisation of the ubiquitin-proteasome system (UPS). These selective covalent ligands can also be utilized in chemoproteomic approaches such as activity-based protein profiling (ABPP) to label specific E3 ligases [155]. These probes serve as molecular tools to (1) investigate E3 ligase regulation (2) discover novel enzyme classes, and (3) capture transient ubiquitylation complexes, thereby enabling structural characterization of enzymatic architectures by trapping E1-E2-E3 components in their active or inactive states. For E3 ligases that contain catalytic cysteines in their binding site, covalent probes have facilitated structural studies of Parkin [64] and uncovered novel ligase mechanisms, such as the RING-Cys-Relay (RCR) class of MYCBP2 E3 ligases [156] and the ATP-dependent giant E3 ligase RNF213 [157]. Even RING E3 ligases that lack catalytic cysteines, have been chemistry-enabled through activity based probes that possess a photocrosslinking moiety to compensate for the absence of a catalytic nucleophile [158].

1.4.3 Targeted Protein Degradation: from chemical probes to proximity inducers

Building on the use of small-molecule chemical probes as E3 ligase ligands, these molecules can also serve as recruitment handles to bring an E3 ligase into proximity with a new target protein, effectively redirecting ubiquitylation activity at will. The ubiquitylated target proteins are then led to the proteasome for degradation similarly to physiological substrates. Leveraging proximity-induced ubiquitylation can thus enable selective degradation of previously *undruggable* proteins, overcoming the limitations of traditional occupancy-based inhibitors [159]. This principle is harnessed by two major categories of molecules: **PROTACs** and **molecular glues**, which act as molecular bridges, enhanced by weak, induced protein-protein interactions that stabilize the ternary complex.

Proteolysis-targeting chimeras (PROTACs) are heterobifunctional molecules that induce targeted protein degradation by linking an E3 ligase to a protein of interest (POI) [Fig. 1.11 (b)]. They consist of

a ligand for the POI or *neosubstrate* at one end, a ligand for an E3 ligase at the other, and a linker that ensures optimal distance for ternary complex formation [Fig. 1.12 (a)].

The pioneering PROTACs in the field are based on immunomodulatory drugs (IMiDs) that recruit the cereblon ligase $CRL4^{CRBN}$ [160] and hydroxy-proline HIF-1α mimetics that recruit the von Hippel-Lindau ligase $CRL2^{VHL}$ [152, 161, 162]. CRBN and VHL ligases are central in PROTAC development due to their well-characterised ligand affinity, binding modes, and flexible linker positioning, which provide effective options for the design of degraders. PROTACs have been successfully used to target a variety of substrates, including kinases such as Bruton's tyrosine kinase (BTK) [163, 164] and receptor tyrosine kinases [165], as well as epigenetic targets such as bromodomain-containing proteins and other chromatin regulators [162, 166]. Through intense research efforts in both academia and industry, we are now at a point where new degraders are being developed for previously difficult-to-target proteins, such as tau [167]. However, challenges persist, as these small molecules often exhibit poor biophysical properties, such as high lipophilicity, low solubility, and limited membrane permeability, that can negatively impact their metabolic stability and overall bio-availability.

Molecular glues, in contrast to PROTACs, also bridge interactions between E3 ligases and a (neo)substrate but exhibit a different mechanism, along with more favourable drug-like biophysical properties and improved cell permeability. Molecular glues are small, monovalent molecules that either mediate new or stabilize weak protein-protein interactions by being sandwiched between E3 ligase and POI affording a "neo-surface" for the interaction. Essentially, both the POI and E3 recruitment moieties - in an analogy with the PROTAC architecture - coexist into the same molecule, with no linker needed [Fig. 1.12 (b)].

Molecular glues can be compared to naturally occurring compounds such as auxin [78] and jasmonate [77], which mediate degron recognition by complementing suboptimal degrons [Fig. 1.11 (a)]. The discovery of molecular glues, in contrast to PROTACs that can be rationally designed by linking two identified ligands with an appropriate linker, is largely serendipitous. The first xenobiotic discovered

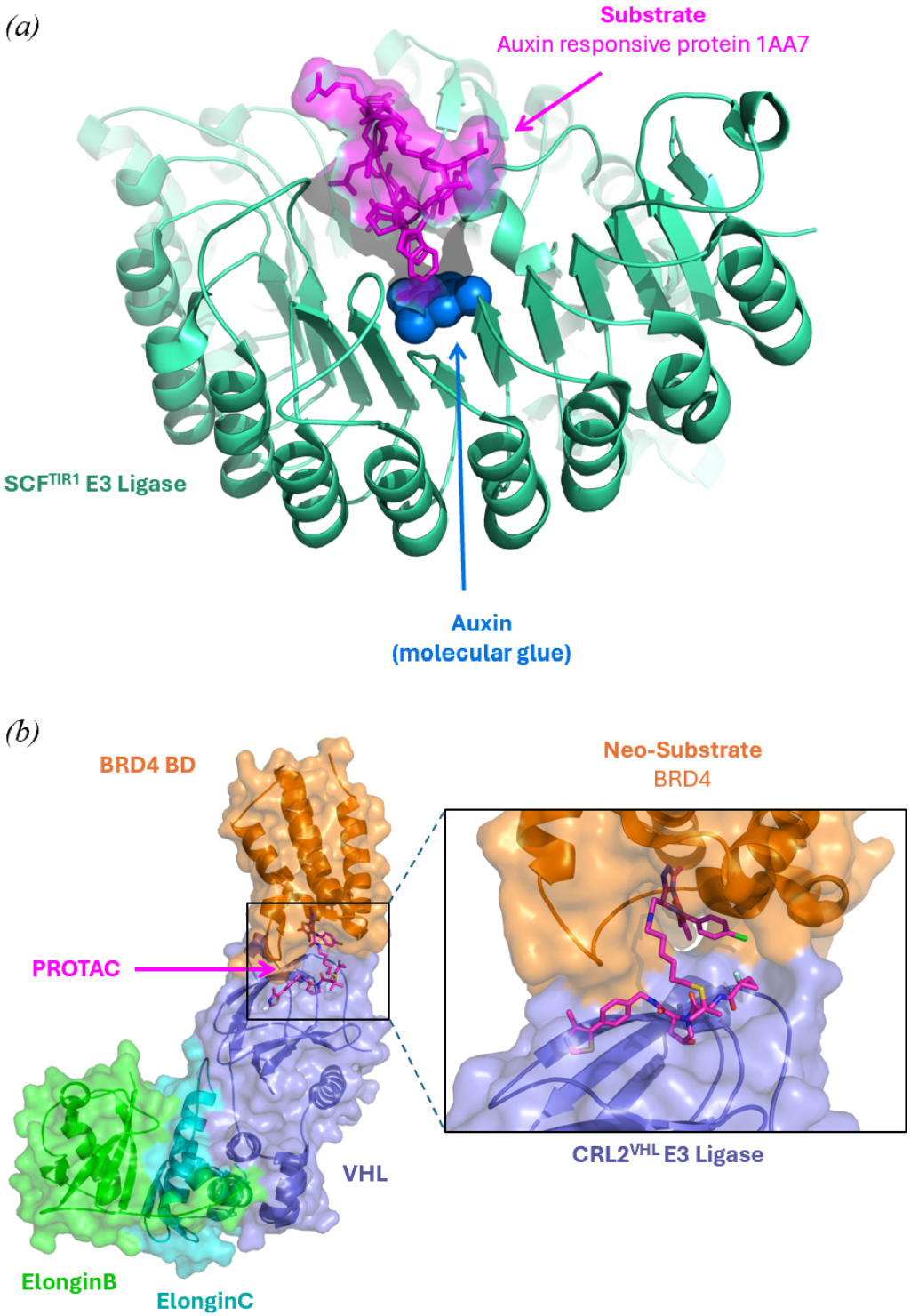


Figure 1.11: **Structural basis of Targeted Protein Degradation: Molecular Glue vs. PROTAC.** (a) Representative structure highlighting the mechanism of action of molecular glues. The molecular glue auxin is mediating the substrate (auxin responsive protein 1AA7) recognition by ligase SCF^{TIR1}. PDB: 2P1N [78] (b) Representative structure highlighting the mechanism of action of PROTACs. PROTAC AT7 recruits the second bromodomain of human BRD4 to the VHL CRL substrate adaptor in complex with ElonginBC. PDB: 7ZNT [168].

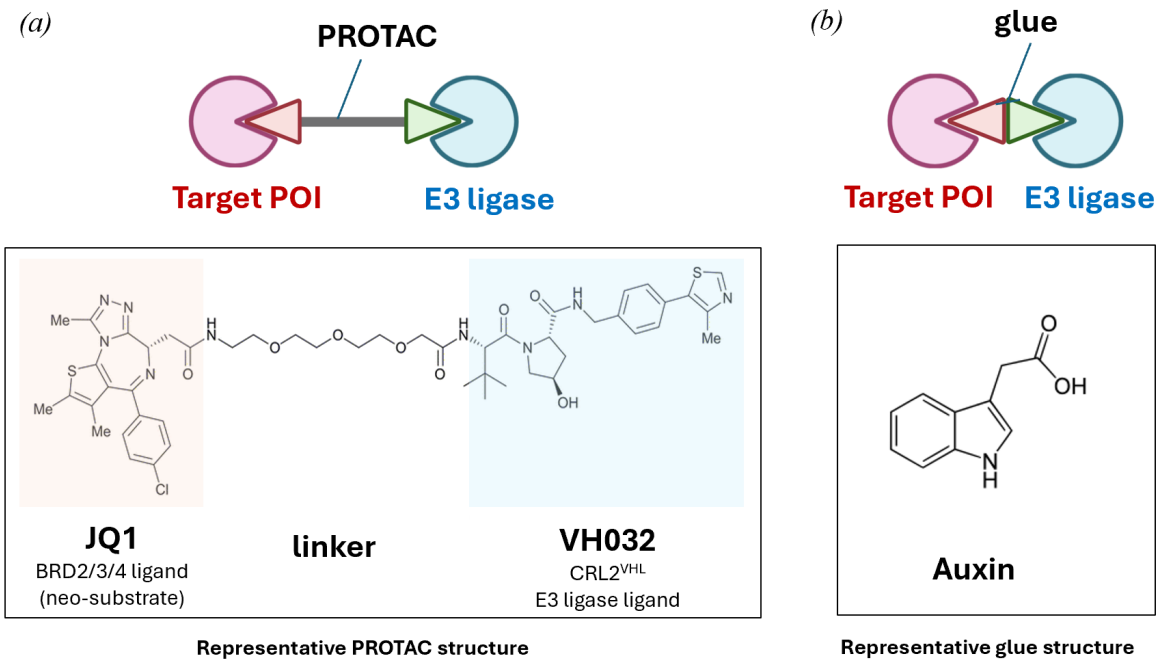


Figure 1.12: **Molecular Glues and PROTACs recruit neosubstrates to E3 ligases with different mechanisms.** (a) PROTACs are bivalent molecules that combine independent warheads against an E3 ligase and a target through a short linker. Representative structure of MZ1 PROTAC, recruiting BRD4 to $CRL2^{VHL}$ ligase [162]. (b) Molecular glues are small compounds that sandwich between an E3 ligase and a protein. Representative structure of a natural molecular glue, phytohormone auxin [78].

to have molecular glue properties was thalidomide; its binding to cereblon (CRBN) [169] causing it's teratogenicity [170]. Thalidomide and its analogues termed immune-modulatory imidide drugs (IMiDs) have been approved for use in multiple myeloma [171]. IMiDs mediate cereblon interactions to transcription factors of the Ikaros zinc finger family leading to their proteasomal degradation [172, 173].

These emerging modalities of protein modulation, differ from traditional occupancy-based agonism/antagonism. Critically, The E3-ligase bound PROTAC or glue can function catalytically, turning over multiple POI molecules even at sub-stoichiometric concentrations. In addition, the duration and effect of the incurred substrate degradation is not associated with the physical presence of the PROTAC/glue but depends on the relative rate of protein resynthesis and proteasomal degradation. In summary degraders operate via an event-driven rather than occupancy-driven pharmacology [174].

Table 1.3: Landscape of Chemical Probes for Kelch domains

Target Protein	**Ligand Discovery Method**	**Study**	**Potency/Efficacy**
KEAP1	Fragment-based drug discovery (FBDD) & pharmacophore screening	[175]	K_d = 2.5 nM (SPR) IC_{50} <15 nM (FP)
KEAP1	Fragment-based deconstruction & reconstruction (FBDR)	[176]	K_i = 0.04-0.5 μM (FP)
KEAP1	Gene expression studies	[177, 178]	Not Determined
KEAP1	Phenotypic screen (covalent inhibitor)	[179]	EC_{50} = 62 nM (translocation)
KEAP1	X-ray crystallographic screening & fragment merging (FBDD)	[180, 181]	K_d = 1.3 nM (ITC) EC_{50} = 12 nM (mRNA expression)
KLHDC2	Structure-based de novo design mimicking substrate degron	[182]	K_d = 95 nM (AlphaLisa)
KLHDC2	Thermal Shift Assay (TSA)	[183]	K_d = 16.9 μM (SPR) IC_{50} = 42 μM (TR-FRET)
KLHL20	Synthetic macrocycle compound, cyclised version of the linear degron	[184]	K_d = 600 nM (FP)

Table 1.4: Landscape of PROTACs for Kelch Substrate Receptors

E3 Ligase	**E3 Ligase Ligand Type**	**Target (Ligand)**	**Linker**	**Linkage Bond**	**Study**
KEAP1	small molecule	BRD4 (JQ1)	alkyl	amide, ether	[185]
KEAP1	small molecule	FAK (TAE684)	alkyl	alkyl amine	[185]
KEAP1	small molecule	BRD4 (JQ1)	glycine	amide, ether	[186]
KEAP1	covalent small molecule	BRD4 (JQ1)	alkyl piperazine	amide	[187]
KLHDC2	small molecule	BRD4 (JQ1)	biphenyl	amide, ether	[183]
KLHDC2	small molecule	BRD4 (JQ1)	alkyl acetamide	amide	[182]
KLHDC2	small molecule	CDK4 (YHJ1039)	alkyl	amide, ether	[188]
KLHL20	macrocycle	BRD4 (JQ1)	PEG	amide	[184]

1.5 Motivation & Objectives of this Thesis

1.5.1 Motivation

Of the over 700 human E3 ligases, only a fraction has been structurally characterised or chemically modulated. PROTAC development has largely focused on ligases with well-defined small-molecule binders, such as CRBN and VHL, limiting the therapeutic scope of targeted protein degradation. As PROTACs advance clinically, challenges like acquired resistance and tissue specificity highlight the need to diversify the E3 ligase toolkit [71]. This thesis aimed to expand the landscape of chemically tractable E3 ubiquitin ligases by structurally and mechanistically characterising two underexplored CRL complexes: $CUL3^{KLHL12}$ and $CUL2/4^{NRBP1}$. These ligases were selected for study as they fulfilled three criteria. They:

(a) **Form atypical assemblies hinting at diverse functional outputs.** KLHL12 dimerises via its BTB domain to form a homotypic CRL complex that can engage co-factors and substrates simultaneously, whereas NRBP1 is proposed to also dimerise to form a heterotypic CRL2/CRL4 assembly. Even though these assemblies have been inferred from cellular data, they have not been structurally resolved. The mechanistic rationale for substrate adaptor dimerisation and its functional consequences have not been fully elucidated.

(b) **Contain putative ligand-binding domains.** KLHL12 engages substrates via its solvent-exposed Kelch domain, which recognises a defined short linear motif. NRBP1, though a pseudokinase, is predicted to retain the canonical kinase fold and thus may contain a substrate-binding pocket amenable to small-molecule targeting.

(c) **Have functional links to pathophysiology.** These ligases are implicated in key cellular processes such as Wnt signaling, collagen trafficking, and stress responses, and are associated with diseases including cancer and neurodegeneration. Developing chemical probes or PROTACs against these E3 ligases would not only expand the toolbox for targeted protein degradation but also enable deeper investigation of their roles in disease in future studies.

1.5.2 Objectives

The chapters of the thesis are organised into two thematic pillars: structural studies of E3 ligase assemblies and development of chemical tools to manipulate ligase function, as outlined in Fig. 1.13. The specific objectives were:

In Chapter 3, to characterise the KLHL12 substrate-binding Kelch domain and its interactions with non-substrates: The KLHL12 Kelch domain engages diverse interactors, raising questions about the mechanism of its specificity. To address this, I examined antagonist recognition through X-ray crystallography and investigated binding determinants in comparison to previously resolved co-factor and substrate co-crystal structures. This work enabled the design of fragment-based ligand discovery assays.

In Chapter 4, to identify novel chemical small-molecule hits for KLHL12: The Kelch domain, a widely used substrate-recognition module in CRL complexes, remains largely unexplored for chemical modulation, as exemplified by the scarcity of selective Kelch-targeting chemical probes [Table 1.3] and functionalised PROTACs [Table 1.4]. In this chapter, I pursued ligand discovery efforts screening both covalent and non-covalent small-molecule libraries for binding to KLHL12 Kelch domain. Building on the findings of peptide binding in the Kelch substrate binding site, I employed *in silico* small-molecule docking, as well as *in vitro* high-throughput AlphaScreening displacement assays to identify competitive inhibitors on the KLHL12 Kelch binding site. A selection of non-covalent compounds, marked as putative KLHL12 binders after a DNA-encoded library (DEL) screen conducted by ZebiAI Therapeutics, Inc. and Google Research, were screened initially. In parallel, a curated set of covalent fragments from an in-house Enamine Ltd. collection, was also screened in collaboration with Ms Francesca d'Amico, Dr. Monique Mulder's group at Leiden University Medical Center (LUMC) - an EUbOPEN[2] partner. As an alternative to the immobilised screening format of the AlphaScreen assay, compounds were screened using a thermal shift assay (TSA) for in-solution

[2]Enabling and Unlocking Biology in the OPEN

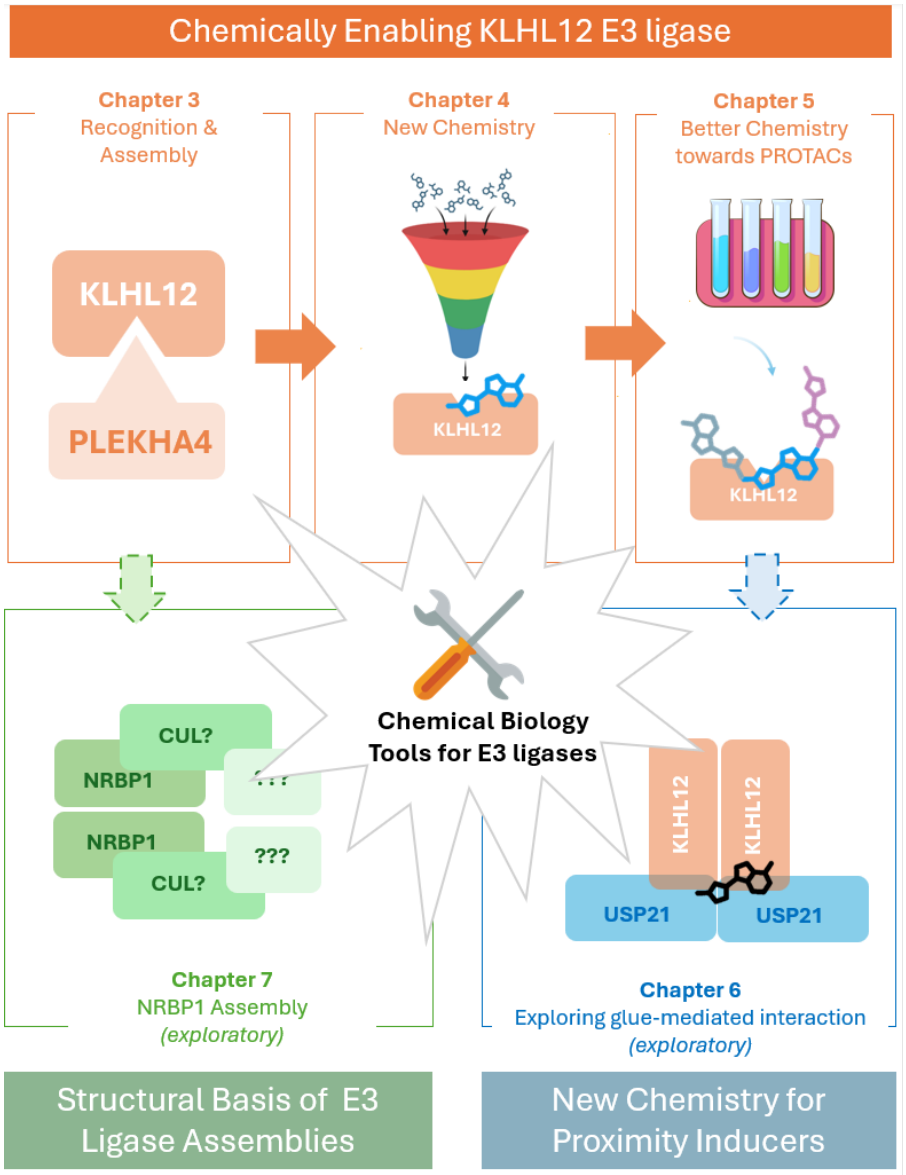


Figure 1.13: **Overview of Thesis Chapters.** The central focus of this thesis is the KLHL12 E3 ligase (**Chapters 3-5**), with exploratory investigations into related E3 systems (**Chapters 6-7**). Initial studies targeted the modular architecture of a putative NRBP1 E3 ligase (**Chapter 7**), though structural resolution proved intractable. Insights from NRBP1 assembly were applied to KLHL12 E3 ligase, whose substrate adaptor also homodimerises to form multimeric assemblies. **Chapter 3** characterised the physiological recognition of the antagonist PLEKHA4 by KLHL12, establishing a basis for ligand development. **Chapter 4** employed *in silico* and *in vitro* screening to identify small molecules targeting KLHL12. These were further optimised in **Chapter 5** with an outlook toward PROTAC development. In parallel with the screening efforts, another avenue for identifying novel chemistry for KLHL12 in **Chapter 6**. **Chapter 6** explored KLHL12's potential for mediating proximity interactions with USP21 through a molecular glue, expanding its utility beyond traditional substrate degradation. Therefore, **Chapters 3** and **7** are unified by their focus on E3 ligase assembly and substrate adaptor recognition; **Chapters 5** and **6** converge on novel small-molecule modalities for proximity induction.

detection of ligand binding. Orthogonal validation of non-covalent hits was performed using biophysical methods, including surface plasmon resonance (SPR) and grating-coupled interferometry (GCI). Covalent hits were validated by intact protein mass spectrometry.

In Chapter 5, to optimise the identified hit and assess it potential as a PROTAC handle for KLHL12: X-ray crystallography was used to determine the binding mode of the previously identified covalent ligand on the KLHL12 Kelch domain. This informed *in silico* techniques, including covalent docking and pharmacophore filtering, to screen for analogues and scaffold-hopping candidates from commercial (Enamine Ltd.) and in-house (LUMC) covalent libraries. Those commercial/in-house hits were assessed for selectivity against other Kelch domains using mass spectrometry analysis, binding affinity via the established AlphaScreen assay, and binding mode via crystallography to guide fragment elaboration strategies and PROTAC linker extension. These findings allowed for preliminary selectivity profiling and structure-activity relationship (SAR) analysis to be established and guided the design and organic synthesis of elaborated fragments as well as fragments primed for PROTAC derivatisation.

In Chapter 6, to investigate non-competitive modes of KLHL12 modulation: In parallel with competitive inhibitor development efforts outlined above, I explored a potential molecular glue-mediated interaction between KLHL12 and the deubiquitinase USP21, aiming to uncover alternative mechanisms for E3 ligase recruitment. This interaction was initially identified by Prof. Dalia Barsyte-Lovejoy's group (SGC Toronto) in a proteomic study but had not been validated *in vitro*. In collaboration with the Toronto team, I attempted to reconstitute the KLHL12-USP21 complex *in vitro*, though limited by the expression and stability of full-length proteins. I used microscale thermophoresis to validate the interaction and generated preliminary cryo-EM data in an effort to structural characterise the complex.

In Chapter 7, to assess chemical tractability of CRL2/CRL4^{NRBP1} ligase: Although this is the final chapter, it represents the first project I undertook, aiming to understand the assembly and mech-

anism of this heterotypic complex and to identify a potential binding site on NRBP1 - a pseudokinase thought to act as a substrate adaptor. Extensive construct design and protein expression efforts, as well as attempts at structural determination using X-ray crystallography, native mass spectrometry, and cryo-EM were employed to reconstitute and structurally resolve the CRL2/CRL4^{NRBP1} ligase. The *in vitro* work presented here was the first to my knowledge to validate NRBP1 interactions previously observed only *in cellulo*, laying the foundation and providing valuable reagents for future structural and functional studies.

Finally, **Chapter 8**, summarises the study's key findings, challenges, and implications for expanding the landscape of chemically targetable E3 ligases. Together, these studies aimed to (i) deepen our understanding of substrate recognition and complex assembly in poorly characterised E3 ligases, (ii) identify ligands suitable for chemical biology applications, and (iii) develop assay platforms to facilitate future targeting of novel ligases.

CHAPTER 2

Materials & Methods

2.1 Production of recombinant human proteins

Constructs for KLHL12 used in crystallisation studies in **Chapter 3** and biophysical measurements in **Chapters 4, 5** were already designed before starting my thesis from past colleagues in my group. ΔBTB-KLHL12-LgBit and ΔBTB-KLHL12-HiBit mammalian constructs in **Chapter 6** were a kind gift from Dr Dalia Barsyte's Group, University of Toronto. Other individual constructs required for specific applications are detailed in the results chapters as they become relevant.

New constructs for protein NRBP1 and interacting partners were designed for follow-up structural and functional studies. Construct design was guided by secondary structure prediction using bioinformatics tools such as Globplot (http://globplot.embl.de/) [189] and Psipred (http://bioinf.cs.ucl.ac.uk/psipred/) [190], as well as with molecular modelling tools such as AlphaFold [191] and Swiss Model [192]. The source of cDNA templates was Mammalian Genomics Collection (MGC), or mammalian constructs ΔBTB-KLHL12-LgBit and ΔBTB-KLHL12-HiBit constructs. The oligonucleotide (primers) were purchased from IDT™. All constructs were cloned using Ligation Independent Cloning as outlined below.

2.1.1 Ligation Independent Cloning.

Full-length and truncated constructs of the proteins of interest were designed and generated following the established Structural Genomics Consortium ligation-independent cloning (LIC) method [193].

The LIC method leverages the exonuclease activity of T4 DNA polymerase to generate sticky ends for ligase-independent annealing. This is achieved by incorporating a 12-16 base pair overhang into PCR primers, which is processed by T4 DNA polymerase to create complementary single-stranded regions between gene insert and linearised vectors for efficient cloning.

A summary of constructs for Sf9 and *E. coli* expression can be found in Tables 7.2, 7.1 and 7.3. Constructs were generated in 96-well format and cloned into *E.coli* and baculovirus expression vectors. Cloning and small scale test expression of NRBP1 *E.coli* constructs was performed by colleague Dong Wang.

PCR amplification. For a summary of oligonucleotides designed for bacterial NRBP1 constructs see Appendix Table A.1, for Sf9 NRBP1 constructs see Appendix Table A.2 and for constructs of interacting partners see Appendix Table A.3. Polymerase chain reactions (PCR) amplification reactions using Herculase II Fusion DNA polymerase (Agilent, 600675) and Q5 High-Fidelity DNA Polymerase (NEB, M0491S) were set-up in a plate-format in parallel under thermocycling conditions outlined in Fig. A.4, and Fig. A.5.

The PCR products were analysed on E-Gel precast agarose gels with SYBR safe gel stain (Invitrogen). For PCR optimisation, additives such as dimethyl sulfoxide (DMSO) or GC Enhancer (NEB) were added to facilitate primer-template annealing. In the case where the template construct and cloning vector shared the same antibiotic selection, the PCR product was further treated with DpnI that specifically digests the template DNA (which is typically methylated) while leaving the newly synthesized PCR product (which is unmethylated) intact. Finally, the successful PCR products were cleaned up using the mini-columns PCR purification kit (Qiagen, 28104). Representative agarose gel analysis of PCR purified amplicons showing NRBP1 baculo constructs and NRBP1 interactors can be found in Fig. A.6.

Annealing and transformation. In-house LIC-adapted vectors containing the sacB gene were linearised by restriction digest [Fig. A.7]. Linearised vectors and inserts treated by T4 DNA polymerase

before they were mixed in a 1:2 molar ratio and heat-shocked transformed into chemically competent cells (produced in-house using the RbCl method [194]) and later plated on 5% sucrose Luria Broth (LB) agar plates supplemented with the vector-appropriate antibiotic.

Colony PCR screening. Single colonies were picked from each plate and resuspended in RNase-free water. The colony resuspension was used as DNA template for PCR amplification by MyTaq™ Red DNA Polymerase (Bioline) in the presence of FBAC1[1] /FBAC2[2] oligonucleotides for insect constructs and pLIC-for[3] / pLIC-rev[4] or ACYCDuetUP1[5] /DuetDOWN1[6]oligonucleotides for bacterial constructs under thermocycling conditions outlined in Fig. A.8. The reactions were analysed in 1.5% agarose gel as summarised for NRBP1 constructs in Fig. A.10 (a) respectively. The successfully uptaken insert DNA of each construct was recovered from small-scale LB cultures that were inoculated with the same resuspended colonies and supplemented with the vector-appropriate antibiotic using the Promega Magnetic beads Plasmid Purification Kit.

2.1.2 Screening and protein production in insect cells

Transposition, bacmid production and PCR screening. Construct DNA was heat-shock transformed in chemically-competent *E. coli* DH10Bac™ (Invitrogen) cells, before it was plated on LB-agar recombinant bacmid selection plates, supplemented with appropriate antibiotics, 40 µg/mL isopropyl 1-thio-β-d-galactopyranoside (IPTG) and 100 µg/mL Blue-GalGlycosynth. The recombinant bacmid DNA was retrieved from small scale 2xLB medium cultures that had been inoculated with white single colonies. Samples of the bacmid DNA were used as templates for amplification by MyTaq™ Red DNA Polymerase (Bioline) [A.9] in the presence of FBAC1 and M13bac_reverse oligonucleotides[7]. The reactions were analysed in 1-2 % agarose gel to confirm amplicon of the expected mass [Fig. A.10 (b)].

[1]TATTCATACCGTCCCACCA
[2]GGGAGGTTTTTTAAAGCAAGTAAA
[3]TGTGAGCGGATAACAATTCC
[4]AGCAGCCAACTCAGCTTCC
[5]GGATCTCGACGCTCTCCCT
[6]GATTATGCGGCCGTGTACAA
[7][CAGGAAACAGCTATGAC]

Sf9 transfection, baculovirus generation, virus amplification and test expression. Recombinant bacmid DNA was transfected into Sf9 insect cells (SFM adapted, Invitrogen) using established methods [195]. The P0 baculovirus (BV) and consequently PI BV were harvested from culture supernatant while the cell pellets from the PI amplification were resuspended in Lysis Buffer[50 mM 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid (HEPES), pH 7.5, 300 mM NaCl, 5% (v/v) glycerol, 10 mM imidazole, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP), sterilised through a 0.2 µm membrane filter] supplemented with 1 µL/mL protease inhibitor cocktail and disrupted by sonication. After centrifugation of the cell lysate, the clarified supernatant was incubated with Ni^{+2}-sepharose beads. The protein was eluted by gradient elution with elution buffers of increasing imidazole concentration [50 mM HEPES pH 7.5, 500 mM NaCl, 5 % glycerol, 30 - 250 mM imidazole]. Eluted protein samples were analysed with Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in MES buffer (Sigma Aldrigh) [Fig. 7.3].

Screening and protein production in insect cells. Constructs were expressed in the eukaryotic baculovirus expression vector system (BEVS) following the established methods of the Structural Genomics Consortium [195]. All operations in parallel in a 96-well format.

2.1.3 Large-scale protein expression and purification

***E. coli* expression.** All proteins in bacterial vectors were expressed in rich media (TB or 2xYT) in BL21(DE3)-R3-pRARE2 cells. For 1 L expression, culture containing the appropriate antibiotic was inoculated with 5 mL overnight culture, and grown at 37°C with shaking at 180 RPM until OD600 reached 1-2 for TB and 0.6 for 2xYT. Induction with 0.4mM-1mM of IPTG and overnight incubation at 18°C with shaking followed. Cells were harvested by centrifugation at 4000 RPM in 4°C for 20min. Pellets were resuspended in binding buffer with 1/1000 SET III protease inhibitor cocktail (Millipore, 539134) and with 1 mM tris-(2-carboxyethyl)-phosphine (TCEP).

Baculoviral expression. Bacmid and P0-PI viruses were generated pas outlined above [see **Chapter 2.1.2**]. Viruses were further amplified from PI to PII in a 50 mL scale and the PII virus was used to

infect Sf9 insect cells at 2x10e^6 /mL. Sf9 cells were grown in Sf900-II SFM medium containing 2% (v/v) at 27°C with 100 RPM shaking for 48 to 72 hours. Cells were harvested by centrifugation at 900xg in 4°C for 20min. Pellets were resuspended in binding buffer with 1/1000 SET III protease inhibitor cocktail.

Mammalian expression. Expi293F™ cells (Life Technologies Limited, A14528) were grown to 2x10e^6 /mL cultured in FreeStyle 293 Expression medium (Thermo Fisher Inc, 2338-018) shaken at 100 rpm in a humidified incubator at 37°C with 8% CO_2. Per 1 L of media, the cells were supplemented with 10 mM sodium butyrate and transiently transfected with 1 mg of sterile DNA and Polyethylenimine (PEI) mixture in 20 mL Opti-MEM I Reduced-Serum Medium (Gibco, 1850613) - amount of PEI adjusted based on transfection efficiency and trials in test expression. The cells were incubated at 30°C or 37 °C for 72 or 48 hours, shaken at 170 rpm at 80% humidity and 8% CO_2. Cells were harvested by centrifugation at 900xg in 4°C for 20 min. Pellets were resuspended in binding buffer with 1/1000 SET III protease inhibitor cocktail. Cells were disrupted with lysis buffer (HEPES 30 mM, 150 mM NaCl, 1 mM TCEP, 5%glycerol, NP40 0.25% supplemented with protease inhibitors.

Expressing biotinylated protein. The AviTag is a short peptide sequence (GLNDIFEAQKIEWHE) that serves as a specific substrate for the BirA biotin ligase, which covalently attaches a single biotin molecule to a lysine residue within the tag [196]. In BirA-expressing cells, this enables site-specific biotinylation of AviTag-fused proteins during expression or in vitro, facilitating applications like affinity purification and protein labeling [197]. BL21(DE3)-R3-BirA cells made in-house were a kind gift by Dr Michael Fairhead. The following modifications to the established protocols above were employed to express biotinylated protein. BL21(DE3)-R3-BirA cells expressing the protein of interest bearing an AviTag were cultured in media supplemented by biotin solution (biotin 0.1 mM in 0.1 mM Bicine Buffer, pH = 8.35) just before induction, and before cell lysis by sonication. Complete biotinylation was confirmed with Q-TOF MS using an Agilent 6530 series accurate mass Q-TOF

LC/MS instrument.

2.1.4 Protein Purification

General protocols. All harvested pellets were resuspended in binding buffer (see below) supplemented with SET III protease inhibitor cocktail (see below) and disrupted by sonication (35% amplitude with pulsing 5 sec on and 10 sec off at 0°C), to yield whole-cell lysate (WCL). Addition of polyethylenimine (PEI) 0.15% followed and the samples were centrifuged at 35.000xg for 50min at 4°C to yield clear supernatant (SN). The supernatant was further applied through 0.4 µm filters. Unless stated otherwise, the purification buffers were the following:

Nickel Binding Buffer: 500 mM NaCl, 5% glycerol, 50 mM HEPES pH7.5, 5 mM imidazole, 0.5 mM TCEP, **Nickel Wash Buffer:** 500 mM NaCl, 5% glycerol, 50 mM HEPES pH7.5, 30 mM imidazole, 0.5 mM TCEP, **Nickel Elution Buffers:** 500 mM NaCl, 50 mM HEPES pH7.5, 5% glycerol, 50 - 250 mM imidazole, 0.5 mM TCEP, **Glutathione S-transferase (GST) Elution Buffer:** 300 mM NaCl, 50 mM HEPES pH7.5, 0.5 mM TCEP, 20 mM reduced glutatione (GSH), **Flag Lysis Buffer:** 50 mM HEPES pH7.5, 150mM NaCl, 5% Glycerol, 1mM TCEP, 0.25% NP-40, DNase I, **Flag Binding and Wash Buffer:** 50 mM HEPES pH7.5, 150mM NaCl, 5% Glycerol, 1mM TCEP, **Flag Elution:** 50 mM HEPES pH7.5, 150mM NaCl, 5% Glycerol, 1mM TCEP, supplemented with 100 µg/mL, **Size exclusion chromatography buffer:** 300 mM NaCl, 50 mM HEPES pH7.5, 0.5 mM TCEP

For interaction studies - such as the NRBP1-DDB1 co-expression or KLHL12-USP21 complex preparation for cryo-EM grids - modified buffers of the above were used at decreased salt concentrations of 150 mM NaCl. After purification, protein samples were analysed in SDS-PAGE gels and the masses compared to standard protein ladders. Protein purification methods outlined below:

Immobilized metal ion affinity chromatography (IMAC). The clarified supernatant of proteins

with *polyhistidine* tags was applied to a pre-equilibrated Ni^{+2}-sepharose gravity-flow column. The column was washed with 10 column volume (CV) binding buffer, followed by 10 CV wash buffer. Bound protein was then eluted with 2 CV elution buffers containing gradient concentrations of imidazole ranging from 50-250 mM.

Glutathione S-transferase (GST) affinity chromatography. The clarified supernatant of proteins with GST tags was applied to a pre-equilibrated Glutathione Sepharose resin gravity-flow column with lysis buffer. The mixture was incubated for 2 h at 4 °C with gentle rolling before being loaded onto a gravity column, washed with 10 CV wash buffer. Elution buffer (10 CV) containing 20 mM freshly prepared reduced glutathione was applied to the resin and the mixture was incubated for 30 min at 4 °C with gentle rolling. The resin was then applied to a gravity flow column and the elution was collected.

FLAG Purification. The clarified supernatant of proteins bearing the FLAG tag (DYKDDDDK) was applied to pre-equilibrated anti-FLAG M2 resin with lysis buffer. The mixture was incubated for 1 h at 4°C with gentle rolling before being loaded onto a gravity column, washed with 20 CV wash buffer, and eluted in 2 CV elution buffer containing 100 µg/mL 3xFLAG peptide.

Size exclusion chromatography (SEC). Protein samples were concentrated using an Amicon Ultra centrifugal filter with an appropriate molecular weight cut-off to up to 5 mL, and loaded onto a pre-equilibrated size-exclusion chromatography column (sephadex 200 HiLoad 16/60 column or sephadex 75 HiLoad 16/60 column) and ran at 1 mL/min flowrate on the AKTA express or BioRad system at 4°C. Fractions were collected and protein was detected by UV absorbance.

Ion-exchange chromatography (IEX). Proteins were purified using an anion-exchange chromatography approach with a Mono Q 5/50 GL (Cytiva) or HiTrap Q SP (Cytiva) columns on an Biorad FPLC system. The column was equilibrated with Buffer A (50 mM Tris pH = 7.4, 2.5% glycerol) at a flow rate of 1 mL/min. The protein sample was loaded onto the column. Unbound proteins were washed out with 10 column volumes (CVs) of Buffer A. Bound proteins were eluted using a linear

gradient of 0-1 M NaCl in Buffer B (50 mM Tris pH = 7.4, 2.5% glycerol, 1 M NaCl) over 20 CVs. Elution was monitored by UV absorbance at 280 nm.

Protein handling. Between each purification stage, 0.5 mM TCEP was added to prevent protein oxidation. Proteins were either stored at 4 °C or on ice during purification. Intact Mass Spectrometry was performed for quality control before assays or crystallisation trials.

Tag cleaving. All constructs used in this thesis bear a tobacco etch virus (TEV) cleavage site EN-LYFQ*(G/S) between the tag and the target. If desired, TEV protease was added to cleave the tag by overnight incubation at 4 °C at a protease to target protein ratio of 1:100 (w/w).

2.2 Co-purification assays

NRBP1-DDB1 Sf9 co-expression. Full length human DDB1 [Uniprot Q16531, 2-1140 a.a.] and human NRBP1 [Uniprot Q9UHY1, 41-535 a.a.] were subcloned into the baculovirus vectors pFB-GST3C-LIC and pFB-NHStIIT-LIC respectively. Sf9 cells at $2x10^6$ cells/mL in Sf-900®II SFM medium were infected with PII DDB1 and NRBP1 BV and incubated at 27 °C, shaking at 100 rpm for 72 hours. Cells (100 mL) were harvested by centrifugation at 1000xg for 25 min and the cell pellet was resuspended in Binding Buffer (50 mM HEPES, pH 7.5, 100 mM NaCl, 5% (v/v) glycerol, 10 mM imidazole, 0.5 mM TCEP, sterilised through a 0.2 µm membrane filter) supplemented with 1 µL/mL protease inhibitor cocktail and disrupted by sonication. The clarified supernatant was split in two: *(i)* one half was incubated with Ni^{+2}-sepharose beads at 18 °C for 1 hour, with end-over-end rotation. The protein was eluted by gradient elution with elution buffers of increasing imidazole concentration [50 mM HEPES pH 7.5, 100 mM NaCl, 5 % Glycerol, 30 - 250 mM Imidazole]. *(ii)* the other half was incubated with GST-sepharose beads with end-over-end mixing for 2 hours. The column was washed with binding buffer, before incubated with elution buffer (50 mM HEPES, pH 7.5, 100 mM NaCl, 5% glycerol) containing 20 mM reduced glutathione.

NRBP1-ElonginBC, TSC22D2/TSC22D4 pull-down experiments. 200 µg of previously purified GST-tagged human TSC22D2 [Uniprot O75157, 146-180 a.a.] and TSC22D4 [Uniprot Q9Y3Q8,

56-90 a.a.] in 50 mM HEPES, pH 7.5, 300 mM NaCl, 5% glycerol were each applied to 25 µL of equilibrated glutathione-sepharose beads. Each reaction was incubated for 2 hours at 4 °C with end-over-end mixing to allow for the GST-TSC22D2/D4 proteins to bind to the resin. The resin was washed with wash buffer (10x bed volume) to remove unbound protein. Equimolar quantity of untagged purified human NRBP1 [Uniprot Q9UHY1, 59-409 a.a.] in complex with ElonginB and ElonginC were added to the mixture and further incubated for 2 hours at 4 °C with end-over-end mixing. The resin was washed with wash buffer, before incubated with elution buffer (50 mM HEPES, pH 7.5, 100 mM NaCl, 5% glycerol) containing 20 mM reduced glutathione. All washing and elution steps were performed by centrifugation at 4 °C, 500xg. Wash buffer consisted of 50 mM HEPES pH 7.5, 300 mM NaCl, 0.5 mM TCEP, 5% glycerol.

NRBP1-ElonginBC, DBB1 pull-down experiments. Full length human DDB1 [Uniprot Q16531, 2-1140 a.a.] was subcloned into the baculovirus vector pFB-GST3C-LIC. Sf9 cells at $2x10^6$ cells/mL in Sf-900®II SFM medium were infected with PII DDB1 BV and incubated at 27 °C, shaking at 100 rpm for 72 hours. Cells (100 mL) were harvested by centrifugation at 1000xg for 25 min and the cell pellet was resuspended in Binding Buffer (50 mM HEPES, pH 7.5, 100 mM NaCl, 5% (v/v) glycerol, 10 mM imidazole, 0.5 mM TCEP, sterilised through a 0.2 µm membrane filter) supplemented with 1 µL/mL protease inhibitor cocktail and disrupted by sonication. 1 mL of clarified lysate was applied to 25 µL of equilibrated glutathione-sepharose beads. The procedure outlined for TSC22D2/D4 pull-down was followed from this point onwards.

NRBP1-ElonginBC, CUL2 pull-down experiments. 100 µg purified FLAG-tagged N-terminal human CUL2 [Uniprot Q13617, 1-384 a.a.] and 100 µg purified untagged human NRBP1-ElonginBC [Uniprot Q9UHY1, 59-409 a.a.] in complex with ElonginB and ElonginC were mixed and incubated together with 50 µL Anti-FLAG M2 Beads (Sigma Aldrigh) for 3 hours at 4 °C, with end-over-end mixing. The resin was washed with wash buffer (3x bed volume) to remove unbound protein, and bead sample was boiled in 1x lameli sample buffer. All washing and elution steps were performed by

centrifugation at 4 °C, 300xg. Wash buffer consisted of 50 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 10% glycerol, Tween20 0.05%.

Western blotting. Western blotting is an analytical technique used for detecting specific proteins in a sample of tissue homogenate or extract. Proteins were first denatured in sample buffer with boiling for 5 minutes, followed by separation via SDS-PAGE analysis. Proteins were then transferred from the gel to polyvinylidene difluoride (PVDF) membrane in transfer buffer (10% tris-glycine stock, 10% methanol). After transfer, the rest of the available protein binding sites on the membrane were blocked with incubating in 2.5% Bovin Serum Albumin (BSA) in phosphate-buffered saline (PBS) supplemented with 0.1% Tween for 1 hr, shaking at room temperature (PBST). The membrane was then probed with specific antibodies either directly conjugated with Horseradish Peroxidase (HRP) or after a two step incubation with primary and then secondary-HRP conjugates. HRP catalyses the oxidation of luminol to 3-aminophthalate, which results in emission of low-intensity light at 428 nm. The membranes were washed with PBST to remove unbound antibody, before being incubated with Enhanced chemiluminescence substrates (Thermo fisher, Pierce, 32209) briefly, and imaged by LAS4000 camera with ChemiDoc Imaging System (Bio-Rad).

Antibodies: GST primary rabbit mAb, Cell Signaling Technology, 91G1; anti-rabbit secondary goat immunoglobulins Horseradish Peroxidase, Dako, P0448; PentaHis HRP Conjugated antibody, QIAGEN, 34460; Flag antibody, Sigma-Aldrich, F1804; GAPDH antibody, Thermo Fisher, MA5-15738

2.3 *In silico* methods

Chemical Clustering. Chemical clustering performed in ICM-Pro, using fingerprints to represent molecules in order to calculate pairwise similarities, and UPGMA (Unweighted Pair Group Method with Arithmetic Mean) as the clustering method.

Library Preparation. Compounds were prepared prior to docking by Schrödinger's LigPrep tool of Maestro (Schrödinger Release 2021-3: LigPrep, Schrödinger, LLC, New York, NY, 2021); OPLS_2005 force-field was applied, structures were optimised, hydrogens were added, pKa, protonation states and

tautomers were predicted using Epik [198] at pH 7.4 $\pm$ 1 and finally, up to 32 possible conformations were generated.

Small molecule docking. The structure was extracted from the PDB and prepared using the "Protein Preparation Wizard" tool of the Schrödinger suite (Schrödinger Release 2021-3: Protein Preparation Wizard; Epik, Schrödinger, LLC, New York, NY, 2021); water molecules were removed, hydrogen atoms were added, bond orders were assigned, before hydrogen bond optimization, and restrained energy minimization using OPLS_2005 force field. The prepared protein was then used for grid generation using the "Receptor Grid Generation" panel of the Glide module of the Schrödinger suite. The protein's binding site was defined by the ligand, and the grid was generated based on the centroid of the binding site's residues. Docking was performed by Glide (Schrödinger Release 2021-3: Glide, Schrödinger, LLC, New York, NY, 2021) [199]. Docking was performed using the prepared library of conformations and the Virtual Screening Workflow protocol of Glide. Molecular docking was performed at two different levels of precision, standard (SP) and the more accurate extra precision (XP) modes using default parameters. Co-crystallised cognate ligand/peptide was docked as well and treated in the same manner as described above, as a benchmark.

Small molecule covalent docking. Covalent docking was performed using ICM-Pro (Molsoft) [200, 201]. The KLHL12 protein from the co-structure with compound **EVD-1** served as the receptor. The protein was prepared by optimizing protonation states, assigning charges, optimizing the hydrogen bond network, and removing solvent molecules. Residues surrounding the covalent ligand were defined as the receptor using default settings. **EVD-1** ligand was removed, and residue C511 was designated for covalent modification via a nitrile reaction to form a thioimidate adduct. Both commercially available fragment libraries and libraries of our collaborators in LUMC, from Dr. Dalia Barsyte's group were filtered for the presence of a nitrile substructure. The Enamine commercial libraries screened were *CLOUD*, *Covalent Fragment Library*, *Covalent Screening Library*, *Covalent Serine Hydrolase Library*, *Cysteine-focused covalent fragments*, *Electrophilic covalent probe library*,

and *FDA approved drugs collection*. Conformer libraries, pre-processed with Schrödinger's LigPrep, were imported into ICM-Pro and docked using the "Virtual Ligand Screening (VLS)" batch docking method. Rigid docking was applied, as conformations were precomputed. Conformers with docking scores above -15 were discarded, and the best-scoring pose for each compound was selected after filtering in KNIME. Docked poses were exported to Schrödinger Maestro, and those replicating **EVD-1** hydrogen bonds with KLHL12 H486 and Y512 were identified as hits. These were taken forward for LC-MS analysis to assess covalent labeling of KLHL12.

2.4 Structure determination of protein macromolecules

2.4.1 X-ray crystallography

Crystallisation. Proteins were concentrated to 8∼13 mg/mL using a 10 kDa centrifugal concentrator (Millipore). For co-crystallisation, peptides or compounds in DMSO were added to the protein solutions and incubated for 1 hr before crystallisation trials. Crystallisation was performed using sitting drop vapour diffusion and commercially acquired (Molecular Dimensions) sparse matrix coarse screens as modified by the Structural Genomics Consortium (SGC) were employed; namely modified versions of the Hampton Research Crystal Screen (HCS), the Hampton Research Index (HIN) screen, the Joint Centre for Structural Genomics (JCSG) screen, the Ligand Friendly Screen (LFS), Morpheus, the Basic Chemical Space screen (BCS), the Midas+ screen, the Stura Footprint and finally, the poly-γ-glutamic acid polymer screen (PGA). Where needed, fine screens were designed sampling conditions around the conditions of the initial crystals obtained from the coarse screens. Initial crystals obtained from coarse screens were used to make seed stocks to help initiate nucleation [202] of different working dilutions in reservoir solution.

The protein-peptide or protein-compounds mixtures were incubated for 1 hr at 4°C, prior to setting up sitting-drop vapour-diffusion crystallization plates using a mosquito HTS Nanolitre Liquid Handler (150 nL drops, protein: mother liquor ratio of 2:1, 1:1 and 1:2). Vitrification in liquid nitrogen followed the crystal cryoprotection with direct addition of reservoir solution supplemented with 25%

ethylene glycol. Diffraction data were collected on beamlines I04, I02 in Diamond Light Source, Didcot, UK. Diffraction data were indexed and integrated using MOSFLM [203], and scaled and merged using AIMLESS [204]. Molecular replacement was performed with Phaser MR [205] in Phenix using KLHL12 apo structure (PDB: 2VPJ) as the model. COOT was used for PLEKHA4 building and manual refinement [206], whereas PHENIX.REFINE [207] or Buster [208] was used for automated refinement. Refined models were validated by MolProbity [209]. Structure figures were made with PyMOL [81].

2.4.2 Cryo-electron microscopy

Cryo-electron microscopy (cryo-EM) grids were prepared by applying 3.5 µL of purified protein samples in 50 mM HEPES, 150 mM NaCl, 0.5 mM TCEP buffer at 0.5 - 3.5 mg/mL onto Quantifoil R1.2/1.3 300-mesh copper or Quantifoil Au R1.2/1.3 200-mesh grids (Quantifoil Micro Tools GmbH) that were freshly glow-discharged for 60s. The grids were then vitrified using a FEI Vitrobot Mark IV (Thermo Fisher Scientific) at 4 °C and 100% humidity, sampling varying blotting times 1-7s and blot force 0-1 (see **Chapters 6.12, 7.9**).

NRBP1 initial screening for condition optimisation; Initial NRBP1-ElonginBC grids [Table 7.6] were assessed on a Talos Arctica microscope (Thermo Fisher Scientific) operating at 200keV. A small screening dataset was collected using automated data acquisition EPU software (Thermo Fisher Scientific) on NRBP1-u010 on a Falcon 4i detector at 150kv nominal magnification (pixel size = 0.94 Å/pix). A total electron dose of 50 $e^-/Å^2$ was applied to collect 181 micrographs, with a total exposure time of 5.33 seconds (1 shot per hole). Defocus values ranged from -3.3 to -1.8 µm, rising with a 0.3 step.

NRBP1 screening of optimised grids; Further NRBP1-ElonginBC grids [Table 7.7] were assessed on a Glacios microscope (Thermo Fisher Scientific) operating at 200keV with a spherical aberration coefficient of 2.7 µM, equipped with a Falcon 3 detector at 120kv nominal magnification (pixel size = 1.2 Å/pix). Automated data acquisition was performed using EPU software (Thermo Fisher Sci-

entific). A total electron dose of 50 $e^-/Å^2$ (26.21 e/px/s , 17.67 e/Å2/sec) was applied to collect 205 micrographs (4 shots per hole). Defocus values ranged from -3.3 to -1.5 µm, rising with a 0.3 step.

NRBP1-ElonginBC full collection; Grid NRBP1-u014 were assessed with EPU on a Krios IV instrument (Thermo Fisher Scientific) at eBIC (Harwell, UK), operating at 300keV, on a K3 detector at 105kv nominal magnification (pixel size = 0.831 Å/pix). Automated data acquisition was performed using EPU software (Thermo Fisher Scientific). A total electron dose of 60 $e^-/Å^2$ was applied to collect 11573 super-resolution micrographs (3 shots per hole). Defocus values ranged from -2.6 to -1.0 µm, rising with a 0.2 step. The micrographs were motion-corrected using Patch Motion Correction in CryoSPARC 3.3.1 (Structura Biotechnology) or MotionCorr within Diamond's automated pipeline. Defocus values were estimated using the Patch CTF function in CryoSPARC or CTFFind4 in Relion (Diamond pipeline). Initial particle picking was performed using the Blob Picker function in CryoSPARC or CRYOLO in Diamond's pipeline, followed by multiple rounds of 2D classification. Polished particles were then used to train a Topaz-based particle-picking model [210], which was subsequently applied for particle extraction. Further 2D classifications were performed on the Topaz-picked particles. 2D templates from neither method generated meaningful *ab initio* models in CryoSPARC. For details of the processing workflow, see Fig. 7.10.

KLHL12-USP12 initial screening for condition optimisation; Purified KLHL12-HiBit-LgBit was incubated with purified USP21 in presence and absence of BAY-805 (100 µM, 1% DMSO for 1 hour at 0 °C. Different molecular ratios and concentrations of the components were sampled [**Chapter 6.2.3**]. Grids were assessed on a Glacios microscope (Thermo Fisher Scientific) operating at 200keV with a spherical aberration coefficient of 2.7 µM, equipped with a Falcon 3 detector at 120kv nominal magnification (pixel size = 1.2 Å/pix). No particles could be detected, thus no collection followed.

2.5 Methods to probe molecular interactions

2.5.1 Surface Plasmon Resonance

Surface Plasmon Resonance (SPR) detects biomolecular interactions by shining plane-polarized light onto a metal surface (chip). Binding events on the surface cause a shift in the reflected light's intensity, which is directly related to changes in the refractive index, providing real-time measurement of molecular interactions. All experiments assessing small molecules' affinity for KLHL12 Kelch domain were performed using a Biacore S200 (GE Healthcare) instrument.

Analysis and sample compartment temperature were maintained at 25 °C. The reference flow cells contained no immobilized protein. Purified biotinylated KLHL12 Kelch domain [278-567 a.a.], alongside negative controls comprising the Kelch domains of KLHL20 [303-602 a.a.], KLHL14 [319-628 a.a.], and KLHL6 [302-610 a.a.], were captured via the biotin tag on a Cytiva Series S Sensor Chip SA in running buffer 50 mM HEPES-NaOH, 150 mM NaCl, 0.05%Tween-20, 0.5 mM TCEP, 2 % DMSOm pH 7.5. Capture levels ranging between 7,000 - 10,000 RU were obtained. DMSO effects were corrected by preparing a range of standards (1.5% - 2.5%). The chip was washed with 50 mM NaOH in 1 M NaCl and between analyte cycles with 1 M NaCl. DVL1 20-mer peptide (LifeTein Inc) at 200 µM was used as a positive control at 20 µM every 30 cycles. Compounds were initially tested at a single concentration (200 µM) and subsequently analyzed using 12-point dose-response curves (0-200 µM) in an assay containing 2% DMSO. The association time was set to 200 s, while dissociation was monitored for 300 s, with a flow rate of 30 µL/min.

The data were reference-subtracted, corrected for DMSO effects, and analyzed using Biacore S200 Evaluation Software. In the case of dose-response curves, one-site steady-state affinity analysis was employed by fitting the equation:

$$R = \frac{R_{\text{max}}}{1 + \frac{K_d}{C}} + RI$$

Where R_{eq}: Binding signal at equilibrium, R_{max}: Binding signal at saturation, maximum binding response, K_d: Dissociation constant, C: Analyte concentration, RI: bulk refractive index contribution.

Our collaborators in Karolinska institute used an adapted protocol for analysing purified single-chain variable fragments (scFv) interaction with biotinylated KLHL12 Kelch domain. Biotinylated KLHL12 Kelch domain was immobilised, while purified scFvs were diluted to 0.6 µg/ml in running buffer (PBS, 0.01% Tween-20) and included in a kinetic measurement using SPR-16 biosensor instrument (Sierra, Bruker). Determination of the affinity constat was done as outlined above.

2.5.2 Thermal Shift Assay

Thermal Shift Assay (TSA) measures the protein's melting temperature (Tm), which reflects the thermal stability of the folded protein. A ligand that binds specifically and stabilizes the protein usually increases the Tm, resulting in a positive thermal shift (ΔTm>0), while a ligand that binds through non-specific hydrophobic interactions may incur aggregation and thus destabilisation (ΔTm<0).

Melting curves of Kelch domain KLHL12 [278-567 a.a.] with and without the presence of compounds were determined by TSA using the Sypro Orange dye (Life Technologies), a Stratagene Mx3005P RT-PCR apparatus (Agilent Technologies, Waldbronn, Germany), and white 96-well PCR-plates. Reactions were set up at 20 µL total volume and 2.5% DMSO concentration. KLHL12 Kelch domain at 4 µM protein, was mixed with compounds dissolved in DMSO at final concentration of 200 µM, positive control peptide DVL1 (AYTVVGGPPGGPPVR), or negative control DMSO, buffered in 50 mM HEPES (pH 7.5), 150 mM NaCl and left to incubate at room temperature for 15 min. Addition of SYPRO Orange dye (Invitrogen) followed to a final concentration of 5x. Fluorescence was monitored using the FAM channel while samples were heated from 25°C to 95°C in 70 cycles at a ramp rate of 1°C/21s. The melting temperature (Tm) was determined by fitting the fluorescence intensity data to a Boltzmann sigmoidal curve using Agilent MxPro software. All experiments were performed in three technical replicates. The sigmoidal plot of the normalized fluorescence intensity values versus temperature was fitted to the Boltzmann equation:

$$Y = Bottom + \frac{Top - Bottom}{\exp\left(\frac{Tm - X}{Slope}\right) + 1} \tag{2.1}$$

Where, **Y:** fluorescence intensity values, **Bottom:** signal corresponding to the denatured state of the protein, **Top:** signal corresponding to the native state of the protein, **X:** temperature (independent variable, in °C), **Slope:** Tm melting temperature, where the protein is 50% denatured.

2.5.3 Grating-coupled interferometry

Grating-Coupled Interferometry (GCI) is an optical biosensing technique that measures biomolecular interactions by detecting phase shifts in *guided light*. Light is coupled into a waveguide through a diffraction grating, and binding events at the sensor surface alter the refractive index, causing measurable phase changes. This label-free method provides high sensitivity and real-time kinetic analysis of molecular interactions. Experiments were performed using the Creoptix WAVEcore instrument, coupled with a temperature control autosampler at 25 °C. Purified biotinylated Kelch domains of KLHL12 [278-567 a.a.] and negative control KLHL6 [302-610 a.a.][8] were immobilised on WAVEchip Prep 4PCH-STA streaptavidin sensor chip for 600s to reach 5.500 RU. The chip was conditioned with borate prior to use. No protein was immobilised on one of the cells as a negative control. Assay buffer was 50 mM HEPES-NaOH pH = 7.5, 150 mM NaCl, 0.05% Tween20, 0.5 mM TCEP, 2 % DMSO. Biotinylated DVL1 peptide (Biotin-GGTTKAYTVVGGPPGGPPVRE) was used as a measure of KLHL12's health every 25 cycles. Initially, compounds were injected over the sensor chip at the single concentration of 200 µM. Subsequent dose-response experiments were performed at 8 concentrations 250, 125, 62.5, 31.25, 15.63, 7.81, 3.91, 1.95 µM. The compounds were allowed to associate for 5s before their dissociation for 20s at a flow rate of 400 µL/min. Data were analysed using WAVEcontrol software; measurements were DMSO corrected and reference subtracted. Kinetics

[8] KLHL6 protein was a gift from Dr Zhuoyao Chen

were fitted with waveRAPID evaluation software using the 1:1 Kinetic equation fit.

2.5.4 Bio-layer Interferometry

Biolayer Interferometry (BLI) is an optical biosensing technique that measures biomolecular interactions in real time by detecting changes in the interference pattern of *reflected light*. When molecules bind to a biosensor tip, the optical thickness of the layer changes, leading to a shift in the interference pattern that is directly proportional to the binding response. Analyses were performed in black 96-well plates (Greinerbio, Germany) on an OctetRED384 instrument (ForteBio, USA). Prior to use, biosensors were soaked in assay buffer (20 µM in 25mM HEPES, pH = 7.5, 100 mM NaCl, 0.05% Tween20 buffer). Biotinylated RFXV containing peptides KCC2 [Uniprot Q9H2X9, 1-32 a.a., LifeTein LLC][9], TSC22D2 and TSC22D4 were immobilised on eight Octet®Super Streptavidin (SSA) Biosensors each. Biocytin (Sigma) was used on an equal number of reference sensors to quench free streptavidin (reference sensors). Analyte proteins NRBP1 [Uniprot Q9UHY1, 59-409 a.a.] and NRBP1 [Uniprot Q9UHY1, 330-530 a.a.] were serially diluted into the assay buffer[10], while positive control protein SPAK [Uniprot Q9UEW8, 446-545 a.a.] measured in a single concentration. BLI assays all performed in the assay buffer as follows: Biotinylated-peptide-loaded sensors were dipped into buffer (initial baseline, 60s), then on analyte (association, 180s) and finally back to buffer (dissociation, 240s). Biocytin-loaded sensors were used as a control of unspecific binding and were subtracted from experimental values.

Data processing using ForteBio Data analysis 9.0 included parallel reference sensor subtraction, baseline aligned by Y-axis, inter-step correction aligned to disassociation and finally Savitzky-Golay Filtering. The dissociation constant K_D was determined by fitting the global response to a 1:1 Langmuir binding model $R = R_{\text{max}} \cdot \frac{\text{Conc.}}{K_D + \text{Conc.}}$. The association rate constant (k_{on}) and dissociation rate constant (k_{dis}) were determined by globally fitting the BLI response data across all concentrations to the kinetic binding model $R(t) = R_{\text{max}} \cdot (1 - e^{-k_{\text{on}} \cdot C \cdot t}) \cdot e^{-k_{\text{dis}} \cdot t}$, where C is the analyte concentration

[9] Biotin-MSRRFTVTSLPPAGPA

[10] from 90, 30, 10, 3.3, 1.1, 0.37, 0.12, 0 µM

and t is time. The global fit was performed by minimizing the residuals between the experimental response and the model using non-linear least squares fitting. The equilibrium association constant (K_a) was calculated as $K_a = k_{\text{on}}/k_{\text{dis}}$, where K_a is the inverse of the dissociation constant (K_D), describing the binding affinity of the interaction.

2.5.5 Microscale Thermophoresis

Microscale Thermophoresis experiments were performed in a Monolith NT.115 instrument (NanoTemper Technologies GmbH). Purified recombinant USP21 catalytic domain [Uniprot Q9UK80, 210-561 a.a.] and ΔBTB-KLHL12-HiBit-LgBit complex [Uniprot Q53G59, 268-567 a.a.] were labelled using a Monolith RED-NHS 2$^{\text{nd}}$ generation Protein Labelling Kit (cat MD-L011) following manufacturer's instructions. DVL1 20mer peptide (GGTTKAYTVVGGPPGGPPVRE) was used as a positive control for KLHL12 binding, while compound BAY-805 dissolved in DMSO was used a positive control for USP21 binding. KLHL12 alone or KLHL12 and BAY-805 (1 μM) combined were added to 100 nM labelled USP21 buffered in 20 mM HEPES, 150 mM NaCl, 5% glycerol, 0.05% Tween 20, 0.5 mM TCEP. The reactions were incubated at room temperature for 15 min and then transferred into premium MST capillaries. Measurements were conducted at 25 °C with 20% excitation power and 40% (medium) MST power. Data were subsequently analyzed using the Monolith Evaluation software using the implemented K_d model (1:1 Langmuir binding model).

2.5.6 AlphaScreen

Amplified Luminescent Proximity Homogeneous Assay (AlphaScreen) is a homogenous assay that allows for high-throughput probing of macromolecular interactions in a competitive binding configuration for discovery of new inhibitors. The screen uses chemically-coated donor and acceptor beads for the generation of luminous signal when the beads are in proximity. Upon excitation by light at 680 nm, the donor bead which is coated with a photosensitiser converts molecular oxygen (O_2) into a singlet oxygen ($^1O_2$), which is transferred to an acceptor bead if there is one within 200 nm, resulting

in light emission at 520 nm - 620 nm. Thus by binding two different molecules of interest to the acceptor and donor beads, their interaction can be assessed by monitoring of the chemiluminescent signal intensity.

The known interaction between KLHL12 and the native substrate peptide DVL1 was exploited for discovering KLHL12 inhibitors. The assays were performed with hexahistidine tagged KLHL12 Kelch domain [268-567 a.a.] and biotinylated DVL1 peptide (Biotin-GGTTKAYTVVGGPPGGPPVRE). Prior to AlphaScreen assay, the optimal concentration of biotinylated DVL1 and histidine tagged protein was ascertained to ensure adequate signal, at 50 nM and 100 nM respectively [Fig. 4.4 (a)].

Firstly, compounds were dispensed into 384-well reaction plates in the Access Laboratory Workstation using an Echo-555 Acoustic Liquid Handler (Labcyte USA). 10 µL of protein and biotinylated peptide in the *appropriate concentrations (see below)* were added to every well of the reaction plates, and incubated with compounds for the *appropriate time (see below)*. AlphaScreen Histidine (Nickel Chelate) Detection Kit, 6760619M (Perkin Elmer, USA) donor and acceptor beads were mixed and prepared at 1:1000 dilution in the assay buffer in dark. 10 µL of beads mixture were added to each well in the reaction plates using combi liquid handler (Thermo fisher) in the dark. The reaction was incubated for a further 1 hour at room temperature in dark. Signals were read using PHERAstar FSX Plate reader (BMG Labtech, Germany) with the following settings: excitation at 680 nm, reading emission at 520-620 nm. Non-biotinylated DVL1 peptide (AYTVVGGPPGGPPVR) was used as a positive control at 200µM. Assay buffer was 25 mM HEPES-NaOH, 100 mM NaCl, 0.05% CHAPS, 0.1%BSA, pH7.5. Data analysis was performed in Graphpad Prism. The %DVL1 displacement was normalised between averages of negative (DMSO, 0% displacement) and positive (DVL1 peptide, 100% inhibition) controls. A non-linear curve that allowed for a variable slope was fitted on logarithmically transformed compound concentrations over normalised response. The concentration where 50% of DVL1 binding is inhibited (IC_{50}) and Hill Slope of each curve were fitted according to equation:

$$Y = Bottom + \frac{Top - Bottom}{10^{HillSlope \cdot (LogIC_{50} - X)} + 1} \tag{2.2}$$

X: log of dose or concentration, **Y:** Response **Top and Bottom:** Plateaus in same units as Y **logIC_{50}:** same log units as X, **Hill Slope**: Slope factor or Hill slope, unitless

In more detail:

Single concentration experiment of selected DEL compounds. Preliminary signal analysis determined optimal concentrations of hexahistidine-tagged KLHL12 and biotinylated DVL1 to be 75 nM for each. Compounds were dispensed in duplicate at 100 µM as final assay concentration. Assay DMSO concentration was 0.2 %. The protein, peptide and compounds were incubated for 1 hour before addition of the beads and further incubation of 1 hour.

IC_{50} determination of selected DEL compounds. Preliminary signal analysis determined optimal concentrations of hexahistidine-tagged KLHL12 and biotinylated DVL1 to be 100 nM and 50 nM respectively. Compounds were dispensed in duplicate at 100 µM as final assay concentration with sequential 1:1 dilutions to generate 11-point dose-response curves. Assay DMSO concentration was 0.2 %. The protein, peptide and compounds were incubated for 1 hour before addition of the beads and further incubation of 1 hour.

IC_{50} determination of selected covalent hits. Signal analysis determined optimal concentrations of hexahistidine-tagged KLHL12 and biotinylated DVL1 to be 150 nM for each. A preliminary experiment to determine concentration ranges was performed where each compound was dispensed in duplicate at 100 µM maximum concentration with sequential 1:1 dilutions to generate 11-point dose-response curves. Assay DMSO concentration was 1%. After this initial experiment, the dose-response curves of observed binders were repeated in triplicates, as 16-point curves. The maximum compound concentration was adjusted to 25 or 200 µM for strong and weaker binders respectively.

The assay DMSO concentration was at 2 % for weaker binders and 0.25 % for strong binders. The protein, peptide and compounds were incubated for 2 hours and 23 hours before addition of the beads and further incubation of 1 hour.

2.6 Mass Spectrometry (MS)

All results were analyzed using the Qualitative Analysis MassHunter Acquisition Data and the Bioconfirm software. Methods of individual instruments are outlined below:

2.6.1 Electrospray Ionization Time of Flight (ESI-TOF)

Reversed-phase chromatography was performed in-line prior to mass spectrometry using an Agilent 1100 HPLC system (Agilent Technologies inc. - Palo Alto, CA, USA). Concentrated protein samples were diluted to 0.02 mg/ml in 0.1% formic acid and 50 µl was injected on to a 2.1 mm x 12.5 mm Zorbax 5um 300SB-C3 guard column housed in a column oven set at 40°C. The solvent system used consisted of 0.1% formic acid in ultra-high purity water (Millipore) (solvent A) and 0.1% formic acid in methanol (LC-MS grade, Chromasolve) (solvent B). Chromatography was performed as follows: Initial conditions were 90% A and 10% B and a flow rate of 1.0 ml/min. After 15 seconds at 10% B, a two-stage linear gradient from 10% B to 80% B was applied, over 45 seconds and then from 80% B to 95% B over 3 seconds. Elution then proceeded isocratically at 95% B for 1 minute 12 seconds followed by equilibration at initial conditions for a further 45 seconds. Protein intact mass was determined using an MSD-ToF electrospray ionisation orthogonal time-of-flight mass spectrometer (Agilent Technologies Inc. - Palo Alto, CA, USA). The instrument was configured with the standard ESI source and operated in positive ion mode. The ion source was operated with the capillary voltage at 4000 V, nebuliser pressure at 60 psig, drying gas at 350°C and drying gas flow rate at 12 L/min. The instrument ion optic voltages were as follows: fragmentor 250 V, skimmer 60 V and octopole RF 250 V. For samples analysed under native conditions, proteins were buffer-exchanged by overnight dialysis into 50 mM ammonium acetate. Samples were introduced via a Gastight syringe (200 L) with a blunt needle, connected to 20 m ID PEEK tubing and a fill port, and analysed as described in

[211].

2.6.2 In-gel digestion MS/MS analysis

The identity of proteins in gel bands excised following SDS-PAGE was confirmed by tryptic digestion and tandem MS (LC-MS/MS). SDS-PAGE gel bands were excised as 1 x 4 mm slices using a gel cutting tip (GeneCatcher, Web Scientific) and stored in 10% MeOH at 4°C. Prior to digestion, the methanol solution was removed and replaced with 100% acetonitrile for 2-5 minutes. The solution was then removed and replaced with 100 µL of 100 mM NH_4HCO_3 (pH 8.0). 1 µL of 1M dithiothreitol was added and incubated at 56°C for 40 minutes. Four microlitres of 1M iodoacetamide was then added and the reaction incubated at ambient temperature in the dark for 20 minutes. A further 1 µL of 1 M dithiothrietol, 200 µL of 100mM NH_4HCO_3 and 1 µL of trypsin solution (sequencing grade, Sigma cat 1 mg/ml in 0.01 M HCl) was then added. Tryptic digestion proceeded at 37°C for 16 hours and was terminated by addition of 3 µL of formic acid. LC-MS/MS was performed using a Dionex U3000 nano HPLC coupled to a Bruker Esquire HCT ion trap mass spectrometer. 1-5 µL of tryptic digest was loaded on to a 200 µm ID x 5 cm PS-DVB monolith column (Pepswift, Dionex Corp.) A linear gradient of 0% B to 15% B was developed over 5 minutes, followed by a second linear gradient from 15% B to 40% B over 2 minutes. The column was washed at 90% for 2 minutes and then equilibrated at 90% B for a further 6 minutes. Solvent A was 2% (v/v) acetonitrile, 0.1% formic acid in water, solvent B was 80% acetonitrile 0.1% formic acid. The flow rate was 2.5 µL/min. The mass spectrometer was operated in positive ion, standard enhanced mode with a scan rate of 8,100 m/z/sec and a scan range of 250 - 1800 m/z. The trap accumulation time was 200 msec and the accumulation target was 200,000 counts. Data dependent peptide fragmentation was performed in Auto MS/MS mode. Compound extraction and peptide deconvolution was performed using the DA data analysis program (Bruker Daltonik). Database searching was performed using the Mascot 2.2.04 search algorithm (MatrixScience) with the following search parameters: Charge states +2, +3; MS tolerance 1.5Da; MSMS tolerance 0.5 Da; UniProt SwissProt database without taxonomic restrictions.

2.6.3 High pressure liquid chromatography (HPLC) MS

Denaturing intact mass spectrometry measurements were performed using an Agilent Infinity II LC system connected to an Agilent 6530B quadrupole-time-of-flight (QTOF) mass spectrometer with a Jetstream electrospray source. An Agilent Poroshell StableBond 300 C3, 2.1 x 12.5 mm column, 5 µm particle size, was used. Protein samples, quantity from 0.5 to 10 µg, were diluted in 0.1% formic acid and between 2 and 20 µL was injected onto the column in a running buffer (solvent A) of 10% methanol, 0.1% formic acid at 1.25 mL/min flow rate. Samples were eluted from the column using a linear gradient to 65% methanol and 20% acetonitrile over 0.6 minutes. The mass spectrometer was in positive ion mode and source parameters were: drying gas temperature 325 °C, flow rate 12 L/min, nebulizer 60 psi, sheath gas temperature 350 °C, flow rate 13 L/min. Mass spectrometry data was acquired between 100 and 3000 m/z.

2.6.4 RapidFire MS

RapidFire-MS analysis was performed using a 1290 Infinity coupled to a 6530 Accurate-Mass Q-TOF LC/MS (Agilent Technologies). The mass instrument is the same as above in *(HPLC) MS*. Samples were loaded directly onto a C4 solid-phase extraction (SPE) cartridge. The cartridge was washed with aqueous buffer (0.1% formic acid in water, 5500 ms, 1.25 mL/min), followed by elution with 85% acetonitrile in 0.1% formic acid (5500 ms, 1.25 mL/min). Eluted samples were analyzed by MS in positive ion mode.

2.7 Chemical Synthesis

2.7.1 General Reagent Information

Commercial materials were used as received (analytical or HPLC grade) without prior purification unless otherwise noted. Water was deionised. All reactions involving moisture and air sensitive reagents were carried out under a nitrogen atmosphere. Room temperature (rt) refers to ambient temperature. Temperatures of 0 °C were maintained using an ice-water bath. Compound structures

and names were generated using ChemDraw v20.1

2.7.2 Chromatography

Reactions were monitored by thin layer chromatography (TLC) and liquid chromatography-mass spectrometry (LC-MS).

TLC. Commercially prepared aluminium plates coated with 60 F254 silica gel were used and visualised using combination UV light source at 365nm/254nm.

LC-MS. Liquid chromatography was performed with Waters System Fluidics Organizer instrument with 515 HPLC pump and Waters Binary Gradient 2545 device, with a Kinetex 5µm EVO C18 100, 100 x 3.0 mm lc column attached. Linear gradient elution using 5-95% of solvent A [93 % Water, 5 % MeCN, 2 % of 0.5 M NH_4OAc pH 6.0)] in solvent B (18% Water, 80% MeCN, 2% of 0.5 M NH_4OAc pH 6.0) was applied, eluting at a flow rate of 2 mL/min. Compound formation and purity was assessed by UV absorbance at 254 nm (Waters UV/Visible Detector 2489), ELSD signal (Waters ELS Detector 2424) or ESI+ TIC (SQ Detector 2).

Flash column chromatography. Pre-packed Biotage Sfär silica columns were used in a Biotage Selekt automated flash chromatography system. The samples were dry-loaded; adsorbed on the silica after solvent evaporation.

2.7.3 Compound Characterisation

NMR spectra. Compounds were dissolved in the appropriate deuterated solvent and the spectra were recorded using a Bruker Avance 400 MHz spectrometer at 298 K. The field was locked by external referencing to the relevant deuteron resonance. Chemical shifts (d) are quoted in parts per million (ppm) and referenced to the residual solvent peak. Proton shifts are quoted to the nearest 0.01 ppm and carbon shifts to the nearest 0.1 ppm. When peak multiplicities are reported, the following abbreviations are used: br = broadened, s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, dd = doublet of doublets, dt = doublet of triplets, dq = doublet of quartets, ddd = doublet of doublet

of doublets, tt = triplet of triplets and qd = quartet of doublets. Coupling constants (J) are quoted in Hz and rounded to the nearest 0.1 Hz. Assignments of protons and carbons were supported by two dimensional NMR experiments (COSY, HSQC, HMBC) or by analogy.

Low resolution mass spectrometry. Recorded by LC-MS on a Water SQ Detector 2 or on an ACQUITY QDa II detector. Data acquisition and processing were performed using Waters FractionLynx software.

High resolution mass spectrometry. Measurements were performed using an Agilent Infinity II LC system connected to an Agilent 6530B quadrupole-time-of-flight (QTOF) mass spectrometer with a Jetstream electrospray source. An Agilent Poroshell StableBond 300 C18 column, 2.1 x 50 mm, 5 µm particle size, was used. Compounds were prepared in 0.1% formic acid/50% acetonitrile/100% acetonitrile at a concentration of 0.5 - 50 ng/µL and 20 - 50 µL were injected onto the column at 0.4 ml/min flow rate in solvent A (5% acetonitrile, 0.1% formic acid). Compounds were eluted with a gradient to 100% solvent B (85% acetonitrile, 0.1% formic acid) over 5 minutes. Mass spectrometry data was acquired between 50 and 1000 m/z and analysed using Agilent MassHunter Qualitative Analysis software.

Glutathione Reactivity Assay.

To assess compound reactivity, compounds at 100 µM were incubated at room temperature with 1 mM glutathione (GSH) in 50 mM HEPES, 20 mM NaCl, pH 7.5, 1.0% DMSO. Compounds were dispensed in different wells of a 96-well plate and each well was analysed with low resolution LC-MS after appropriate incubation time. The percentage of compound remaining, over GSH-adduct, was quantified using Waters FractionLynx software based on peak area. The natural logarithm of the results were fitted to a linear regression, and half-life was calculated as $t_{1/2}$ = ln(2)/-slope.

CHAPTER 3

Structural insights into KLHL12 interactor binding

3.1 Introduction

Kelch-like protein 12 (KLHL12) is the substrate-specific adapter of a BCR (BTB-CUL3-RBX1) E3 ubiquitin ligase complex; KLHL12 recruits cullin 3 through its N-terminal BTB domain and engages the substrate or other target-specific co-adaptors through the C-terminal Kelch β-propeller repeat domain, as previously discussed in **Chapter 1.2**. Previous work in our lab led by Chen *et al.* mapped the binding site of KLHL12 aforementioned known interactors to a conserved "PGXPP" motif [Table 1.1] which is essential for both their recruitment and degradation [72]. In the same study the linear peptide degron motif of DVL1 was co-crystallised with KLHL12 Kelch domain, revealing a binding mode in conflict with the binding mode of a DVL3-derived peptide reported in a contemporaneous study by Zhao *et al.* [114].

Even though the "PGXPP" consensus motif can be found in all known KLHL12 direct interacting partners, making it likely that co-factors and antagonists would adopt similar binding modes as those reported for DVL substrates, there was previously no solved structure of the Kelch domain of KLHL12 in complex with a co-factor or antagonist to verify that hypothesis. Moreover, since the crystallised DVL3 peptide [114] was N-terminally truncated and missing the first out of the four prolines compared to the DVL1 crystallised peptide [72], it was unclear if the two alternative binding modes were preferred by different subgroups of binding partners or the former was an artefact as a consequence of the truncation [Fig. 1.9].

In this chapter, I employed X-ray crystallography to assess other "PGXPP" peptides for their binding to KLHL12. I determined the crystal structure of KLHL12 Kelch domain in complex with the recognition motif of antagonist PLEKHA4 at 1.95 Å resolution. The structure revealed a peptide binding mode comparable to that previously reported for DVL1 [72] and also to the concurrently resolved PEF1 complex, suggesting that all other known interactors of KLHL12 with the same proline-rich sequence would also adopt a similar binding mode. The finding posits interesting questions on how KLHL12 allows for both substrate and co-factor binding in parallel, at the same binding site.

3.2 Results

3.2.1 Structure determination of KLHL12-interactor complexes

Analysis of the crystal contacts of apo KLHL12, which had been previously crystallised by our group [106], revealed that the substrate binding site is occluded. More specifically, blades I and V of the single KLHL12 molecule in the ASU (coloured yellow) form hydrogen bonds and hydrophobic interactions with blades I, V and VI of its symmetry molecule (coloured orange) [Fig. 3.1]. The residues involved in these crystal-packing contacts, namely Q293, Y434, Y528 and L533, are also involved in the substrate binding of DVL1 [Fig. 1.9 (c)]. Since the binding site is not accessible, this crystal form is not suitable for peptide soaking. Therefore, I focused my efforts on identifying new conditions for co-crystallisation of KLHL12 and peptide interactors.

Moreover, previous efforts in our lab to co-crystallise KLHL12 with a peptide from human PEF1 were successful. Colleague Ashley Ramdass prepared and applied Kelch domain samples to a few coarse screens after incubation with the co-factor peptide PEF1 [UniProt Q9UBV8, 15-25 a.a., GQAPGAPPGSY]. Well-diffracting crystals were identified from a coarse screen reservoir solution (0.1 M Bis-Tris pH 5.5, 200 mM NaCl, and 30% PEG4K), yielding a 1.88 Å structure of the KLHL12-PEF1 complex [Table 3.2]. Since no information was known on how antagonist binding on KLHL12 Kelch domain compared to substrate and co-factor binding, I proceeded to co-crystallise KLHL12 with an antagonist peptide.

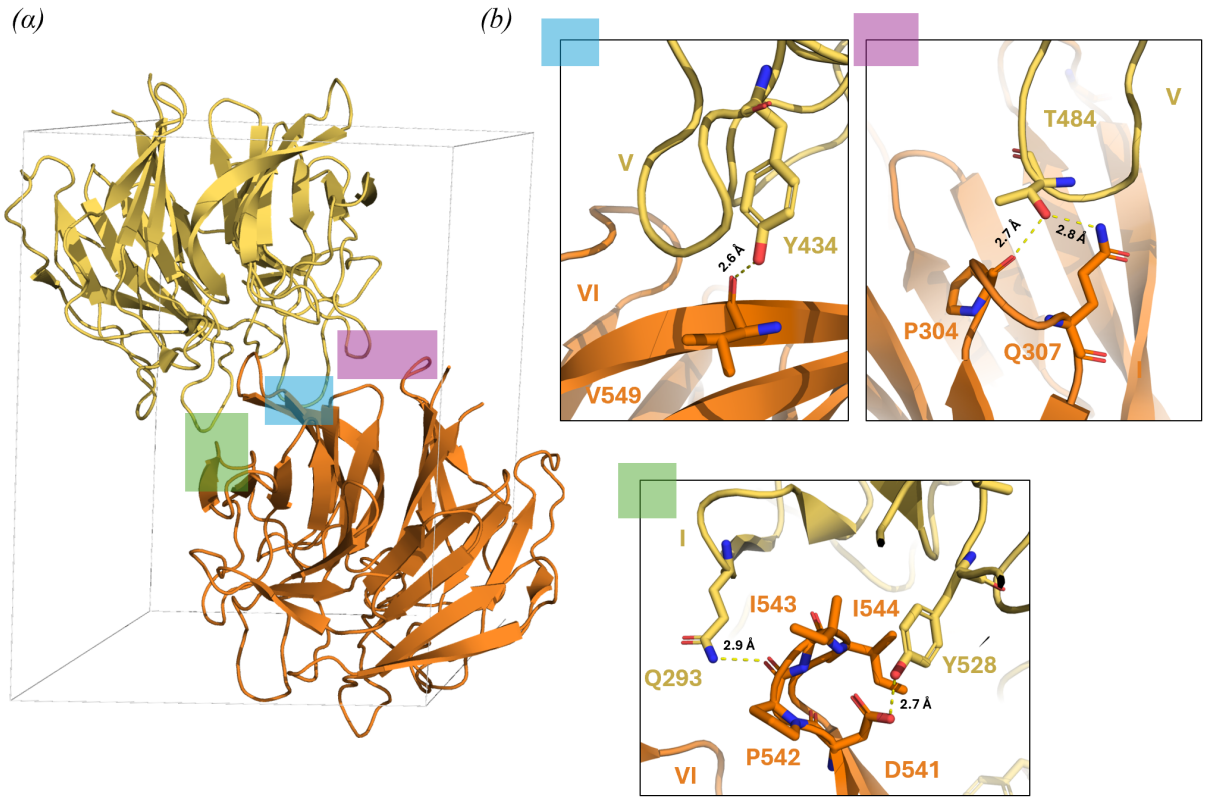


Figure 3.1: **Crystal packing of apo KLHL12** (a) Ribbon model based on apo structure of KLHL12 (PDB: 2VPJ). The crystal has one molecule in the asymmetric unit (ASU) which is coloured yellow. Its symmetry molecule is coloured orange. The unit cell is traced with light gray lines. Contacts are marked with green, blue, and purple boxes. (b) Zoom-in on the crystal contacts following the colour scheme in (a). The Kelch repeat blades are numbered in roman numerals. Dashed lines represent hydrogen bonds. From the crystal packing it is evident that the binding site is occluded and this crystal form is not suitable for ligand soaking.

To start, the KLHL12 Kelch domain [268-567 a.a.] was expressed in *E. coli* and purified using Ni^{2+}-sepharose resin. The N-terminal hexahistidine tag was cleaved with TEV protease, and the protein was further purified by size exclusion chromatography. The final sample, concentrated to 9.8 mg/mL, was incubated with the antagonist peptide PLEKHA4 [UniProt Q9H4M7, 174-184 a.a., PGGPG-GPPEVS], containing the previously identified PGXPP binding motif [72]. Initial co-crystallisation trials of KLHL12 in complex with PLEKHA4, across various coarse screens and temperatures, did not yield crystals. To refine the approach, I analyzed the crystal hits from KLHL12-PEF1 and designed focused fine screens sampling the chemical space around the conditions of interest. The fine screens were also cross-seeded with diluted micro-seed stocks from KLHL12-PEF1 crystals at various dilutions [Table 3.1]. Fine screen trials were set up using the vapour diffusion method with sitting drops. Crystals of various morphologies were obtained at 20 °C in precipitant solutions containing variable molecular weight PEGs and salts, such as NaCl or potassium thiocyanate (KSCN), in acidic pH [Fig. 3.2].

Diffraction data of selected crystals were then obtained by unattended data collection on the I04 beamline at Diamond Light Source, Didcot, UK. All diffracting crystals of the KLHL12-PLEKHA4 complex originated from a reservoir solution containing 0.1 M citrate pH 5.5, 0.4 M NaCl, 27.1% PEG8K [Fig. 3.2]. After initial processing, the best diffraction dataset was taken forward for further structure refinement as no significant pathologies, other than moderate anisotropy were identified. Anisotropy was accounted for prior to molecular replacement using the automated autoPROC and STARANISO Diamond pre-processing pipeline. Molecular replacement was performed by Phenix.phaser-MR using the apo KLHL12 structure [PDB: 2VPJ] as a search model. The structure was refined at 1.95 Å resolution with Phenix.refine. The refinement strategy consisted of grouped B factor restrained refinement in the early cycles, while Translation-Libration-Screw-rotation (TLS), non-crystallographic symmetry (NCS) and occupancy refinement were incorporated in the later stages. Crystallographic data collection and refinement statistics of the deposited structure are shown in Table 3.2.

KLHL12-PLEKHA4 and KLHL12-PEF1 co-crystals share the same space group of P 1 2_1 1, with the first having four complexes in the ASU (average RMSD of 0.348 Å over 312 C_α atoms), and the latter having (RMSD of 0.914 Å over 288 C_α atoms) [Fig. 3.3]. All the expected binding sites were occupied by the respective peptides in both crystals [Fig. 3.9 (a), (b)].

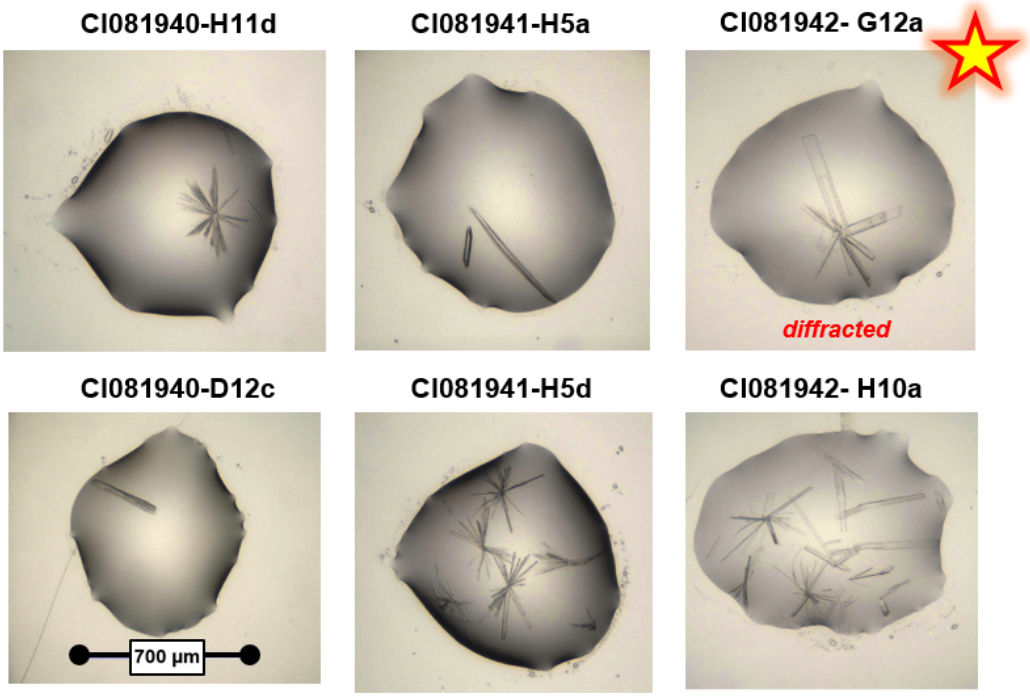
Table 3.1: Summary of KLHL12-PLEKHA4 fine screen designs

Coarse Screen	Well	Well Conditions	Designed Fine Screen	Fine Screen Identifier
PEF1 JCSG7, 20 °C, seeded with PLEKHA4 HSC3-H8c	**C1**	20% PEG8k, 0.2M NaCl, 0.1M citrate pH 4.2	PEG8K 10% - 30%, NaCl 0.05M-0.4M, citrate pH 3-5.5	ED-JCSG7-C1
PEF1 JCSG7, 20 °C, seeded with PLEKHA4 HSC3-H8c	**A2**	20% PEG3k, 0.1M citrate pH 5.5	PEG3k 10%-30%, citrate pH 4-6.2	ED-JCSG7-A2
PEF1 JCSG7, 20 °C, seeded with PLEKHA4 HSC3-H8c	**G9**	30% PEG2k MME, 0.1M KSCN	PEG2k MME 20% - 40%, KSCN 0.05M-0.3M	ED-JCSG7-G9
PLEKHA4 JCSG7, 20 °C, seeded with PLEKHA4 HSC3-H8c	**H8**	25% PEG3350, 0.2M NaCl, 0.1M bis-tris pH 5.5	PEG3350 15% - 35%, NaCl 0.05M-0.4M, bis-tris pH pH 5.5	ED-JCSG7-H8

3.2.2 Structural observations on binding mode of KLHL12 interactors

Even though peptides were occupying all available binding sites in the ASU of both the PLEKHA4 and PEF1 complexes, the peptide sequences were only resolved around the "PGXPP" key interacting motif. For cofactor PEF1 peptide, residues 15-GQAPGAPP-22 were modelled. The PEF1 peptide adopts a U-shaped turn conformation that is stabilised by an intramolecular hydrogen bond between the carbonyl of P18 and A20 amide, in accordance with findings for DVL1 [72]. PEF1 peptide adopts two distinct conformations between the two chains in the ASU [Fig. 3.5 (b), (c)]. Residue Q16 which is solvent facing, either extends out of the pocket (chain C) or points into it (chain D), a fact suggesting this alternate conformation being a truncation artifact. Chain D also lacked resolvable density for G15. The side chain of A17 is rotated nearly 180 °C to that in chain C [Fig. 3.5 (b), (c)].

In comparison, only PLEKHA4 residues 177-PGGPP-181 could be modelled in three of the peptide chains in the ASU, while electron density for P181 was not observed in the fourth chain [Fig. 3.4 (b)]. For two of the molecules in the ASU, there is also non-connecting electron density proximal to



Identifier	Fine Screen	Peptide	Temp (°C)	Seeds Source
CI081940	ED-JCSG7-G9	PLEKHA4	20	PEF1 JCSG-H08c 1:10
CI081941	ED-JCSG7-A2	PLEKHA4	20	PEF1 JCSG-H08c 1:10
CI081942	ED-JCSG7-C1	PLEKHA4	20	PEF1 JCSG-H08c 1:10
CI081777	ED-JCSG7-H8	PLEKHA4	20	PEF1 JCSG-H08c 1:500

Figure 3.2: **Optimisation of KLHL12-PLEKHA4 co-crystallization using microseeding.** Representative KLHL12-PLEKHA4 crystals (top panel) from co-crystallisation fine screens (below).Seed stock was prepared by crushing KLHL12-PEF1 crystals and diluting 1:10 or 1:500 (v/v) in seed buffer 0.1 M Bis-Tris pH 5.5, 200 mM NaCl, and 30% PEG4K. Initial KLHL12-PEF1 crystals were observed in well H08C of coarse screen JCSG, in reservoir solution 0.1 M Bis-Tris pH 5.5, 200 mM NaCl, and 30% PEG4K. The scale is shown relative to average drop diameter which is 0.7 mm. The crystal yielding diffraction is marked with a yellow star.

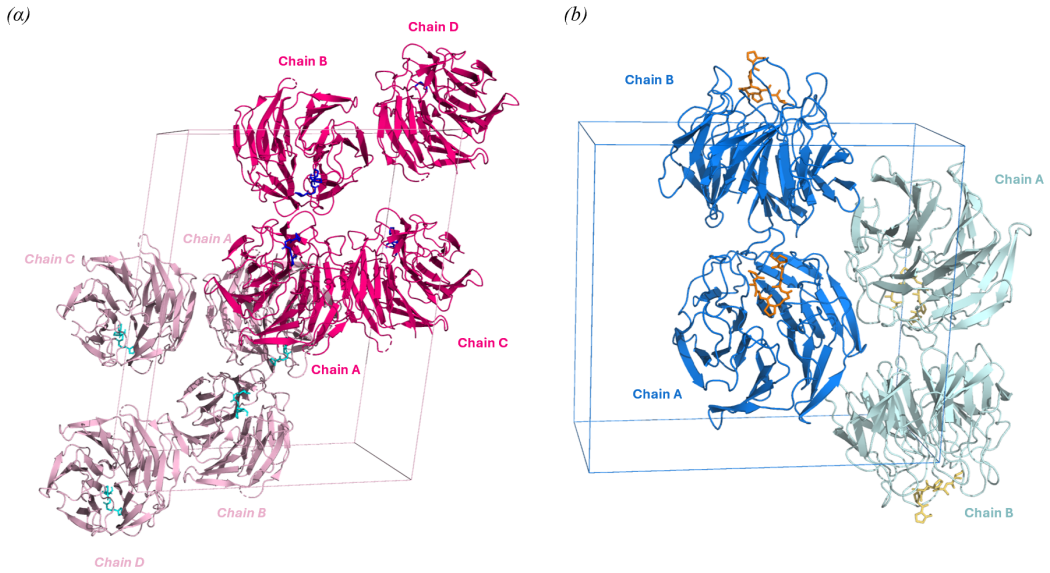


Figure 3.3: **Asymmetric Units of KLHL12-Interacting Peptide Co-Crystals** (a) Unit cell of the KLHL12/PLEKHA4 co-crystal, with a single asymmetric unit (ASU) highlighted in magenta and the symmetry-related ASU in pink. PLEKHA4 peptides are shown in blue and cyan. (b) Unit cell of the KLHL12/PEF1 co-crystal, with a single ASU highlighted in blue and the symmetry-related ASU in cyan. PEF1 peptides are shown in orange and yellow.

the N-terminal P177. Even though the remaining peptide residues 175-GG-176 could be modelled convincingly [Fig. 3.4 (c)], in the previously reported structure of KLHL12 Kelch domain in complex with DVL1 [72] that density's region was occupied by solvent. Due to this ambiguity, the density was left unmodeled. The structure suggests that the unresolved residues N-terminal (174-PGG-176) and C-terminal (182-EVS-184) to the key interacting motif of PLEKHA4 are flexible and/or do not make strong stabilising contacts with KLHL12 [Fig. 3.4 (b)]. Since the successfully resolved residues overlaid well with the other peptide structures, it is possible that PLEKHA4 also adopts a U-turn conformation as PEF1 and DVL1, held in place with a similar intramolecular bond.

Both peptides seem to adopt a common binding mode, comparable to reported DVL1, whereby the "PGXPP" motif accounts for the majority of the resolved residues and is responsible for most of the interactions with all 6 blades of KLHL12. In all three structures, the peptides are tethered to the shallow pocket by a hydrogen bond between the phenolic hydroxyl group of Y512, and the carbonyl of the first proline in the PGXPP motif (P18, P177, and P658 for PLEKHA4, PEF1, and DVL1 respectively). Peptides PLEKHA4 and PEFl are also binding with two further hydrogen bonds not observed

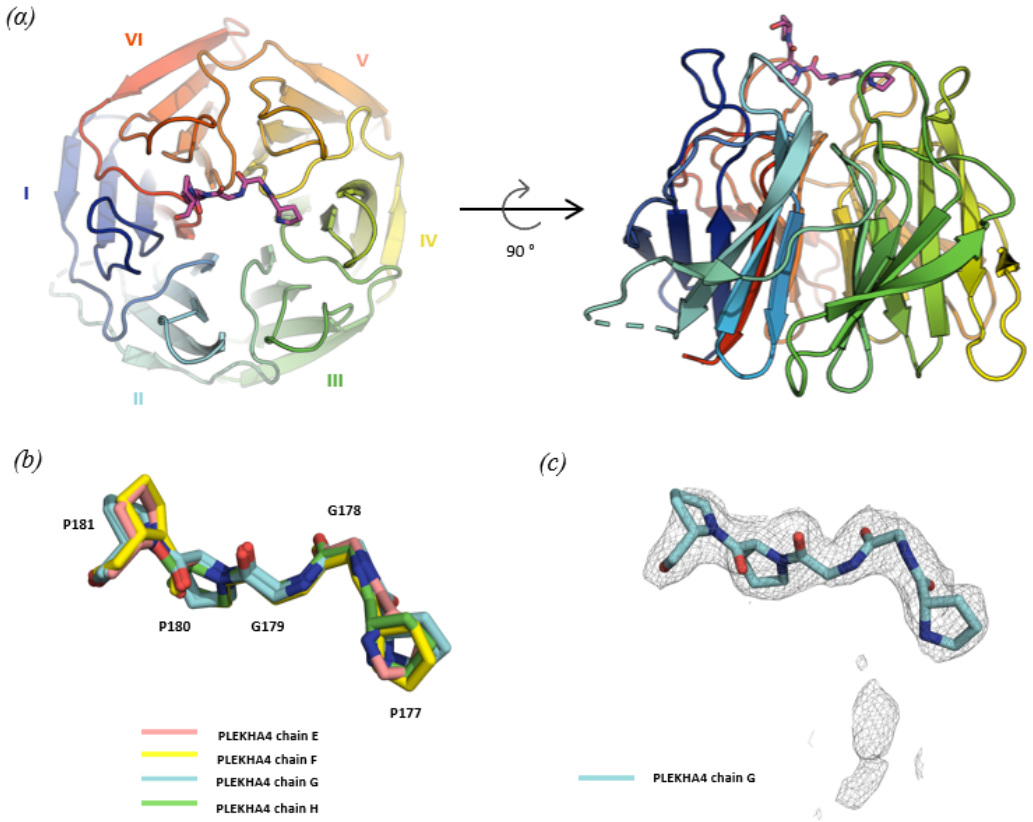


Figure 3.4: **1.95 Å structure of KLHL12 Kelch domain bound to PLEKHA4 peptide** (a) Top view and side view of the structure of KLHL12 Kelch (rainbow ribbon) in complex with PLEKHA4 peptide (coloured in magenta). Blades I to VI of the Kelch repeats are labelled accordingly. The six blades formed by the six Kelch repeats are labelled I - VI, and the four antiparallel β-strands that form each blade are labelled A-D. The binding groove for the peptide is formed by the six BC extended loops. (b) Superposition of the four PLEKHA4 chains in the same asymmetric unit, stick representation. (c) 2Fo-Fc electron density map (gray mesh) for the PLEKHA4 peptide of the best resolved chain, contoured at 1 σ. Ambiguous density N-terminally of the peptide, that could not be definitely resolved also shown.

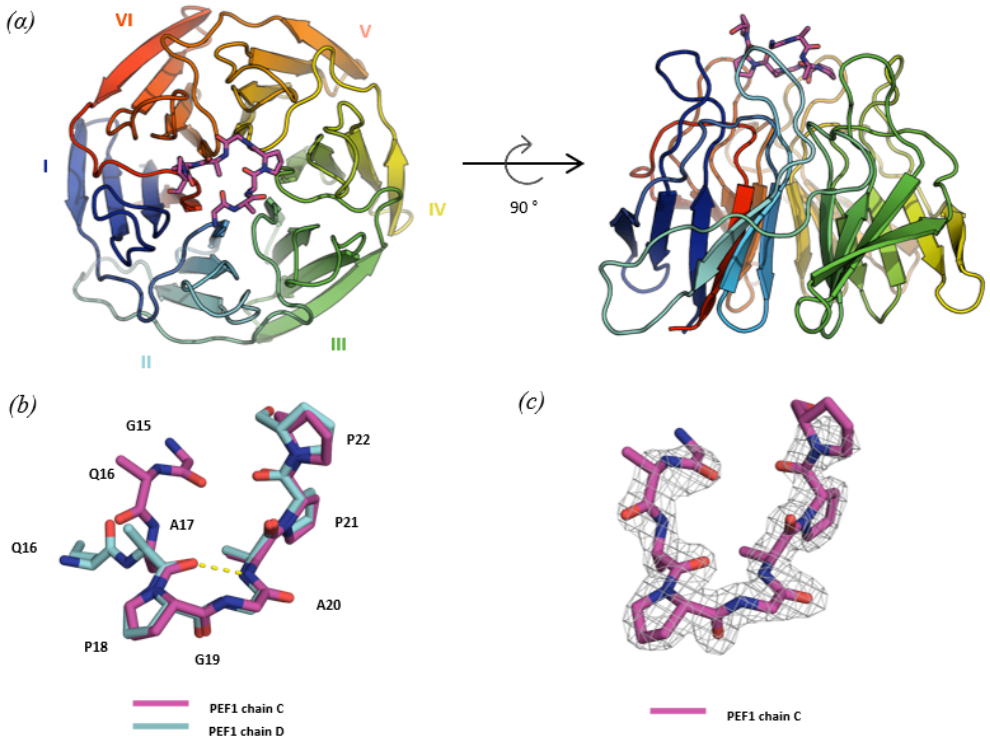


Figure 3.5: **1.88 Å structure of KLHL12 Kelch domain bound to PEF1 peptide** (a) Top view and side view of the structure of KLHL12 Kelch (rainbow ribbon) in complex with PEF1 peptide (coloured in magenta). Blades I to VI of the Kelch repeats are labelled accordingly. (b) Superposition of both PEF1 chains in the same asymmetric unit, stick representation. (c) 2Fo-Fc electron density map (gray mesh) for the PEF1 peptide of the best resolved chain, contoured at 1 σ. Structure determined by colleague, A. Ramdass.

in DVLl: (i) between H486 and the PGXPP motif's glycine (G19 and G178 for PEFl and PLEKHA4 respectively), and (ii) between KLHL12 Q293 and the final proline of the peptide consensus motif [Fig. 3.6].

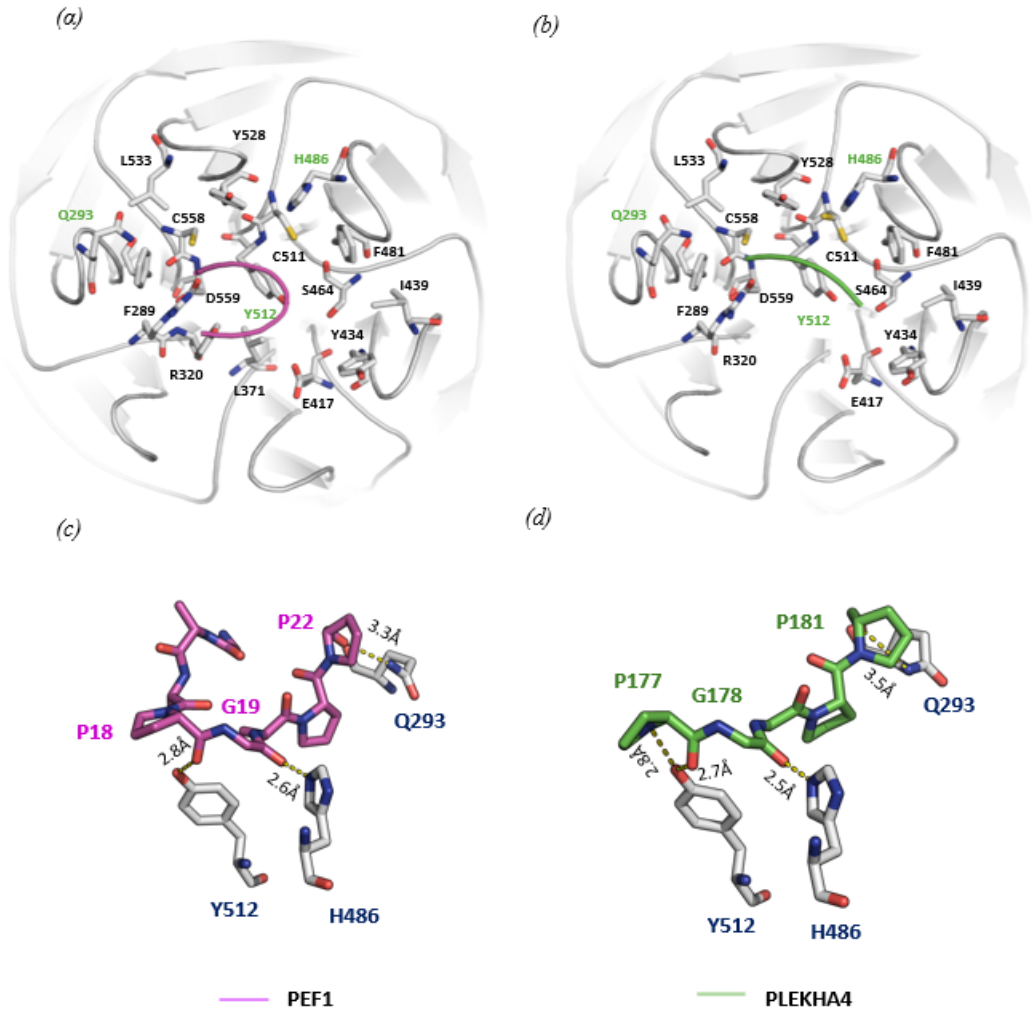


Figure 3.6: **Overview of the main interactions that tether PEF1 and PLEKHA4 peptides in the KLHL12 Kelch domain binding site**. (a) Overview of the KLHL12 residues binding PEF1. PEF1 peptide coloured in magenta. (b) Overview of the PLEKHA4-binding residues in KLHL12. PLEKHA4 peptide coloured in green. KLHL12 is coloured white and its interacting residues are labelled in black; those involved in hydrogen bonding are labelled in green. (c) PEF1 peptide and (d) PLEKHA4 hydrogen bond interactions with Kelch binding site, shown in yellow dashed lines. The poor density in the N-terminus of PLEKHA4 did not allow for modelling the first two residues of the motif, thus the full U-shaped conformation of the peptide cannot be seen.

Peptide binding to the KLHL12 Kelch domain pocket is further stabilised by extensive hydrophobic interactions including CH/π and van der Waals interactions. The first proline of the consensus motif makes two van der Waals/dipole-dipole interactions with the hydroxyl group of KLHL12 Y434 (Blade

IV) [Fig. 3.7, 3.8 (a)]. The motif's second proline nestles between F289, Y528, L533 and Y512 (Blades I, VI), forming an aryl-aromatic ring-ring stacking interaction with KLHL12 Y528 and F289 [Fig. 3.7, 3.8 (b)]. The extent of these hydrophobic interactions can be assessed by analysis of buried surface areas, calculated by the EMBI-PISA server [115] and summarised in Fig. 3.7, 3.8 (c) and (d). Both cases seem to follow DVL1's paradigm of contacting all six blades of Kelch domain, with the exception of PLEKHA4's minimal contact of blades II-III, which is likely a result of the unmodelled residues N-terminally to the PGXPP motif.

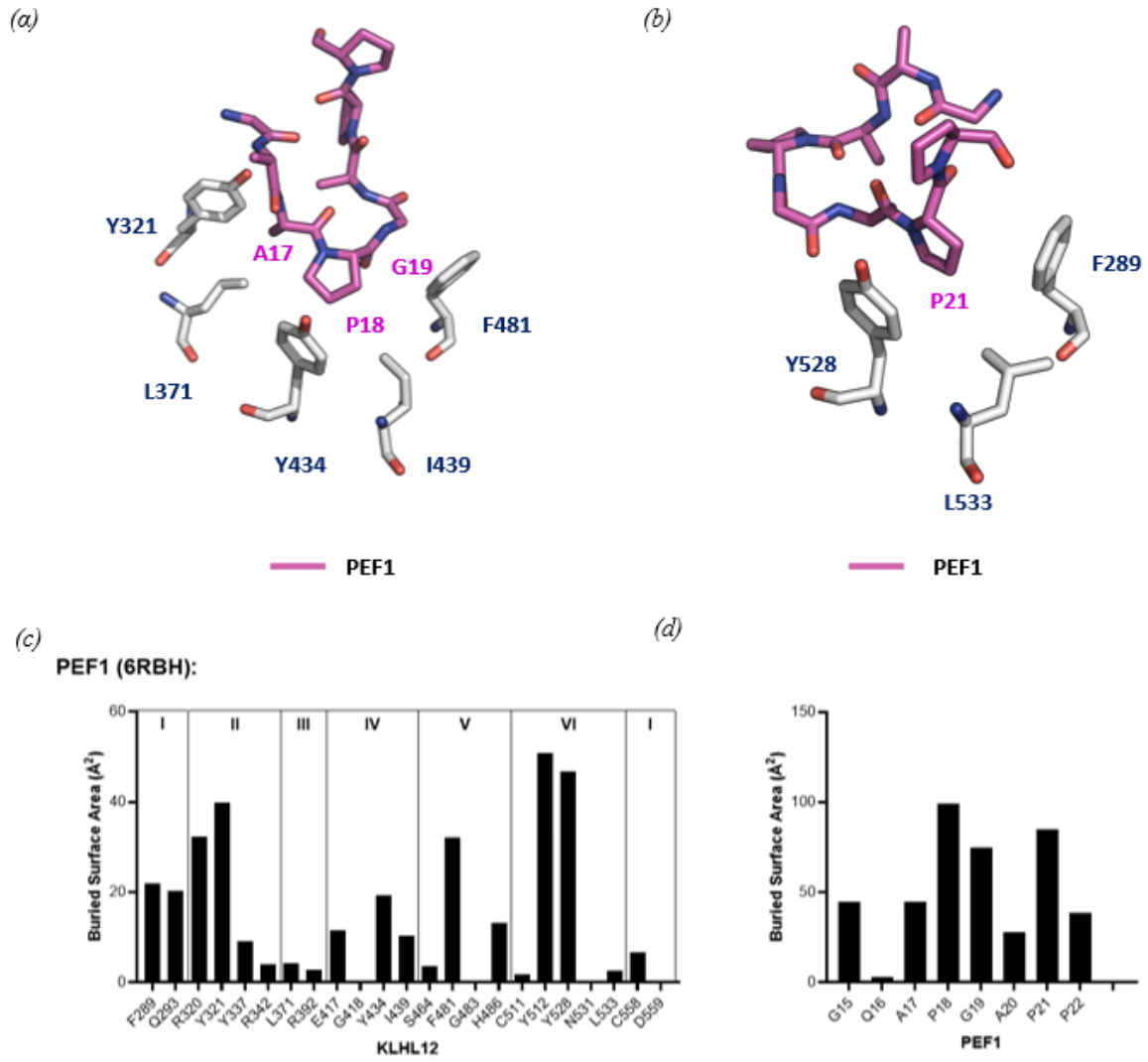


Figure 3.7: **Overview of hydrophobic interactions of PEF1 peptide nested in the KLHL12 Kelch domain binding site**. Hydrophobic interactions between KLHL12 Kelch domain and PEF1 residues 17-19 (a) and KLHL12 and PEF1 residue 21 (b). Buried surface area of residues involved in the PEF1-KLHL12 interaction; (c) for analysis of interacting residues in the KLHL12 Kelch domain chain A, and (d) PEF1 chain D. Buried surface areas were calculated by the EMBI-PISA server [115].

The BSA analysis of the Kelch domain interaction surface identified the critical residues for recruitment of the peptide consensus PGXPP motif, as Y528 and Y512 [Fig. 3.7, 3.8 (c)]. Similar analysis of peptide residues posits the first and second prolines of the motif as the main contributors of hydrophobic interactions with KLHL12. In contrast, the variable "X" residue (alanine for PEF1, glycine for DVL1 and PLEKHA4) exhibits lower BSA values, and its identity seems to be determined by merit of having a small enough side chain that would allow (i) for the peptide to adopt the critical U-turn conformation and (ii) bind to the shallow surface sterically unhindered. Both alanine and glycine side-chains seem to be tolerated, with the former directed into the centre of the U-turn conformation and Kelch domain surface [Fig. 3.10].

In both structures, the last C-terminal proline of the PGXPP degron motif (P22 and P181 for PEF1 and PLEKHA4 respectively) does not make direct contacts with the Kelch domain, as it extends away from the binding interface. It also appears to be more flexible as the electron density was only partially resolved in both PEF1 and PLEKHA4 structures [Fig. 3.4, 3.5 (c)].

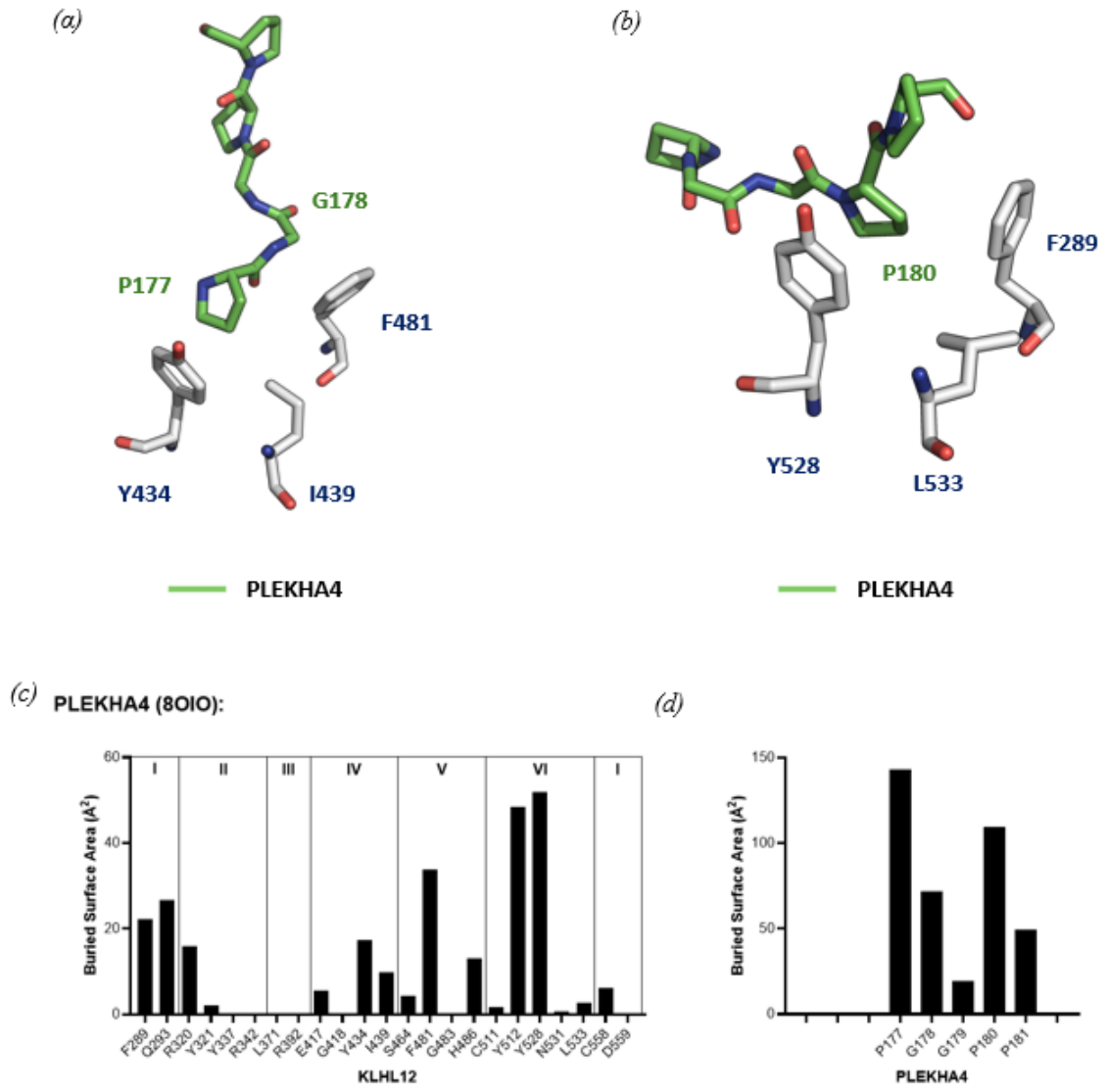


Figure 3.8: **Overview of hydrophobic interactions of PLEKHA4 peptide nested in the KLHL12 Kelch domain binding site**. Hydrophobic interactions between KLHL12 Kelch domain and PLEKHA4 residues 177-178 (a) and KLHL12 and PLEKHA4 residue 180 (b). Buried surface area of residues involved in the PLEKHA4-KLHL12 interaction; (c) for analysis of interacting residues in the KLHL12 Kelch domain chain D, and (d) PLEKHA4 chain G. Buried surface areas were calculated by the EMBI-PISA server [115].

Table 3.2: Summary of crystallographic data for KLHL12 and PEF1, PLEKHA4 complexes

Parameter	**KLHL12/PEF1**	**KLHL12/PLEKHA4**
Wavelength (Å)	0.9763	0.9795
Resolution range (Å)	72.83 - 1.88 (1.948 - 1.88)	68.08 - 1.95 (2.02 - 1.95)
Space group	P 1 21 1	P 1 21 1
Unit cell		
a, b, c (Å)	45.0, 78.4, 73.8	80.5, 73.1, 103.7
α, β, γ (°)	90, 99.2, 90	90, 98.8, 90
Total reflections	278,317	
Unique reflections	40,016 (3858)	85,892 (8315)
Multiplicity	6.9	7.0
Completeness (%)	97.20 (94.30)	99.57 (97.09)
Mean I/sigma(I)	6.2	26.7
Wilson B-factor (Å^2)	26.91	29.59
R-merge	0.2826 (3.8425)	0.154 (1.209)
R-measure	0.306	0.167
R-pim	0.115	0.063
CC1/2	0.991	0.997
Reflections used in refinement	40,015 (3858)	85,892 (8315)
Reflections used for R-free	1986 (171)	4317 (404)
R-work	0.209 (0.347)	0.201 (0.346)
R-free	0.254 (0.364)	0.251 (0.392)
Number of non-hydrogen atoms	4803	9175
Macromolecules	4443	8747
Ligands	24	58
Solvent	348	390
Protein residues	591	1155
RMSD (bonds, Å)	0.008	0.008
RMSD (angles, °)	0.98	0.99
Ramachandran favored (%)	95.71	96.46
Ramachandran allowed (%)	4.12	3.45
Ramachandran outliers (%)	0.17	0.09
Rotamer outliers (%)	0.21	0.21
Clashscore	14.77	11.04
Average B-factor (Å^2)	33.70	35.77
Macromolecules	33.44	35.65
Ligands	29.28	45.62
Solvent	37.08	37.51

3.3 Discussion

KLHL12 Kelch domain binds proteins that bear a PGXPP motif. Here I proved by resolving the PLEKHA4-KLHL12 complex structure at 1.95 Å that not only the motif, but also the binding mode is conserved across all tested interactors, including the substrate, antagonist and co-factor. This work demonstrated that PLEKHA4 is a competitive inhibitor of KLHL12 ligase.

Peptides from both the antagonist PLEKHA4 and co-factor PEF1 bind to the same pre-formed shallow hydrophobic pocket in the KLHL12 Kelch domain formed by the six BC loops - one from each of the six blades I-VI - as previously described for substrate DVL1 [72]. Structural alignment of the KLHL12 Kelch domain apo structure [PDB: 2VPJ, [106]] and the PLEKHA4-bound, and PEF1-bound complexes showed they are almost identical [RMSD 0.636 Å over 285 C_α atoms, and 0.794 Å over 289 C_α atoms respectively], similar to the reported findings comparing apo and DVL1-KLHL12 complex [RMSD 0.621 Å over 285 C_α atoms]. The low RMSD values of apo and bound alignments suggest that KLHL12 affords a mostly rigid protein surface that already exhibits high shape complementarity with the interactor loops and does not undergo significant conformational changes. On the peptides' side, the multiple prolines conformationally restrict the backbone, also exhibiting relative geometrical rigidity. Therefore, both the peptide and KLHL12 come to the interaction preformed and the binding event seems to be an enthalpic rather than an entropic event, as also evidenced by the extensive hydrogen bonding and van der Waals interactions analysed in Fig 3.2.2. This hypothesis could be tested in future experiments using isothermal calorimetry (ITC).

There are three KLHL12 residues in the Kelch domain binding site that undergo small conformational changes upon PLEKHA4 and PEF1 binding, Q293, F289 and Y528 as shown in Fig. 3.9 (d). Interestingly, those are three of the residues that contribute most to the binding. The sidechain of Q293 of blade I flips 40°C inwards towards the binding site to engage in hydrogen binding with third proline of the PGXP**P** motif. Additionally, residues F289 (blade I) and Y528 (blade VI) shift slightly, at 7.7 and 5.2 degrees respectively, inwards to nestle around the second proline of the PGX**P**P motif. These

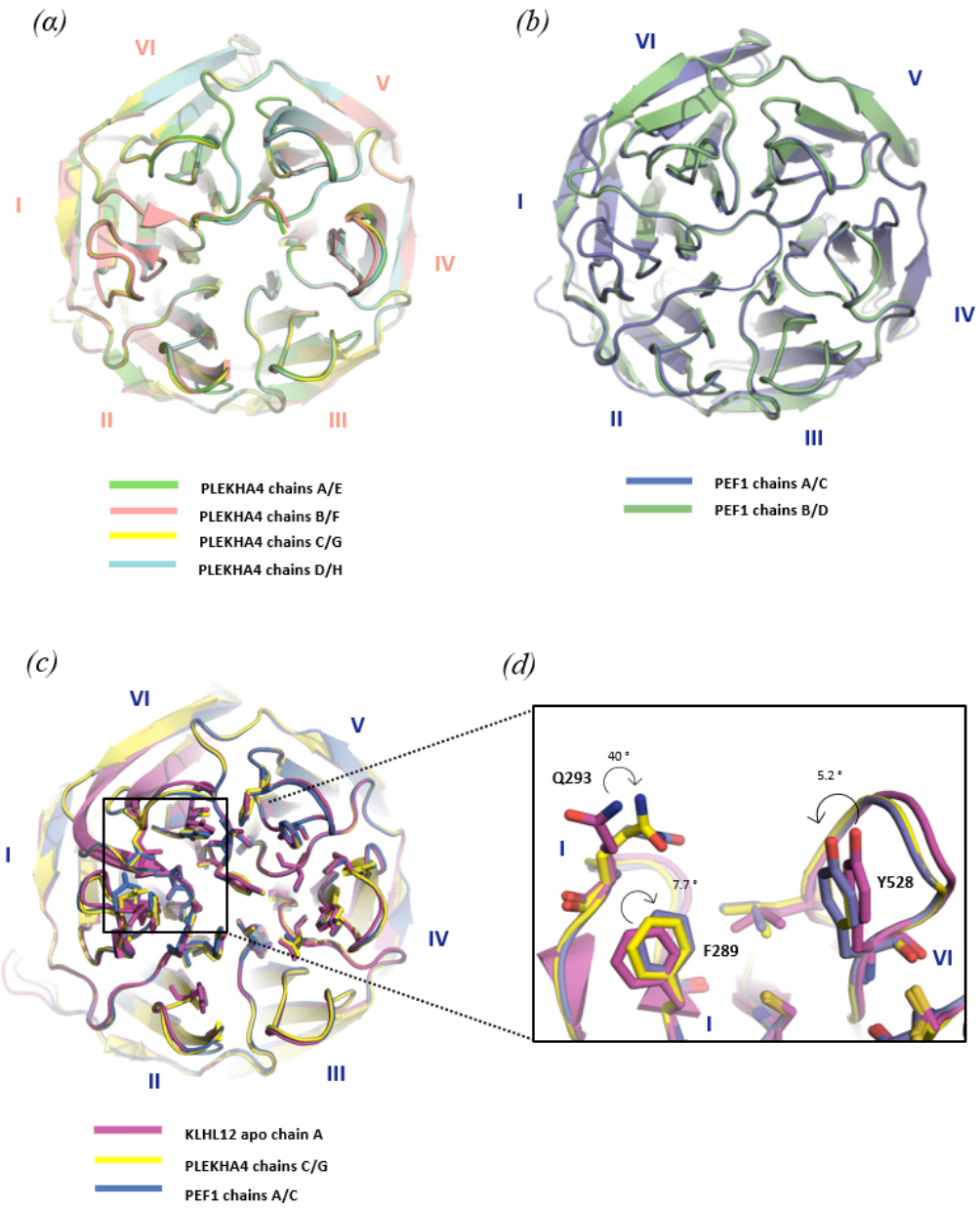
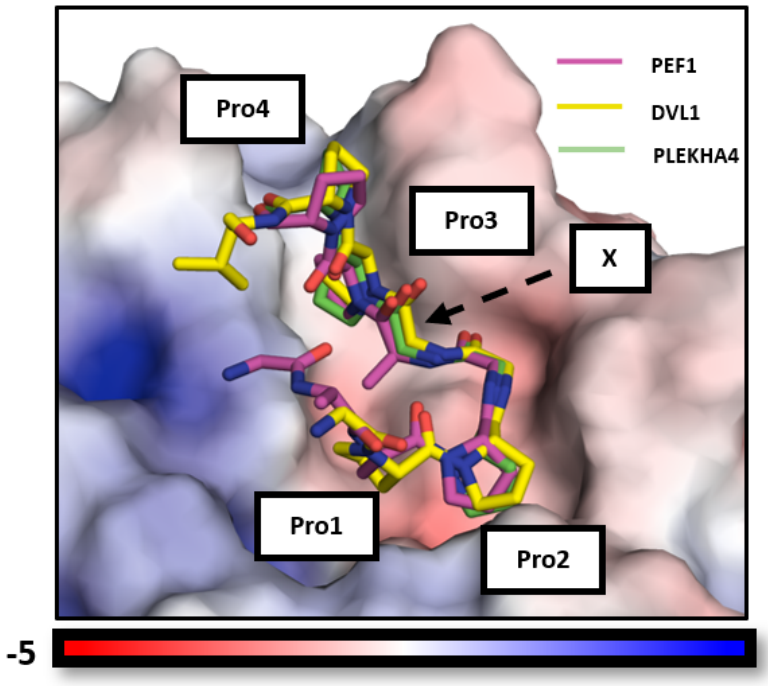


Figure 3.9: **Comparison of KLHL12 Kelch domain between the bound and apo states**. (a) Superposition of four KLHL12-PLEKHA4 complexes in the same asymmetric unit, ribbon representation. (b) Superposition of two KLHL12-PEF1 complexes in the same asymmetric unit, ribbon representation. (c) Superposition of KLHL12 Kelch domain in apo and PLEKHA4- or PEF1-bound states. (d) Zoom-in on residues in the binding site that differ in position between the apo and bound states.

small shifts are consistent for both PLEKHA4 and PEF1 peptides and perhaps positions the ring planes optimally for the proline-π (planar, equivalent to π-π stacking) and edge-to-face interaction (T-shaped geometry, equivalent to T-stacking).



KLHL12 Interactor	KLHL12 Binding Motif	Relation to KLHL12
DVL1	657-P**PG**G**PP**-662	Substrate
PLEKHA4	176-G**PG**G**PP**-181	Antagonist
PEF1	17-A**PG**A**PP**-22	Co-adaptor
Consensus	**PGXPP** ◆ ◆◆ 2 3 4	

Figure 3.10: **Binding mode of KLHL12 interacting partners**. Alignment of antagonist PLEKHA4 (this study, in green), co-adaptor PEF1 (this study, in magenta) and substrate DVL1 (PDB: 6TTK, in yellow) peptides in complex with KLHL12 Kelch domain. All peptides share the same binding motif (consensus "PGXPP") and the same binding mode. Rendering and alignment in PyMol.

This recognition motif residing in a disordered proline-rich loop seems to be of importance to the binding affinity and binding mode. Prolines, unlike other amino acids, have a fixed angle of rotation around their Cα-C_O bond, and in this case, they seem to be necessary to introduce kinks and extend N- and C-termini away from the protein interface. This is supported by the observation that even though C-terminal prolines of the PGXP**P** motif do not make any contacts with KLHL12, they greatly

contribute to binding, as shown by deletion studies by Chen *et al.* [72]. The prolines' role in changing termini directionality not only serves for avoiding steric hindrance with blade II of KLHL12, but also allows the peptides to adopt this twisted U-turn conformation forming contacts with all six blades of KLHL12 to maximise interactions.

The structural alignment of the antagonist PLEKHA4 and the co-adaptor PEF1 from this study revealed that both peptides adopt the conformation described for the substrate DVL1 [72], rather than the one reported for DVL3 [114]. It is likely that the divergent binding mode observed for DVL3 is an artifact of the N-terminally truncated DVL3 peptide sequence, which lacks the proline residue preceding the conserved "PGXPP" motif (labeled as **Pro1** in Fig. 3.10), rather than an alternative binding mode sampled by a subgroup of KLHL12 binding partners. This hypothesis is further supported by the fact that other peptides lacking a proline at the Pro1 position (PLEKHA4, PEF1) do not favour this alternative conformation [Fig. 1.9]. In conclusion, this study demonstrates that all the tested KLHL12 interactors not only share the same KLHL12 binding site but also adopt the same binding mode, consistent with the reported conformation of the substrate DVL1 [Fig. 3.10].

It has been shown that CUL3 and the substrate adaptor KLHL12 require two calcium-binding proteins, PEF1 and ALG2, to form a target-specific co-adaptor complex for the recognition of the alternate substrate SEC31 [95]. McGourty *et al.* reported that PEF1 binds to ALG2, while ALG2 associates with SEC31 [95]. Despite multiple attempts to co-crystallise KLHL12 and SEC31, including fine screens and micro-seeding, the structure of this complex remains elusive. Since the SEC31 protein contains the critical PGXPP motif in its C-terminal domain (957-GPGAPP-962), which is known to bind KLHL12 [95], I speculate that its binding mode would be similar to the one described above [Fig. 3.7, 3.8]. Supporting this hypothesis, Jin *et al.* reported that a double mutation in the KLHL12 substrate-binding site (F289A, G290A) abolished SEC31 recruitment, and large COPII vesicle formation [94]. Similarly, recent reports located the KLHL12 binding motif of substrate Lunapark to it's N-terminal cytosolic region (202-SAPGGPPER-210) which I also speculate might follow the

paradigm of the binding mode described above [102].

The binding mode being shared between the tested substrates and co-factors has interesting implications on the putative assembly of $CUL3^{KLHL12}$-substrate complexes. The dimerisation of KLHL12 provides a plausible mechanism to account for the simultaneous binding of co-factors and substrates previously proposed [95]. KLHL12 homodimerises through its BTB domain such that its E3 ligase complex contains two Kelch domains. Thus, one KLHL12 protomer can bind via its Kelch domain to the co-adaptor PEF1, while the other can bind to the matched substrate SEC31 [Fig. 1.7]. This mechanism implies that ubiquitylation levels of the substrate are titrated by the co-adaptor's concentration. In this proposed model, if SEC31 levels are low, while PEFl levels high, then two PEFl molecules would bind to both of the KLHL12 monomers, leaving the substrate SEC31 un-ubiquitylated. With rising substrate levels, SEC31 could compete with one PEFl molecule for binding, and thus become ubiquitylated. This model would explain previous observations that increasing PEF1 concentrations impair the ability of $CUL3^{KLHL12}$ to ubiquitylate SEC31 *in vitro* and reduce SEC31-KLHL12 interaction in cells [95]. E3 activity that is responsive to the changing protein levels of a substrate/co-adaptor pair has been observed before on the KEAP1[1]-NRF2[2] system [84].

In addition to the known and validated interactors that were the focus of this study, it is likely that many more PGXPP-containing KLHL12 interactors remain to be discovered. Querying the human proteome for those proteins that contain the known recognition element for KLHL12 binding, yielded 258 unique human proteins. Membrane-associated proteins accounted for the largest proportion at 36% of the dataset, with approximately a third of those localised in the cell-membrane and a quarter in cytoplasmic vesicles [Fig. 3.11 (a)]. This localisation pattern is consistent with KLHL12's known interactions with the membrane-tethered PLEKAH4, with Wnt pathway secondary effectors DVLs that localise near the trans-membrane Frizzled receptors, and finally, with endoplasmic SEC31 and Lunapark proteins [72, 90, 102, 114]. Identified PGXPP-containing proteins whose localisation or

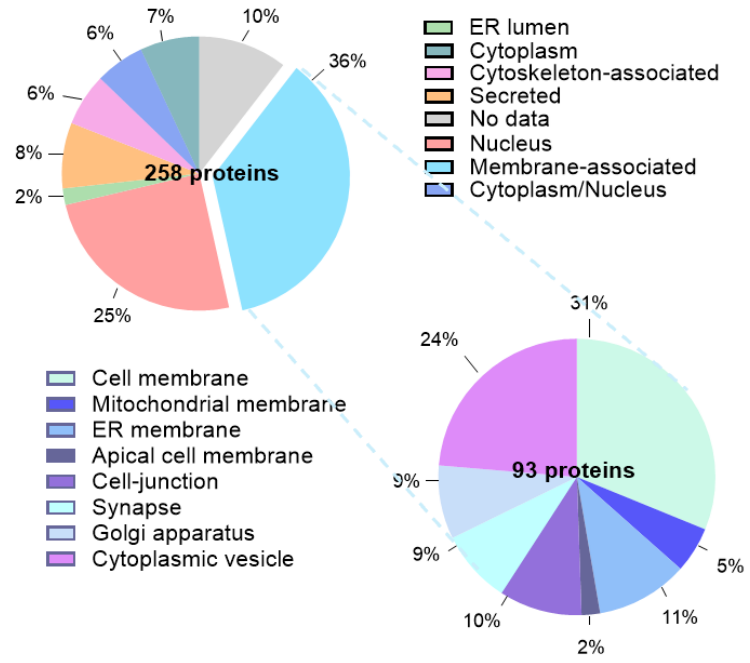
[1] Kelch ECH-associating protein 1
[2] nuclear factor erythroid 2-related factor 2

function was relevant to KLHL12's roles in ubiquitination, vesicular trafficking, and Wnt signalling are presented in Fig. 3.11.

Notable candidates included the E3 ubiquitin ligase Deltex-1 (DTX1) and the substrate adaptor α-arrestin 1, also known as arrestin domain-containing protein 1 (ARRDC1), both of which localise to vesicular compartments and regulate Notch signalling. Notch is a key developmental pathway involved in cell fate specification, neurogenesis, lymphogenesis, and myogenesis [212–214]. These hits suggest KLHL12 may function beyond Wnt signalling, potentially contributing to Notch pathway regulation. Beyond the established KLHL12 substrate SEC31, a core component of the COPII vesicle coat that facilitates ER-to-Golgi trafficking, multiple proteins associated with endoplasmic reticulum exit sites (ERES) were also identified. These include SEC23-interacting protein (p125), SEC16A, and TANGO1, all of which play critical roles in the formation of vesicles accommodating large cargo such as procollagen [215, 216]. Additionally, several hits localised to cell-cell junctions, reminiscent of the ER three-way junction localisation of known KLHL12 substrate Lunapark [102]. These included Afadin and cadherin family members, which participate in the E-cadherin/β-catenin complex involved in homotypic and heterotypic adherens junctions. These junctional proteins also associate with cytoskeletal elements such as actin and nectin complexes, also identified as potential PGXPP-containing proteins i.e. synaptopodin 2-like protein [217–219].

While no direct experimental evidence currently supports physical interactions between KLHL12 and these newly identified proteins, their spatial and functional relevance suggests avenues for future investigation. Such studies may reveal broader roles for KLHL12 in developmental signalling and membrane trafficking beyond its established functions. Having gained an insight on the binding site of KLHL12 and the binding mode of its partners is an invaluable tool for development of displacement-based assays that led me to explore new chemical matter for KLHL12 in the next chapters. Identifying selective KLHL12 inhibitors would provide a useful chemical probe of the Wnt pathway and aid in uncovering its role in cell differentiation and proliferation.

(a)



(b)

Human Protein	Localisation	Motif	Function
RBBP6	Nucleus and centrosome	576 –PGVPP-580	
RNF217	Single pass cell protein	189-PGAPP-193	E3 ubiquitin-protein ligase
DTX1	Endocytic vesicles	238-PGGPP-242	
α-arrestin 1	Cell membrane and extracellular vesicles	301-PGAPP-305	E3 ubiquitin-protein ligase substrate adaptor
USP43	Nucleoplasm	493-PGGPP-497	Deubiquitinase
DAAM2	Extracellular exosome	562-PGGPP-566	Key regulator of the Wnt signalling pathway
SEC23-interacting protein	ERES	142-PGAPP-146	Regulates secretory cargo traffic from ER to the Golgi apparatus
SEC16A	ERES	2005-PGVPP-2009	
TANGO1	ERES	1819-PGVPP-1823	
cPLA2-δ	Perinuclear membranes (ER, Golgi)	6-PGGPP-10	Calcium-dependent phospholipase
Synapsin-1	Synaptic vesicles	540-PGAPP -544	Regulates neurotransmitter vesicles trafficking
Connexin-46	Gap junctions	288-PGAPP-292	Structural component of lens fibre gap junctions
Afadin	Cell-cell adherens junctions (AJs)	1701-PGAPP-1705	Belong to the cadherin-catenin system connecting heterotypic cell junctions
Cadherin 7	Single pass membrane	722-PGAPP-726	
PECAM-1	Single pass cell protein	350-PGAPP-354	Cell adhesion molecule required for leukocyte migration in inflammation
Synaptopodin 2-like protein	Cytoskeleton	168-PGAPP-172	May associate with actin to modulate actin-based shape

Figure 3.11: **Many PGXPP-containing proteins in the human proteome may be potential KLHL12 interactors**. A proteome-wide search of the UniProtKB database (excluding isoforms) was performed using ScanProsite [220] to identify the PGXPP motif, where X is a small amino acid (Ala, Gly, Ser, Cys, Val). After filtering for reviewed entries, 258 unique human proteins containing this motif were identified. (a) Pie charts showing the subcellular localisation of identified proteins, as annotated in UniProt. (b) Table listing selected PGXPP-containing proteins of interest based on functional or localisation relevance to known KLHL12 pathways.

CHAPTER 4

Targeting KLHL12: Identification of Small Molecule Binders

4.1 Introduction

The aim of this chapter was to discover a potent small molecule inhibitor selective for KLHL12 Kelch domain over the β-propeller fold of other members of the Kelch family. Since the binding site and interactions with known KLHL12 partners were elucidated in **Chapter 3**, I utilised this knowledge to set up competition assays, in addition to assays detecting direct binding. Initially an AlphaScreen assay was conducted to identify compounds capable of displacing the moderately bound DVL1 peptide. Although AlphaScreen utilises fluorescent dyes, detecting binding indirectly and is therefore susceptible to false positives, it was selected as the primary screening method because it is compatible with assay miniaturisation, enabling high throughput screening. In addition, the displacement format was suitable for identifying competitive inhibitors that bind to the substrate recognition site, which aligned with the goal of this approach.

As an orthogonal validation of binding, I used label-free non-competitive techniques that relied on protein immobilised on a surface and compound flowing freely in solution, namely surface plasmon resonance (SPR), gated coupled interferometry (GCI), and biolayer interferometry (BLI). These techniques allow for kinetic fitting of the observed direct binding and estimation of thermodynamic binding affinity. For compounds containing a covalent warhead, mass spectrometry was also used to characterise labelling efficiency. *In silico* tools, such as docking and pharmacophore filtering, were used in virtual screening as complementary techniques to prioritise compounds from online com-

mercial libraries (Enamine) and in-house collections. These methods were expected to enrich the selection with candidates more likely to exhibit favourable binding to the target.

In this chapter I identified and characterised a novel covalent fragment-hit that exhibited selectivity for KLHL12 over other Kelch family members. The molecule binds in the peptide-binding site of KLHL12, disrupting DVL1 substrate binding. The fragment is conjugating KLHL12 once to form a mono-adduct, indicating the interaction is based on affinity rather than chemical reactivity. This finding forms the starting point for further fragment development into a PROTAC handle as detailed in the next chapter, aiming to add KLHL12 to our armory of chemistry-enabled E3 ligases.

4.2 Investigating DNA-encoding library hits from ZebiAI

A collaboration with ZebibAI Therapeutics, Inc. and Google Research had previously shortlisted 199 small-molecules as predicted hits selective for the KLHL12 Kelch domain, after performing a **virtual** DNA-encoded chemical library (DEL) screen [221]. The shortlisted compounds were not the off-DNA equivalents to the actual DEL screening hits, but rather commercially available analogs selected by machine learning. The "hits" had therefore not undergone any true experimental validation, which I attempted, as outlined below.

4.2.1 *In silico* filtering for initial compound selection

KLHL12 Kelch domain in complex with native DVL1 (PDB 6TTK) was chosen as protein receptor for docking experiments. B-factor analysis of each protein chain showed that chain C had the lowest average value, as visualised in Fig. 4.1. Since B-factors quantify the atomic displacement or vibration and reflect the thermal motion and static disorder of atoms in the crystal, chain C with the lowest average would have the most well-defined binding site and could minimise docking uncertainty. In addition, DVL1 peptide's termini (chain G) were not resolved, therefore capping them was necessary to avoid non-physiological interactions and mimic real protein conditions [therefore, the N-terminus was acetylated (-$COCH_3$) and a methyl amide group was introduced to the C-terminus (-$CONHCH_3$)]. Hydrogens were added to the protein complex (chains C/G), and the hydrogen-bond network of the

protein's surface was optimised for the ligand interaction.

(a)

PDB 6TKK	Average B-factor (Å^2)
Chain **A**	21.43
Chain **A (BS)**	21.28
Chain **B**	29.11
Chain **B (BS)**	24.43
*Chain **C***	*17.44*
*Chain **C (BS)***	*16.14*
Chain **D**	35.19
Chain **D (BS)**	33.86

(b)

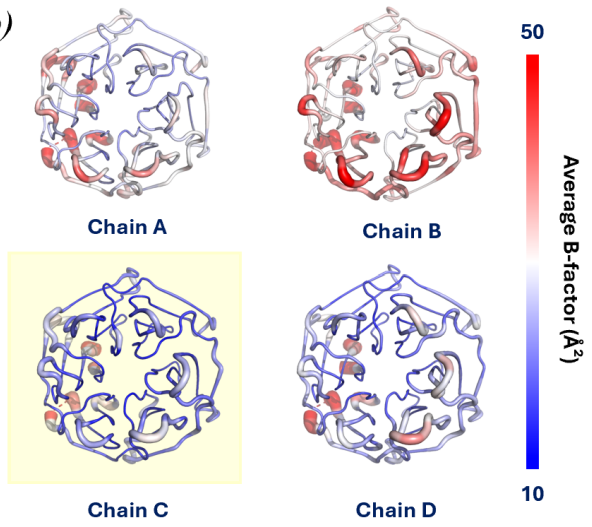


Figure 4.1: **Average B-factor analysis of KLHL12-DVL1 complex**. (a) Average B-factors of each chain of the KLHL12-DVL1 structure PDB: 6TTK [72]. Residues around 5 Å of ligand DVL1 are denoted as the binding site "BS". Chain C had the lowest B-factor average and was selected for docking studies as the protein receptor. (b) KLHL12 Kelch domain protein chains A-D are visualized in PyMOL using the "putty" representation; where the tube thickness and colour correspond to the B-factor values, indicating structural flexibility. Regions with lower B-factors (rigid) are shown in blue, while more flexible regions with higher B-factors are represented in red. Intermediate flexibility is visualized in shades of white. This visualization aids in identifying dynamic regions within the protein, which may be relevant for ligand binding or conformational changes.

199 small molecule hits and cognate peptide DVL1 peptide (as benchmark) were prepared similarly; whereby hydrogens were added, pKa and protonation states and tautomers were predicted at pH 7.4 ± 1 and all possible conformations were generated. 463 conformers were generated for the 199 compounds overall, and 32 conformers for peptide DVL1. The docking grid was created based on the centroid of KLHL12's binding site, as defined by residues interacting with DVL1 ligand residues 657-PPGGP-661. Docking settings were optimised, based on their capacity to recapitulate the co-crystallised DVL1 pose. Docking studies were performed using Schrödinger's Glide (Maestro 2023), and binding affinities were ranked based on GlideScores (in kcal/mol), with lower values indicating stronger predicted binding [method details in **Chapter 2.3**]. The best docking score was the one belonging to the re-docked peptide at -6.9 kcal/mol while hit compounds' ranged between -6.0 to 0.3 kcal/mol, indicating binders with lower binding affinity than the moderate binding affinity of DVL1.

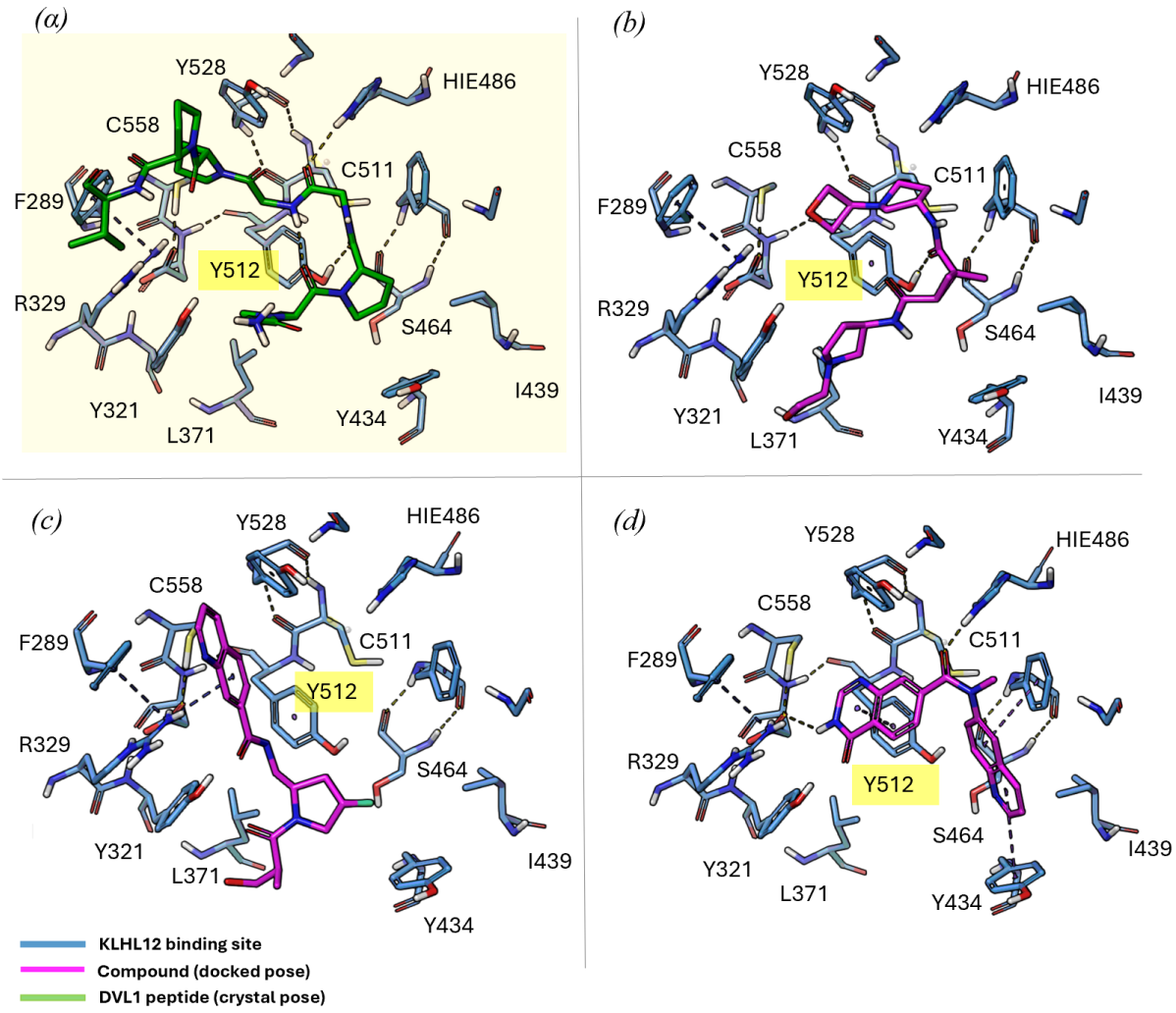


Figure 4.2: **Representative docked poses of DEL hits**. DEL hits selected from ZebiAI were docked on KLHL12 Kelch domain [PDB: 6TTK, chain C] using Schrödinger's Glide (Maestro 2023). (a) Crystal pose of DVL1 peptide in the KLHL12 binding site. Since this is the bencharm ligand, it is highlighted in yellow. Of note the hydrogen bond between DVL1 and KLHL12 Y512. Docked pose of one example DEL hit (b) recapitulating a hydrogen bond interaction with Y512, (c) not interacting with Y512, and (d) exhibiting π-π stacking with Y512.

37 compounds were selected for initial testing based on multiple criteria, including *docking* and emodel scores. The docking score ranks the different docked poses based on their predicted binding affinity, with the lower value denoting higher binding affinity. The emodel score provides an estimate of the binding affinity by combining the docking score with additional terms that account for internal ligand strain and complex desolvation effects. Emodel score takes into account how likely the conformation of the docked pose is to be adopted, and is important for assessing multiple chemical species (i.e. tautomers, alternate protonation states) of the same chemical identity. Other than the nominal value of scoring functions, the poses were filtered in Maestro for their ability to recapitulate the important hydrogen bond interaction with Y512 [Fig. 4.2 (a)]. The best docked poses of a quarter of the DEL library (50 unique molecules) did engage with Y512 in a hydrogen bond [Fig. 4.2 (b)], while other molecules interacted with Y512 in a different manner, such as π-π stacking [Fig. 4.2 (d)]. In addition, care was taken for selected molecules that represented the chemical diversity of the ligand set. Finally, commercial availability was also taken into account. A summary of the chemical clusters and the representative DEL hits that were purchased from Enamine Ltd., Fig. 4.3.

4.2.2 Screening of selected DEL hits

Since I was primarily interested in molecules competing for binding at the KLHL12 Kelch domain's substrate binding site, a displacement AlphaScreen assay was utilised as a first validation step. Amplified Luminescent Proximity Homogeneous Assay (AlphaScreen) uses chemically-coated donor and acceptor beads for the generation of luminous signal when the beads are in proximity. The known interaction between KLHL12 and the native substrate peptide DVL1 was exploited for discovering inhibitors that would disrupt that binding leading to measurable signal attenuation. Hexahistidine tagged KLHL12 Kelch domain [268-567 a.a.] was purified as described previously in **Chapter 3.2.1** and incubated with biotinylated "PGXPP" DVL1 peptide. Free "PGXPP" DVL1 peptide was used as a positive control in a concentration where full displacement of biotinylated DVL1 from KLHL12 would be ensured (200 μM).

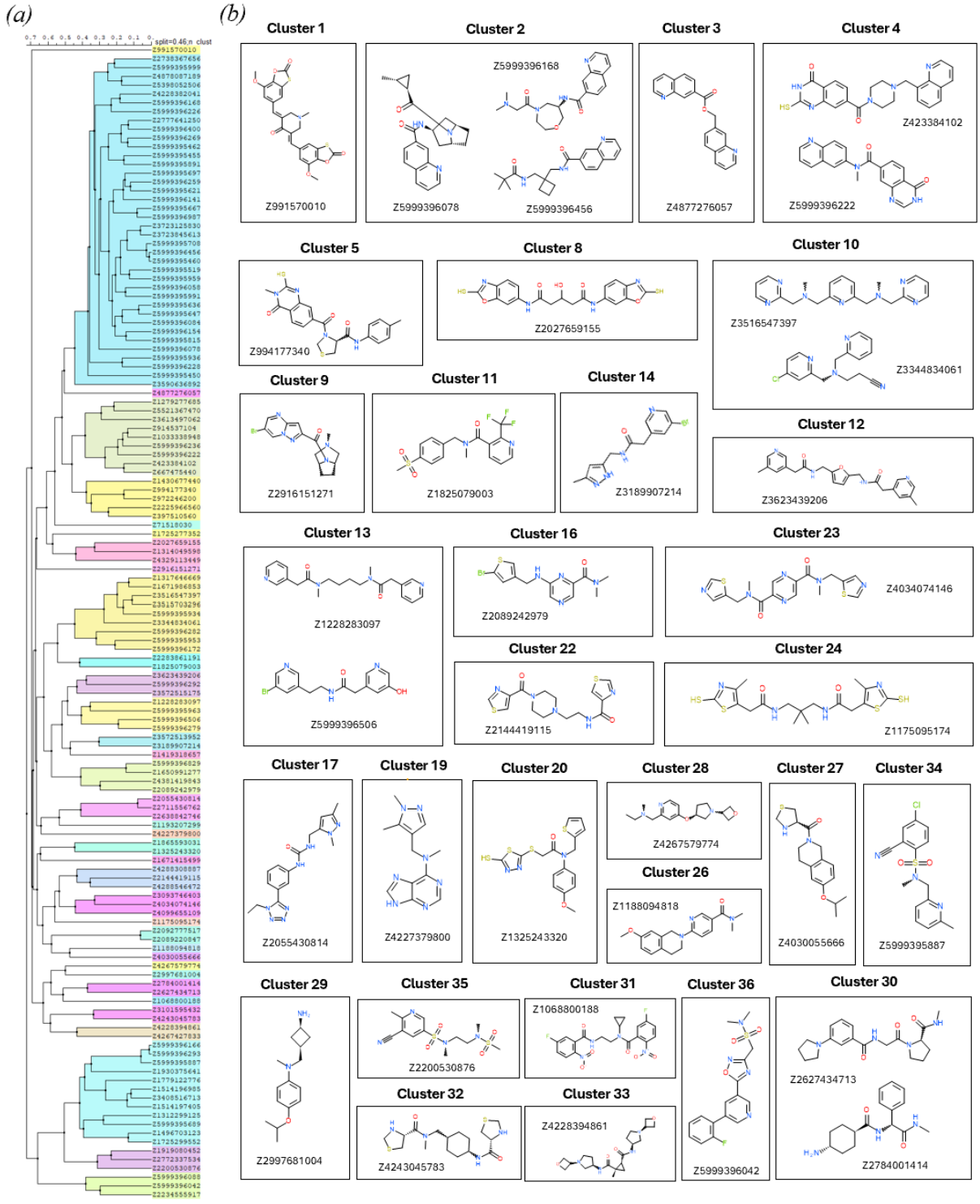


Figure 4.3: **Selection of a subset of DEL hits representing chemical diversity of the collection**. (a) Chemical clustering performed in ICM-Pro (Molsoft), split = 0.46, n = 36 clusters; colour represents cluster identity (b) Selected compounds for initial testing. Compound names follow the supplier's (Enamine Ltd.) naming convention.

(a)

His-tag-KLHL12-p061 (nM) \ Biotin-DVL1 (GGTTKAYTVVGGPPGGPPVRE) nM	800	400	200	100	50	25	12.5	6.25	3.125	1.5625	0
800	887224	800508	859370	767030	857166	816848	695552	453017	211774	57133	551
400	899745	843125	864861	851960	888041	849794	676932	374376	147326	27189	1178
200	842137	819375	866951	844987	877705	844455	659205	337934	109668	18753	931
100	807975	764313	829863	820135	**805353**	761140	580013	243124	63403	8645	1824
50	707484	659908	677806	703874	687439	670947	412148	139384	27018	4636	988
25	463562	433580	424973	424289	436335	411388	190228	43643	6498	1235	1463
12.5	217056	150005	151126	147782	145540	132639	35986	7315	1539	646	342
6.25	33440	14896	13927	11096	9785	8132	760	1539	703	323	513
3.125	3743	760	589	1007	1197	589	684	817	684	665	1482
1.5625	494	608	931	532	988	1026	1349	1064	1197	342	950
0	1539	1995	266	475	627	1919	551	836	855	323	1007

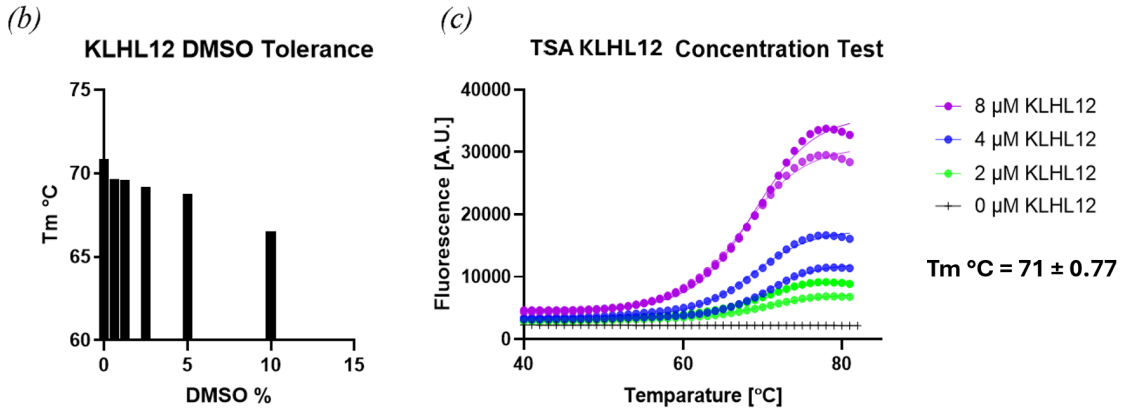


Figure 4.4: **Overview of preliminary assay steps**. (a) Determination of His-KLHL12:biotin-DVL1 ratio that affords adequate signal, while conserving samples for the AlphaScreen assay. Representative data; repeated for every different protein or peptide batch. (b) Determining DMSO tolerance for KLHL12 comparing TSA measurements of KLHL12 in increasing DMSO concentration in assay buffer. Representative data; repeated for every different protein batch. (c) Determining ideal KLHL12 concentration to afford appropriate assay window for the thermal shift assay. The standard deviation (SD) from the mean of two technical replicates is reported.

Preliminary tests to assess KLHL12's tolerance to increasing concentrations of DMSO, revealed that the protein's thermal stability remains unchanged up to 2.5 % DMSO, with an apparent melting temperature, Tm = 70 °C. Slight destabilisation of 1 °C was observed at 5% DMSO [Fig. 4.4 (b)]. Therefore, all assays were performed in < 2.5 % DMSO concentrations. The thermostability of the protein also allowed all incubation steps to be performed at room temperature. The 37 prioritised DEL hits were dispensed for 11-point dose-response curves with maximum concentration of 100 μM and 0.5 % DMSO final concentration in two technical duplicates. The average Z' (mean) = 0.87 and Z'(median) = 0.90 values across three 384-plates suggested that the assay can distinguish between signal and noise and is suitable for screening.

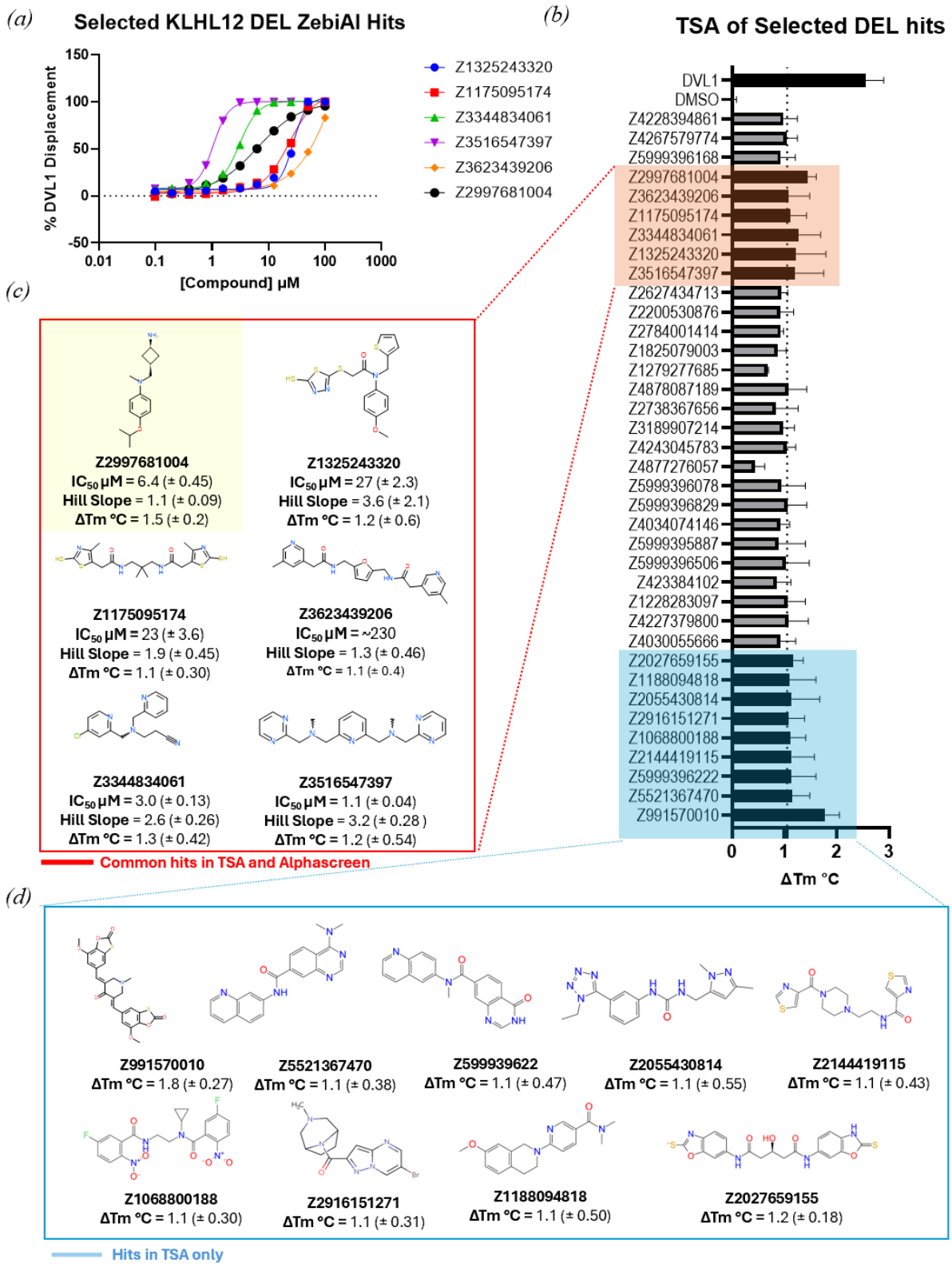


Figure 4.5: **AlphaScreen and TSA hits from initial DEL selection**. *Figure caption continues in the following page.*

Figure 4.5: (a) IC_{50} determination of the best hits from initial selected DEL hits. The compounds' potency was assessed using DVL1 peptide displacement from KLHL12 via AlphaScreen experiments, as described in **Chapter 2.5.6**, and quantified through time-dependent IC_{50} titration curves. The %DVL1 displacement was normalised between averages of negative (DMSO, 0% displacement) and positive (DVL1 peptide, 100% inhibition) controls. A non-linear curve that allowed for a variable slope was fitted on logarithmically transformed compound concentrations over normalised response in Graphpad Prism. Error bars represent the ± standard deviation (SD) from the mean of two technical replicates. (b) Thermal Shift Assay (TSA) of initial DEL selection compounds at 200 µM over KLHL12 Kelch domain at 4 µM. Error bars represent the ± standard deviation (SD) from the mean of three technical replicates. Dashed line represents 3σ level of the measurement. Compounds that showed thermal stabilisation $\Delta Tm > 3\sigma$ were considered positive hits, and are represented as black bars. Those that did not incur statistically significant Tm shifts are represented in grey bars. Physiological substrate 20mer DVL1 peptide was the positive control, while DMSO was the negative control. (c) Compounds shaded in orange were the common hits between the Alpha-Screen assay and TSA assay. (d) Compounds shaded in cyan were TSA hits but did not show DVL1 displacement in AlphaScreen.

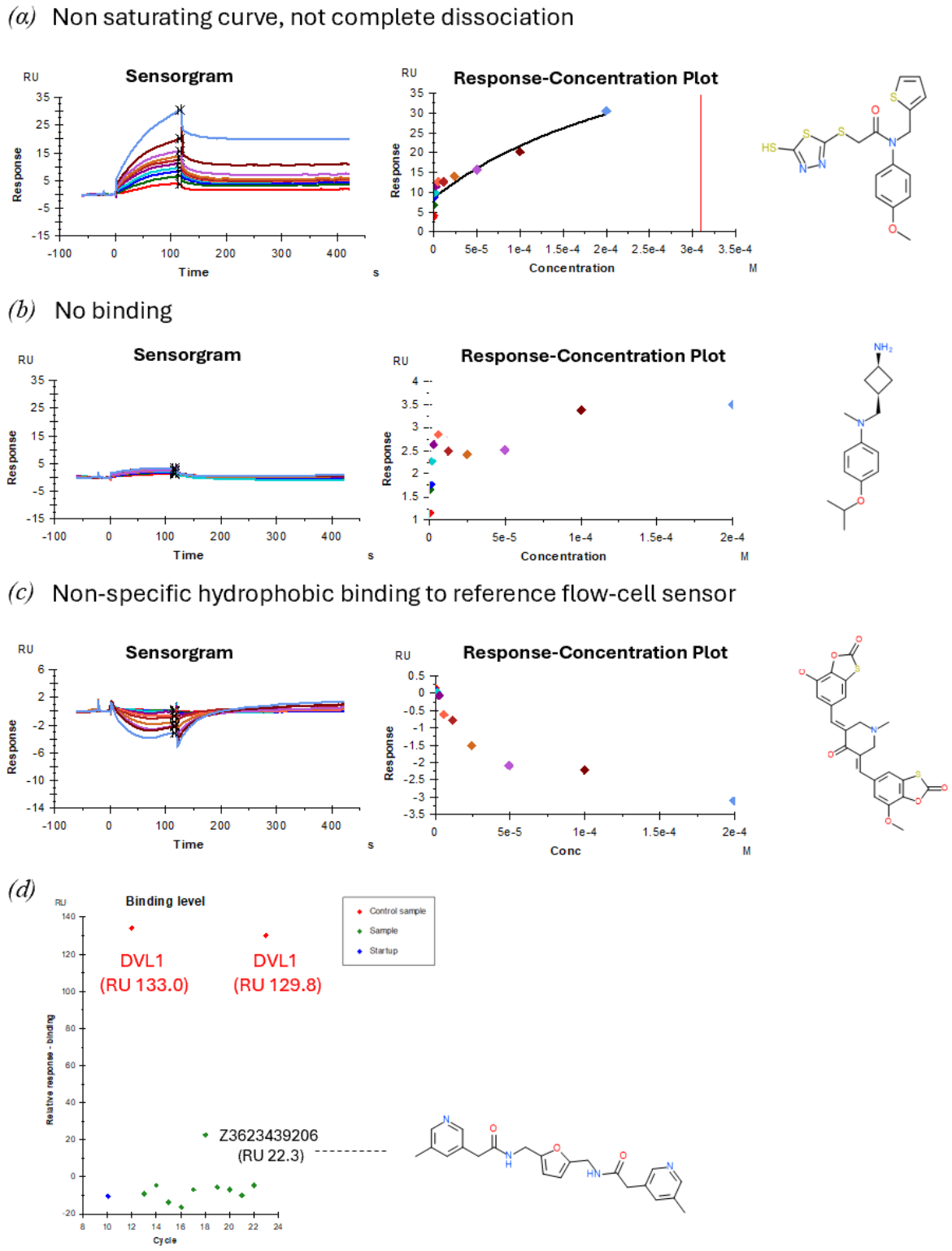


Figure 4.6: **Orthogonal validation of DEL hits; representative SPR results**. *Figure caption continues in the following page.*

Figure 4.6: Dose-response sensorgrams and Response-Concentration plot from an SPR experiment assessing the binding of selected compounds to immobilized biotinylated KLHL12 Kelch domain on a streptavidin sensor. Sensorgrams show the real-time binding response at increasing concentrations of selected compounds (0.10 - 200 µM, 12-point curves). (a) Compound Z1325243320 exhibited a non-saturating binding curve, failing to reach equilibrium steady-state, with incomplete dissociation. This suggested a non-specific interaction, potentially due to compound aggregation and multiple compound molecules binding to the surface. (b) Compound Z2997681004 showed no detectable binding to immobilized KLHL12. (c) Compound Z991570010 displayed a negative response, indicative of non-specific hydrophobic interactions with the reference flow-cell sensor, further suggesting non-specific hydrophobic binding. (d) Summary of relative response-binding at a single concentration (200 µM) for selected compounds over immobilised KLHL12 Kelch domain on a streptavidin sensor. The DVL1 substrate peptide was used as a positive control and is highlighted in red. DVL1 showed the expected binding response, confirming that the KLHL12 Kelch binding site remained unobstructed: $R_{\text{max}} = R_{\text{immo}} \left(\frac{M_{\text{KLHL12}}}{M_{\text{DVL1}}} \right) = 9780 \times \left(\frac{32873}{2241} \right) = 143$, $R_{\text{observed}} = 133$. Only compound Z3623439206 showed detectable binding with an RU of 22.3.

Overall, six compounds successfully displaced DVL1 from KLHL12's binding site with IC_{50} values ranging from 1 μM to 235 μM. Compound structures are summarised in red rectangle in Fig. 4.5 (a) and (c). Even though the results were encouraging at first glance, hill slope analysis revealed that some of the tightest binding hits, such as Z3516547397 (IC_{50} = 1.05 μM) and Z3344834061 (IC_{50} = 3.04 μM), displayed hill slope values 3.22 and 2.61 respectively, suggesting positive cooperativity. Since compound binding to possible multiple binding sites or allosteric effects would not incur DVL1 displacement and thus would not be detected, this finding might be indicating aggregator compounds and artificial cooperativity instead. This would explain the unusually high binding affinity that accompanies the high hill slopes. However, compound Z2997681004 (IC_{50} = 6.36 μM) exhibits non-cooperative binding likely in a single-site interaction (hill slope = 1.14), making it the most promising compound of this series (highlighted in yellow in Fig. 4.5 (c)).

To compliment the AlphaScreen's displacement methodology, a thermal shift assay (TSA) that allows for direct observation of binding was also assessed as an alternative high-throughput assay. The benchmark DVL1 ligand induced thermal stabilisation of 2.5 °C which did not afford a big assay window, reflected on the marginal Z'(mean) value of 0.5. Overall, fifteen compounds stabilised KLHL12 significantly - exhibiting $\Delta Tm > 3\sigma$ - and none showed a destabilising effect [Fig. 4.5 (b)]. Interestingly, all compounds that appeared as hits in the AlphaScreen assay, were also significant in this assay (shaded orange), in addition to nine compounds that had not previously shown any displacement of DVL1 (shaded blue).

4.2.3 Extending screening of DEL hits

Since most of the identified hits at that point exhibited positive cooperativity and high molecular symmetry (Fig. 4.5), there was concern that these could be pan-assay interfering compounds that induced aggregation through nonspecific binding. As a result, a broader selection of the remaining untested, prioritized ZebiAI DEL hits were purchased. To increase the diversity of chemical matter in our hits, seventy-one untested compounds from the commercially available virtual DEL hits were

hand-selected from different clusters to avoid assay interference substructures. This approach was taken, as purchasing all the prioritized ZebiAI hits would have been cost-prohibitive.

The extended collection (108 compounds, including previous selection of 37) was screened in a second round of AlphaScreen assay in two technical replicates at the single concentration of 100 µ M, and final DMSO concentration of 0.5%. The results could be analysed with confidence as the Z'(mean) = 0.89 and Z'(median) = 0.97 assay values denoted a reliable and sensitive assay. As expected, all six of the previously observed AlphaScreen hits reproduced, however two new hit compounds, Z5999396172 and Z2638842746 displaced DVL1 at 86 % and 46% respectively (coloured in red, Fig. 4.7 (a)). Interestingly, compound Z5999396172 was the only one to induce Tm shift of 1.09 °C, making it the most promising compound of the extended collection.

4.2.4 Orthogonal validation of AlphaScreen hits

To validate the binding of hits previously identified by AlphaScreen, direct-binding, non-displacement methods like Surface Plasmon Resonance (SPR) and Grating-Coupled Interferometry (GCI) were employed. Both are label-free real-time biophysical techniques that measure molecular interactions by detecting changes in the refractive index near a sensor surface, where the protein of interest is immobilised. Even though the techniques both measure refractive index changes, they employ different optical detection mechanisms (plasmon resonance vs. interferometry). Using them both reduces the likelihood of systematic biases and strengthening confidence in the assay outcomes.

The KLHL12 Kelch domain (a.a. 278-567) was expressed in BirA-transformed *E. coli* cells grown in rich media supplemented with biotin, enabling *in vivo* biotinylation of the protein's C-terminal AviTag by the BirA ligase. After cell lysis, the protein was captured by its N-terminal hexahistidine tag, using Ni^{2+}-sepharose resin. The N-terminal hexahistidine tag was cleaved using TEV protease and the protein was further purified by size exclusion chromatography.

The biotinylated KLHL12 Kelch domain was immobilised on streptavidin sensor chips for both SPR

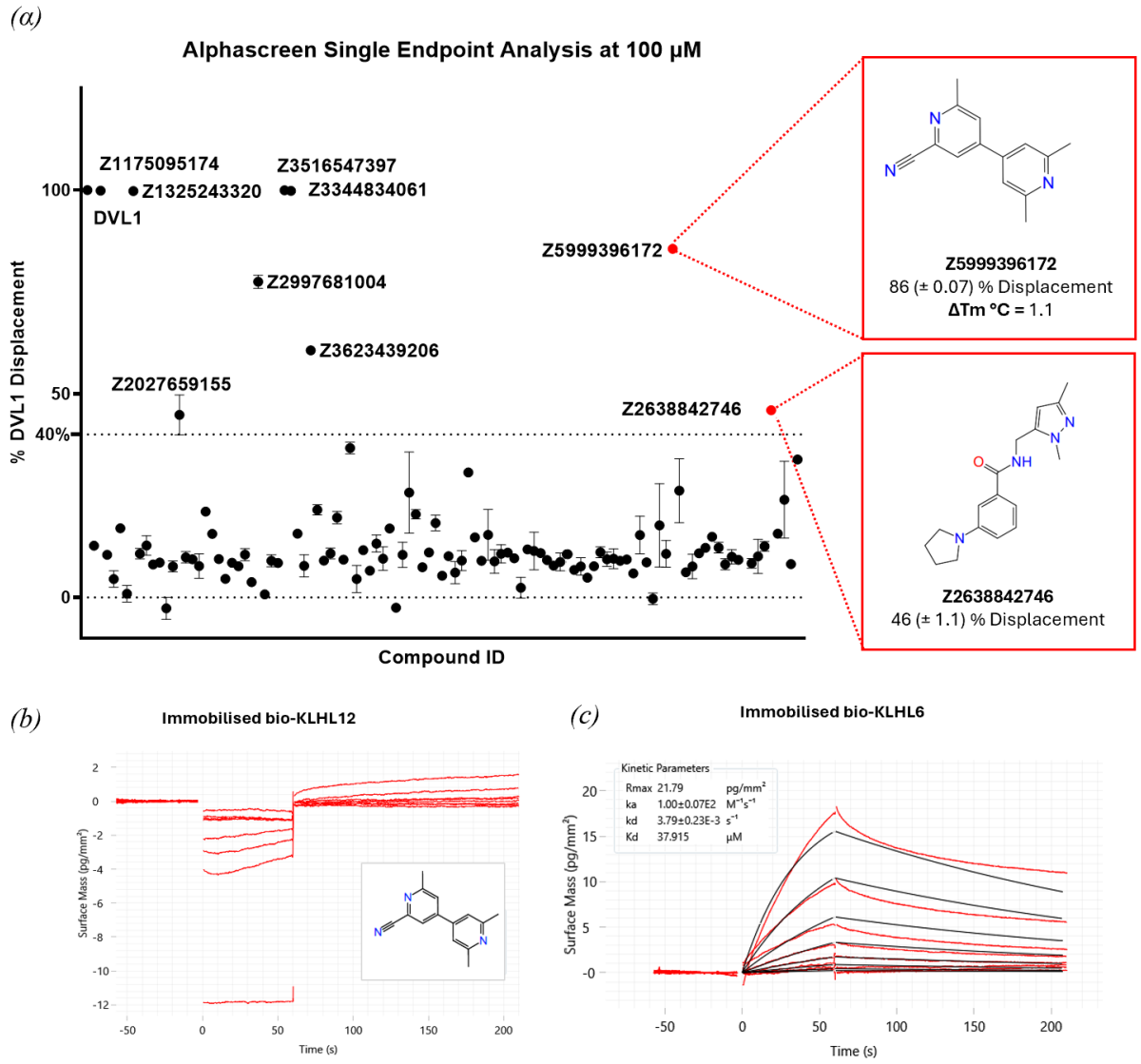


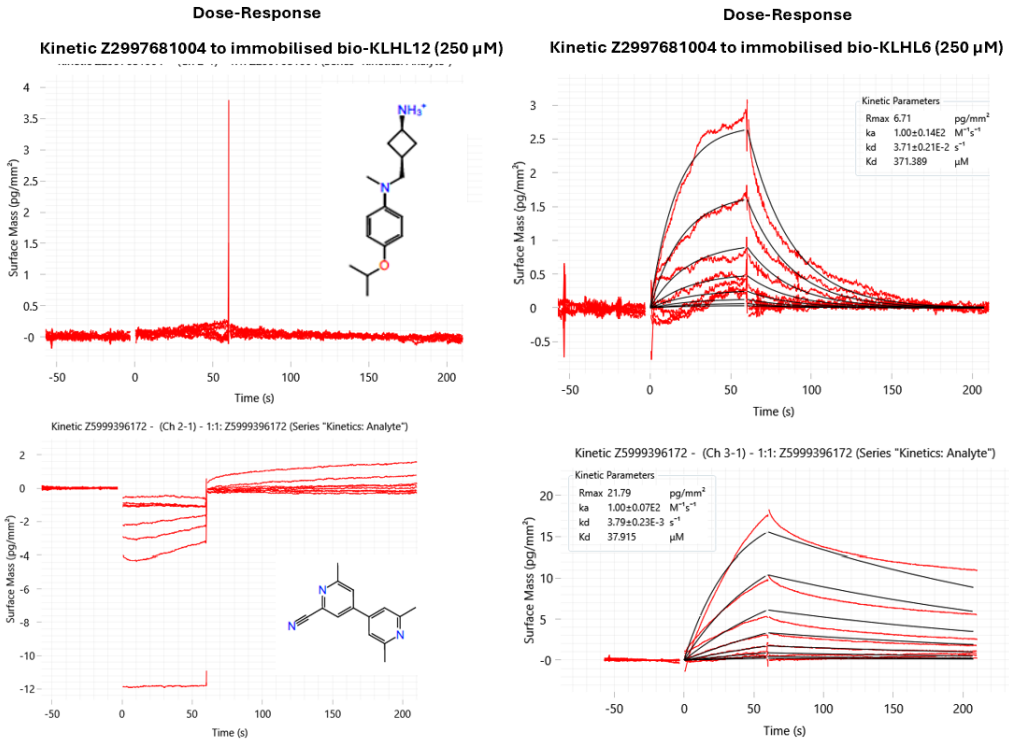
Figure 4.7: **Screening of extended DEL library**. (a) Testing of an extended selection of DEL hits. Single endpoint analysis of DVL1 displacement from KLHL12 Kelch domain binding site with compound at 100 µM after. Error bars represent the ± standard deviation (SD) from the mean of two technical replicates. Compounds that incurred more than 40% DVL1 displacement were considered as hits. Compounds highlighted in red were previously unobserved hits. (b) Representative GCI sensorgram showing real-time binding of **Z5999396172** to immobilized protein **KLHL12** at varying concentrations (8 point dose-response curve 1.9 - 250 µM). Non-specific interaction to the reference sensor was observed. (c) Representative GCI sensorgram showing real-time binding of **Z5999396172** to immobilized negative control protein **KLHL6** at varying concentrations (8 point dose-response curve 1.9 - 250 µM). Association and dissociation phases are depicted, with experimental data shown as solid black lines and global kinetic fits represented by red lines. The binding kinetics were analysed using a **1:1 Langmuir kinetic equation fit** binding model, yielding a dissociation constant (Kd) of **37.9** µM. Curve fitting was performed using waveRAPID evaluation software to extract kinetic parameters, including association (ka) and dissociation (kd) rate constants.

and GCI experiments in 50 mM HEPES, 150 mM NaCl, 0.05% Tween20, 0.5 mM TCEP, 2% DMSO running buffer. Kelch domains of other members of the family, namely biotinylated KLHL20 a.a. 303-602, KLHL14 a.a. 319-628 and KLHL6 a.a. 302-610) were also immobilised and compounds were tested against them too, to assess their selectivity. To confirm the ability of KLHL12 to act as a substrate receptor when immobilised, DVL1 peptide at 200 μM concentration was used as a positive control. In both assays the DVL1 control exhibited signal at the expected response units (RU) confirming KLHL12's binding site was unobstructed [Fig. 4.8].

Association and dissociation phases are depicted, with experimental data shown as solid black lines and global kinetic fits represented by red lines. The binding kinetics were analysed using a **1:1 Langmuir kinetic equation fit** binding model, yielding a dissociation constant (Kd) of **37.9** μM. Curve fitting was performed using waveRAPID evaluation software to extract kinetic parameters, including association (ka) and dissociation (kd) rate constants.

Dose-response and single-endpoint experiments up to high fragment concentrations of 250 μM did not to show any direct binding to KLHL12 using either SPR or GCI assay. The most promising compound from the initial DEL hit selection (compound Z2997681004, IC_{50} = 6.36 μM, Fig. 4.5 (c)) did not interact with KLHL12 in either SPR or GCI assay, but showed weak binding to negative-control KLHL6 (K_D = 371 μM) instead. Similarly, compound Z5999396172, the best hit from the extended DEL hit collection (Fig. 4.7 (a)), exhibited strong binding to KLHL6 and non-specific interaction to the reference sensor. A summary of representative SPR and GCI curves can be found in Fig. 4.6 and 4.8 respectively. Overall, none of the virtual DEL hits exhibited an interaction with the binding site of KLHL12 that could be validated as summarised in Table 4.1.

(a) Not binding to **KLHL12** but binding to negative control **KLHL6**



(b) Selective binding to **KLHL12** over negative control **KLHL6**

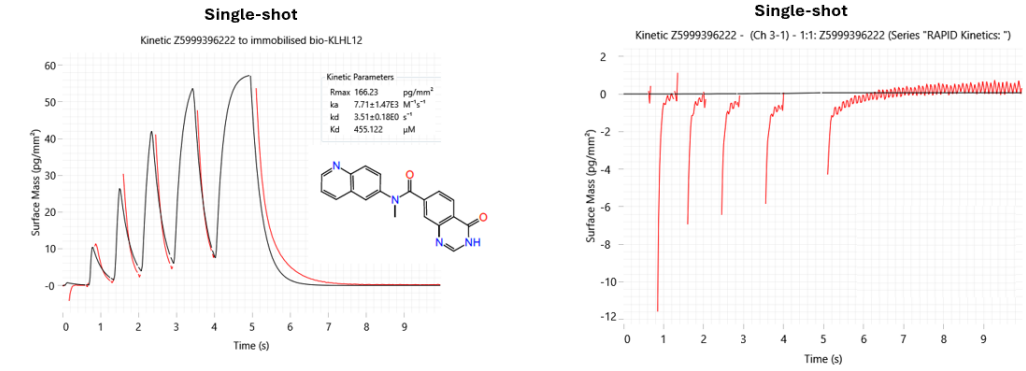


Figure 4.8: **Orthogonal validation of DEL hits; representative GCI results**. (a) Representative GCI sensorgram showing real-time binding of tested compounds to immobilized **KLHL12** and negative control protein **KLHL6** at varying concentrations (8 point dose-response curve 1.9 - 250 µM). Compound Z2997681004 (a modest binder in AlphaScreen assay IC_{50} = 6.36 µM) did not bind to KLHL12, but showed weak binding affinity to negative-control KLHL6 (371 µM) instead.
repeated single-shot injections of positive control DVL1 binding to immobilised KLHL12. The response is plotted against time (seconds). WaveRAPID kinetic data obtained using the Creoptix WAVEdelta system.

Table 4.1: Summary of virtual DEL hits validation

Compound ID	△Tm °C	AS % (100 µM)	AS IC_{50} µM (Slope)	SPR Commemts	GCI Comments
Z3516547397	1.2 (± 0.54)	99.8 ± 0.08	1.1 (3.2)	Sticking to reference chip	No binding
Z1325243320	1.2 (± 0.58)	99.8 +/- 0.01	27 (3.6)	Non-saturating curve, not complete dissociation	N/A
Z3344834061	1.3 (± 0.42)	99.9 ± 0.01	3.1 (2.6)	No binding	N/A
Z1175095174	1.1 (± 0.30)	100 ± 0.03	23 (1.9)	Tighter binding to KLHL20	
Z2027659155	1.2 (± 0.18)	44.8 ± 4.9	N/A	No binding	No binding
Z3623439206	1.1 (± 0.41)	60.7 ± 0.93	230 (1.3)	Sticking to reference chip	N/A
Z991570010	1.8 (± 0.27)	33.9 ± 0.19	N/A	No binding	No binding
Z2997681004	1.5 (± 0.16)	77.5 ± 1.6	6.4 (1.1)	No binding	No binding
Z5999396172	No shift	85.6 ± 0.07	N/A	No binding	No binding
Z2638842746	No shift	46.0 ± 1.1	N/A	No binding	No binding
Z5521367470	1.1 (± 0.38)	7.4 ± 0.70	N/A	N/A	No binding
Z2055430814	1.1 (± 0.55)	9.8 ± 1.4	N/A	N/A	No binding
Z2144419115	1.1 (± 0.43)	9.4 ± 0.37	N/A	N/A	No binding
Z1068800188	1.1 (± 0.30)	12.7 ± 0.10	N/A	N/A	No binding
Z1188094818	1.1 (± 0.50)	10.5 ± 0.78	N/A	N/A	No binding

4.3 Identification of covalent fragment hits for KLHL12

4.3.1 Cysteine tractability assessment

Since the search for traditional occupancy-based binders did not yield a selective, high-affinity binder, the potential for covalently targeting KLHL12 was investigated. The Kelch domain of KLHL12 contains six cysteines, two of which, Cys558 and Cys511, are located on the surface of the binding site [highlighted in green, Fig. 4.9 (a)]. Structural analysis of the KLHL12-DVL1 complex suggests that covalently targeting either of these cysteines should inhibit peptide binding. The suitability of these cysteines for covalent modification depends on their accessibility and reactivity. While their accessibility is evident, their reactivity required further evaluation. To predict the likelihood of these cysteine residues forming covalent bonds, an *in silico* analysis was performed using ICM Pro software (Molsoft). The tool assessed various parameters, including the distance to the protein surface (distToSurface), solvent-accessible surface area of the cysteine side chain (saSC), reactivity score (a higher score indicates greater reactivity), proximity to positively charged residues (nearNPlus), distance to histidine residues (nearHip), and electrostatic potential at the sulfur atom (esp1) [Table 4.2]. Two of the cysteines were predicted to be reactive - Cys539 on the side of the β-propeller fold, and Cys511 in the known substrate binding site [Fig. 4.9 (b)].

The prediction of two reactive and accessible cysteines was also corroborated experimentally by in-

cubation with the alkylating agent iodoacetamide (IAM), that irreversibly reacts with all reactive cysteines to form a 58 Da covalent adduct. After four hours of incubation with IAM (20 μM) all unlabelled KLHL12 had been converted to the mono- or bi-labelled adduct (+58 Da and +116 Da respectively). After 24 hour incubation IAM had likely induced KLHL12 denaturation and all six cysteines of the protein had been labelled [Fig. 4.9 (c)]. Taking both experiments into account, it can be inferred that KLHL12 has two tractable cysteines for covalent inhibition with at least one of them being on the peptide recognition site allowing for competitive inhibition.

Table 4.2: Cysteine Reactivity Prediction in the KLHL12 Kelch domain

Cysteine	distToSurface	saSC	NearNplus	nearHip	esp1	score	reactivity
C348	0.09	1.40	4.44	9.81	-5.85	0.04	
C563	0.13	7.30	9.93	12.72	-3.22	0.14	
C558	0.00	8.39	4.29	9.05	-10.94	0.34	
C430	0.00	8.64	4.66	10.66	1.30	0.43	
C539	0.00	20.09	5.39	16.08	-4.83	0.66	reactive
C511	0.00	17.73	4.97	3.56	-2.23	0.78	reactive

4.3.2 Validation of covalent hits following AlphaScreen selection

An in-house collection of 2,000 covalent compounds compiled from various cysteine-targeting covalent warheads including acrylamides, epoxides, vinyl sulfones, nitriles and chloroacetamides [Fig. 4.10], had been screened in two concentrations (100 μM and 12.5 μ M) for their ability to displace DVL1 peptide (as outlined previously in AlphaScreen section 4.2.2) by colleague Ms Karly Buchan. After analysing these KLHL12 hits for their selectivity against historical data for KLHL20 and KLHL6 inhibition[1], I selected a subset of potent and relatively non-promiscuous hits to orthogonally validate using mass spectrometry. Most of the selected compounds were acrylamides, while nitrile, epoxide and vinyl sulfones were also present, as summarised in Fig. 4.11.

Intact mass spectrometry analysis of adduct formation on purified KLHL12 Kelch domain was the assay of choice to analyse covalent labelling, because both potency (quantification of % labelling) and selectivity (number of protein adducts formed) could be assessed. Since the compound collection

[1]Homogeneous Time-Resolved Fluorescence (HTRF) assays displacing substrate peptides performed by Ms Karly Buchan

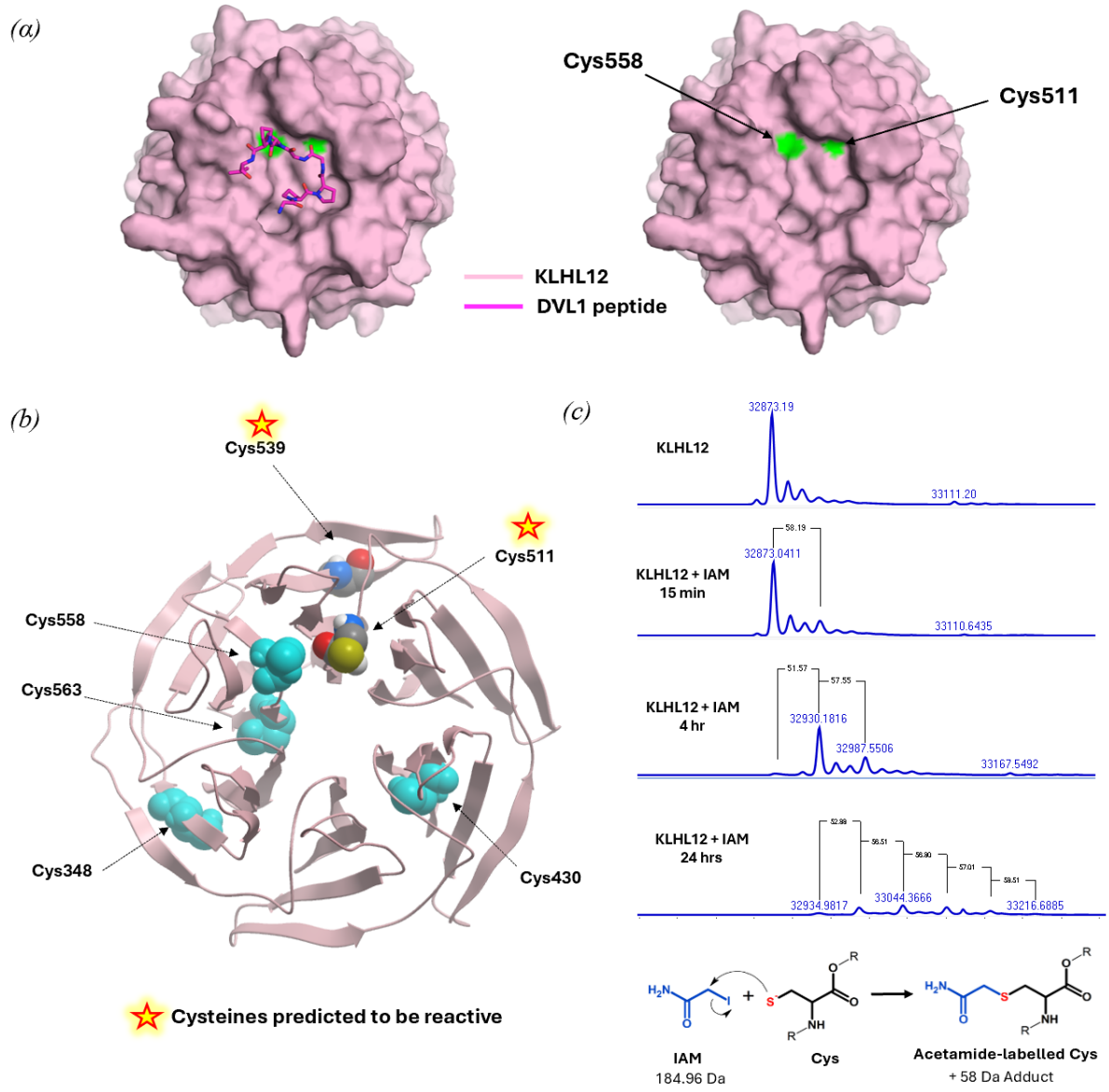


Figure 4.9: **Assessment of KLHL12 cysteine reactivity**. (a) DVL1 peptide (in magenta) on the binding site of KLHL12. The two cysteines on the surface of the binding site are coloured in green. PDB: 6TTK [72]. (b) Computational prediction of KLHL12 cysteine reactivity done in ICM-Pro (Molsoft). KLHL12 cysteines are presented as cyan spheres. Cysteines predicted to be reactive are denoted with a yellow star. (c) Experimental assessment of KLHL12 cysteine reactivity following incubation with iodoacetamide (IAM, 20 µM). Samples were analysed by mass spectrometry, and the deconvoluted mass spectra are presented. Covalent modification of KLHL12 by IAM results in a mass shift of +58 Da per labeled cysteine, consistent with a nucleophilic substitution (S_N2) reaction. The reaction mechanism is illustrated below the mass spectra.

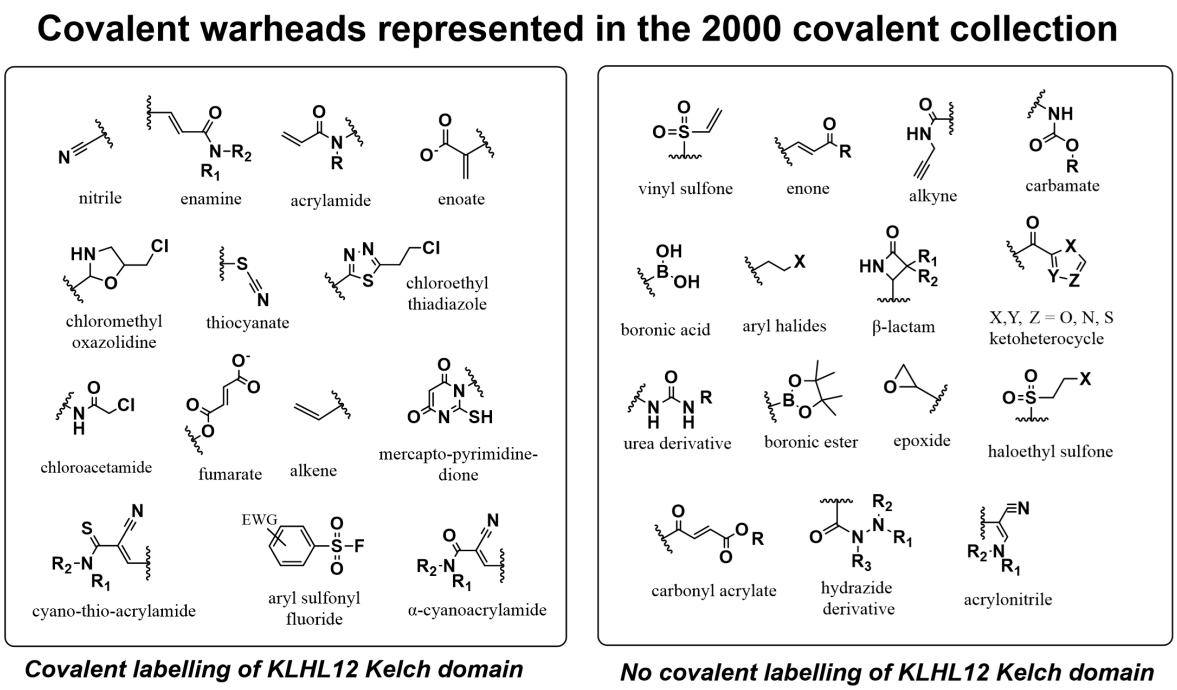


Figure 4.10: **Chemical diversity of covalent warheads in the screened collection**.

was consisted of fragments, the incubation times extended up to 24 hours and the compounds were in ten-fold excess to the protein (KLHL12 Kelch domain 10 μM, fragment 100 μM). Only compound Z2182010384 formed a measurable mono-protein adduct via nucleophilic addition of the nitrile warhead [Fig. 4.12 (d)]. The covalent reaction progressed slowly, with only 12% of the total KLHL12 protein labeled after 3 hours, and 21% labeled even after 24 hours [Fig. 4.12 (a)]. Notably, the reaction rate and labeling efficiency remained unchanged in the presence of reducing agent (TCEP 0.5 mM), suggesting the cysteine sight is structurally or chemically protected. In an additional experiment, the ability of the compound to chemically react with denatured KLHL12 was assessed; addition of 0.2% formic acid to KLHL12 to induce unfolding prior to compound incubation. Even after 24 hours, no labelling was observed on the unfolded protein [Fig. 4.12 (b)]. These data suggest that compound Z2182010384 exhibited low intrinsic reactivity and selectively targeted a single cysteine residue, likely through affinity, rather than solely relying on its reactivity.

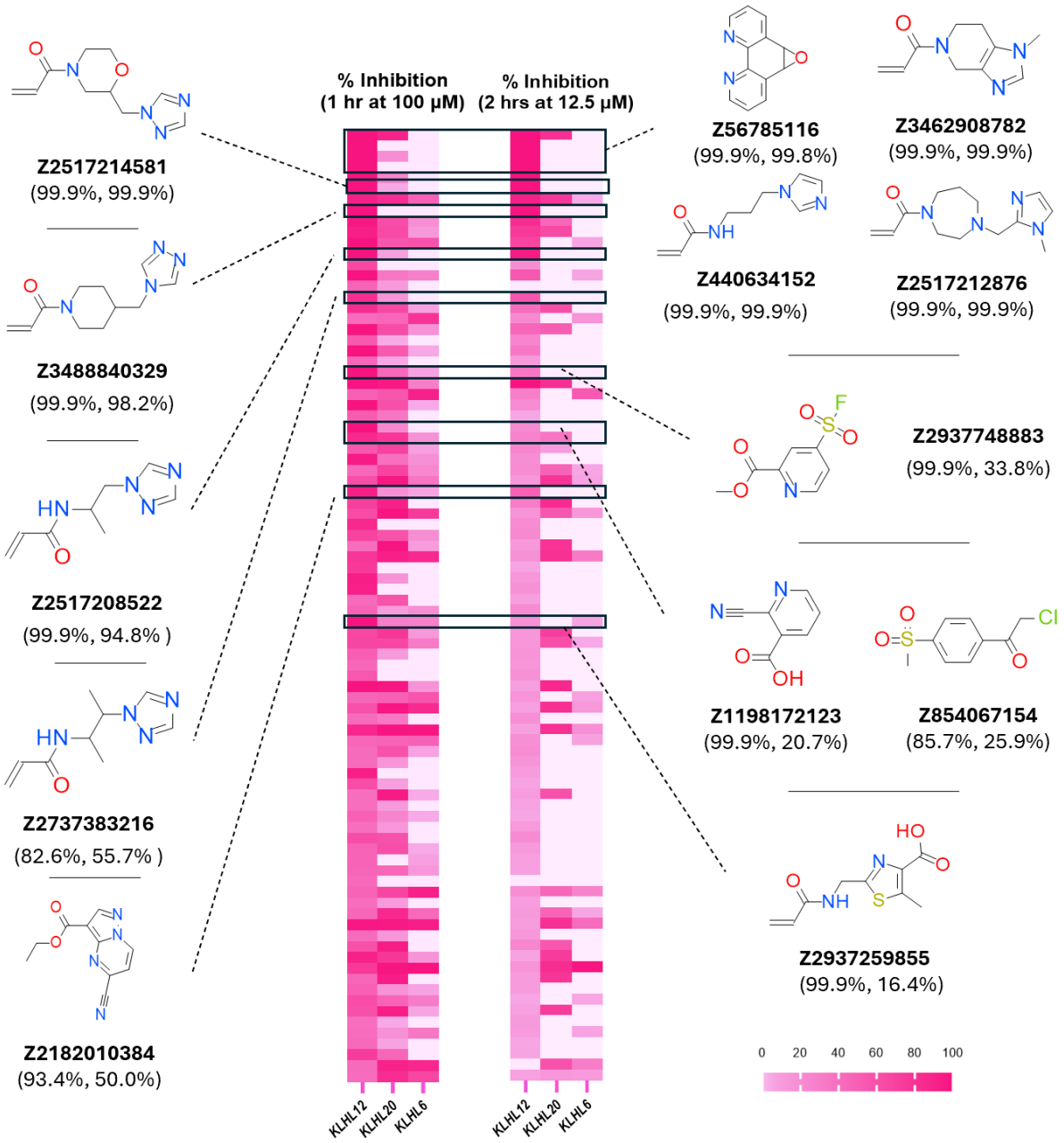


Figure 4.11: **Heatmap of covalent hits targeting KLHL12, KLHL20, and KLHL6**. Summary of historical screening results from an in-house covalent library, performed by colleagues prior to this study. An AlphaScreen assay displacing DVL1 peptide was performed for KLHL12. Homogeneous time-resolved fluorescence (HTRF) assays displacing substrate peptides were performed for KLHL20 and KLHL6. The most potent 88 compounds inhibiting KLHL12 are represented on the heatmaps. Compounds selected for further orthogonal validation are highlighted in black rectangles. In parentheses (%KLHL12 inhibition, %KLHL20 inhibition).

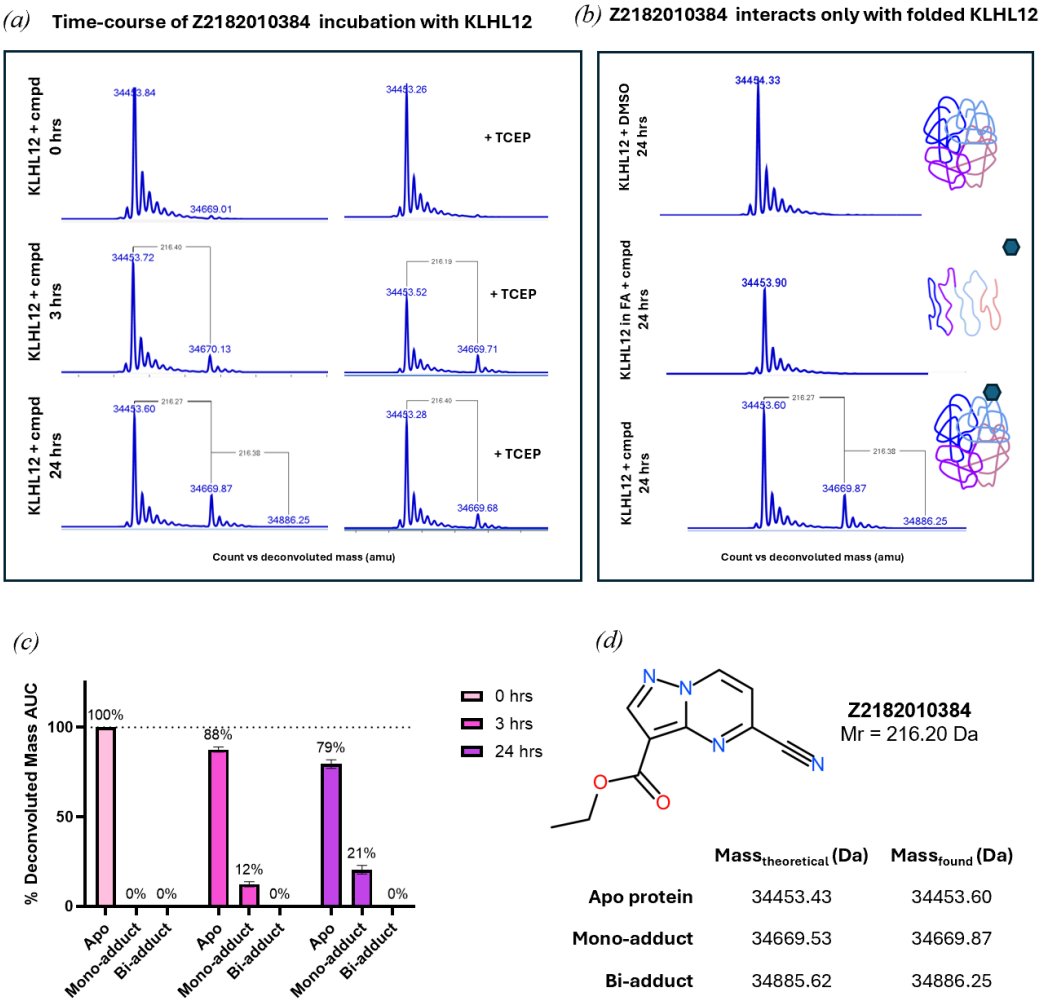


Figure 4.12: **Covalent hit selectively targets one cysteine of KLHL12**. (a) Mass-spectrometry analysis of adduct formation after Z2182010384 incubation with KLHL12 over the course of 24 hrs. Experiment performed in the absence of TCEP (left) and in the presence of 0.5 mM TCEP. KLHL12 at 10 µM and compound at 100 µM. (b) Mass-spectrometry analysis of Z2182010384 adduct formation on folded and unfolded KLHL12. KLHL12 was denatured with 0.2% formic acid before or after compound addition and incubation for 24 hrs. (c) Quantification of adduct formation based on the deconvoluted mass peak's area under the curve (AUC), using three biological repeats. (d) Hit compound Z2182010384 structure and molecular weight.

4.4 Discussion

Based on the insights into the binding site and native interactions of the KLHL12 Kelch domain, as described in the previous chapter, I leveraged the native PGXPP interaction to screen our in-house compound collections for potential small molecule hits. To date, no small molecules have been reported in the literature that bind or inhibit KLHL12. In this chapter, I investigated both covalent and non-covalent modes of inhibition, employing a range of biophysical methods, including direct binding detection and competitive displacement assays. From these efforts, I identified some potential false positive hits using AlphaScreen assay that were not validated with orthogonal assays, alongside a true covalent fragment binder. This covalent fragment represented the first reported chemical matter to target KLHL12 by selectively forming a mono-adduct on the KLHL12 Kelch domain.

No chemical ligand for KLHL12 has been reported so far. However, similar screening and chemical optimisation methods to those that were employed here have also been successful to detect ligands for other Kelch E3 ligases. An initial high-throughput TSA screen, followed up with validation with SPR and TR-FRET assay yielded a selective probe that was further developed into a PROTAC for the Kelch domain of KLHLDC2 [183]. A different approach also managed to ligand KLHDC2 based on computational methods in search of compounds that would mimic the cognate substrate degron, SelK, anchored by its C-terminal di-glycine motif, followed by alphaLISA displacement assay to measure affinity [[182]].

An interesting approach reported by Farrell *et al.* for liganding KLHL20, included the design of a synthetic macrocycle compound [184] that was a cyclised version of the linear LPDLV recognition motif of the protein [222]. This approach, even though an attractive one since we also knew the PGXPP recognition motif of KLHL12, hinged on the initial high binding affinity of the linear LPDLV motif (K_D > 5 μM), that could be further improved by cyclisation to 600 nM by rigidifying the structure and reducing entropic cost [Fig. 4.13 (b)]. In the case of PGXPP binding to KLHL12, the initial binding affinity is moderate (IC_{50}= 11.1 μM), making this approach less suitable.

Structure-based approaches to discover initial fragment chemical matter via crystallographic screens have also been successful for targeting the Kelch domain of ligases - such as the low nanomolar inhibitors of KEAP1 developed from initial micromolar fragments [180, 181]. However, in the KLHL12 apo structure, crystal packing occludes the substrate-binding site, preventing fragment soaking [Fig. 3.1]. An alternative crystal form, potentially accessible via co-crystallisation with a binding partner such as an antibody, nanobody, or single-chain variable fragment (scFv), would be required [223].

To this end, I produced biotinylated KLHL12 Kelch domain, which was used by Opher Gileadi's team (Karolinska Institute) in a phage display screen, yielding nine scFvs with nanomolar affinity [data not shown]. Mapping their epitopes, via AlphaScreen displacement or co-crystallisation, could identify non-blocking binders suitable for generating new crystal forms in the future. A similar approach also utilising naive *in vitro* phage display screens could be used to select nanobodies, which may offer advantages over scFvs for co-crystallisation due to their smaller size, rigidity, and higher stability [224].

My initial attempts to target the KLHL12 binding site with non-covalent inhibitors arose from a machine learning selection of DEL hits in a pilot study in collaboration with ZebiAI. The hits were then prioritised with *in silico* docking, which suggested they would be moderate binders, all weaker than the native DVL1 peptide. Even though the subsequent AlphaScreen assay showed some promising hits, no direct binding was observed in subsequent SPR and GCI experiments. The unusually high affinity (low micromolar) of those initial hits, in combination with hill slopes > 1 hinted to potentially aggregator compounds exhibiting artificial cooperativity. However, since detergent was present in the initial AlphaScreen assay and inspection of melting curves did not show broader or irregular thermal transitions, there was no evidence of compound-induced aggregation.

Upon examination of the hit molecules' structures however an additional mechanism to explain the high false positive rate could hinge upon the existence of 2-picolylamines, (bi)-pyridines and pyrimidines bridged by ethyl amine or ethyl amide linkers in a symmetrical arrangement [Fig. 4.5 (c)].

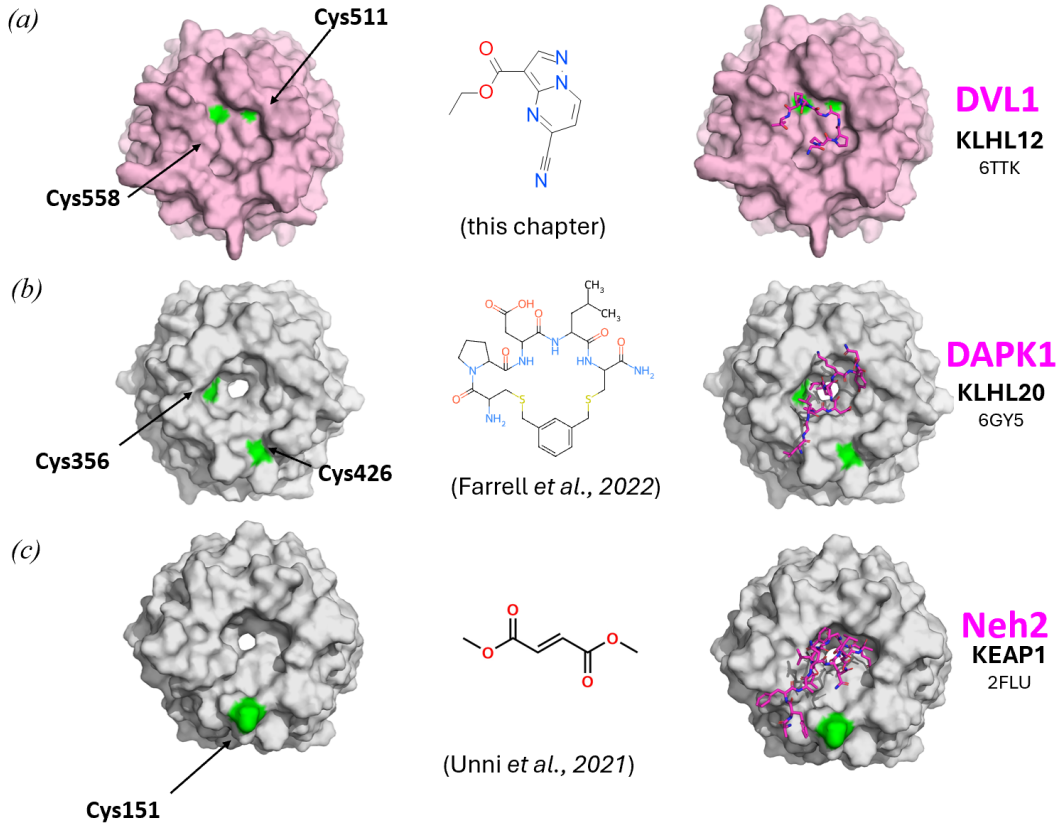


Figure 4.13: **Overview of compounds that competitively inhibit cognate peptide binding on Kelch binding pockets**. (a) DVL1 peptide (in magenta) on the binding site of KLHL12 [PDB: 6TTK [72]]. (b) DAPK1 peptide (in magenta) on the binding site of KLHL20 [PDB: 6GY5 [222]]. Synthetic macrocycle structure that interacts non-covalently [184]. (c) NEH2 peptide (in magenta) on the binding site of KEAP1 [PDB: 2FLU [225]]. Dimethyl fumarate (DMF) structure that covalently targets KEAP1 cysteine [226]. The cysteines on the surface of the substrate binding sites are coloured in green.

These features are frequent in combinatorial libraries that rely on common synthons, and have been recently classified as common AlphaScreen frequent hitters by Schorpp *et al.*, who proposed new substructure filters for them as assay interference compounds [227]. The authors suggest that these moieties can form bifurcated hydrogen bonds with XH-donating groups (where X = N,O) and coordination complexes with metal ions [227]. It is thus possible that these initial hits can chelate both the metal-coated AlphaScreen beads and the tandem imidazole rings of His_6-tagged KLHL12 as shown in Fig. 4.14 (a) and (b) respectively. Adsorption on the metal surface would prevent protein binding to acceptor beads, leading to decreased fluorescence and identification of false-positive hits.

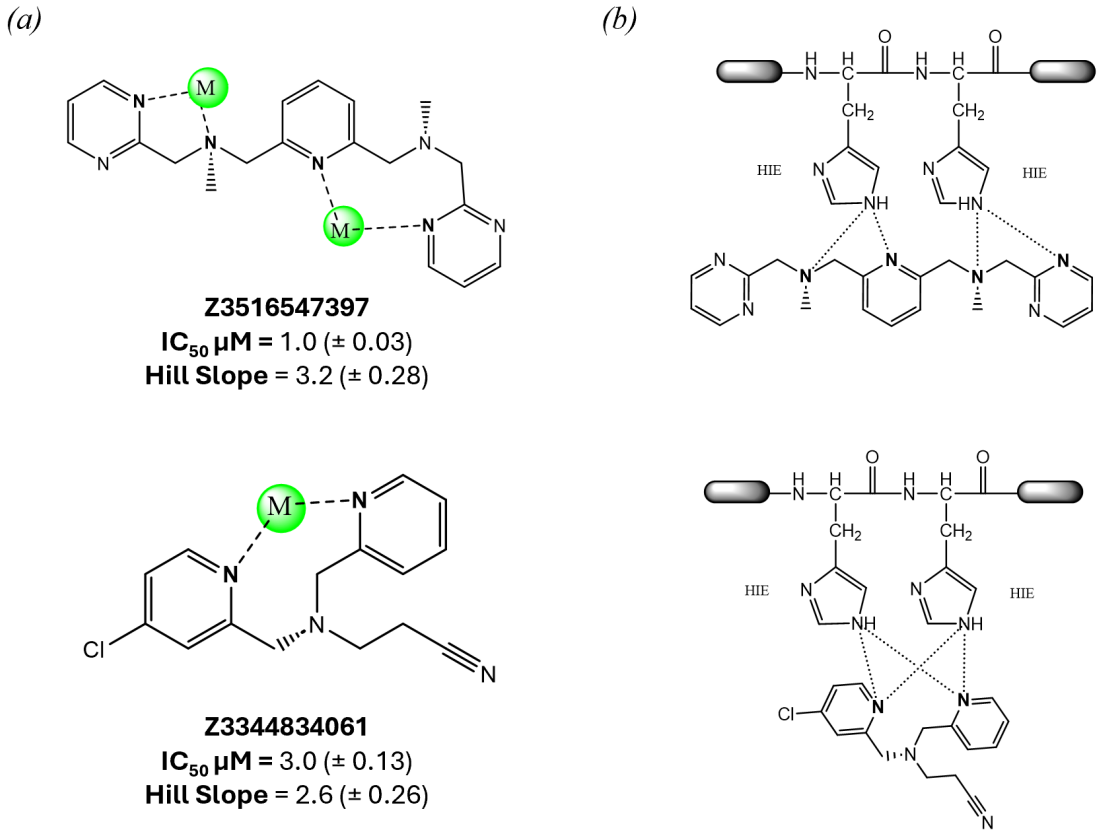


Figure 4.14: **Possible mechanism of metal chelation by false-positive AlphaScreen hits**. (a) Hypothetical chelating complexes formed by the compounds Z3516547397 and Z3344834061 with Ni^{+2} ions on the acceptor bead surface (coordinate bonds represented by dashed bonds). (b) Hypothetical hydrogen bond interactions of the compounds Z3516547397 and Z3344834061 with KLHL12 Kelch domain His_6-tag. Mechanisms adapted from Schorpp *et al.* [227].

Interestingly, no other type of hits emerged from the initial DEL screen. This may reflect intrinsic properties of KLHL12 as a target or could stem from the method used to select hits. ZebiAI applied

a graph convolutional neural network (a machine learning classifier) trained on experimental DEL screening data from proprietary libraries for KLHL12 binding. However, the compounds I validated were not the original experimental hits from these proprietary libraries; instead, they were virtual predictions generated by projecting chemical features (chemical fingerprints) onto the commercially available Enamine REAL library [228]. Because the Enamine REAL library is a large combinatorial collection assembled from commercially reactive building blocks, it is possible that certain structural biases were introduced. For example, the palindromic symmetrical nature of the (P)PGXPP motif recognised by KLHL12 might have caused the machine learning model to favor symmetrical molecules, which may exhibit unintended chelating properties. Therefore, the virtual hits could suffer from bias due to the combinatorial assembly of the library rather than represent diverse, experimentally validated binders which remain unknown as they are proprietary.

A recent study on efforts to discover ligands for 16 members of the WDR family, another class of β-propeller structures [229], also used the same machine learning algorithms to assess DEL hits in collaboration with ZebiAI. The authors reported first-in-class small molecule ligands for 7 of the 16 WDR domains screened, thus exhibiting < 50% success rate [230]. This pilot DEL hit finding campaign could therefore be more useful in family-wide screening campaigns to increase the possibility of a validated compound.

As previously described for binding of cognate peptides in **Chapter 3.3**, KLHL12 affords a shallow broad surface, rather than a deep groove, to accommodate the U-shaped expanded conformation of PGXPP peptides. Both surface and peptide are preformed and rigid, with no conformational change to induce fit upon binding. Even the physiological DVL1 substrate has weak to moderate affinity (IC_{50} of 11.1 μM for the best peptide binder), making the chance of identifying a more potent hit from a naive screen unlikely. The properties of the substrate-binding pocket combined with the high thermostability of the KLHL12 Kelch domain (Tm = 71°C) made common hit-finding techniques, such as TSA, less suited for this protein. For comparison, the cognate ligand only gave 2.55 °C

stabilisation and the average hit only 1 °C, in contrast with a literature example of a similar campaign to ligand KLHDC2 Kelch domain whose cognate peptide SELK elicited a 15 °C shift [183].

To circumvent these issues, covalent inhibition was explored as an alternative approach to targeting the shallow binding pocket of KLHL12, aiming for covalent inhibition of the single predicted reactive cysteine in the binding pocket. Screening of an in-house compound collection afforded a nitrile-containing compound hit that exhibits a favourable selectivity profile; even at the high concentration of 100 μM 93.4% inhibition of KLHL12 is observed in contrast with only 18.3% KLHL20 inhibition and negligible inhibition of KLHL6 (5.0%). Lowering the concentration at 12.5 μM improves the selectivity further with no observed inhibition for KLHL20 and KLHL6, while the inhibition of KLHL12 remains significant (49.66%) [Fig. 4.7].

Interestingly, even though almost complete KLHL12 inhibition is observed after 1 hour of incubation in AlphaScreen displacement assays, there is only 12% of labelling observed in the MS experiments even after 3 hours of incubation [Fig. 4.12]. This could be explained by the different protein:compound ratio, suggesting the thousandfold excess of compound in the former case is pushing the reaction to completion compared to the tenfold excess employed in the latter case (100 nM:100 μM for AlphaScreen assay and 10 μM: 100 μM for MS).

An alternative hypothesis is that the compound is recognised and tethered to KLHL12 with non-covalent interactions and occupies it before the covalent addition is precipitated, suggesting that the interaction is affinity driven rather than reactivity driven. This hypothesis is also supported by the observation that Z2182010384 does not label the reactive cysteine of the denatured unfolded KLHL12 even after 24 hours of incubation [Fig. 4.12 (b)], suggesting reactivity alone is not the driving force of the interaction. A relevant observation is that among the covalent library screened - which included more reactive and promiscuous warheads like acrylamides and chloroacetamides - the only validated hit was a nitrile, a compound with significantly lower intrinsic reactivity [231]. In contrast, a recent covalent probe targeting Cys151 residue of KEAP1 BTB domain bears the more potent electrophile

aryl-sulfonyl moiety that acts as a good leaving group [179].

Covalent probes have gained prominence for their ability to access challenging protein targets beyond the reach of non-covalent ligands. Numerous studies have thus aimed to map reactive, and potentially ligandable, cysteines across the human proteome. Seminal work by the Cravatt group developed the isoTOP-ABPP (isotopic Tandem Orthogonal Proteolysis-Activity-Based Protein Profiling) chemoproteomic platform, combining an iodoacetamide-based probe with biotin enrichment and isotopic tags for quantitative mass spectrometry analysis [232, 233]. This foundational method has since inspired many variants employing diverse electrophiles to identify reactive cysteines amenable to covalent liganding [233–242].

Although these studies collectively identified thousands of reactive cysteines and highlighted several in Kelch domains (such as KEAP1, KLHDC3, KLHL20, KBTBD4), KLHL12 was not globally or strongly labelled. KLHL12 BACK domain cysteines C161 and C173 were detected but not enriched relative to DMSO controls in an initial broad ABPP screen using 285 electrophiles in HCT116 and HEK293T cells [241]. However, more recent studies with improved proteome coverage and sensitivity, including targeted assessments of E3 ligase ligandability in the context of targeted protein degradation, revealed significant enrichment of KLHL12 BACK domain C185, modest enrichment of C123, C147, C248, and low enrichment of Kelch domain C348 and C558 [234, 240, 243, 244]. Notably, the relevant Cys511 was moderately enriched only in one recent comprehensive study by Yang *et al.*, which profiled a total of 38,450 cysteines from 8,274 human proteins using 192 different electrophilic compounds [234].

Mass spectrometry analysis in this chapter demonstrated that compound Z2182010384 (hereafter **EVD-1**) covalently modifies only one of two computationally predicted reactive cysteines on the surface of KLHL12 Kelch domain, via its heteroaryl nitrile warhead. Given its strong displacement of DVL1 from the binding site, **EVD-1** is likely targeting Cys511, as the only predicted reactive cysteine within the Kelch domain binding site. The modest enrichment of Cys511 in global chemoproteomic

screens may reflect its lower intrinsic reactivity or limited surface accessibility in the absence of a ligand, suggesting that **EVD-1** engages this site selectively through structural complementarity rather than generalized cysteine reactivity. This suggests that **EVD-1** functions as a selective covalent probe for KLHL12, capable of targeting functionally relevant, but not hyperreactive, cysteine residues. The hit molecule described here will serve as a starting point for chemical probe development. The further development and characterisation **EVD-1** and efforts for its elaboration will be described in the following chapter.

CHAPTER 5

Targeting KLHL12: Chemical Probe Development

5.1 Introduction

This chapter describes the chemical and structural characterisation of covalent fragment Z2182010384, hereafter referred to as **EVD-1**, identified in **Chapter 4** as a selective binder to the substrate-binding site of KLHL12. No small-molecule ligands targeting the Kelch domain of KLHL12 have been reported to date, although several BTB-Kelch family members, such as Keap1, have been successfully targeted by covalent and non-covalent ligands [176, 180, 245], as summarised in Table 1.3. KLHL12 is known to regulate Wnt signalling through ubiquitination of Dishevelled (DVL1) [89], making it a mechanistically tractable but chemically underexplored E3 adaptor.

To optimise initial hit **EVD-1**, a hit-to-lead strategy was employed combining structure-based and ligand-based computational methods guided by crystallography. First, co-crystallisation studies with KLHL12 confirmed covalent engagement of C511 by **EVD-1**, consistent with the hypothesis of this residue being uniquely reactive in the substrate-binding site (see Section 4.4). The binding pose informed subsequent molecular docking of close analogues, used to predict relative binding orientations and rank synthetic or commercial analogues by predicted affinity. In parallel, ligand-based pharmacophore screening was employed, a method widely used for scaffold hopping, to identify iso-functional molecules with similar spatial arrangements of key features but different molecular backbones or scaffolds in order to expand beyond **EVD-1**'s chemotype [246].

Given the time and resource costs of early synthetic campaigns, I prioritised catalog-based structure

activity relationship (SAR) by screening **commercially available analogues** for KLHL12 binding and substrate displacement. These assays were based on covalent mass spectrometry and AlphaScreen methods established in **Chapter 4**. Additional compounds from Dr. Monique Mulder's **proprietary library** (Leiden University Medical Center) were tested by Francesca d'Amico using the same protocols, expanding our SAR around **EVD-1**. Subsequently, a synthetic route was developed to prepare **non-commercial analogues** extending from **EVD-1**. Structural analyses revealed potential exit vectors that could be leveraged for heterobifunctional degraders (e.g. PROTACs), prompting an exploration of synthetic feasibility for linker attachment. Extending across such vectors is a recognised strategy in fragment-to-degrader transitions [247, 248].

The rationale for this strategy was informed by precedent in targeting other E3 ligase substrate receptors. Covalent and reversible inhibitors of Keap1's Kelch domain have been developed using structure-guided methods [175, 249], and scaffold-hopping from teratogenic thalidomide to lenalidomide and pomalidomide led to immunomodulatory cereblon binders with improved therapeutic index [169, 250]. Recent docking-guided optimisation has also yielded potent WD40 inhibitors for DCAF1, a substrate receptor of CRL4 [135]. These efforts collectively validated the integrated strategy employed in this chapter, combining structural biology, cheminformatics, and medicinal chemistry to advance KLHL12-targeting ligands.

5.2 Results

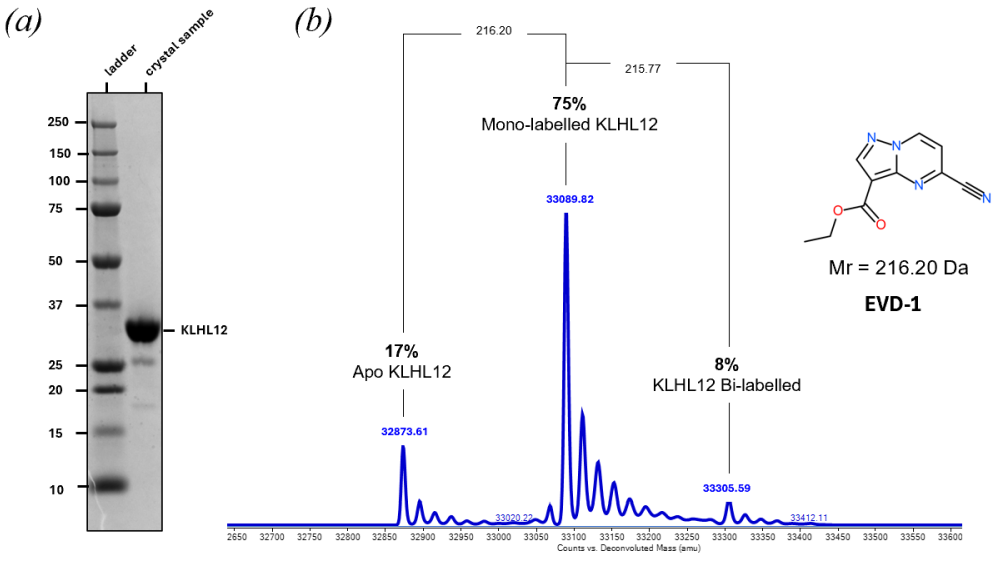
5.2.1 Structure determination of KLHL12-covalent hit complex

As previously discussed [**Chapter** 3.2.1] the crystal form of KLHL12 apo crystal does not allow for soaking, thus compound co-crystallisation was needed. The KLHL12 Kelch domain (M268-E567) was expressed in *E. coli* and captured using Ni^{2+}-sepharose resin. The N-terminal hexahistidine tag was cleaved using TEV protease and the protein was further purified by size exclusion chromatography and a second Ni^{2+}-sepharose resin to remove nickel-binding contaminants [Fig. 5.1 *(a)*] .

The sample was concentrated to 12.6 mg/mL (382 µM). Previous studies shown in Fig. 4.12 *(a)*

showed low labelling efficiency - less than a quarter of the protein was labelled after 24 hours even at a tenfold compound excess. Thus, the concentrated protein was incubated with three-fold excess of compound at 1 mM (2% DMSO concentration which is tolerated see Fig. 4.4).

After 3 hours of incubation at 4 °C, the formation of covalent adduct was confirmed with mass spectrometry. Interestingly, a higher labelling efficiency was observed with 75% of the protein being labelled once and 8% twice, possible precipitated by the higher concentrations [Fig. 5.1 *(b)*]. At a final protein concentration of 10.1 mg/mL, coarse crystal trials were set up using the vapour diffusion method with sitting drops, as summarised in Fig. 5.1 *(c)*.



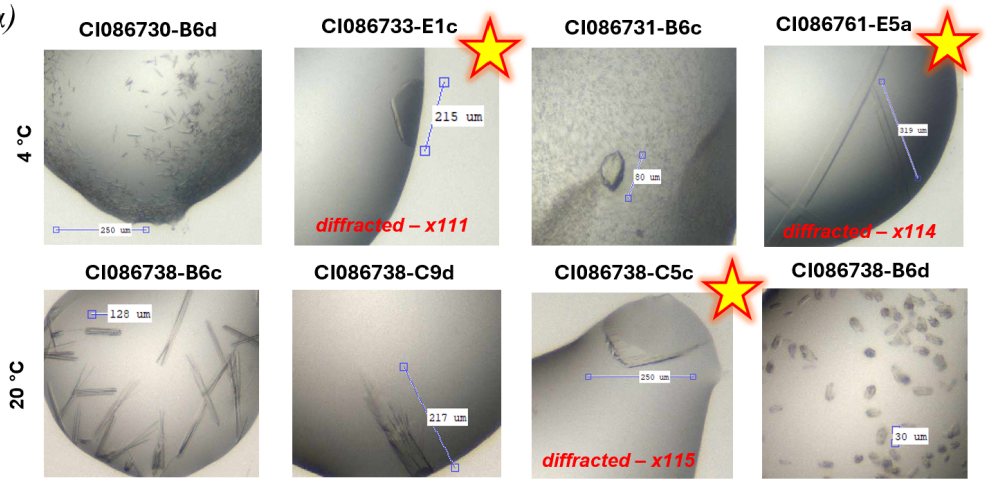
(c)

Plate Identifier	Protein Conc. (mg/ml)	Cmpnd	Coarse Screen	Temp (°C)	Hit Wells
CI086734	10.13		BCS5	20	-
CI086731	10.13		BCS5	4	B6c
CI086763	10.13		HCS3	4	-
CI086737	10.13		HCS3	20	-
CI086738	10.13		HIN3	20	B6d, C5c, C9d, E1a
CI086730	10.13		HIN3	4	B6d
CI086417	10.13		JCSG+ ECO	20	-
CI086733	10.13		JCSG+ ECO	4	E1c
CI086762	10.13	1 mM	LFS ECO	4	-
CI086735	10.13	**EVD-1**	LFS ECO	20	-
CI086761	10.13		MIDASPLUS	4	E5a
CI086739	10.13		Morpheus	4	-

Figure 5.1: **Co-crystallisation trials of KLHL12 and covalent hit** (a) SDS-PAGE analysis of KLHL12 crystal sample, (b) mass spectrometry analysis prior to crystal trials shows primarily mono-labelled KLHL12, (c) summary of coarse screen in the presence of 1 mM **EVD-1**

Crystals of various morphologies ranging from fine needles to large plates were obtained at both 20

(a)



(b)

Identifiers	Conditions	Space Group	Outer Shell Resolution (Å)	Cell Dimensions	Comments
x111	1M Sodium citrate tribasic dihydrate, 0.1M MES pH6.5	P 21 21 21	2.76 - 2.47	142.39, 150.98, 179.18, 90.00, 90.00, 90.00	moderate anisotropy, selected for refinement
x114	25% w/v SOKALAN® CP 45	P 2 2 21	7.29 - 7.42	96.07, 139.87, 237.61, 90.00, 90.00, 90.00	poor resolution, Wilson factor > 200
x115	60% tascimate	P 21 21 21	3.05 - 3.32	139.97, 151.19, 178.47, 90.00, 90.00, 90.00	moderate anisotropy

Figure 5.2: **Representative crystal hits of EVD-1 in complex with KLHL12.** (a) Representative crystals of **EVD-1** co-crystallised with KLHL12 from coarse screens (b) Summary of diffraction datasets collected. Best diffracting crystals denoted with yellow stars.

°C and 4 °C as summarised in Fig. 5.2 *(a)*. Diffraction data of selected crystals were then obtained by unattended data collection on the I03 beamline at Diamond Light Source, Didcot, UK. Three diffracting datasets, namely x111, x114 and x116, originated from crystals from three reservoir solutions of a JCSG+ ECO, MIDAS+, and HIN3 coarse screens respectively [Fig. 5.2 *(b)*]. The numerous crystals originating from the synthetic polymer SOKALAN (i.e. x114), yielded poor diffraction datasets, at > 7 Å resolutions. After initial processing, the best diffracting dataset x111 that grew from a solution of 1 M sodium citrate tribasic dihydrate, 0.1 M MES pH 6.5 at 4 °C was selected for further structure refinement.

The diffraction dataset exhibited some anisotropy, similar to the KLHL12-PLEKHA4 structure solved previously, but no other significant pathologies. The combined datasets of two crystal sweeps were indexed, integrated, scaled and merged by the 2x xia2 3di (multi) Diamond automated pre-processing pipeline. Molecular replacement was performed by Phenix.phaser-MR using the apo KLHL12 structure [PDB: 2VPJ] as a search model. The crystal was phased in P 2_12_1 2_1 space group with nine protein molecules in the ASU [Fig. 5.3 *(a)*].

The asymmetric unit (ASU) contains seven protein molecules arranged in a planar ring formation, while two additional protein molecules, chains A and C, extend axially above and below the plane at an angle [Fig. 5.3 *(a)*]. The space group $P2_12_12_1$ features three 2_1 screw axes, each aligned with one of the three crystallographic axes. Each screw axis combines a 180° rotation (2-fold symmetry) with a translation of half the unit cell length along the corresponding axis. This symmetry is illustrated in Fig. 5.3 *(b)*, where the contents of two unit cells are shown, revealing the helical, upward arrangement of interlocking rings.

Density corresponding to the **EVD-1** could be observed in all 9 KLHL12 protein molecules in the ASU. Notably the compound mediated crystal contacts between the chains in the ASU forming a daisy-chain and stabilising the ring formation [Fig. 5.3 *(a)*]. For example, the hit bound to chain A makes crystal contact with Chain B, and the hit molecule bound to chain B in turn makes crystal

contacts with Chain G.

The structure was refined to 2.59 Å resolution with Phenix.refine. The refinement strategy consisted of grouped B factor restrained refinement in the early cycles, while Translation-Libration-Screw (TLS) and non-crystallographic symmetry (NCS) non crystallographic symmetry, alongside occupancy refinement were incorporated in the later stages. Compound bond restraints were generated with Phenix.elbow and used in every cycle of refinement. The latest cycles of refinement were performed in Buster [208] using individual B-factors, non crystallographic symmetry restrains and TLS. Crystallographic data collection and refinement statistics for the structure are shown in Table 5.2. Further refinement should be performed to address rotamer and Ramachandran outliers that are a consequence of crystal packing between the nine protein chains in the ASU.

5.2.2 Covalent hit EVD-1 selectively targets C511 on KLHL12 binding site

Electron density for **EVD-1** compound was present in all nine protein chain showing covalent labelling of C511 in the substrate binding site of KLHL12 [Fig. 5.4 *(a)*]. There was no density observed near neighbouring cysteine C558. All nine KLHL12 chains and cysteines adopt the same conformation [Fig. 5.4 *(b)*]. The thiol group of C511 has nucleophilically attacked the carbon of the nitrile group forming a thioimidate product. The fragment extends axially away from the binding site, with the ethyl ester group extending to the solvent.

The crystal structure reveals that, in addition to its covalent bond with the protein, the compound is anchored through three hydrogen bonds [Fig. 5.5 *(a)*]. The imino group of the thioimidate adduct serves as both a hydrogen bond donor and acceptor, forming interactions with the amide backbones of Y512 and S464 (2.5 Å and 3 Å respectively). Additionally, the ester carbonyl oxygen acts as a hydrogen bond acceptor, interacting with the side chain of HIE H486 (2 Å)[1]. The planar aromatic scaffold is sandwiched between aromatic residues Y512 and F481, with the pyrazole ring engaging in

[1]A HIE histidine refers to a histidine residue where the εnitrogen (Nε2) of the imidazole side chain is protonated, while the δ-nitrogen (Nδ1) is unprotonated. Protonation states were assigned to the histidine using Maestro "Protein preparation" that selects protonation states based on optimising hydrogen bond network

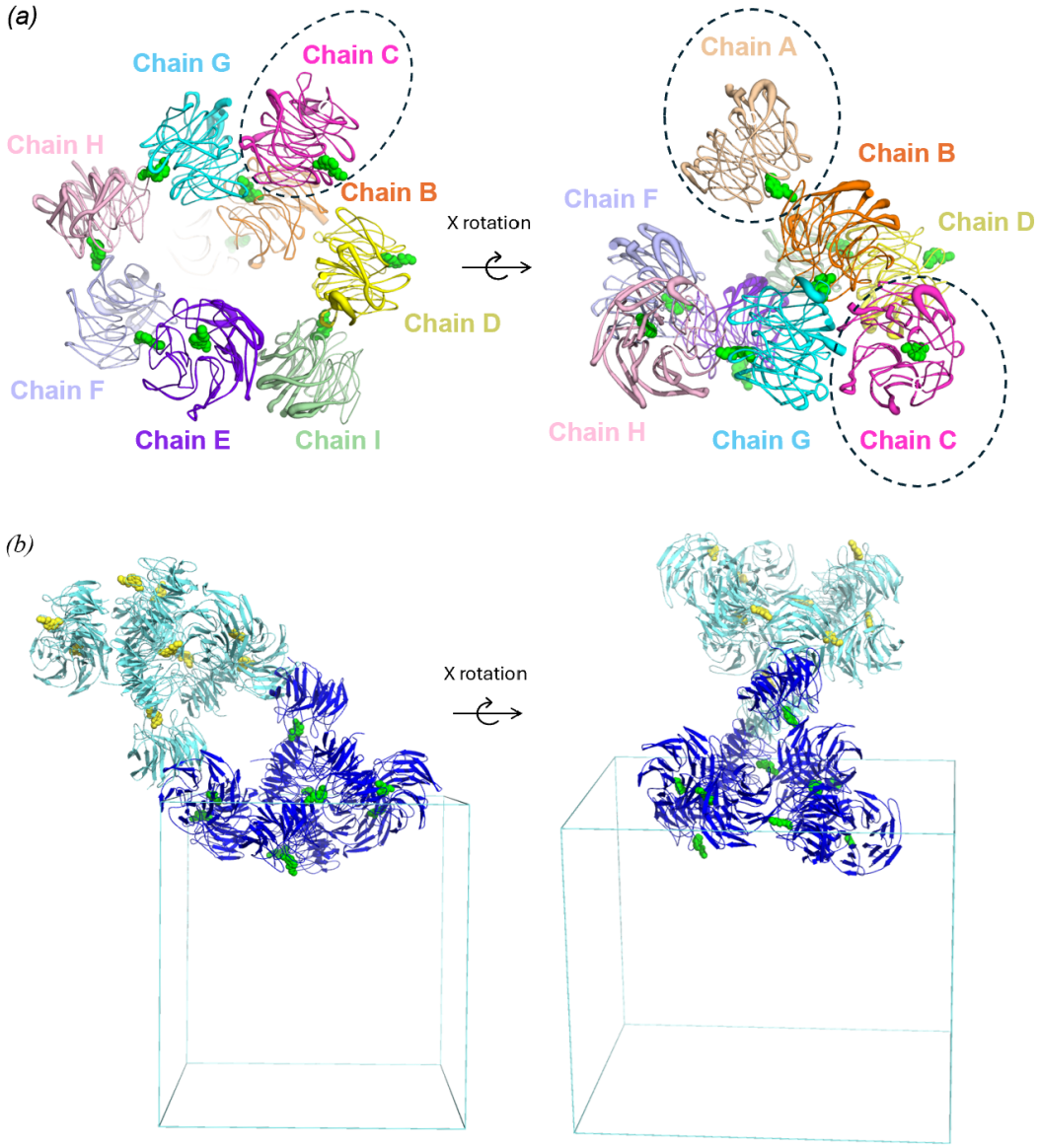


Figure 5.3: **Asymmetric unit of the EVD-1 and KLHL12 co-crystal complex.** (a) Nine protein KLHL12 chains are present in the asymmetric unit; chains named A-I. Proteins chains are visualised in PyMOL using the 'putty' representation; the tube thickness corresponds to B-factors, with thicker regions indicating higher atomic displacement (greater flexibility) and thinner regions indicating lower displacement (greater rigidity). Seven protein molecules are arranged in a planar ring formation, while two additional protein molecules, chains A and C, extend axially above and below the plane at an angle (circled with dashed lines). Compound **EVD-1** is present in all nine KLHL12 protein chains, and is shown in green sphere representation. (b) The unit cell is depicted as a cyan box. KLHL12 protein molecules within a single unit cell are coloured blue, with **EVD-1** represented as green spheres. KLHL12 molecules from adjacent unit cells are colored cyan, with their corresponding **EVD-1** ligands shown as yellow spheres.

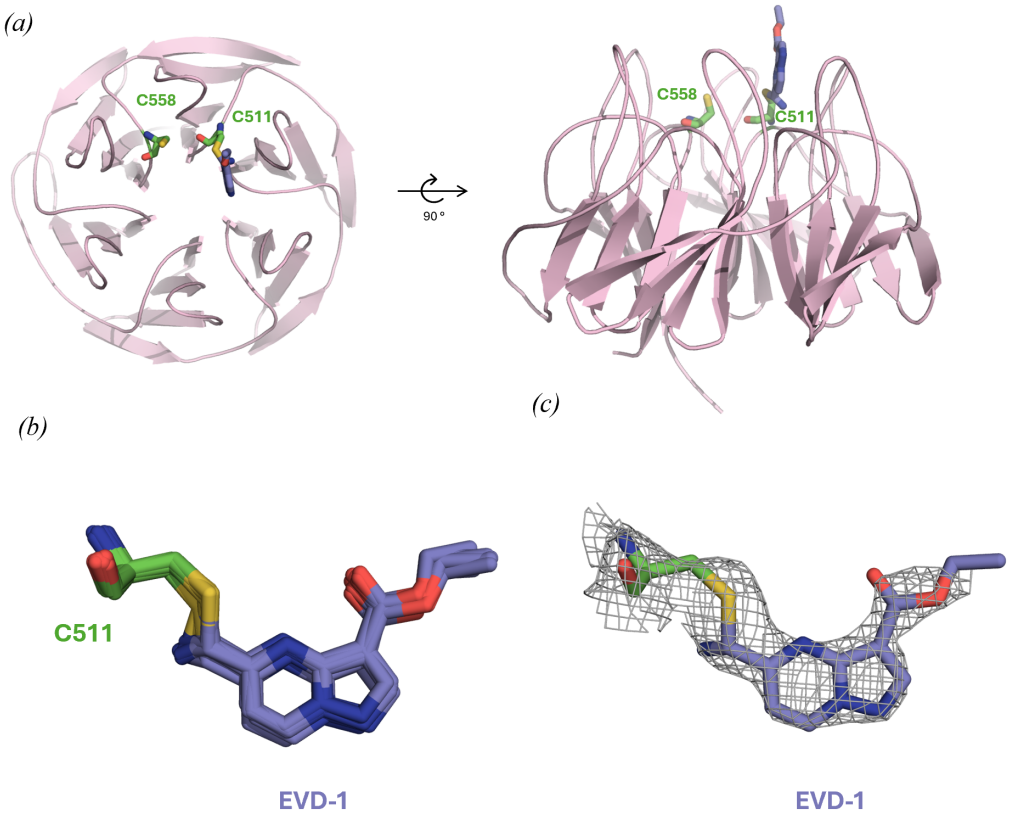


Figure 5.4: **2.6 Å structure of covalent hit EVD-1 bound to KLHL12 Kelch domain** (a) Top and side view of the structure of KLHL12 Kelch (pink) in complex with covalent hit **EVD-1** (coloured in purple). The two cysteines in the binding site are coloured in green. (b) Superposition of the nine **EVD-1** compounds in the same asymmetric unit, stick representation. (c) 2Fo-Fc electron density map (grey mesh) for **EVD-1** of the best resolved chain, contoured at 1 σ.

π-π stacking with F481 and Y512 at distances of 3.9 Å and 3.5 Å, respectively [Fig. 5.5*(b)*]. These distances correspond to the centroid-to-centroid separation between the aromatic rings, indicating favourable stacking interactions within the typical π-π interaction range (3.3 - 4.0 Å). The ethyl group of the ester extends outwards to the solvent and does not contribute to any interactions with KLHL12.

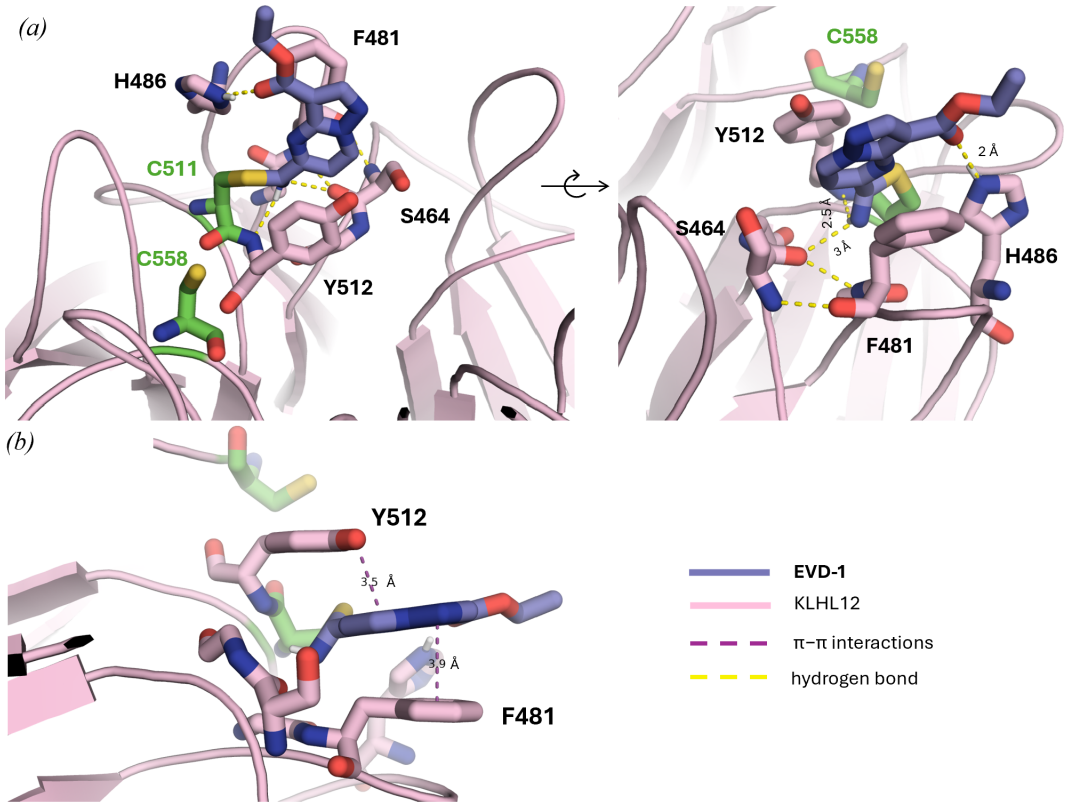


Figure 5.5: **2.6 Å structure of covalent hit bound to KLHL12 peptide binding site.** (a) and (b) Overview of covalent hit (coloured purple) bound to KLHL12 (coloured pink). Cysteines are coloured in green. Hydrogen bond interactions in the KLHL12 substrate binding site are shown by yellow dashed lines. π-π interactions are shown in magenta dashed lines.

5.2.3 *In silico* screening for compound selection

The identified hit served as a starting point for optimization toward the development of a chemical probe. To rapidly explore the chemical space surrounding the fragment, commercial libraries from Enamine were screened to assess the structure activity relationship (SAR) through readily available compounds. Concurrently, libraries synthesized in-house by our collaborators at LUMC were screened to explore potential scaffold hops and identify novel chemical matter.

As an initial step, the libraries were prepared, and conformations were generated following a protocol similar to that described for the ZebiAI hits previously [**Chapter 4.2.1**], utilizing the Maestro LigPrep tool integrated within KNIME. The libraries were first filtered to retain only compounds containing a cyano group, prioritising those with an electron-withdrawing atom in proximity. This filtering was achieved using SMARTS-based workflows in KNIME, based on the hypothesis that the presence of a nearby heteroatom is a key determinant of chemical reactivity. The filtered set of nitrile-containing compounds was subsequently screened using two complementary approaches: *(i)* pharmacophore-based filtering and *(ii)* covalent docking to KLHL12.

(i) Pharmacophore-Based Filtering: The crystal structure of the covalent hit revealed its binding mode and facilitated the formulation of a hypothesis regarding the key features determining activity. The planar aromatic nitrile was essential for forming two hydrogen bonds with Y512 and S464 via the thioimidate adduct. Additionally, the aromaticity of the compounds contributed to π-π stacking interactions, and the carbonyl oxygen of the ethyl ester serves as a hydrogen bond acceptor. These were encoded as pharmacophoric features and used to filter the prepared libraries using Maestro Phase.

(ii) Covalent Docking: Covalent docking was performed at C511. The hit was docked in the apo KLHL12 Kelch domain as a benchmark to assess the model's predictive power. Libraries containing precomputed conformations of the ligands were imported as a chemical table and docked using ICM-Pro. The binding modes of the docked compounds were evaluated to prioritize candidates. Only compounds with docking scores higher than the initial hit, whose binding pose maintained key hydrogen-bond interactions with H486 and Y512, were taken forward.

Finally, compounds identified as hits through both approaches - those satisfying the pharmacophore hypothesis and meeting the binding mode constraints - were selected for experimental testing. These compounds were subsequently assessed for their ability to form covalent adducts with KLHL12 using LC-MS. An overview of the workflow is provided in Fig. 5.6.

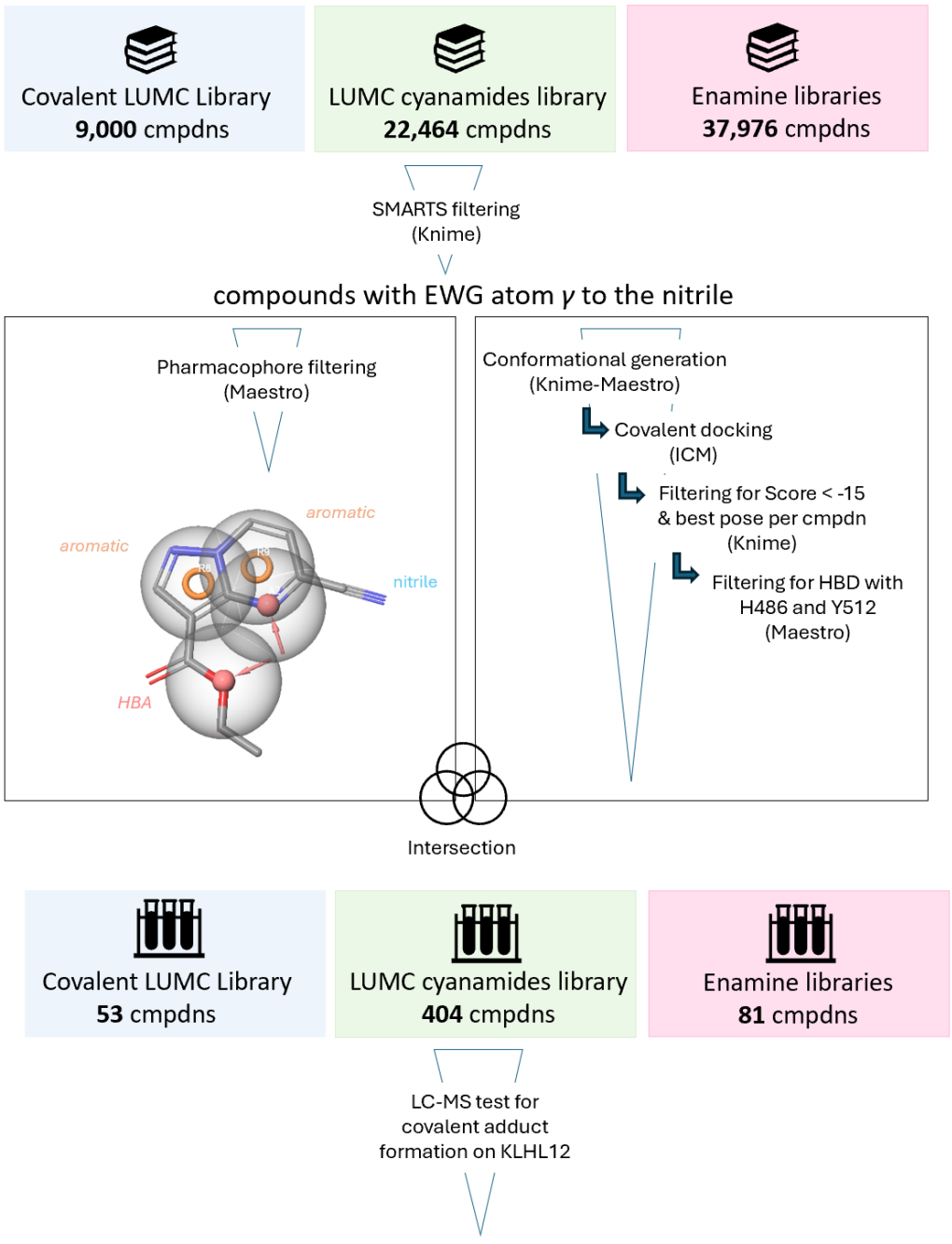


Figure 5.6: **Overview of *in silico* ligand-based screening workflow**

5.2.4 Covalent adduct formation assessment of potential new scaffolds

A total of 53 compounds from the *covalent* library and 404 compounds from the *cyanamide* LUMC library were selected and tested. The LUMC screening team evaluated the compounds using LC-MS to assess covalent adduct formation with KLHL12 after 24 hours of incubation at 100-fold molar excess.

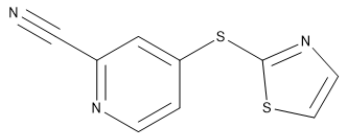
From the 53 compounds tested in the *covalent* library, only one hit, **LUMC-1**, was identified, exhibiting less than 10% covalent adduct formation. **LUMC-1** featured a single-ring pyridine scaffold with a nitrogen atom $\gamma$to the nitrile, similar to the **EVD-1**. Additionally, **LUMC-1** included a nitrogen atom substituting the carboxyl group of the **EVD-1**, but retaining its hydrogen bond acceptor properties [Fig. 5.7 *(a)*]. The low efficiency of covalent labelling suggests that the proximal heteroatom alone is insufficient for binding and that the extended aromaticity of the fused pyrazolopyrimidine ring in the **EVD-1** is critical for binding.

In contrast, from the 404 cyanamide compounds tested, six hits were identified with overall labelling efficiencies below 50%. Among these, **LUMC-2** demonstrated the highest labelling efficiency, reaching 50% adduct formation [Fig. 5.7 *(b)*]. Although the reactivity of these hits was lower than the **EVD-1**, this finding is significant as it represents the first set of high molecular weight compounds, rather than fragments, capable of being accommodated in the binding pocket. Notably, this includes the first non-aromatic nitrile compounds with aminopyrrolidine and aminopiperidine groups.

5.2.5 Covalent adduct formation assessment of hit SAR analogues

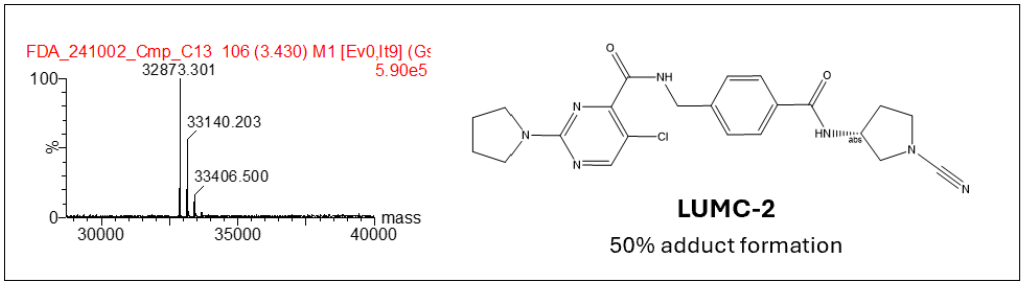
In total 81 compounds obtained from Enamine were assessed for their ability to form covalent adducts on KLHL12 Kelch domain compared to the **EVD-1** - hereafter named **1** to **81**. Compounds were selected to either bear systematic changes to the core and substitutions, or sample chemically neighbouring scaffolds. Thus, initial structure activity relationships by catalog could be established alongside previous attempts to scaffold hop to a molecule with higher labelling efficiency.

(a) **LUMC covalent library pharmacophore hit**



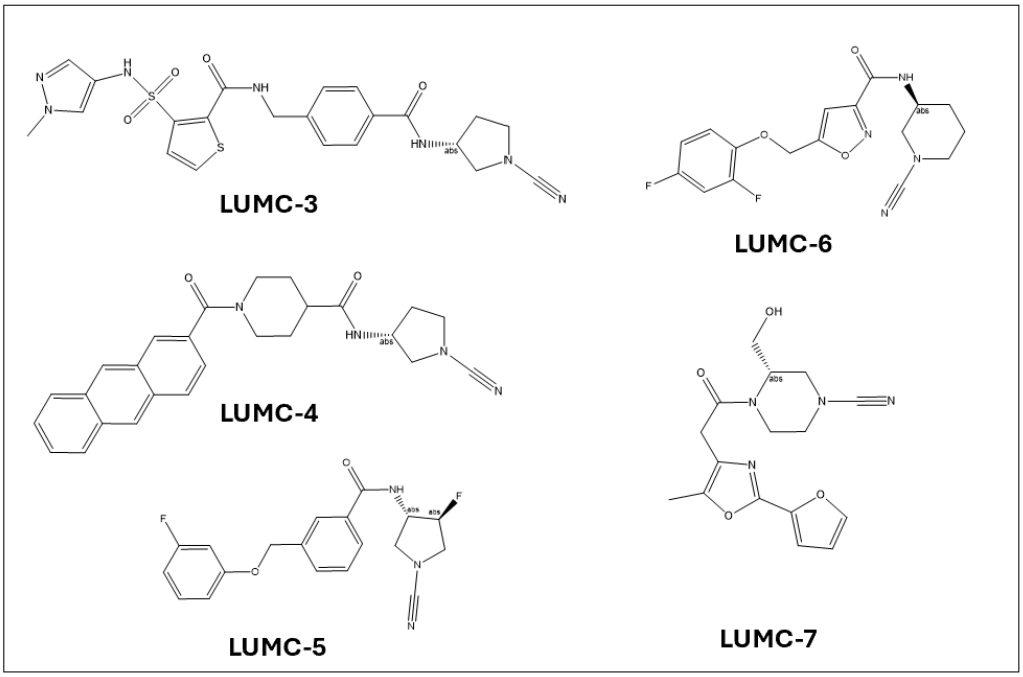
LUMC-1
<10% adduct formation

(b) **LUMC cyanamides library pharmacophore hits**



LUMC-2
50% adduct formation

<15% adduct formation



LUMC-3

LUMC-6

LUMC-4

LUMC-5

LUMC-7

Figure 5.7: **Validated hits from *in silico* screening of LUMC libraries.** A selection of *in silico* hits were validated by LC-MS at the LUMC by Francesca d'Amico. Here are presented those that formed a covalent adduct with KLHL12 Kelch domain. (a) **LUMC-1** was the only *in silico* hit of the LUMC covalent collection that covalently modified KLHL12 Kelch domain. (b) Compound s**LUMC-2** - **LUMC-7** were *in silico* hits of the LUMC cyanamide collection. **LUMC-2** was the only one showing considerable adduct formation, labelling 50% of overall KLHL12 Kelch domain, as quantified by deconvoluted mass peak height.

The reaction efficiency was measured using mass spectrometry monitoring the rate of conjugation in three time points; 10 min, 3 hrs and 24 hrs at room temperature. The formation of the covalent adduct was monitored by detecting a mass shift corresponding to the compound's molecular weight in addition to the protein mass. In parallel, the compound's potency was measured with DVL1 displacement α-screen experiments as previously described, and quantified with time-dependent IC_{50} values.

Overall, 21 of the total 81 compounds formed an adduct with KLHL12 as summarised in Fig. 5.8. The majority of these compounds exhibited slower reaction rates and lower protein labelling percentages compared to the **EVD-1** compound. However, five compounds (highlighted in yellow) demonstrated high labelling efficiency, reacting more quickly and labelling a greater percentage of protein. Interestingly, all five compounds share a common scaffold; a six-membered aromatic ring fused with a five-membered heterocyclic ring that bears the nitrile warhead. This new scaffold is almost an inversion of the **EVD-1**'s cyanopyrazolo[1,5-a]pyrimidine core, where the nitrile warhead is a substitution of the six-membered ring instead.

5.2.5.1 Effect of γ-nitrogen and ethyl ester modifications

A single heteroatom substitution γ to the nitrile, within the fused ring system, is sufficient to abrogate activity. In compound **17** the hit's pyrazolo pyrimidine ring has been replaced by a pyridine ring. This heterocyclic modification rendered **17** unable to label KLHL12 after 3 hours of incubation, in contrast to the hit compound **EVD-1**, which achieved 80% overall labelling. After 24 hours, **17** showed only 11% labelling, whereas the **EVD-1** maintained 78% overall labelling [Fig. 5.9 *(a)*]. This finding suggests that the nitrogen atom γ to the nitrile is essential for reactivity.

The effect of ethyl ester substitution was also investigated. Fragment **20**, which lacks the ethyl ester group, exhibited less than half the labelling efficiency of the hit **EVD-1**, achieving only 32% mono-adduct formation after 3 hours compared to the hit's 76%. This observation was consistent with AlphaScreen assay results, whereby DVL1 peptide is displaced from KLHL12 substrate binding site, as previously described in **Chapter 4.3.2**. The **EVD-1** displayed an IC_{50} of 0.46 µM after 3 hours

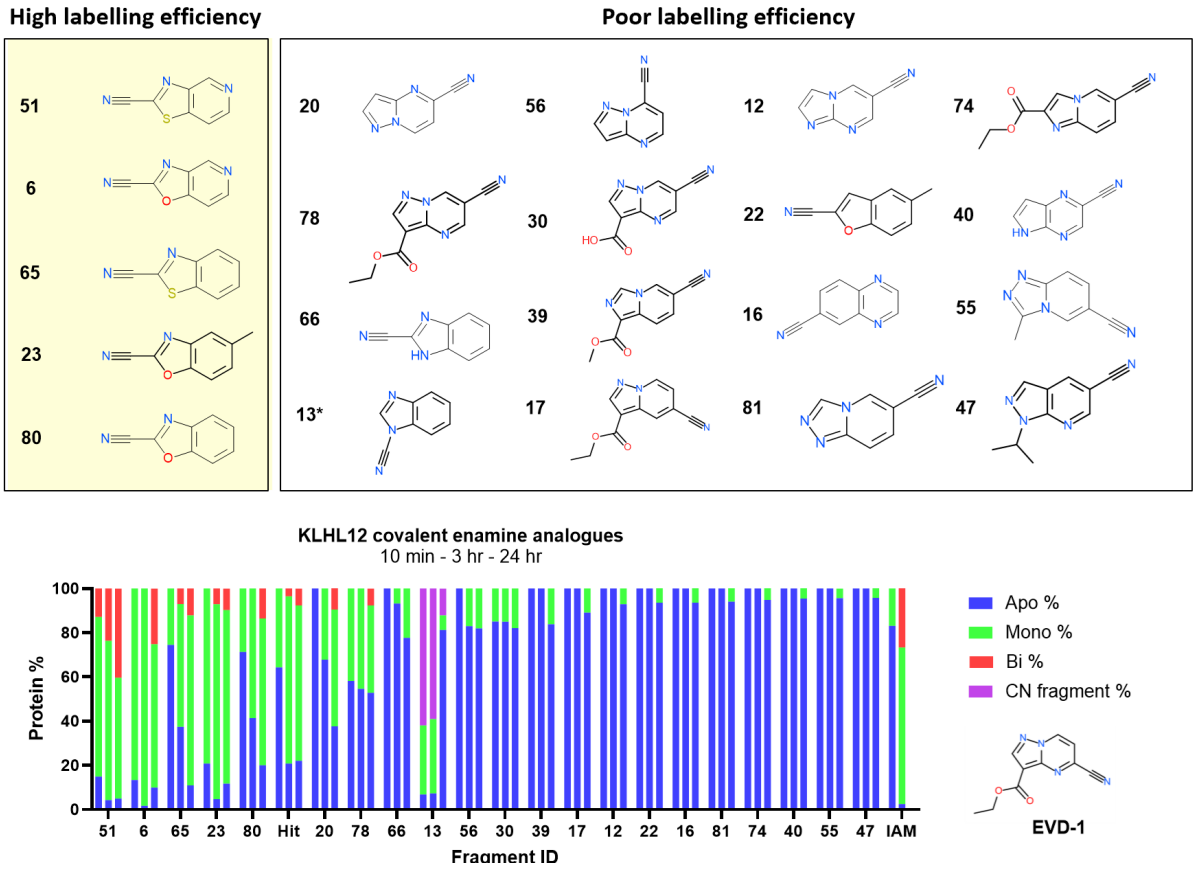


Figure 5.8: **Covalent adduct formation assessment of hit SAR analogues.** Compounds were dissolved in DMSO were dispensed in a plate at 100 µM, at a 50 fold excess of KLHL12 Kelch protein in assay buffer 2 µM in 50 mM HEPES pH 7.5 and 150 mM NaCl. The reaction mixtures were incubated at 25°C before being quenched with 1% formic acid. Measurements were taken after 10 minutes, 3 hours and 24 hours incubation (one independed repeat for each time-point). The percentage of covalent labelling was calculated by comparing the area under the curve (AUC) of the adduct peak to the total protein signal. Extracted ion chromatograms were integrated using Agilent MassHunter software. Compounds' labelling efficiency decreases from left to right. IAM = iodoacetamide, is a reactive alkylating agent for cysteines used here as a positive control. IAM was incubated for 10 minutes and 3 hours only, as incubation of 24 hours negatively affected the stability of the protein. DMSO was used as negative control.

of incubation and 0.16 μM after 24 hours, whereas fragment **20** required 24 hours of incubation to achieve an IC_{50} of 12.6 μM [Fig. 5.9 *(b)*].

Interestingly, despite its reduced overall KLHL12 labelling, fragment **20** showed a slight increase in bi-adduct formation, reaching 10% overall compared to **EVD-1** 8% after overnight incubation [Fig. 5.9 *(a)*]. That suggests that the presence of ethyl ester substitution enhances but is not essential for activity. It would also seem that it's presence might be needed to fine-tune cysteine selectivity possibly through steric hindrance effects.

To evaluate whether the two features - (i) proximal heteroatoms to the nitrile group, and (ii) the ethyl ester - could independently confer affinity and reactivity, or whether their effects were specific to the pyrazolopyrimidine core of the hit, a series of compounds were tested. These compounds retained the positions of the nitrile and ethyl ester groups individually or in combination but varied in their fused cores.

Compounds that retained the nitrile and carboxylate groups but lacked the γ-heteroatom, such as fragments **26**, **70**, **79**, were, as expected, not reactive [Fig. 5.10 *(a)* left panel]. Similarly, compounds lacking both the proximal heteroatom and carboxylate group, regardless of the fused ring system, such as fragments **81**, **55**, **59**, also showed no reactivity [Fig. 5.10 *(a)* central panel]. Interestingly, compounds with various scaffolds that retained the γ-heteroatom relative to the nitrile group but lacked the ethyl ester, such as fragments **15**, and **29** [Fig. 5.10 *(a)* right panel], did not react, unlike compound **20**. This suggests that elimination of the ester group is tolerated only within the hit's specific scaffold.

These findings indicate that the presence of a γ-heteroatom alone is insufficient to confer reactivity across all scaffolds. The pyrazolopyrimidine scaffold of the hit appears to possess unique properties that favor interaction with KLHL12. Furthermore, the ethyl ester is necessary for activity but only in conjunction with a proximal heteroatom, highlighting the interdependence of these features for reactivity.

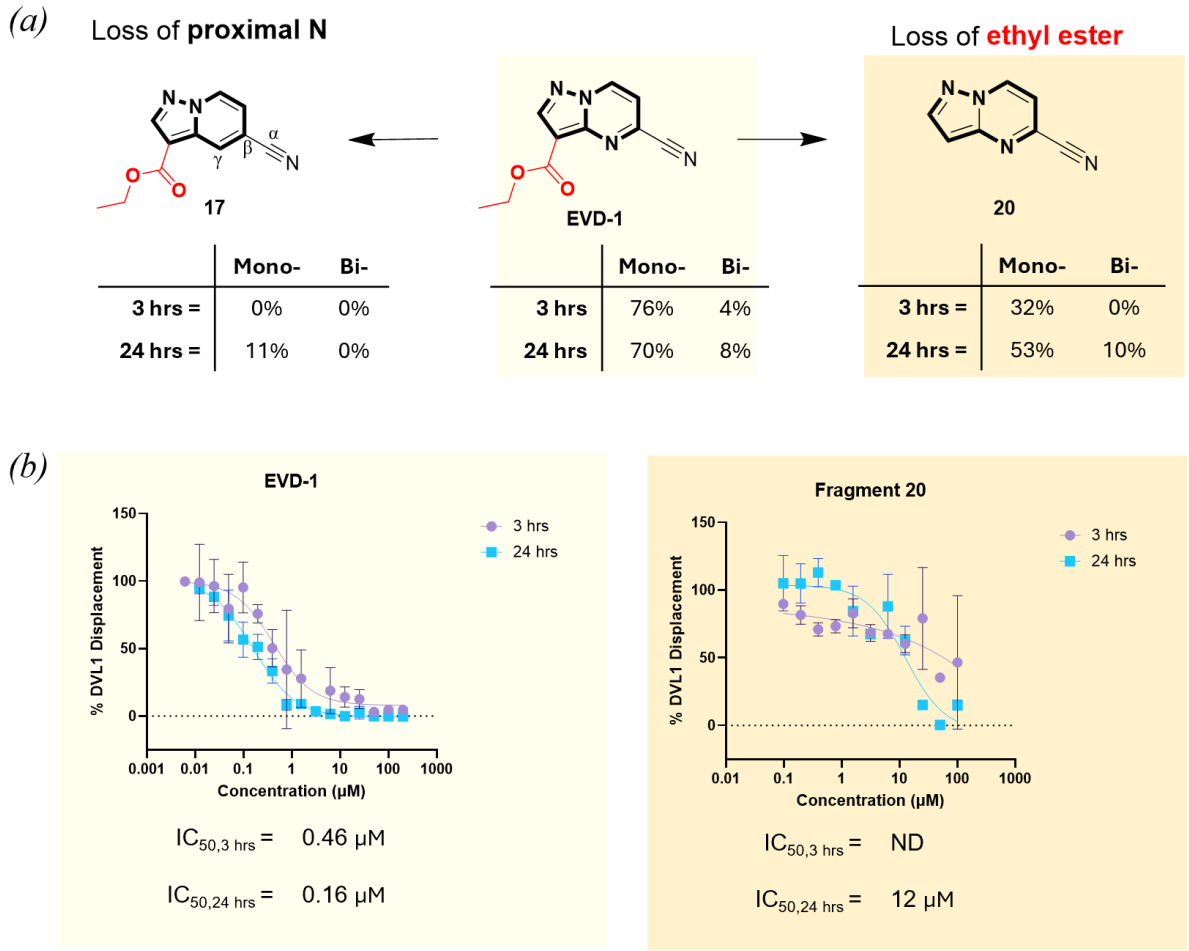


Figure 5.9: **Effect of γ-nitrogen and ethyl ester modifications** Structure-activity relationship (SAR) analysis of compounds assessing the impact of γ-nitrogen and ethyl ester modifications on binding. (a) The covalent reaction efficiency of the compounds on the KLHL12 Kelch domain was measured using mass spectrometry by monitoring the rate of conjugation at 3 and 24 hours at room temperature. Compounds (100 μM) were incubated with KLHL12 Kelch protein (2 μM) at a 50-fold molar excess. The percentage of adduct formation was determined based on the area under the curve (AUC) of the deconvoluted mass peaks relative to each sample, as shown in Fig. 5.8. 'Mono-' refers to mono-alkylation of KLHL12 by the compound, while 'bi-' indicates bi-alkylation. (b) The compounds' potency was assessed using DVL1 peptide displacement from KLHL12 via AlphaScreen experiments, as described in **Chapter 2.5.6**, and quantified through time-dependent IC_{50} titration curves. The %DVL1 displacement was normalised between averages of negative (DMSO, 0% displacement) and positive (DVL1 peptide, 100% inhibition) controls. A non-linear curve that allowed for a variable slope was fitted on logarithmically transformed compound concentrations over normalised response in Graphpad Prism. Error bars represent the ±standard deviation (SD) from the mean of two technical replicates. Compounds were obtained from Enamine Ltd.

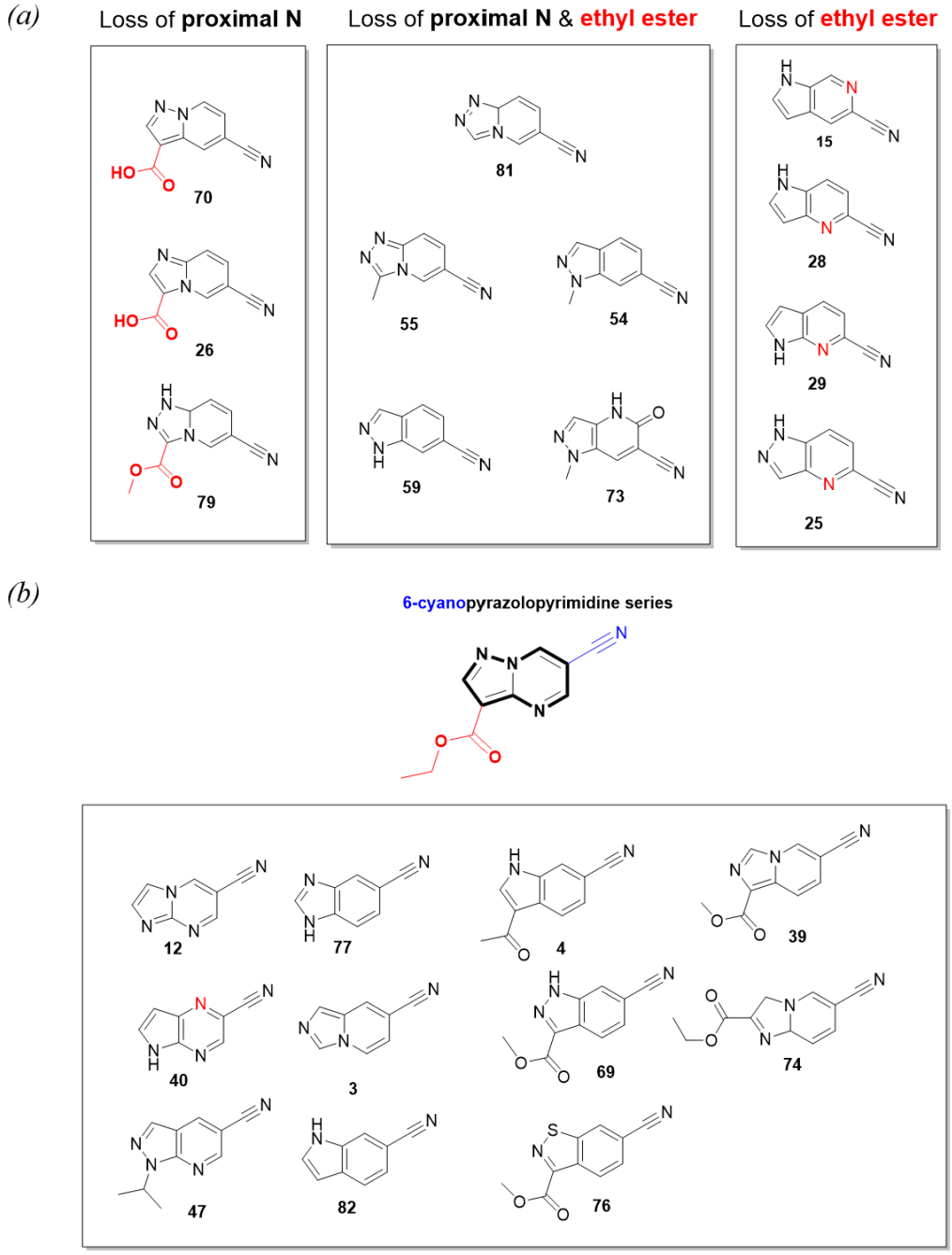


Figure 5.10: **Further Structure-Activity Relationship (SAR)** Structure Activity Relationship of (a) the effect of a combination of modification of essential features (b) further modification on the 6-cyanopyrazolopyrimidine series. None of the compounds formed covalent adducts with the KLHL12 Kelch domain, suggesting modifying a combination of features at once is not tolerated. The samples were analysed using mass spectrometry after KLHL12-compound incubation of 10 min, 3 hours, and 24 hours, as previously described in section 5.2.5. Compounds were obtained from Enamine Ltd.

5.2.5.2 Effect of modifications on position of nitrile warhead and ester group

The relative positions of the nitrile covalent warhead on the hit's pyrazolopyrimidine heterocyclic core scaffold were also investigated with a series of compounds summarised in Fig. 5.11. Any nitrile group position change resulted in non-reactive compounds or incurred more than 100 fold loss of potency, with best IC_{50} for these analogues at 500 μM [Fig. 5.11 *(b)*].

Shifting the nitrile group from position 5 to position 6 of the hit's pyrazolopyrimidine core (fragment **78**, Fig. 5.8) was tolerated but resulted in a 50% reduction in labelling efficiency. When this positional change was combined with the conversion of the ethyl ester to a carboxylic acid (fragment **30**), monoalkylation was significantly reduced, while bi-adduct formation remained unaffected. Shifting the nitrile to position 7 of the hit's scaffold however was not well tolerated with compound **56** achieving only 18% protein labelling even after 24 hours. Interestingly, removal of the bridge nitrogen γ to the shifted nitrogen renders fragment **49** non-reactive, similar to our findings discussed above (compound **17**). Furthermore, the 8-cyanotriazolopyridine compound series (fragment **62**) was also non-reactive, despite the presence of a γ nitrogen to the nitrile.

Additionally, since the 6-cyanopyrazolopyrimidine compounds retained some activity, albeit reduced compared to the hit, a series of compounds with alternative heterocycles was tested [Fig. 5.10 *(b)*]. None of these compounds exhibited reactivity, regardless of the presence of an ester or carboxylate group, a proximal heteroatom to the nitrile, or any combination of these features. This suggests that the electronic properties of the hit's pyrazolopyrimidine core are crucial for activity.

In addition to shifting the nitrile position, nitrile and ethyl ester groups were shifted simultaneously on the hit's pyrazolopyrimidine scaffold as summarised in Fig. 5.12 *(a)*. None of the compounds were reactive, suggesting the relative positions of the two groups are important for activity.

After systematically varying the pyrazolopyrimidine core, I explored the possibility of shifting the nitrile warhead to the 5-membered ring, replacing the position originally occupied by the ethyl ester

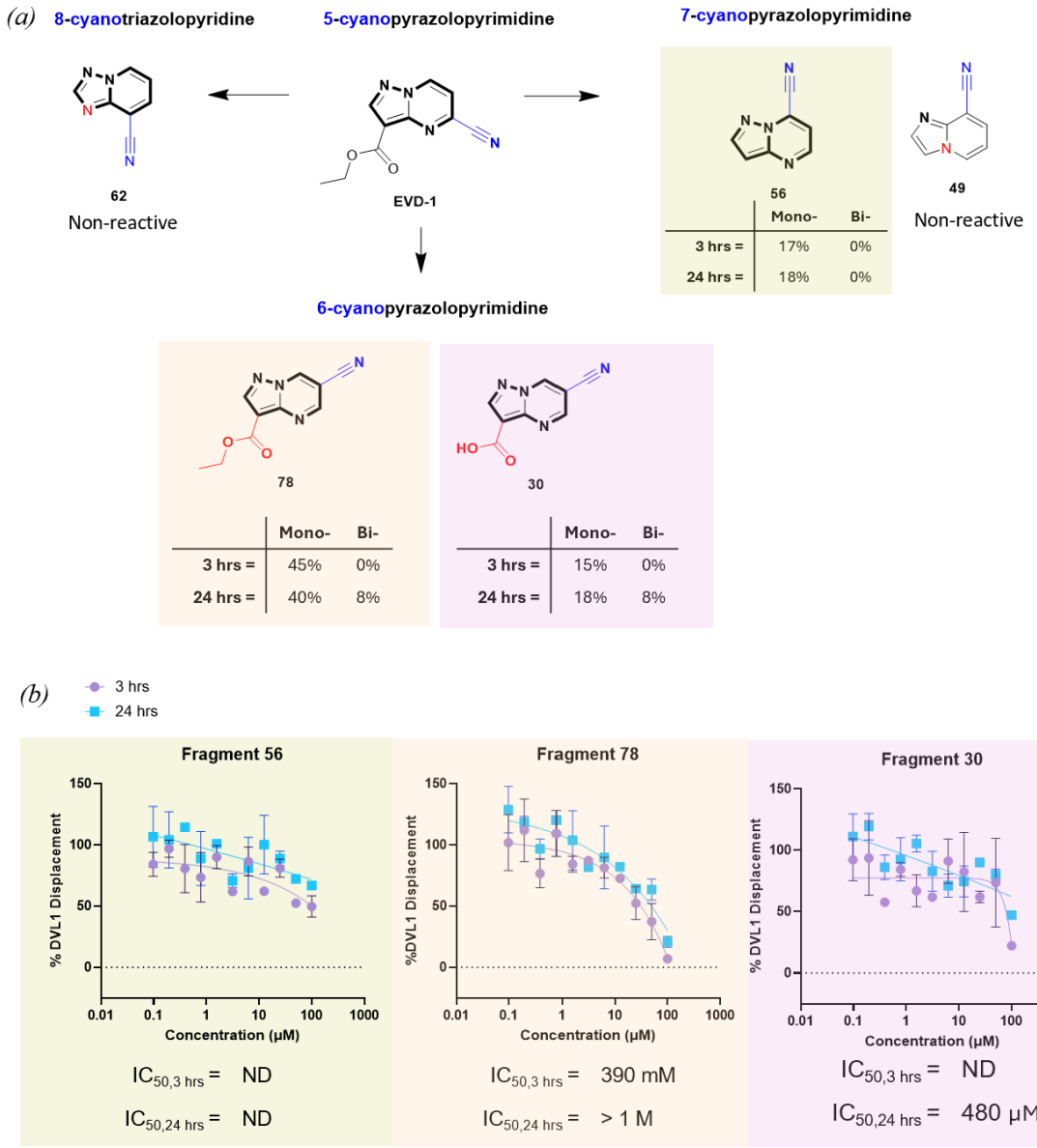


Figure 5.11: **Effect of the position of the nitrile covalent warhead within EVD-1 pyrazolopyrimidine heterocyclic core scaffold.** Structure-activity relationship (SAR) analysis of compounds assessing the impact of the position of the nitrile covalent warhead on **EVD-1** pyrazolopyrimidine heterocyclic core scaffold. (a) The covalent reaction efficiency of the compounds on the KLHL12 Kelch domain was measured using mass spectrometry by monitoring the rate of conjugation at 3 and 24 hours at room temperature. Compounds (100 μM) were incubated with KLHL12 Kelch protein (2 μM) at a 50-fold molar excess. The percentage of adduct formation was determined based on the area under the curve (AUC) of the deconvoluted mass peaks relative to each sample, as shown in Fig. 5.8. 'Mono-' refers to mono-alkylation of KLHL12 by the compound, while 'bi-' indicates bi-alkylation. (b) The compounds' potency was assessed using DVL1 peptide displacement from KLHL12 via AlphaScreen experiments, as described in **Chapter 2.5.6**, and quantified through time-dependent IC_{50} titration curves. The %DVL1 displacement was normalised between averages of negative (DMSO, 0% displacement) and positive (DVL1 peptide, 100% inhibition) controls. A non-linear curve that allowed for a variable slope was fitted on logarithmically transformed compound concentrations over normalised response in Graphpad Prism. Error bars represent the ±standard deviation (SD) from the mean of two technical replicates. Compounds were obtained from Enamine Ltd.

extension, which had not been previously examined. Based on our understanding of the importance of the γ heteroatom, I also introduced nitrogen heteroatoms to investigate heterocyclic systems with the same geometry of fused 5- and 6-membered rings. None of those compounds where the 3-carboxylate group was substituted with 3-cyano warhead showed any reactivity [Fig. 5.12 *(b), (c)*]. Interestingly, the proximity of a heteroatom to the nitrile did not rescue activity (i.e. compounds **5**, **46**, **71** and **24**) in this series of compounds, suggesting the hit's KLHL12 targeting is not only driven by reactivity but by affinity as well.

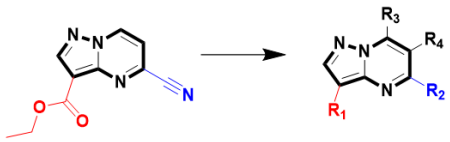
Consistent with this observation, none of the tested naphthalene-based scaffolds or 5-membered heterocyclic rings fused with benzyl rings, as summarized in Fig. 5.13, formed covalent adducts with the protein. This supports the hypothesis that the nitrogen atom at position 4 of the pyrazolopyrimidine scaffold exerts/offers optimal electronic properties which govern the nitrile's chemical reactivity.

5.2.5.3 Scaffold inversion affords more reactive compounds

To further probe the effect of proximal heteroatoms to nitrile reactivity, a series of compounds where the cyano group has been moved to a five-membered ring with one or two proximal heteroatom(s) was tested. An overview of their potency and structures can be found in Fig. 5.14.

Fragments **51** and **6**, thiazolo- and oxazolo [4,5-c]pyridines respectively, form a single cysteine adduct labelling almost all of KLHL12 protein in the first 10 minutes, however a second adduct is formed after extended incubation (3 hrs and 24 hrs), suggesting high compound reactivity (IC_{50} of 0.79 and 1.6 μM respectively). Removal of the pyrimidine's nitrogen yielded equivalent benzo[d]oxazole and benzo[d]oxazole based compounds (**65**, **32** and **80**) that retained the faster and to a higher completion mono additions, but eliminated the non-selective labelling of the second cysteine. Since their potency was also retained (**80**, IC_{50} = 0.79 μM), I can conclude the nitrogen is not an critical for binding. Compound **23**, with benzo[d]oxazole core, exhibited the most favourable profile of this series of compounds, while retaining high potency (0.52 μM). Compound **23** monoalkylates 80% of the KLHL12 protein within the first 10 minutes, compared to 36% monoalkylation observed with

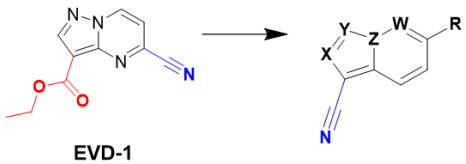
(a)



EVD-1

Frag. No	R_1	R_2	R_3	R_4	Mass Adduct
EVD-1	$COOCH_2CH_3$	CN	CH	CH	N/A
35	CN	CH	CH	CH	N/A
43	CN	CH	$COOCH_2CH_3$	$CH_2CH_2CH_2OH$	N/A
44	CN	CH	$COOCH_2CH_3$	CH	N/A

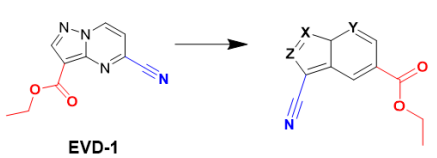
(b)



EVD-1

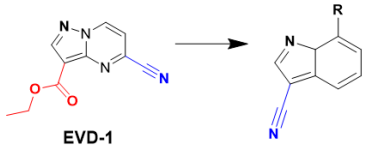
Frag. No	X	Y	Z	W	R	Mass Adduct
5	N	N	N	CH	CH	N/A
10	CH	NH	C	N	CH_3	N/A
36	CH	NH	C	N	COOH	N/A
46	N	NH	C	CH	COOH	N/A
50	CH	NH	C	CH	COOH	N/A
52	CH	NH	C	CH	$COOCH_2CH_3$	N/A
71	N	NH	C	CH	CH	N/A

(c)



EVD-1

Frag. No	X	Y	Z	Mass Adduct
19	NH	CH	CH	N/A
21	NH	N	CH	N/A
24	NH	CH	N	N/A
63	S	CH	CH	N/A
68	N(CH3)	CH	CH	N/A



EVD-1

Frag. No	R	Mass Adduct
11	$COOCH_3$	N/A
48	$CONHCH_3$	N/A

Figure 5.12: **Effect of the relative positions between nitrile and ethyl ester on varying heterocyclic cores** (a) Structure-activity relationship (SAR) analysis of compounds assessing the impact of the relative positions between nitrile and ethyl ester on the hit's pyrazolopyrimidine core (b), and (c) on varying 5:6-membered fused heterocycles. None of the compounds formed covalent adducts with the KLHL12 Kelch domain. The samples were analysed using mass spectrometry after KLHL12-compound incubation of after 10 min, 3 hours, and 24 hours, as previously described in section 5.2.5. Compounds were obtained from Enamine Ltd.

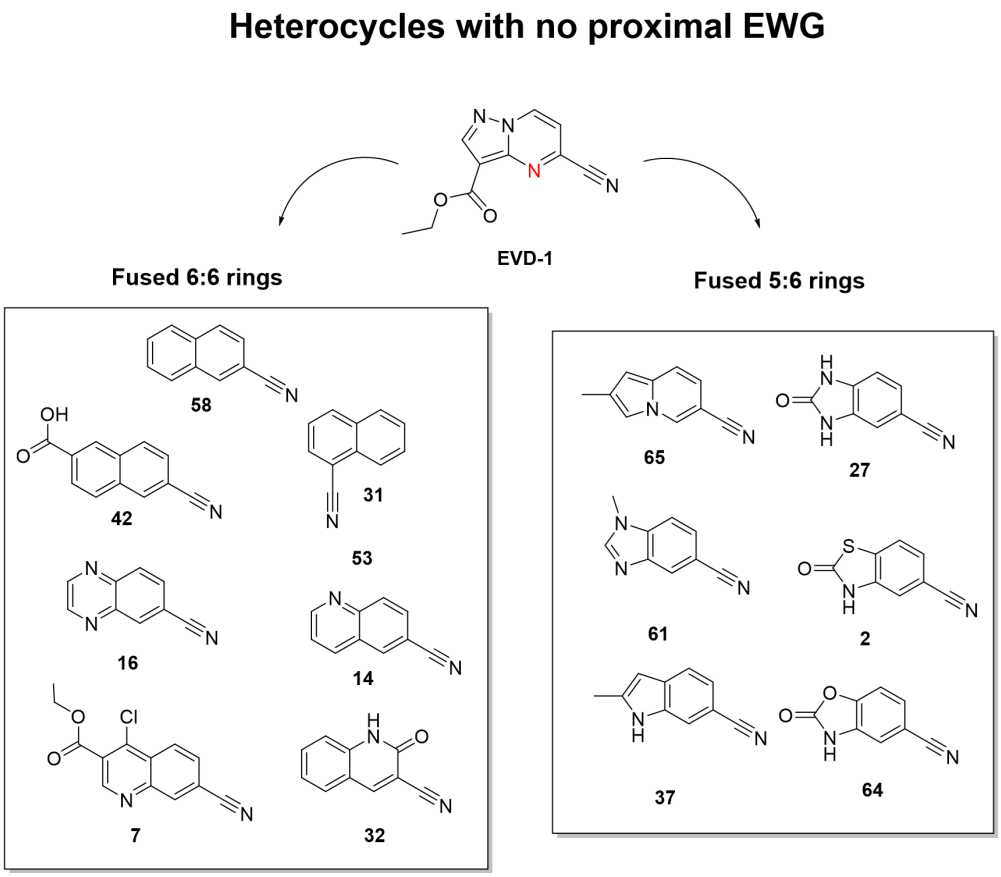


Figure 5.13: **Structure Activity Relationships of nitriles without a proximal heteroatom.** None of the tested naphthalene-based scaffolds or 5-membered heterocyclic rings fused with benzyl rings, formed covalent adducts with the KLHL12 Kelch domain, highlighting the essential proximal heteroatom. The samples were analysed using mass spectrometry after KLHL12-compound incubation of 10 min, 3 hours, and 24 hours, as previously described in section 5.2.5. Compounds were obtained from Enamine Ltd.

EVD-1. This indicates that **23** reacts significantly faster. However, it is not less selective, as after 24 hours of incubation, both **Compound 23** and the **EVD-1** show comparable levels of monoalkylation (79% and 70%, respectively) and bialkylation (9.7% and 7.6%, respectively).

All of the thiazolo- and oxazolo- 5 membered fused rings mentioned above, including compound **23**, show high potency with IC_{50} values ranging from 0.42 to 3.5 μM, comparable to the hit [Fig. 5.14 (a)]. Interestingly, this capacity to displace DVL1 is abrogated upon addition of a hydrogen donor atom (NH, CH) γ to the nitrogen (i.e. fragment **66**, see Fig. 5.14 *(a)*). An overview of the IC_{50} curves are summarised in Fig. 5.15.

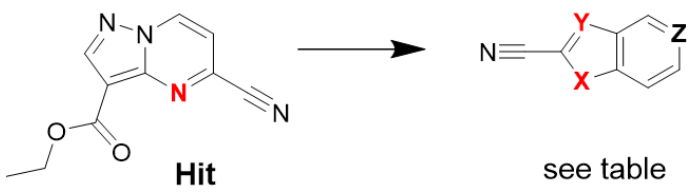
None of the compounds where the cyano group had a single γ heteroatom showed any potency or labelling efficiency [Fig. 5.14 fragments (a) **22**,**45**, (b) **60**, **18**, **57**]. Notably, the benzo[d]imidazole scaffold of compound **13**, where the nitrile warhead was substituted on a nitrogen atom instead of a carbon atom, rapidly monoalkylates KLHL12. During this process, the cyano group remained attached to the protein, while the rest of the molecule acted as an effective leaving group [Fig. 5.8, cyano labelling coloured in purple]. Aromatic cyanamides were therefore not further explored as potential new scaffolds, in contrast with aryl cyanamides as discussed previously [Fig. 5.2.4].

5.2.6 Structure determination of KLHL12 in complex with an alternative covalent compound scaffold

Co-crystallisation was attempted with three alternative scaffolds-fragments **6**, **51**, and **23**-which exhibited higher labelling efficiency. Each fragment was incubated at 1 mM with KLHL12 protein, at concentrations of 10.8 mg/mL for fragment **6**, 15.5 mg/mL for fragment **51**, and 12.7 mg/mL for fragment **23**. Mass spectrometry analysis prior to crystallisation trials revealed that fragments **6** and **51** labeled 66% of the protein once and 33% of the protein twice. In contrast, fragment **23**, which features a methyl substitution on a benzo[d]oxazole core, predominantly labeled the protein once, resulting in the formation of a monoadduct [Fig. 5.16 *(a)*]. Coarse screen crystal trials were set up using the vapour diffusion method with sitting drops, as summarised in Fig. 5.16 *(b)*.

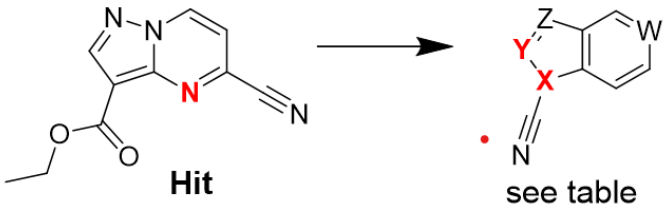
Cyano warhead moved to 5-membered ring

(a)



Frag. No	Y	X	Z	$IC_{50,3\ hrs}$	$IC_{50,24\ hrs}$
51	N	S	N	0.80 µM	0.79 µM
23	N	O	CH_3	1.2 µM	0.52 µM
65	N	S	CH	0.32 µM	0.42 µM
80	N	O	CH	0.24 µM	0.32 µM
6	N	O	N	3.5 µM	1.6 µM
66	N	NH	CH	> 1000 µM	180 µM
22	CH	O	CH_3	N/A	N/A
38	CH	NH	CH	N/A	N/A
41	CH	NH	COOH	N/A	N/A
45	CH	O	CH	N/A	N/A
72	CH	$N(CH_3)$	CH	N/A	N/A

(b)



Frag. No	X	Y	Z	W	$IC_{50,3\ hrs}$	$IC_{50,24\ hrs}$
13	N	CH	N	CH	0.045 µM	0.10 µM
18	CH	N	S	N	N/A	N/A
57	CH	CH	O	CH	N/A	N/A
60	CH	N	S	CH	N/A	N/A

Figure 5.14: **Structure Activity Relationships of nitriles on 5-membered rings.** Compounds designed with the nitrile warhead (a) between two heteroatoms of 5-membered rings, and (b) proximally to one heteroatom of 5-membered rings. The compounds' potency was assessed using DVL1 peptide displacement from KLHL12 via AlphaScreen experiments, as previously described, and quantified through time-dependent IC_{50} titration curves (two technical repeats). Compounds that showed no adduct formation in initial mass spectrometry assay [Fig. 5.8], were not tested for DVL1 displacement, hence the values N/A = not available). Compounds were obtained from Enamine Ltd.

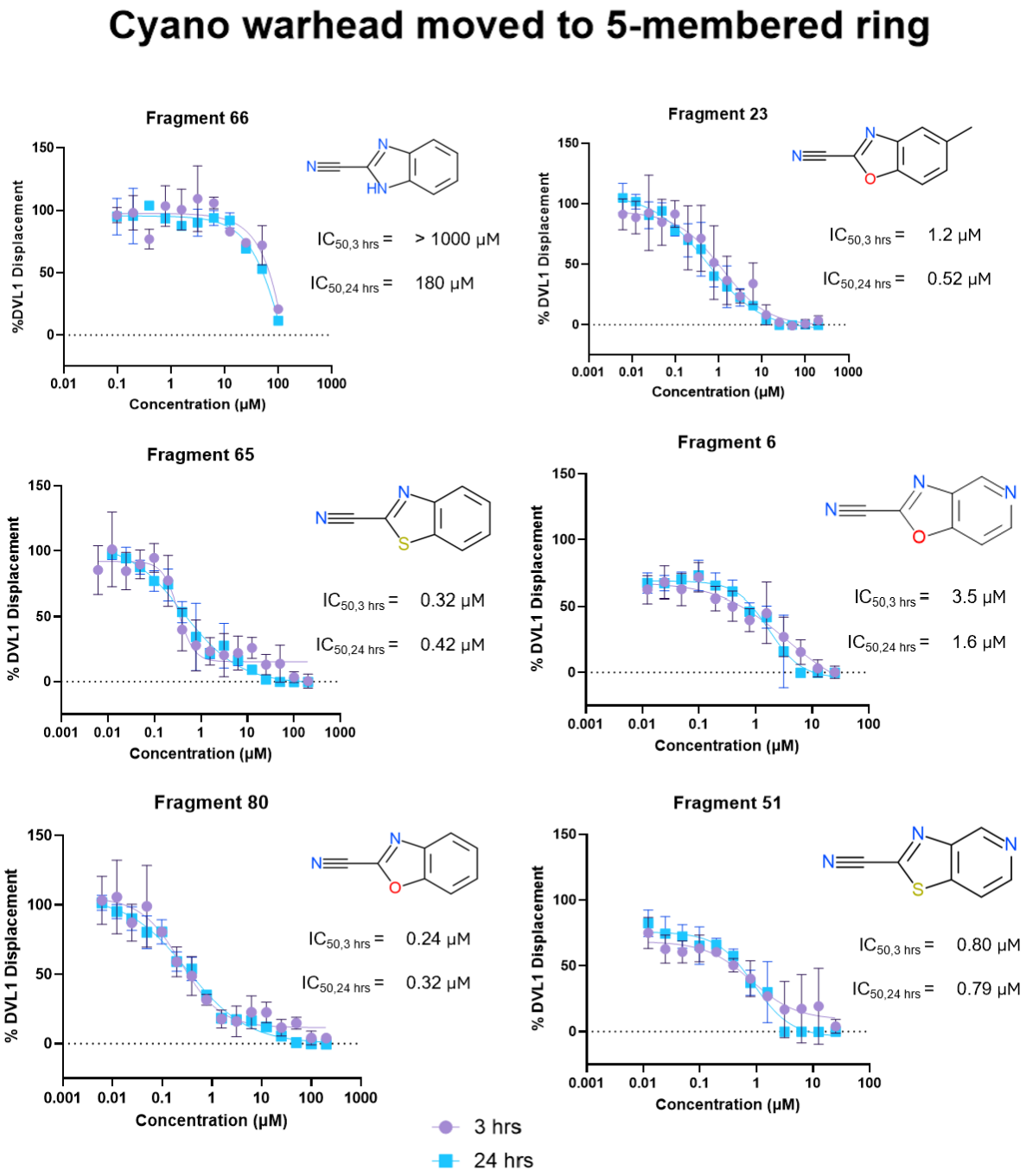
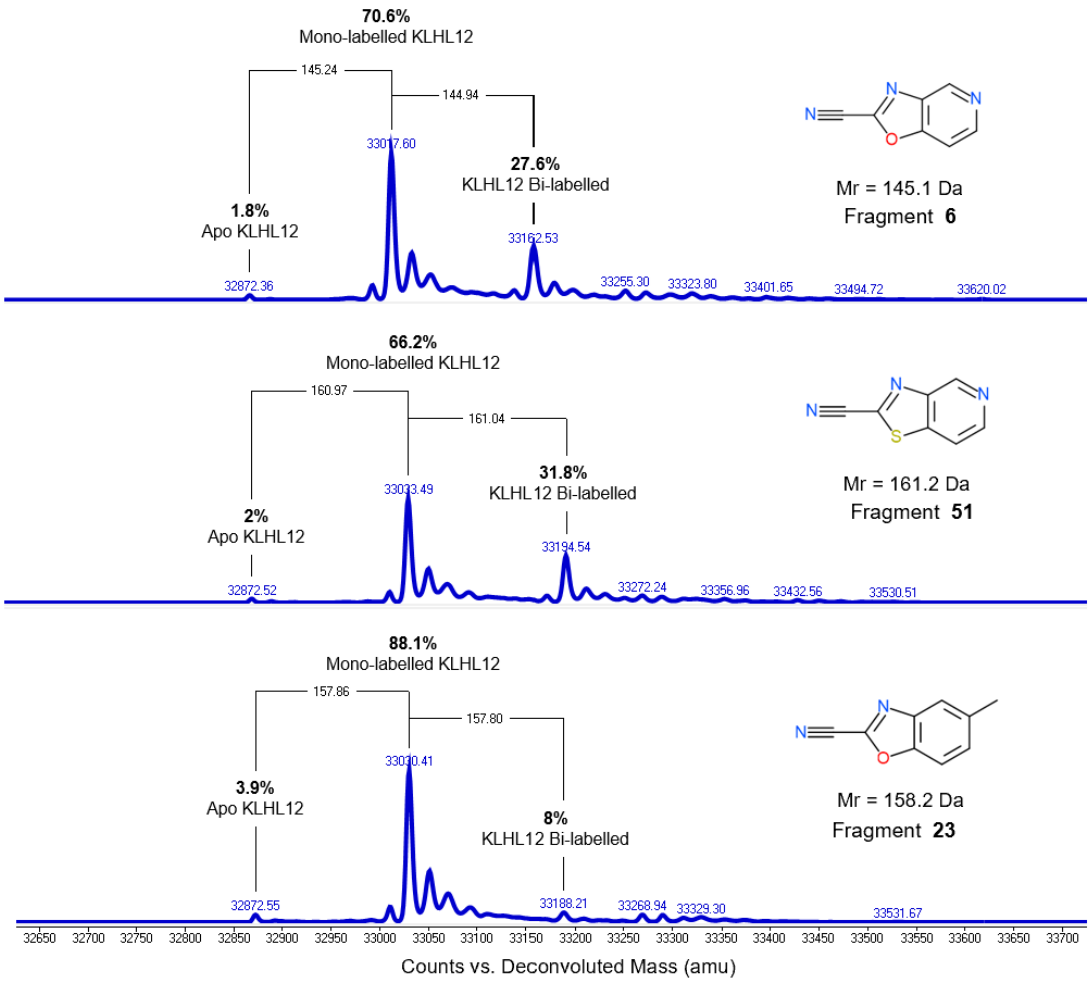


Figure 5.15: **IC_{50} curves of 5-membered ring nitriles.** The compounds' potency was assessed using DVL1 peptide displacement from KLHL12 via AlphaScreen experiments, as described in **Chapter 2.5.6**, and quantified through time-dependent IC_{50} titration curves. The %DVL1 displacement was normalised between averages of negative (DMSO, 0% displacement) and positive (DVL1 peptide, 100% inhibition) controls. A non-linear curve that allowed for a variable slope was fitted on logarithmically transformed compound concentrations over normalised response in Graphpad Prism. Error bars represent the ±standard deviation (SD) from the mean of two technical replicates. IC_{50} curves for fragments incubated with KLHL12 and the DVL1 peptide for 3 hours are shown in purple, while those incubated for 24 hours are shown in cyan. Compounds were obtained from Enamine Ltd.

(a)



(b)

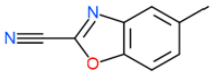
Plate Identifier	Protein Conc. (mg/ml)	Cmpnd	Coarse Screen	Temp (°C)	Hit Wells
CI088437	12.67		JCSG+ ECO	4	-
CI088436	12.67		JCSG+ ECO	20	-
CI088424	12.67		HIN3	20	B6c, B6d
CI088425	12.67		HIN3	4	-
CI088428	12.67	1 mM (Fragment **23**)	MIDASPLUS	4	E5c, F6c, G1d, G2d, H10c, H10d

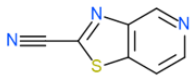
Plate Identifier	Protein Conc. (mg/ml)	Cmpnd	Coarse Screen	Temp (°C)	Hit Wells
CI088434	15.5		JCSG+ ECO	4	-
CI088435	15.5		JCSG+ ECO	20	-
CI088422	15.5		HIN3	20	-
CI088421	15.5		HIN3	4	-
CI088427	15.5	1 mM (Fragment **51**)	MIDASPLUS	4	-

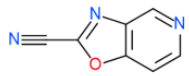
Plate Identifier	Protein Conc. (mg/ml)	Cmpnd	Coarse Screen	Temp (°C)	Hit Wells
CI088432	10.75		JCSG+ ECO	4	-
CI088433	10.75		JCSG+ ECO	20	-
CI088438	10.75		HIN3	20	-
CI088439	10.75		HIN3	4	-
CI088426	10.75	1 mM (Fragment **6**)	MIDASPLUS	4	D10a, G8a

Figure 5.16: **Co-crystallisation trials of KLHL12 and 5-membered ring nitriles** (a) Mass-spectrometry analysis of compounds' adduct formation on KLHL12 prior to setting up crystal trials, (b) summary of coarse screen crystal trials in the presence of 1 mM compound.

Of the three fragments, only fragment **23** produced crystals of sufficient size and quality for mounting and testing. This fragment was the only one that yielded a homogeneous sample of monoalkylated protein KLHL12, which potentially did not impede the crystallisation process. Co-crystals of fragment **23** were observed in both 4 °C and 20 °C as shown in Fig. 5.17. The 4 °C crystals were observed in reservoir solutions containing the polymer SOKALAN, similarly to the hit crystals described above, however these crystals were sensitive to cryoprotectant addition (25% ethylene glycol). An alternative crystal form diffracted at 2.6 Å resolution based on a reservoir condition containing 0.49 M sodium phosphate monobasic (NaH_2PO_4), 0.91 M potassium phosphate dibasic (K_2HPO_4) incubated at 20 °C.

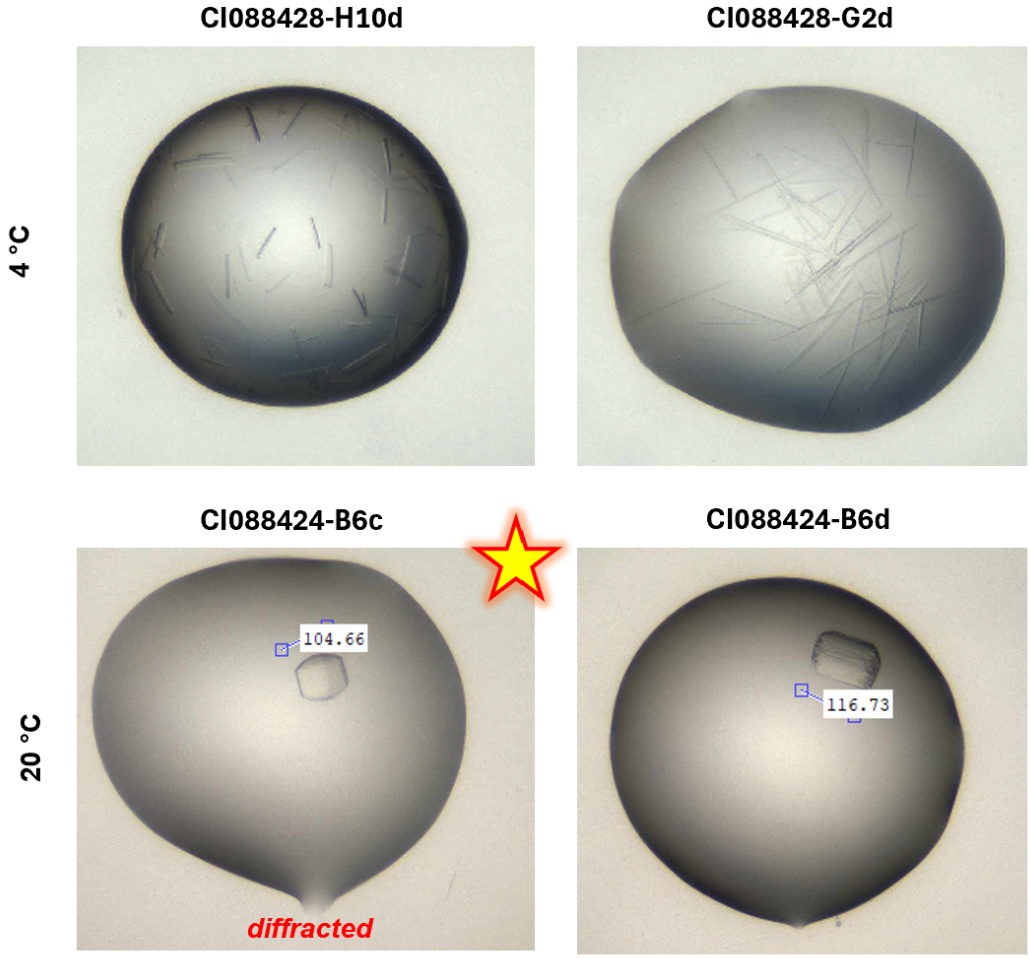


Figure 5.17: **Representative crystal hits of fragment 23-KLHL12 complex.** Representative KLHL12-covalent hit crystals from co-crystallisation coarse screens. Crystal yielding diffraction marked with a yellow star.

This crystal form exhibited moderate anisotropy as detailed in table 5.1, that could not be addressed with refinement in Phenix as described in the **EVD-1** co-crystal in **Chapter 5.2.1**. The two datasets of two crystal sweeps were indexed, integrated, scaled and merged by the 2x xia2 3di (multi) Diamond automated pre-processing pipeline. The two datasets from two crystal sweeps were combined in Pointless, before ellipsoidally cutting the data using the STARANISO anisotropy and Bayesian estimation server [251]. Molecular replacement was performed by Phenix.phaser-MR using the **EVD-1** KLHL12 covalent structure (excluding the bound hit) as a search model. Refinement of the truncated data to account for anisotropy was performed in Buster [208] using individual B-factors in earlier stages of refinement, and non crystallographic symmetry restrains alongside Translation-Libration-Screw (TLS) restraints in final cycles. The covalent linkage descriptions and compound restraints were generated with Grade [252].

The crystal was phased in P $2_12_12_1$ space group with nine KLHL12 protein molecules in the ASU. Similarly to the previous covalent hit, seven molecules were arranged in a planar ring formation with chains G and A extend axially above and below the plane [Fig. 5.18 *(a)*]. The previously observed compound-mediated *daisy-chain* that stabilises the interlocking rings formation was also seen here [Fig. 5.18 *(b)*]. Finally, the structure was refined to 2.60 Å resolution. Crystallographic data collection and refinement statistics of the structure are shown in Table 5.2.

Table 5.1: Anisotropy of KLHL12-covalent compounds co-crystals

Co-crystallised Compound	**High Resolution Limit**	**Diffraction Limit *a****	**Diffraction Limit *b****	**Diffraction Limit *c****
Hit	2.58	3.44	2.57	2.68
23	2.60	4.70	4.86	2.59

5.2.7 Structural observations on the binding mode of an alternative fragment scaffold

As expected, electron density for fragment **23** was present in all nine protein molecules, with covalent labelling of C511 to form a thioimidate product in the substrate binding site of KLHL12. Fragment **23** is tethered to KLHL12 through the imino group of the thioimidate adduct; it serves as both a hydrogen bond donor and acceptor, forming interactions with the amide backbones of Y512 and S464 (2.5 Å

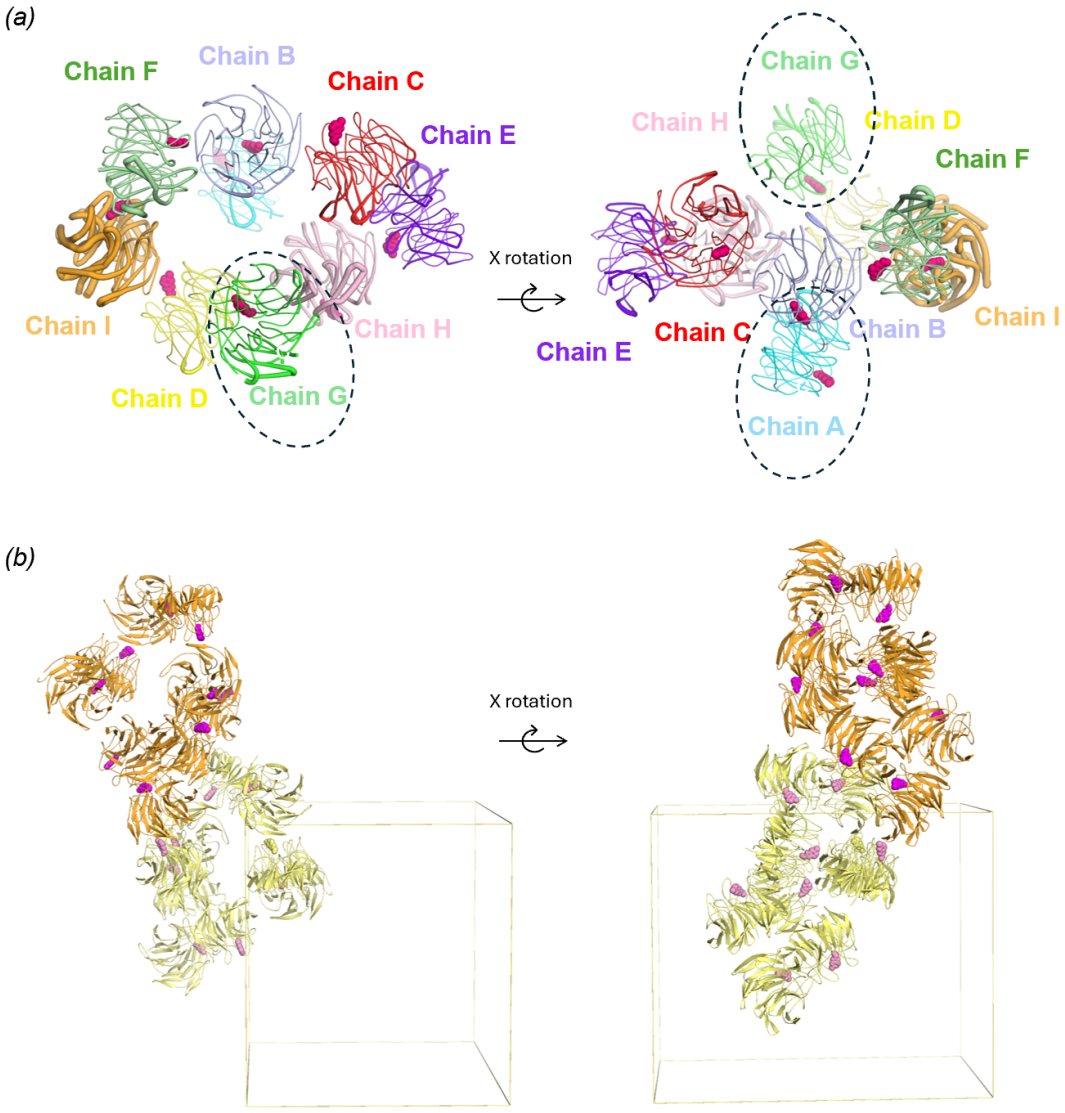


Figure 5.18: **Asymmetric unit of the fragment 23 and KLHL12 co-crystal complex.** (a) Nine protein KLHL12 chains are present in the asymmetric unit; chains named A-I. Proteins chains are visualised in PyMOL using the 'putty' representation; the tube thickness corresponds to B-factors, with thicker regions indicating higher atomic displacement (greater flexibility) and thinner regions indicating lower displacement (greater rigidity). Seven protein molecules are arranged in a planar ring formation, while two additional protein molecules, chains G and A, extend axially above and below the plane at an angle (circled with dashed lines). Fragment **23** is present in all nine KLHL12 protein chains, and is shown in green sphere representation. (b) The unit cell is depicted as a yellow box. KLHL12 protein molecules within a single unit cell are coloured yellow, with fragment **23** represented as pink spheres. KLHL12 molecules from adjacent unit cells are colored orange, with fragment **23** shown as magenta spheres.

and 3.4 Å respectively) [Fig. 5.19 *(a)*]. The planar aromatic scaffold is sandwiched between residues Y512 and F481, with the oxazole ring engaging in π-π stacking with Y512 (3.9 Å) and the phenyl ring with F481 [Fig. 5.19 *(b)*]. All nine EVD-1 molecules adopt the same conformation when bound to C511 [Fig. 5.19 *(c)*].

Comparing the binding modes of **EVD-1** and the inverted scaffold of **23**, we observe that the planar thioimidate adduct formed after reaction with the cysteine contributes two hydrogen bonds in both compounds. However, **EVD-1** is further stabilised by a third hydrogen bond involving its ester group, which is absent in fragment **23** and thus does not engage H486. While both compounds are sandwiched between the same two aromatic residues, **EVD-1** engages F481 in π-π stacking through the 5-membered ring and Y512 with the 6-membered nitrile-bearing ring. In contrast, fragment **23** π-π stacks with F481 through the 6-membered ring and with Y512 via the 5-membered oxazole ring, which directly binds to the cysteine, as shown in Fig. 5.20.

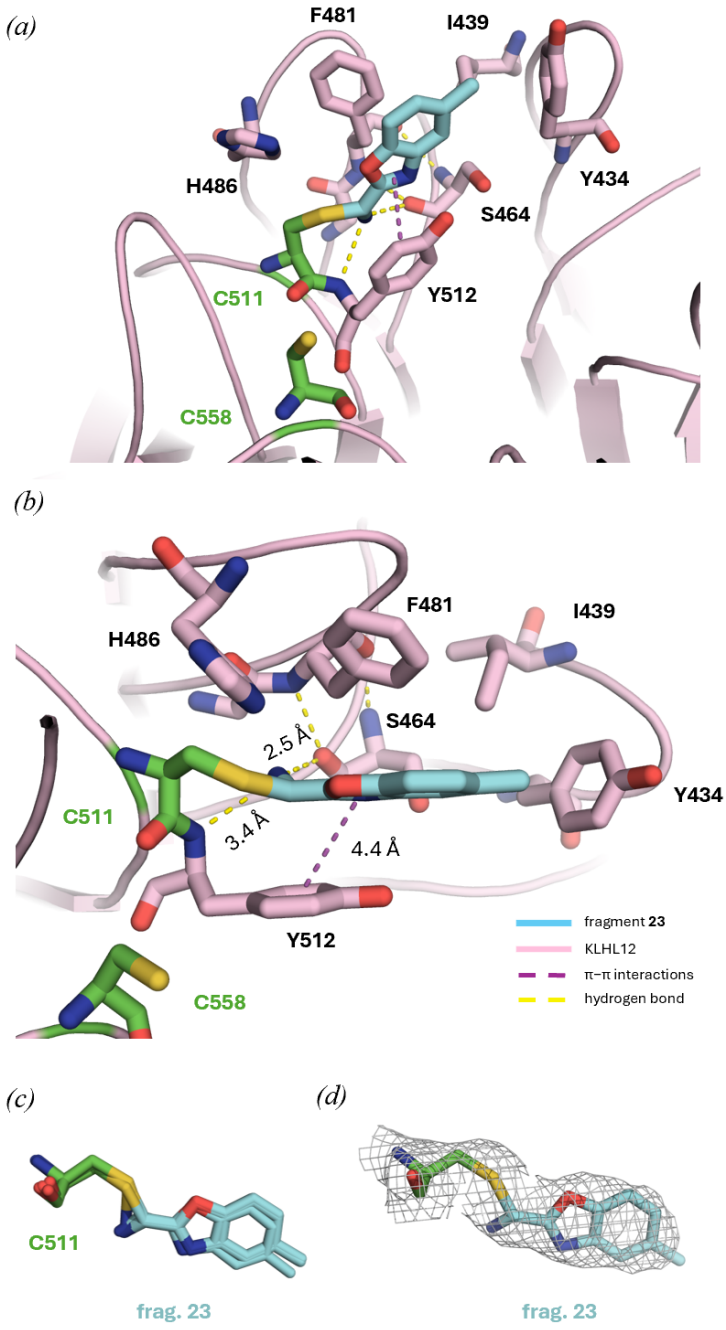


Figure 5.19: **2.6 Å structure of fragment 23 bound to the KLHL12 substrate binding site.** (a) and (b) Overview of fragment **23** (coloured cyan) bound to KLHL12 (coloured pink). Cysteines are coloured in green. Hydrogen bond interactions with the Kelch domain binding site shown by yellow dashed lines. π-π interactions are shown in magenta dotted lines (c) Superposition of all 9 compounds in the same asymmetric unit, stick representation (d) 2Fo-Fc electron density map (gray mesh) for fragment 23 of the best resolved chain, contoured at 1 σ.

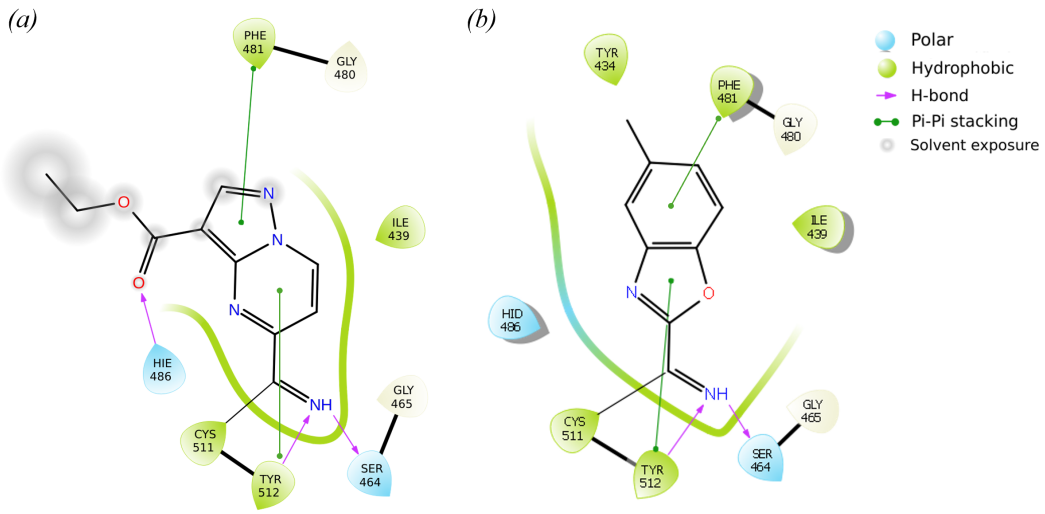


Figure 5.20: **KLHL12-compound interaction diagram.** Interaction diagram showing KLHL12 binding to (a) hit compound **EVD-1** and (b) fragment **23**. Hydrogens were added, the hydrogen bond network was optimised, and the KLHL12 protein-compound interaction diagram was calculated and generated using Maestro, Schrödinger Suite.

Table 5.2: Crystallographic data for KLHL12-**EVD-1** and KLHL12-**23** complexes

Parameter	**KLHL12/fragment 23**	**KLHL12/EVD-1**
Wavelength (Å)	0.97625	0.97627
Resolution range (Å)	49.60 - 2.74 (2.99 - 2.74)	49.70 - 2.60 (2.62 - 2.60)
Space group	P 21 21 21	P 21 21 21
Unit cell		
a, b, c (Å)	140.9, 151.2, 178.9	142.5, 151.1, 179.3
α, β, γ (°)	90, 90, 90	90, 90, 90
Spherically cut data	**xia2 dials multi**	**AUTOPROC**
Total reflections	2060152 (90448)	1572853 (80892)
Unique reflections	74897 (3213)	112630 (5589)
Multiplicity	27.5 (28.2)	14.0 (14.5)
Completeness (%)	99.40% (85.96)	100 (100)
Mean I/sigma(I)	5.2 (0.4)	7.7 (0.6)
Wilson B-factor (Å^2)	58.250	80.32
R-merge	0.518 (7.586)	0.343 (5.335)
R-measure	0.528 (7.723)	0.370 (5.732)
R-pim	0.100 (1.447)	0.137 (2.092)
CC1/2	0.992 (0.393)	0.995 (0.364)
Reflections used in refinement	-	119184 (2384)
Reflections used for R-free	-	5973 (115)
R-work	-	0.268
R-free	-	0.276
Ellipsoidally cut datasets	**STARANISO**	
Total reflections	1542277 (80011)	-
Unique reflections	57131 (2860)	-
Multiplicity	27 (28)	-
Completeness spherical (%)	56.3 (12.3)	-
Completeness ellipsoidal (%)	95.2 (78.1)	-
Mean I/sigma(I)	7.5 (1.6)	-
Wilson B-factor (Å^2)	70.95	-
R-merge	0.403 (2.694)	-
R-measure	0.424 (2.789)	-
R-pim	0.112 (0.723)	-
CC1/2	0.990 (0.657)	-
Reflections used in refinement	57131 (1143)	-
Reflections used for R-free	19039 (765)	-
R-work	0.263	-
R-free	0.276	-
Number of non-hydrogen atoms	19,590	19,701
Macromolecules	19,433	19,500
Ligands	108	148
Solvent	49	53
Protein residues	2,610	2,615
RMSD (bonds, Å)	0.005	0.010
RMSD (angles, °)	0.72	0.06
Ramachandran favored (%)	96.43	95.46
Ramachandran allowed (%)	3.22	4.23
Ramachandran outliers (%)	0.35	0.31
Rotamer outliers (%)	2.41	2.75
Clashscore	5.29	4.48
Average B-factor (Å^2)	79.88	93.10
Macromolecules	80.10	93.15
Ligands	83.00	94.12
Solvent	24.76	70.78

5.3 Selectivity and reactivity characterisation

5.3.1 Selectivity over a panel of Kelch domain E3 ligases

Despite the initial indication of good selectivity of **EVD-1**, for the Kelch domain of KLHL12 over those of KLHL20 and KLHL6, the selectivity experiments were conducted following a 1-hour incubation at 100 μM [as discussed in 4.11]. Given the slow-reacting nature of such compounds, a more comprehensive assessment over a broader range of Kelch family members and extended incubation times was warranted.

To address this, both **EVD-1** and fragment **23** (100 μM) were incubated with protein (2 μM) at a 50-fold molar excess for 3 hours. The samples were subsequently denatured with formic acid and analysed by mass spectrometry [for details of method refer to **Chapter 2.6.3**]. The extent of ligase adduct formation, quantified as the area under the curve (AUC) of the deconvoluted mass peaks of labelled KLHL12, was visualised as a heatmap in 5.21 (a).

The results revealed that **EVD-1** reacted exclusively with KLHL12, labelling 22% of KLHL12 with a single adduct while showing no detectable binding to other Kelch domains of E3 ligases, despite their solvent-exposed cysteines being accessible for adduct formation (highlighted in green in Fig. 5.21 (b)). In contrast, fragment **23** exhibited broad reactivity, forming both mono- and bi-adducts across multiple E3 ligases. Notably, it labeled KLHL12 at 45% efficiency but also reacted with KLHL20 (100%), KBTBD10 (60%), and KBTBD8 (74.5%), indicating high cysteine reactivity but poor selectivity.

5.3.2 Glutathione reactivity testing

To distinguish between a compound's ability to selectively bind its intended target versus promiscuous reaction with any available thiol, glutathione (GSH) reactivity testing was employed. The compounds were incubated at 100 μM with glutathione (GSH) at its physiological concentration of 1 mM and under physiological pH = 7.4, at room temperature and the samples were analysed by HPLC-MS

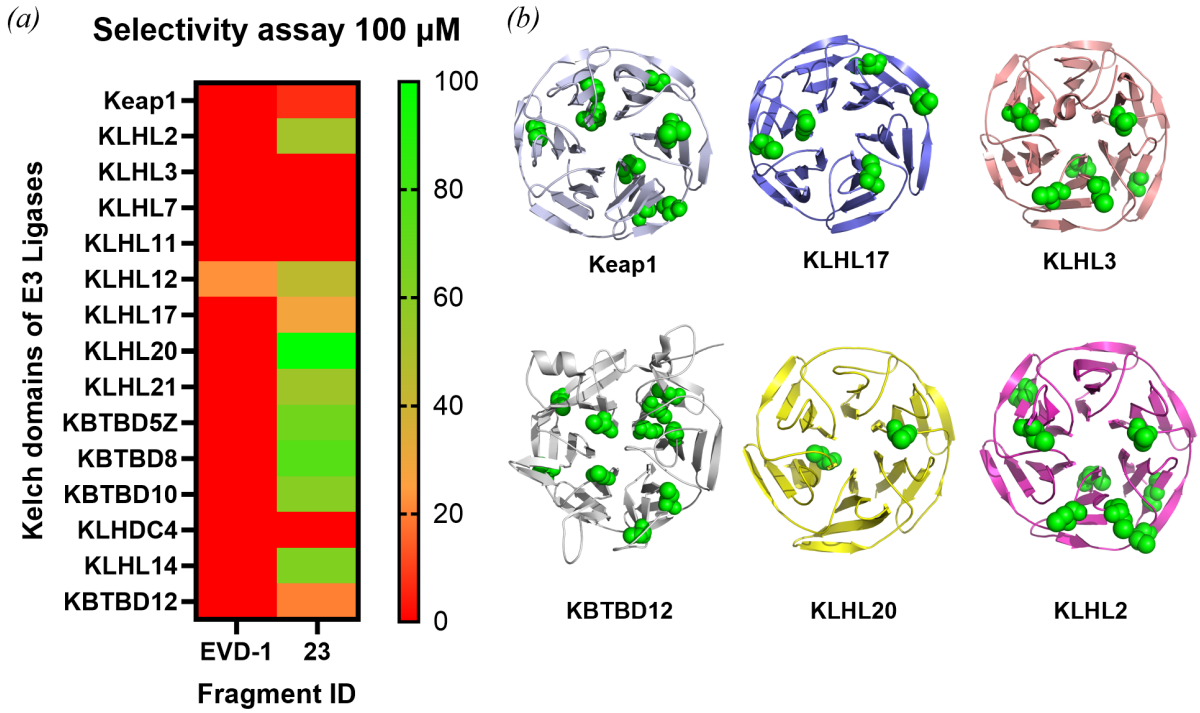


Figure 5.21: **Selectivity assay for covalent EVD-1 and fragment 23.** (a) Original hit **EVD-1** and more reactive scaffold fragment **23** (at 100 µM) were incubated for 3 hours with a range of E3 ligases (2 µM) containing Kelch domains with solvent exposed cysteines to assess their selectivity for KLHL12. The total % of alkylated protein (mono- or bi- labelled by the compound) was plotted as a heatmap in (a), based on the deconvoluted mass peak's area under the curve (AUC) relative to each sample (one independent repeat). (b) Representative structures of Kelch domains of other Kelch-containing E3 ligases that were screened in the selectivity panel in *(a)*.Cysteines are shown as green spheres. KEAP1 PDB: 4IFJ, KLHL17 PDB: 6HRL, KLHL3 PDB: 4CH9, KBTBD12 PDB: 7QZQ, KLHL20 PDB: 8CIA, KLHL2 PDB: 2XN4

after appropriate incubation time. The percentage of compound remaining, over GSH-adduct, was quantified using ultraviolet absorption peak area under the curve. The natural logarithm of the results was fitted to a linear regression, and half-life was calculated as $t_{1/2}$ = ln(2)/-slope.

A marked difference in the half-lives of the two compounds was revealed, with the **hit** exhibiting a half-life of 55 hours and fragment **23** a significantly shorter half-life of 6 hours [Fig. 5.22]. The **hit** compound labels a small portion of GSH even after 24 hours of incubation, reflecting the lower intrinsic reactivity of the pyrazolopyrimidine scaffold. This reduces the likelihood of promiscuous covalent modifications of solvent accessible reactive cysteines on off-target proteins by the hit. However, this reduced reactivity might compromise its ability to efficiently form covalent adducts with the target protein under physiological conditions. On the other hand, **23**'s rapid depletion suggests a higher intrinsic reactivity toward thiol groups, reflecting the greater electrophilic character of the oxazole ring. While this enhances covalent binding efficiency, it raises concerns about off-target reactivity.

Fragment **17**, derived from **EVD-1** but lacking the nitrogen γ to the nitrile group, was used as a negative control. No GSH modification was observed after 24 hours of incubation, consistent with the low intrinsic reactivity of the pyridine ring. Interestingly, prior experiments revealed 11% KLHL12 adduct formation under the same conditions [**Chapter 5.2.5.1**], suggesting that features of the **EVD-1** other than its reactivity can engage KLHL12 in a target-specific manner, albeit with low binding efficiency.

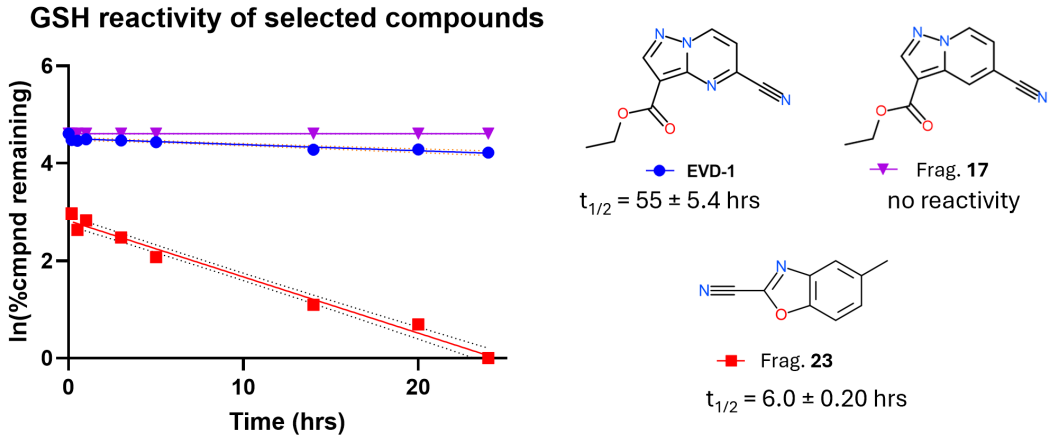


Figure 5.22: **Glutathione reactivity assay of selected compounds.** Compound reactivity with glutathione measured by HPLC-MS to determine the half life of inactivation. Dotted lines represent 95% confidence level of linear fit, based on three technical repeats. To assess compound reactivity, compounds at 100 µM were incubated at room temperature with 1 mM glutathione (GSH) and analysed with LC-MS after appropriate incubation time. The percentage of compound remaining, over GSH-adduct, was quantified using ultraviolet absorption peak area under the curve in Waters FractionLynx software. The natural logarithm of the results were fitted to a linear regression, and half-life was calculated as $t_{1/2}$ = ln(2)/-slope. Original hit **EVD-1** is plotted in blue, inactive negative control fragment **17** is plotted in purple, and alternative more reactive covalent scaffold is plotted in red.

5.4 Synthesis of hit elaborations

Structural analysis of the co-crystallized fragments revealed an opportunity for fragment extension into a proximal hydrophobic pocket, that is currently unexploited. This pocket is formed by residues Y434, I439, G436, and L371. Even though both cores could form the basis of extended fragment, the original hit scaffold OF **EVD-1** was selected due to the existence of the ethyl ester affording a readily accessible exit vector for PROTAC development. Examination of the KLHL12-hit co-crystal structure suggests that position 7 of the pyrazolopyrimidine core offers a viable extension vector for fragment elaboration [Fig. 5.23 (a)].

Covalent docking studies were conducted with a series of designed analogues bearing modifications at position 7 of the pyrazolopyrimidine, including benzylsulfonyl (elaboration 1), methyl benzyl ether (elaboration 3), and indolesulfonyl (elaboration 2) extensions. The results indicated that these modifications could be accommodated within the hydrophobic pocket, potentially enabling interactions with residue Y434 through π-π stacking [Fig. 5.23 *(b)*]. The incorporation of the sulfonyl moiety into the design, instead of the ether or thioether, was intended to introduce a fixed angle that could conformationally constrain the flexible ether angle and increase the potential for it being in plane with Y434 phenyl ring.

A multistep synthesis route was designed with the potential to access multiple ether, thioether, and sulfone elaborations at position 7 of the pyrazolopyrimidine core of **EVD-1**, utilizing a key dichloro intermediate [Scheme 1]. The route begins with the formation of the pyrazolopyrimidine ring via cyclisation of commercially available ethyl 3-imino-2,3-dihydro-1*H*-pyrazole-4-carboxylate and diethyl malonate, yielding the cyclised diketone (step **a**). The diketone is then chlorinated to form the dichloro intermediate (step **b**). At this stage, the desired elaborations can be introduced at position 7 through selective nucleophilic aromatic substitution (S_NAr) of the dichloro intermediate with the appropriate thiol or alcohol (steps **c**, **f**). Care must be taken during this step to minimize the formation of undesired bis-substituted side products. Finally, cyanation of the remaining chloro position and

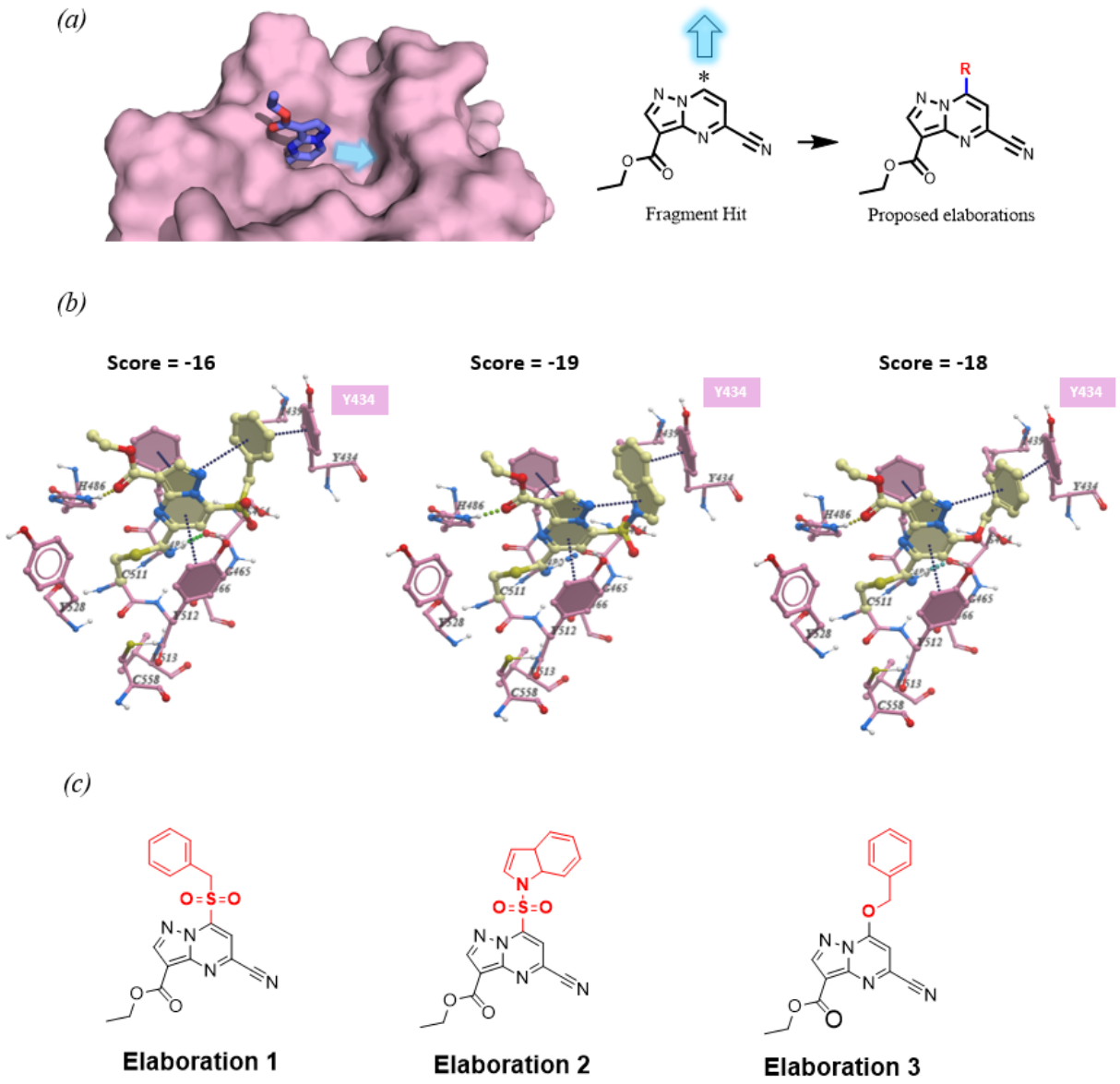
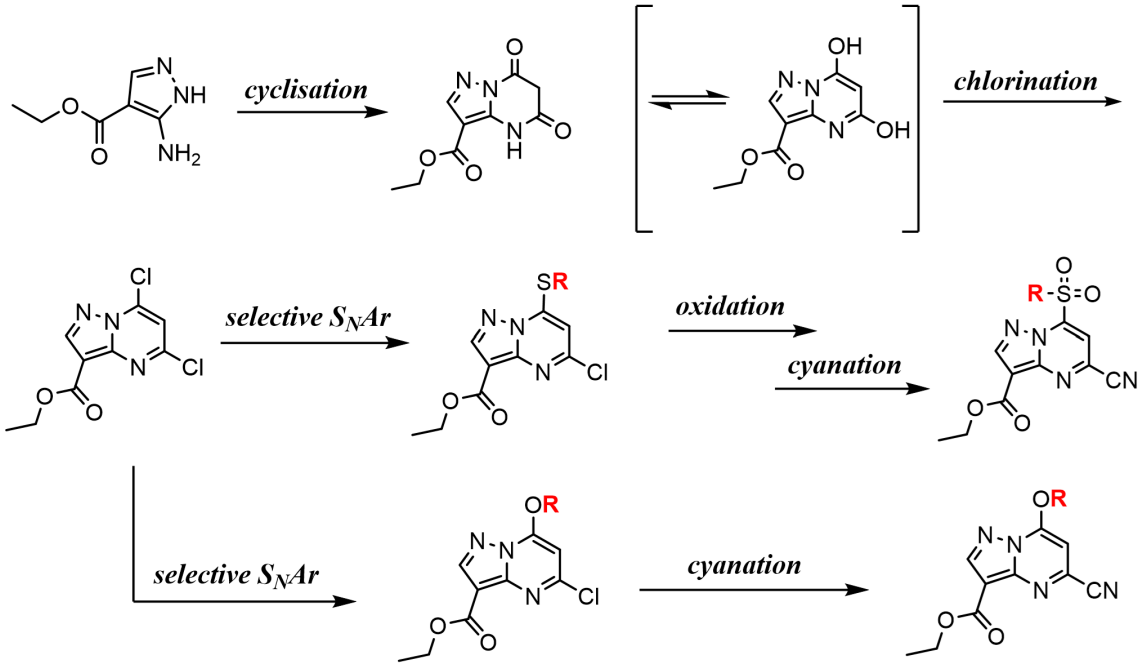


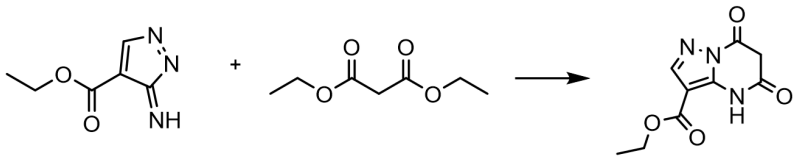
Figure 5.23: **EVD-1-KLHL12 Kelch domain co-crystal reveals accessible elaboration vector.** (a) Hit co-crystal reveals nearby hydrophobic pocket currently unexplored by the hit. Elaboration vector denoted with arrow coloured in blue. Potential alkyl substitution at position 7 of the pyrazolopyrimidine core, coloured in red. (b) Proposed elaborations from synthetic accessible vector based on the hit scaffold. Docking experiments show additional interaction with Y434 is possible. Docking in ICM Pro (Molsoft). ((c) Molecular structure of the proposed elaborated compound, with the elaboration highlighted in red. The docking score of each compound is comparable to the core **EVD-1** docking score of -20. π-π stacking interactions between aromatic rings are denoted with black dashed lines.

optional oxidation of the thioether to the corresponding sulfone, if required, affords the final extended fragments (steps **g**, **d**, **e**).



Scheme 1: **General synthetic route to access hit elaborations**

5.4.0.1 Synthesis of key dichloro intermediate



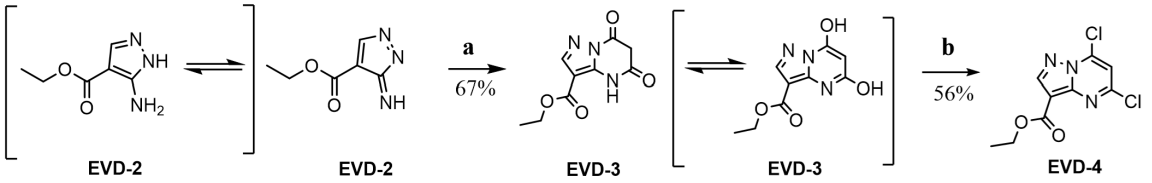
EVD-2 **EVD-5** **EVD-3**

Solvent	T (°C)	Base	t (hrs)	Yield
EtOH	80	NaOEt (Na^0)	24	-
DMA	120	KO*t*Bu	24	30%
EtOH	80	KO*t*Bu	24	-
DMF	120	Cs_2CO_3	6	38%
EtOH	120	21% w/w NaOEt	24	67%

Table 5.3: Cyclisation conditions used for the formation of **EVD-3**

The initial cyclisation and chlorination reactions had to fulfill two criteria: i) they had to be scalable to furnish multigram quantities of the desired product and ii) should be amendable to simple purification techniques avoiding tedious chromatographic methods. Thus, various conditions were tested as summarised in Table 5.3. High temperatures > 91.5 °C and strong bases were necessary to form the pyrazolopyrimidine core of **EVD-1** hit in both protic and aprotic solvents. **EVD-3** was successfully obtained by reacting at 90-120 °C with sodium ethoxide in ethanol, achieving a good yield of 67% [Scheme 2, **step a**]. The subsequent chlorination of **EVD-3** was performed using $POCl_3$ in acetonitrile at similarly high temperatures of 100 °C, yielding the key intermediate **EVD-4** at scale (4.8 g, 56%).

Both **EVD-3** and **EVD-4** were purified without the need for column chromatography; instead, simple direct drop crystallisation followed by filtration sufficed. This, combined with a reasonable overall yield, made the process suitable for scale-up, leading to obtaining 4.8 g of key dichloro intermediate **EVD-2**, sufficient for exploring elaborations as discussed below.

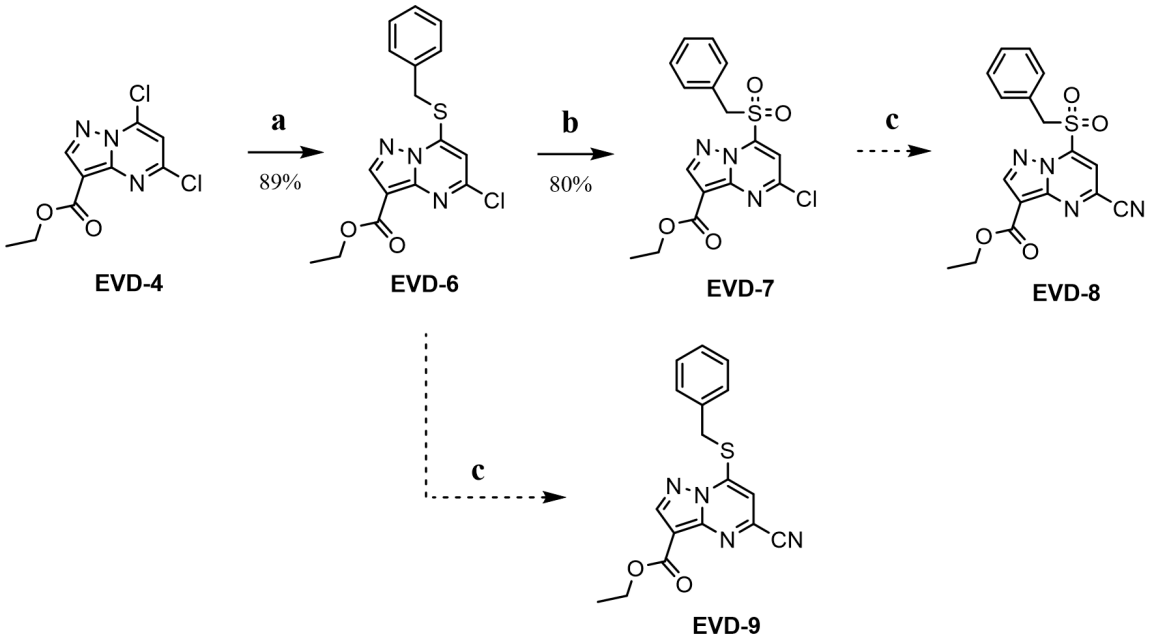


Scheme 2: **Synthesis of key intermediate for hit elaborations** *a) diethylmalonate (1.5 eq), 21% w/w NaOEt (1.5 eq), EtOH (0.89 M), 90 °C, 24 hrs, b) $POCl_3$ (4 eq), MeCN (1 M), pyridine (1 eq), 50 °C, 10 min then 100 °C, 5 hrs*

5.4.0.2 Synthesis of fragment elaborations

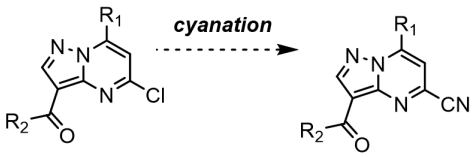
The dichloro intermediate **EVD-4** was firstly reacted with benzyl thiol to yield the substituted 7-benzylthio product **EVD-6** [Scheme 3, step **a**]. In order to achieve selective nucleophilic aromatic substitution only on position 7 and not have 5,7-bisubsituted products, the reaction was performed on ice and was not allowed to progress further than 2 hours. **EVD-6** was then subsequently oxidised to yield sulfone **EVD-7** [Fig. 3, step b].

Subsequent cyanation attempts, as summarised in Table 5.4, on the 5-chloro **EVD-6** and **EVD-7** analogues to yield final compounds **EVD-9** and **EVD-8** respectively proved to be non-trivial [Scheme 3, step c, d]. Initially, palladium-catalysed cyanation attempts using zinc cyanide at high temperatures (120-150 °C) and microwave irradiation were employed. After no conversion was observed, a different approach proposed by Buchwald *et al.* was employed [253]. This approach features mild temperature conditions (< 40 °C) in largely aqueous media (THF/H_2O 1:5) to avoid water soluble cyanide anions poisoning the THF-soluble palladium catalyst, forming inactive off-cycle species, thus inhibiting product formation. However, no conversion was observed as well, possibly because both **EVD-6** and **EVD-7** showed limited solubility in the aqueous solvent system.



Scheme 3: **Attempted synthesis of thioether and sulfonyl hit elaborations** *a) BnSH (1 eq), Et_3N (1 eq), MeCN (0.31 M), 0 °C, 2 hrs b) MCPBA (4.5 eq), DCM (0.1 M), rt, 16 hrs, c) $Zn(CN)_2$ (0.66 eq), tBuXPhos Pd G3 (0.036 eq), H_2O (0.2 M), THF (1 M), 50 °C, 24 hrs*

Synthesis of the elaborated ether compounds was also attempted in a similar manner to that described for the thioethers above. Benzyl alcohol was deprotonated with NaH and reacted with the dichloro intermediate **EVD-4** to yield the 7-substituted product **EVD-11** [Scheme 4, step a]. Despite the stoichiometric, dropwise addition of the deprotonated alcohol and the reaction being performed at 0 °C, the majority of the product consisted of the 5,7-bisubstituted compound **EVD-10**. The desired

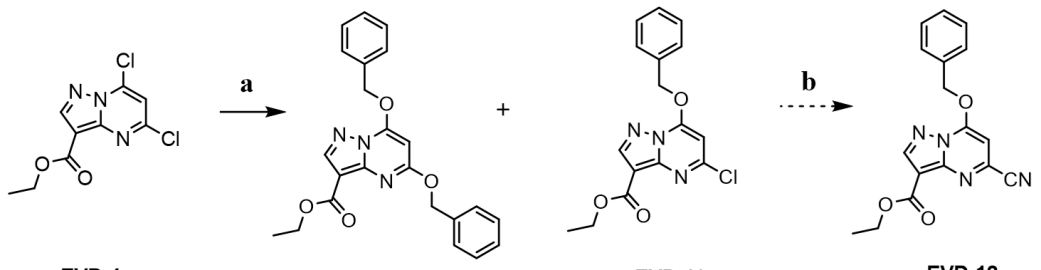


P (product)

R1	R2	Solvent	T (°C)	CN^- source	t (hrs)	Cat	P	Yield
-H	-OEt	Dioxane/H_2O	100	$K_4[Fe(CN)_6]\cdot 3\,H_2O$, KOAc	24	Xphos Pd /Xphos	EVD-1	-
-H	-OEt	DMF	110	$ZnCn_2$	48	$Pd(PPh_3)_4$	EVD-1	50%
-H	-OEt	THF/H_2O	50	$ZnCn_2$	24	tBuXPhos Pd	EVD-1	60%
-SBn	-OEt	NMP	150 (MW)	$ZnCn_2$	3	$Pd(PPh_3)_4$	EVD-9	-
-SBn	-OEt	DMA	120 (MW)	$Zn/ZnCn_2$	3	$Pd(dppf)Cl_2$	EVD-9	-
-SBn	-OEt	THF/H_2O	50	$ZnCn_2$	24	tBuXPhos Pd	EVD-9	-
-OBn	-OEt	THF/H_2O	50	$ZnCn_2$	24	tBuXPhos Pd	EVD-12	-
-H	-NHEt	THF/H_2O	50	$ZnCn_2$	24	tBuXPhos Pd	EVD-15	88%
-SO_2Bn	-OEt	THF/H_2O	50	$ZnCn_2$	24	tBuXPhos Pd	EVD-8	-
-SO_2Bn	-OEt	DMA	110	$ZnCn_2$	24	$Pd(PPh_3)_4$	EVD-8	-

Table 5.4: Conditions used for the cyanation of arylchlorides

monosubstituted product **EVD-11** was subsequently purified by column chromatography. Attempts to cyanate **EVD-11** to obtain the final compound **EVD-12** were unsuccessful, similar to the findings described above [Scheme 4, step **b**].



EVD-4 | EVD-10 56% | EVD-11 8% | EVD-12

Scheme 4: **Attempted synthesis of ether hit elaborations** *a)BnOH (1.1 eq), NaH (1.5 eq) THF (0.96 M), 0 °C, 3 hrs, b) $Zn(CN)_2$ (0.66 eq), tBuXPhos Pd G3 (0.036 eq), H_2O (0.2 M), THF (1 M), 50 °C, 24 hrs*

5.4.1 Synthesis of hit analogues for PROTAC functionalisation

Both the initial hit **EVD-1** and fragment **23** could serve as PROTAC handles for KLHL12 Kelch recruitment. However, due to the low intrinsic reactivity and long half-life (55 hours) of **EVD-1**, as

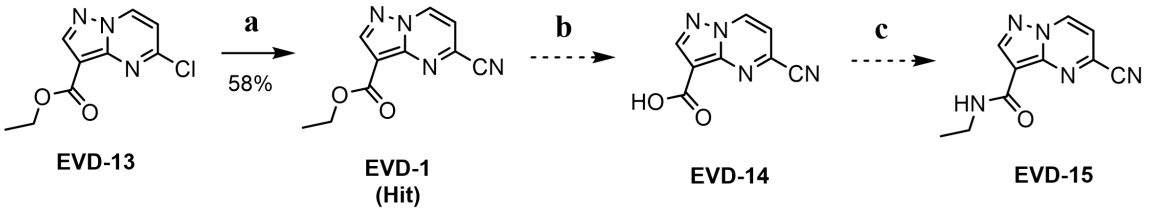
demonstrated in the glutathione assay [Fig. 5.22], along with its high selectivity for KLHL12 over other Kelch domains of E3 ligases [Fig. 5.21], **EVD-1** holds greater prospective value as a PROTAC handle. Consequently, synthetic efforts were focused on **EVD-1**.

The crystal structure of KLHL12 in complex with **EVD-1** hit revealed that the ester group extends into the solvent, providing an exit vector for potential coupling to a PROTAC linker [Fig. 5.5]. For this reason, synthetic routes for corresponding *amide* **EVD-15** and *acid* **EVD-14** of the hit **EVD-1** were designed. The incorporation of an amide group could facilitate amide coupling methodologies, which are widely used to create amide linkages with PROTAC linker libraries.

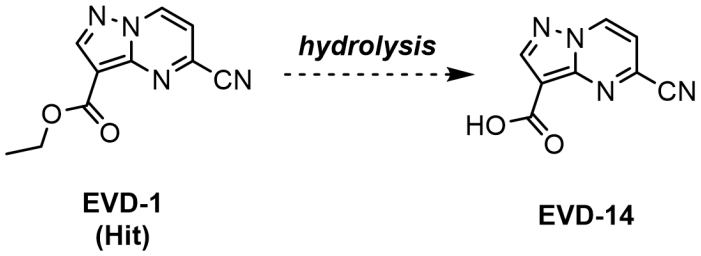
Initially, a synthesis scheme was devised where both free acid and amide analogues would be accessible [Scheme 5]; the hit would be re-synthesised from its 5-chloro commercially available precursor [step a], before ester hydrolysis to the free acid yielding **EVD-14** [step b] and amide coupling to afford final ethyl amide **EVD-15** [step c].

The initial cyanation reaction was successful affording the resynthesised hit **EVD-1**. **EVD-1** was accessible via both palladium-catalysed cyanation using zinc cyanide at high temperatures (110 °C) in organic solvent, and through the milder method in aqueous media outlined above. The subsequent ester hydrolysis (step b) proved to be difficult, despite various attempts as summarised in Table 5.4. Saponification using base LiOH or nucleophilic deprotection using trimethylsilyl iodide under neutral conditions [254] resulted in compound decomposition. Due to the instability under basic conditions, an alternative mild method using $Sn(CH_3)_3OH$ was attempted [255]. Although this resulted in ester cleavage, the ethoxy ion reacted with the cyanide group to afford byproduct 5-(ethoxy(imino)methyl)pyrazolo[1,5-a]pyrimidine-3-carboxylic acid **EVD-16** as seen in Table 5.5. Due to the inherent reactivity of the nitrile warhead and consequent side-product formation, this route was de-prioritised and an alternative one was attempted.

A second route to access the amide analogues starting from commercially available carboxylic acid instead of the ester was designed, as seen in Scheme 6. The carboxylic acid would be first reacted with



Scheme 5: **Attempted synthesis of acid and amide analogues** *a) $Zn(CN)_2$ (0.66 eq), tBuXPhos Pd G3 (0.036 eq), H_2O (0.2 M), THF (1 M), 50 °C, 24 hrs b) $LiOHxH_2O$ (1.5 eq), THF (0.29 M), H_2O (0.72 M), rt, 2 hrs, c) ethylamine (1 eq), EDCI (1.3 eq), DMF (0.15 M), rt, 24 hrs*



Solvent	T (°C)	Reagent	t (hrs)	Yield
THF/H_2O	r.t.	LiOH·H_2O	2	-
DCE	80	TMSI	24	-
DCM	80	$Sn(CH_3)_3OH$	24	-

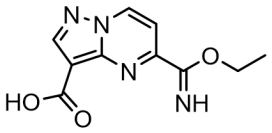


Table 5.5: Conditions used for the hydrolysis of ester **EVD-1**

ethylamine to yield **EVD-18** [step **b**] before being cyanated affording hit amide analogue **EVD-15** [step **c**].

Initial attempts to utilise amide coupling reagents like HATU/DIPEA and DMAP/EDCI did not lead to amide formation [Table 5.6], but led to nucleophilic aromatic substitution on the 5-chloride instead, suggesting the need for further activation of the free acid. Acid transformation to the acyl chloride reactive intermediate **EVD-17** before subsequent reaction with ethylamine afforded amide **EVD-18**. Finally, **EVD-18** was cyanated using the previously established aqueous method to yield amide analogue **EVD-15** [Scheme 6, step c]. A mixture of the desired product **EVD-15** (67% by 1H-NMr) and starting material **EVD-18** (33% by 1H-NMR) was obtained. The presence of **EVD-15** and **EVD-18** was also verified by high resolution mass spectrometry (HRMS) [see Experimental procedures 9.0.1]. Further optimisation of reaction purification is needed to yield pure amide **EVD-15**.

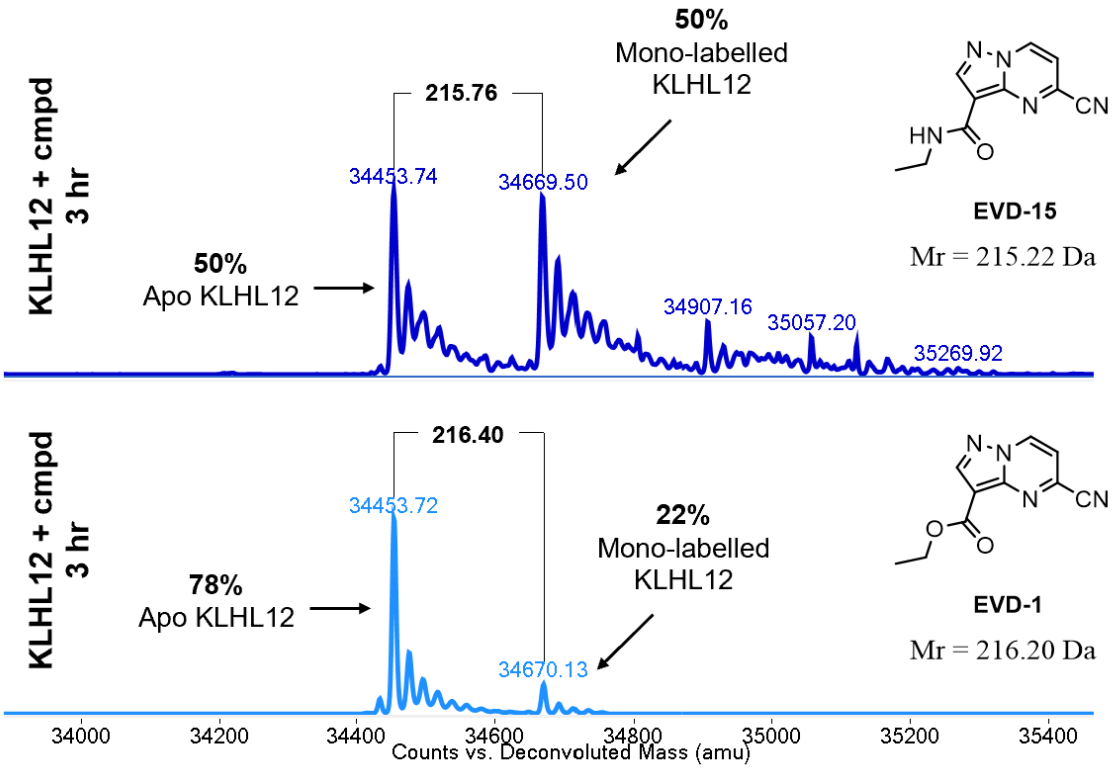
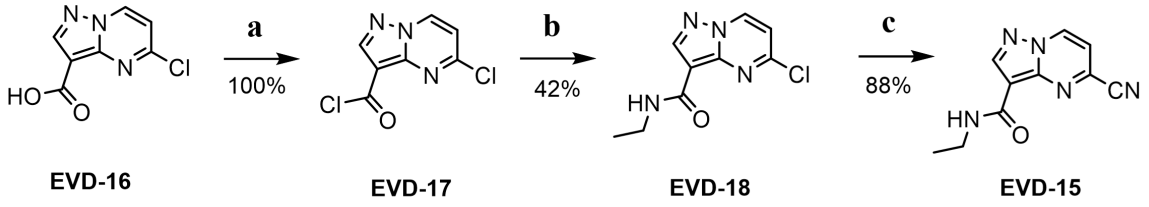
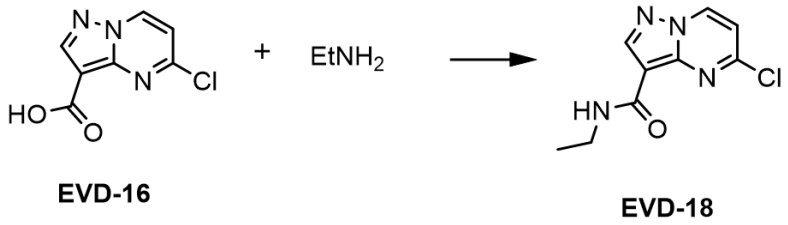


Figure 5.24: **Amide analogue EVD-15 shows greater labelling efficiency than ester hit EVD-1** Mass-spectrometry analysis of amide **EVD-15** efficiency of KLHL12 adduct formation, compared to ester hit **EVD-1**. Compounds (100 µM) were incubated with KLHL12 Kelch domain (2 µM) for 3 hours at room temperature, before mass analysis under denaturing conditions. Quantification based on % area under the curved (AUC) of deconvoluted mass peaks.



Scheme 6: **Synthesis of amide analogue** *a) $SOCl_2$ (54 eq), DCM (0.84 M), 80 °C, 3 hrs b)ethylamine hydrochloride (1.5 eq), Et_3N (3 eq), DCM (0.42 M), rt, 24 hrs c) $Zn(CN)_2$ (0.66 eq), tBuXPhos Pd G3 (0.036 eq), H_2O (0.2 M), THF (1 M), 50 °C, 24 hrs*

The mixture containing 67% amide **EVD-15** was tested for its ability to covalently modify KLHL12 in comparison to the original ester hit, **EVD-1**. Compounds (100 µM) were incubated with the KLHL12 Kelch domain (2 µM) for 3 hours at room temperature before mass analysis under denaturing conditions. The amide analogue, **EVD-15**, exhibited a higher covalent labelling efficiency of KLHL12, modifying 50% of the protein after 3 hours, compared to the 22% labelling observed with the ester hit, **EVD-1**. This finding suggests that amide substitution of the ester not only preserved affinity but enhanced it, making this handle a viable candidate for functionalization in a PROTAC.



Solvent	T (°C)	Reagents	t (hrs)	Yield
DMF	r.t.	HATU, Hünig's base	24	-
DCM	r.t.	DMAP, EDCI	24	-
DCM	(i) 90 (ii) r.t.	(i) $SOCl_2$ (ii) Et_3N	24	41.5%

Table 5.6: Amide coupling conditions used in the synthesis of **EVD-18**

5.5 Discussion

Previous screening efforts identified a covalent binder, **EVD-1**, that selectively targets a cysteine residue within the KLHL12 Kelch substrate-binding site. However, the specific cysteine targeted and the binding mode of **EVD-1** remained unknown. To my knowledge, this represents the first reported small-molecule binder for KLHL12. Therefore, exploring the structure-activity relationship (SAR) of this novel **EVD-1** scaffold was essential to evaluate its potency and selectivity profile, given that this chemotype had not been previously investigated for KLHL12 targeting.

SAR trends were established by comparing the *labelling efficiency* of the KLHL12 Kelch domain using mass spectrometry, and the *binding affinity* using time-depended IC_{50} values determined by substrate DVL1 displacement in AlphaScreen assays. Furthermore, optimising efficient target labelling while minimizing promiscuous reactivity was crucial for covalent binders. Thus, electrophilicity was assessed by reacting fragments with glutathione, an abundant cellular thiol mimicking the cell's reductive environment. Beyond establishing traditional SAR for chemical probe development, I also aimed to determine whether **EVD-1** could serve as a viable starting point for proximity-inducing modalities, such as PROTACs, by assessing its suitability for ligand-based recruitment of KLHL12. To address these questions an integrated approach combining structural biology, biochemistry, and organic synthesis to characterize the novel chemical matter was necessary.

The 2.6 Å co-crystal structure of KLHL12 in complex with the original hit revealed that **EVD-1** selectively modified cysteine 511 (C511) within the substrate-binding site, thereby effectively inhibiting native peptide interactions. Structure-activity relationship (SAR) analyses indicated that binding was driven not only by reactivity but also by target specificity. The pyrazolopyrimidine core of **EVD-1** demonstrates low intrinsic reactivity (half-life: 55 hours) and remained stable at cellular levels of glutathione (GSH) [Fig. 5.22]. Furthermore, **EVD-1** exhibited notable selectivity, reacting exclusively with the KLHL12 Kelch domain among 15 other Kelch domains from various E3 ligases tested [Fig. 5.3.1]. Additionally, **EVD-1** exhibited sub-micromolar apparent IC_{50} values, even at short incubation

times (3 hours) highlighting the compound's potency. Such highly potent compounds are typically effective at low cellular concentrations, thereby reducing the likelihood of polypharmacological effects of **EVD-1** use *in cellulo*.

The KLHL12-**EVD-1** co-crystal structure also allowed for employing *structure-based* approaches, such as covalent docking, and *ligand-based* approaches, such as generation of pharmacophore hypothesis used as a filter for virtual screening of commercial libraries and our collaborators in LUMC ligand collection. This leverage of computational tools enabled the establishment of *"SAR by catalogue"*, eliminating the need for time-consuming initial organic synthesis. Such combination of structure-based docking and ligand-based pharmacophoric approaches have been successfully used for fragment-guided discovery of monoacidic KEAP1-Nrf2 interaction inhibitors that mimic Nrf2 C-terminal diGly degron motif [175].

Exploration of the structure-activity relationships of **EVD-1** and its analogues is summarized in Fig. 5.28, highlighted the impact of γ electron-withdrawing effects on nitrile reactivity. This analysis led to the identification of a modified scaffold with increased reactivity, fragment **23**. Fragment **23** demonstrated sub-micromolar inhibitory effects, and a 2.7 Å co-crystal structure with KLHL12 revealed that the new scaffold adopted a binding mode similar to that of **EVD-1**.

Comparing the co-crystal structures of Kelch domain KLHL12 in complex with **EVD-1** and fragment **23** reveal the fragments' common binding mode and interactions with the Kelch substrate binding site. Surprisingly, despite its small size, the **EVD-1** forms three hydrogen bonds in addition to its covalent bond with KLHL12, two of which involve the thioimidate group. The linear geometry of both the nitrile and thioimidate adduct, combined with the imine group's dual hydrogen bond donor-acceptor properties, is critical for target affinity [Fig. 5.20]. This explains why more reactive covalent warheads such as chloroacetamides and Michael acceptors (e.g., α-β-unsaturated carbonyl compounds) were not predominant KLHL12 binders from our original covalent collection screening [**Chapter 4**].

Notably, no density was observed near neighbouring C558 or C539 on the side of the β-propeller

despite both cysteines being readily accessible and solvent-facing, and the latter predicted to be reactive [**Chapter 4.9**]. A quantum mechanical computational study on nitrile-based cysteine protease inhibitors [256], proposed a reaction mechanism, whereby the thiolate group of cysteine reacts via a nucleophilic addition reaction with substrate to form a covalent linkage. The formed intermediate then abstracts a proton from a nearby histidine residue to give the final product complex. Thus, a possible explanation for the selective binding to C511 and not C558 could be that the proximal H486 acts as a proton source for covalent intermediate with a similar mechanism [Fig. 5.25].

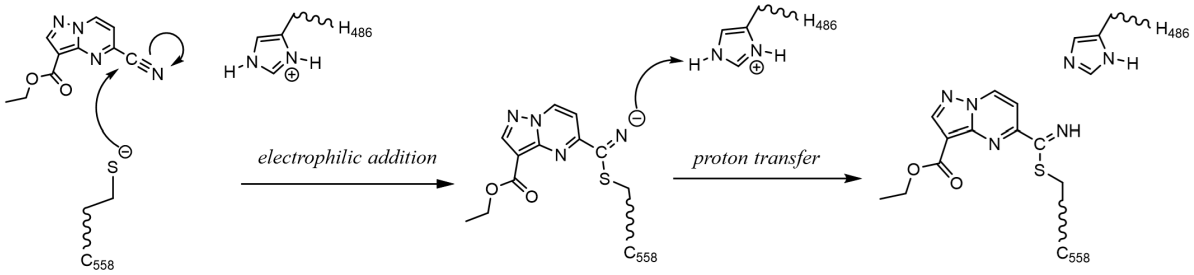


Figure 5.25: **Potential reaction mechanism of covalent bond formation of EVD-1.** The thiolate group of KLHL12 cysteine C511 reacts with the nitrile of **EVD-1** forming a covalent linkage. The resulting intermediate then abstracts a proton from the neighboring histidine H486 residue, yielding the final thioimidate product. The proximity of histidine H486 to C511 may explain the selectivity for adduct formation over C558. This putative reaction mechanism is adapted from the proposed mechanism by Quesne *et al.* [256].

As discussed previously [**Chapter 3.3**], KLHL12 presents a predominantly rigid protein surface with high shape complementarity to its binding partners and undergoes minimal conformational changes [72]. To determine whether compound binding mimics native substrate peptide interactions, we compared binding site residues in the hit and fragment **23** complexes with the DVL1-bound complex. All residues maintain conserved conformations except Y512, which undergoes a 23.5° rotation of the aromatic plane inwards, towards the bound **EVD-1** or fragment **23** [Fig. 5.26] to accommodate both covalent compounds. This subtle aromatic ring movement allows the compound to be sandwiched in-plane between Y512 and F481. The minimal conformational change, coupled with hydrogen bonding, draws clear parallels between compound and peptide binding.

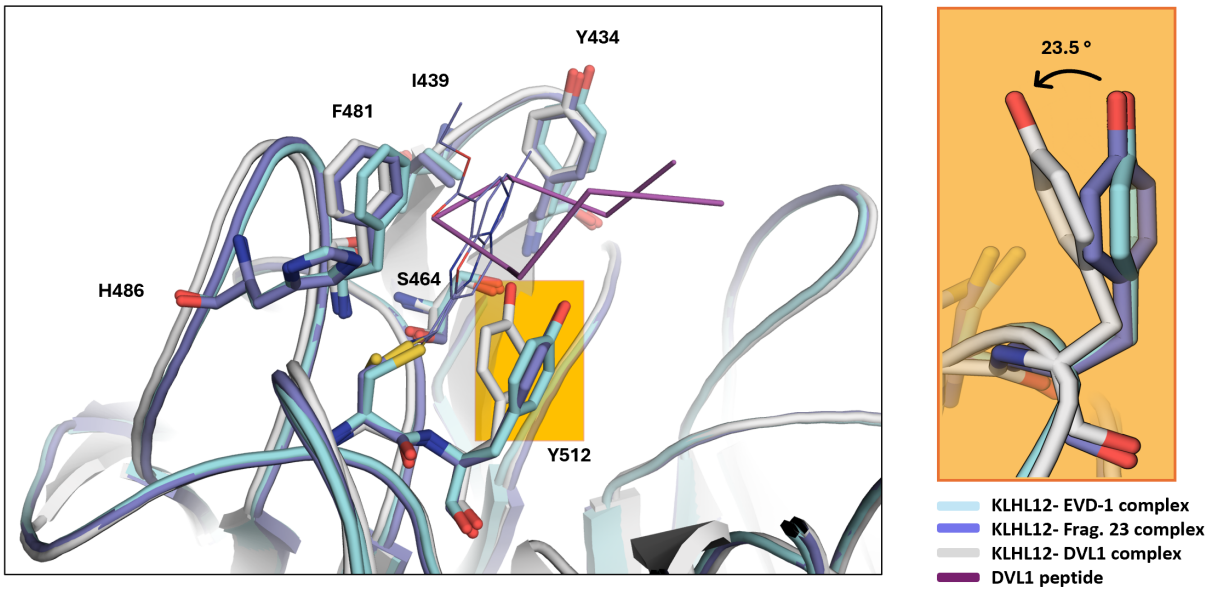


Figure 5.26: **Compound binding induces Y512 shift in otherwise rigid KLHL12 binding site.** Structural alignment of the Kelch domains bound to DVL1 (white), **EVD-1** (cyan), fragment **23** (purple). Residues involved in binding site interactions are shown as thick sticks. Bound compounds **EVD-1** and fragment **23**, and DVL1 C alpha chain are shown as thin sticks. Compound binding for both hit and fragment **23** induces a 23.5 ° rotation of the aromatic ring plane inwards, towards the bound compounds, compared to DVL1-bound structure.

Interestingly, in addition to the conformationally constrained nature of the KLHL12 binding site and its selective interaction with PGXPP-containing proteins that adopt a conserved binding mode [**Chapter 3**], both the KLHL12 substrate-recognition motif (PGXPP) and the receptor binding site itself show strong evolutionary conservation across multiple species [Fig. 5.27]. Most notably, the reactive cysteine targeted by **EVD-1** (C511), along with neighboring cysteines C539 and C558 are highly conserved, supporting their potential functional relevance. Furthermore, since C511 is conserved in preclinical species such as mouse and rat, **EVD-1** and its derivatives are likely to exhibit similar reactivity and pharmacological behaviour *in vivo*, facilitating future animal studies aimed at assessing the phenotypic consequences of KLHL12 inhibition.

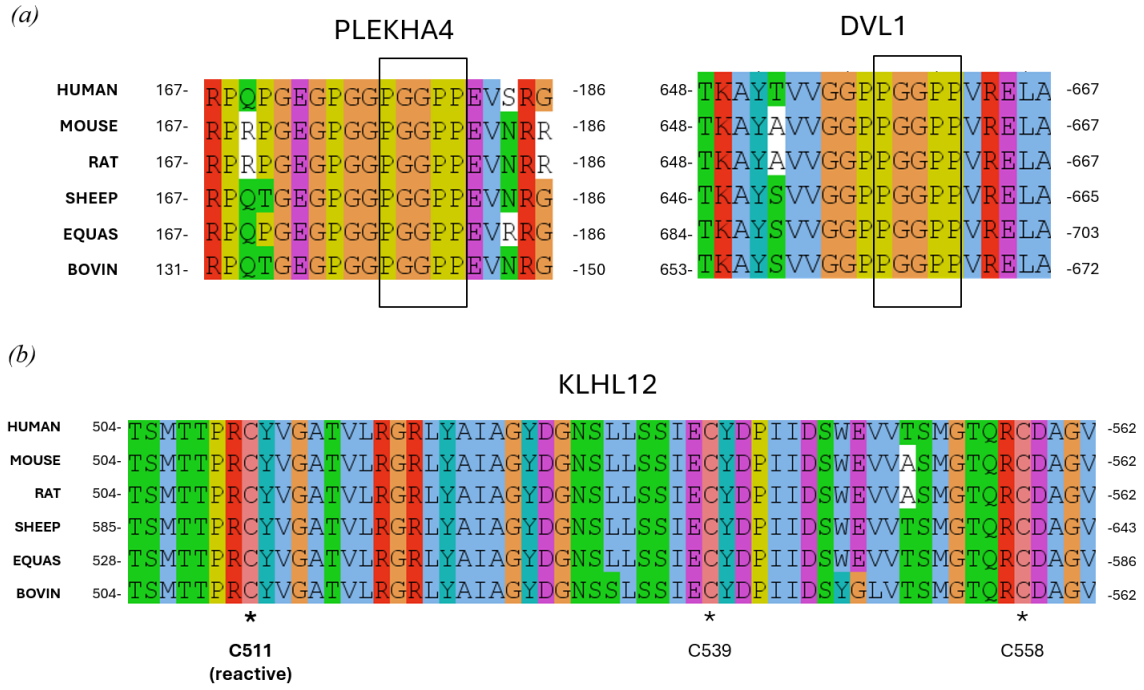


Figure 5.27: **KLHL12 interactor recognition mechanism is evolutionary conserved**. (a) Multiple sequence alignment of KLHL12 antagonist PLEKHA4 and substrate DVL1 proteins across a range of mammalian organisms. The "PGXPP" motif involved in physiological substrate and antagonist recognition shows high evolutionary conservation. (b) Alignment of KLHL12 substrate-binding Kelch domain region highlighting conserved cysteine residues, including reactive C511 (**EVD-1** binding site, marked with an asterisk), and additional accessible cysteines C539 and C558. Conservation of both the substrate recognition motif and the small-molecule binding site across species suggests functional relevance and evolutionary constraint. Multiple Sequence Alignment performed using Clustal Omega [257] and visualised with JalView [258]. Residues are coloured by Clustal X colouring.

The oxazole-based fragment **23** has a shorter half-life (6 hours) compared to **EVD-1** (55 hours), though it still falls within acceptable ranges. For comparison, the reaction half-life of ibrutinib (1

mM), a first-in-class Bruton's kinase inhibitor used in the treatment of multiple lymphomas, in the presence of 10 mM GSH, is 54.5 minutes [259]. Non-optimized KEAP1 covalent hits, which were not developed into probes, exhibit even shorter half-lives, all below 2 hours [179]. However, subsequent analysis of fragment **23** selectivity, following incubation with a panel of E3 ligases, revealed that fragment **23** reacted with the majority of the tested proteins, labelling them with higher efficiency than KLHL12. Literature studies on covalent nitrile warheads support the observation that oxazole nitriles exhibit high reactivity, as electron-withdrawing heteroatoms near the nitrile warhead promote polarization through both inductive and resonance effects, thereby reducing electron density at the cyano carbon [260]. This promiscuity of fragment **23** reduced its value as a chemical probe tool for biological studies and increased concerns about its potential toxicity, as it could inhibit unrecognized pathways and pose risks for drug development. Thus, my efforts for further study focused on the non reactive and selective **EVD-1**.

The co-crystal structures also reveal potential fragment extension vectors targeting a hydrophobic pocket in the vicinity. Fragment extension is essential in hit-to-lead development to enhance potency and selectivity, particularly over other homologous Kelch family members and human proteins bearing solvent-accessible reactive cysteines. To this end, I proposed elaboration strategies and developed a synthetic route to assess their impact on KLHL12 interaction, selectivity, and cross-reactivity with the cysteine-containing proteome.

Furthermore, the co-crystal structures suggest an exit vector for **EVD-1**, which could enable potential PROTAC functionalisation. I assessed the synthetic accessibility and proposed a route for generating amide analogues of the ester group, facilitating amide-coupling chemistry for linker conjugation.

Testing the amide analogue **EVD-15** for its ability to covalently label KLHL12 revealed an enhanced labelling efficiency compared to the ester-containing hit **EVD-1**. This finding suggests that amide substitution not only preserved affinity but also enhanced it, positioning this handle as a viable candidate for functionalization in a PROTAC. The amide analogue **EVD-15** was specifically synthesized

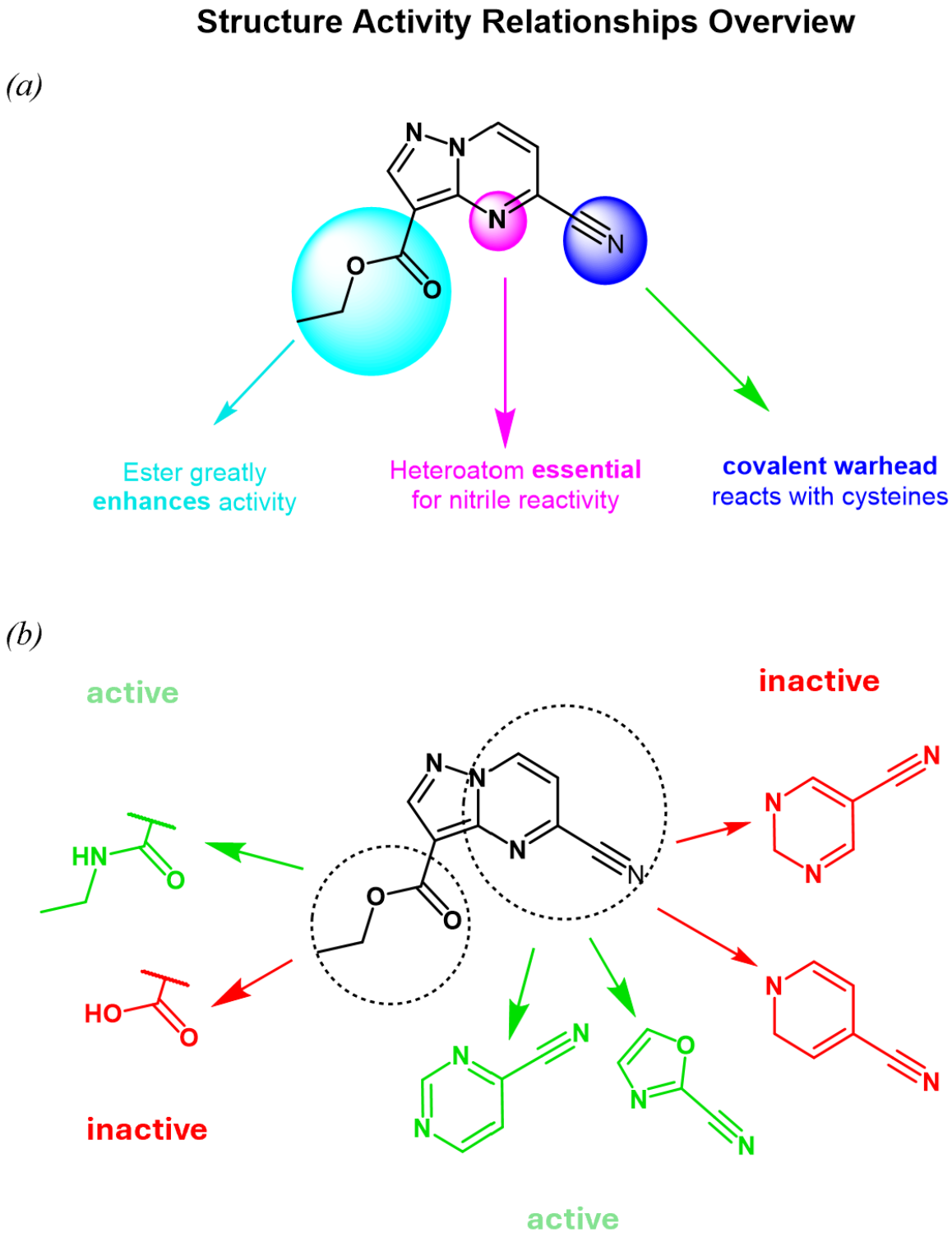


Figure 5.28: **Structure activity relationship overview for EVD-1**

to facilitate PROTAC development, as amide coupling chemistry is commonly employed for linker attachment. The observed increase in labelling efficiency may be attributed to the amide carbonyl oxygen forming a more stable hydrogen bond with H486. Unlike esters, amides exhibit greater resonance stabilisation, which increases the partial negative charge on the carbonyl oxygen, making it a stronger hydrogen bond acceptor. Additionally, the planarity of the amide group enhances its ability to align with hydrogen bond donor H486, whereas the increased flexibility of esters may hinder optimal hydrogen bonding. This improved interaction likely contributes to the enhanced affinity of **EVD-15** for KLHL12.

Notably, screening of Leiden University's library identified expanded analogues approaching PROTAC-relevant molecular weights, suggesting that KLHL12 can accommodate such extended structures [Fig. 5.7]. These findings establish a strong foundation for further optimization of **EVD-1** and **EVD-15** as chemical probes and PROTAC handles, enabling the targeted degradation of KLHL12, an E3 ligase that has remained largely unexplored from a chemical biology perspective.

CHAPTER 6

Exploring molecular glue mechanism in KLHL12-USP21 assembly

6.1 Introduction

In parallel with inhibitor development efforts that led to the discovery of **EVD-1** as described in **Chapters 4** and **5**, I also explored an alternative path that could lead to identification of novel chemical matter for KLHL12 approaching it from a different angle; not that of competitive inhibition on the substrate binding domain of the $CUL3^{KLHL12}$ ligase, but exploring novel opportunities for targeting the multi-protein molecular assembly with cullins. These efforts are presented in this chapter. While the development of **EVD-1** ultimately proved more tractable, the findings reported here represent the first *in vitro* exploration of a potential molecular glue compound for KLHL12.

This work started after SGC Toronto's recent report of a novel inhibitor for deubiquitinase USP21, BAY-805 [261]. During their efforts on characterising the probe with mass spectrometry analysis of transient interactors of USP21, it was observed that BAY-805 induced a significant enrichment of $CUL3^{KLHL12}$ compared to experiments analysing interactors in its absence (unpublished data, bioID quantitative mass proteomics experiment from Gstaiger group, ETH Zurich). The BAY-805 mediated USP21-KLHL12 interaction was further validated in cells using NanoBRET experiments, expressing NanoLuc-USP21 and HaloTag-KLHL12 fusion proteins (unpublished data, Dr Victoria Vu, Dalia Barsyte Lovejoy Group, SGC Toronto). In this experiment, in case of BAY-805-induced proximity of the two fusion proteins, resonance energy transfer was induced from NanoLuc to Halo Tag ligand (HL) bound to a HaloTag (HT) emitting a fluorescence signal. This interaction was observed only

with full-length KLHL12, whereas deletion of the BTB, BACK, or Kelch domains abolished binding. This suggested that no single domain is solely responsible for enabling the interaction between BAY-805 and USP21.

Since $CUL3^{KLHL12}$ ligase dimerises via the BTB domain at the N-terminus of KLHL12 [106], we hypothesised that the BTB-BACK domain serves a critical scaffolding role in dimerisation rather than directly interacting with USP21 or the probe. To test this hypothesis, constructs were designed in which the BTB-BACK domain was replaced with either HiBIT peptide tag or LgBIT, which bind with 1 μM affinity to reconstitute NanoLuc, thereby functioning as surrogate dimerisation domains [Fig. 6.1]. Interestingly, this engineered dimer could also engage USP21 in the presence of BAY-805 in cells, suggesting that the dimerised Kelch is essential for the interaction and BTB-BACK has a scaffolding role [Fig. 6.1 *(a)*].

In this chapter, in collaboration with SGC, University of Toronto, I investigated whether USP21 inhibitor BAY-805 [261] recruits homodimeric $CUL3^{KLHL12}$ to the deubiquitinase. Since no direct interaction between USP21 and KLHL12 has been previously reported, nor are they linked in known biological pathways, this chapter explores the hypothesis that BAY-805 acts as a molecular glue bridging KLHL12 and USP21. In order to facilitate exploration of both of these systems, high-yielding protein constructs of preferably full-length NRBP1 and KLHL12 and their interacting partners were needed. Ligation independent cloning was performed to access both full-length constructs, as well as engineered HiBIT-LgBIT KLHL12 Kelch domain dimers, and their expression levels were then assessed in bacterial *E.coli*, and eukaryotic insect Sf9 or mammalian Expi293 cells. Protein purification protocols for novel protein targets were established. Microscale thermophoresis that detects the motion of fluorescent molecules in solution was used to assess binding affinity of the putative KLHL12-USP21 complex. Purified protein complexes were also used to set up crystallisation trials and cryo-EM grids in order to uncover the structural basis of BAY-805-mediated KLHL12-USP21 assembly.

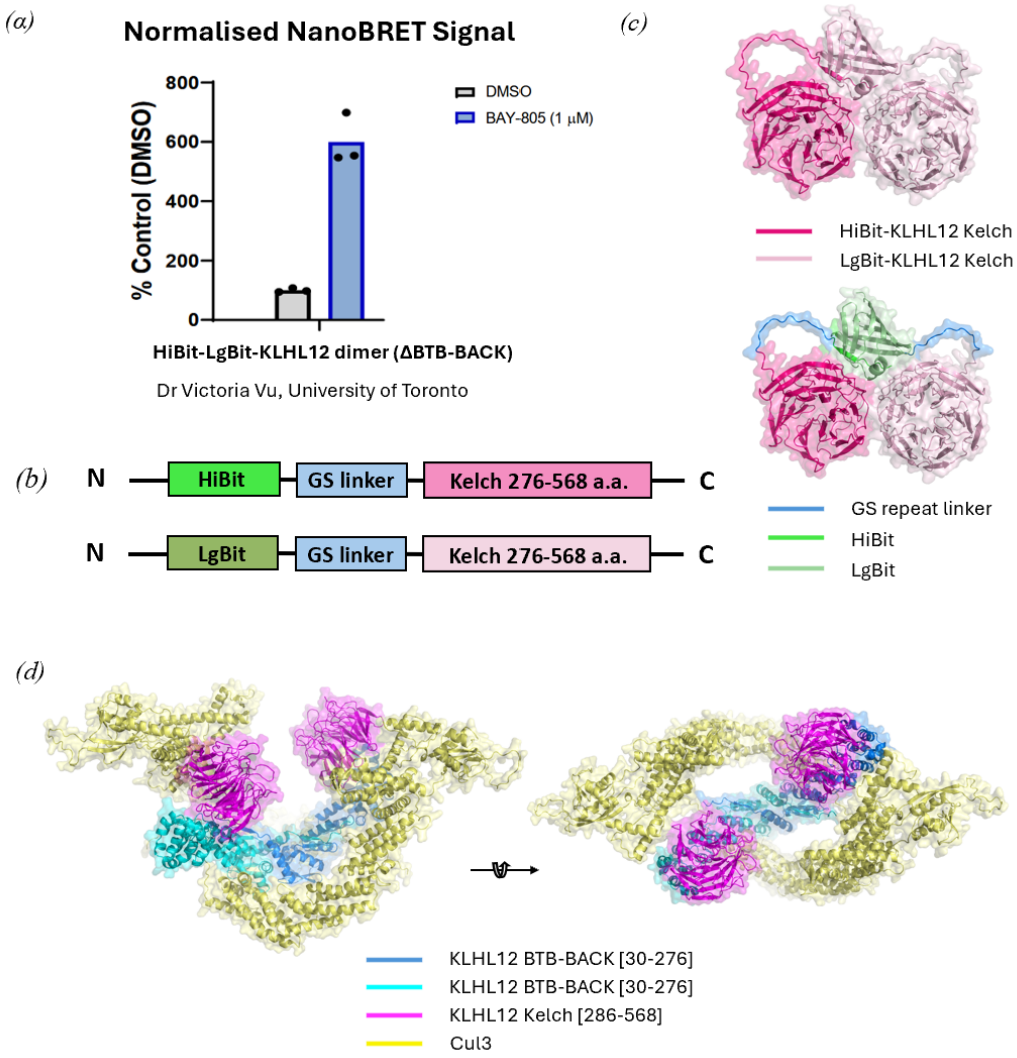


Figure 6.1: **USP21 interacts with engineered HiBIT-LgBIT KLHL12 dimer even in the absence of BTB-BACK.** (a) Normalized BRET signal shows that HiBIT-LgBIT dimer interacts with KLHL12 in the presence of BAY-805, by Dr Victoria Vu, Dalia Barsyte Lovejoy Lab, (b) Construct design of HiBIT-LgBIT, (c) Model of ΔBTB-BACK HiBIT-LgBIT KLHL12 reconstituted dimer by AFv3, (d) Model of BTB-BACK-Kelch containing physiological dimer of $CUL3^{KLHL12}$ by AFv3.

6.2 $CUL3^{KLHL12}$ E3 ligase associates with USP21 in the presence of BAY-805

6.2.1 Expression and purification trials of dimeric KLHL12

Since dimeric KLHL12 was necessary for USP21 interaction, extensive protein expression and purification trials were conducted in search for a construct enabling recovery of the protein from the soluble cell lysate.

Previous attempts in our lab, prior to my DPhil studies had failed to yield full length soluble KLHL12 in bacterial *E. coli* BL21 cells and insect Sf9 expression systems. Thus, full-length KLHL12 (1-568 a.a.), N-terminally tagged with tandem cleavable 3xFLAG and hexahistidine tags, was cloned into vectors suitable for Expi293 cell transfection for mammalian expression: (i) *pcDNA* under the CMV promoter, (ii) *pLEXm* under the stronger chicken β-actin promoter [262], both allowing transient expression, and (iii) *pHTBV* and (iv) tetracycline-induced *MC260* (pHL-based) [263], which enable stable transfection.

Systematic optimisation of transient transfection and expression conditions of the above vectors in mammalian Expi293 cells was performed to ascertain which yielded better recovery of protein in the soluble fraction after test purification. The effect of the following conditions was assessed: (i) polyethylenimine (PEI) to DNA ratio during transfection, (ii) expression temperature, and (iii) expression duration in small-scale expression in 1 mL of Expi293 cells as detailed in methods section 2.1.3. The cells were centrifuged and the pellets were lysed in 200 µL lysis buffer containing detergent NP40 0.25% that is disruptive to the cell membranes, followed by centrifugation to remove the cell debris. Equal amount of clarified lysate samples of each tested condition were analysed on SDS-PAGE gel and probed with α-FLAG antibody [Fig. 6.2 (a)]. Qualitative comparison of band intensity across these constructs revealed that KLHL12 in *pHTBV* vector transected at 1:6 DNA:PEI ratio and expressed in 30 °C for 72 hours had the highest recovery of protein in the soluble fraction after test purification [Fig. 6.2 (a)].

Due to persistently low protein yields, further optimisation was carried out using the same small-

scale procedure in 3 mL Expi293 cell cultures. Many RING-type E3 ligases are subject to auto-ubiquitination, a self-regulatory mechanism that targets them for proteasomal degradation, often resulting in poor yields or rapid turnover in heterologous expression systems. For example, F-box proteins like Grr1p and HECT ligases such as Rsp are known to undergo ubiquitin-mediated degradation as a result of their self-ubiquitinylation [264, 265], while RING ligases such as SIAH1 and TRAF6 further require homodimerisation prior to self-ubiquitination [266, 267]. Similar mechanisms may account for the low recovery of KLHL12, particularly if adaptor-induced destabilisation of the Cullin-RING complex is at play. To avoid KLHL12 degradation via self-catalysed auto-ubiquitinylation precipitated due to over-expression, the addition of the neddylation inhibitor MLN4924 [268] and the proteasome inhibitor MG132 [269] were tested, but had no effect, indicating that proteasome or neddylation inhibition does not rescue low expression levels [Fig. 6.2 (b)].

Western blot qualitative analysis of equal volumes of clarified supernatant from the 3 mL small-scale cultures suggested that KLHL12 is more abundantly recovered from lower temperatures and shorter expression durations [Fig. 6.2 (a)]. In these experiments, normalisation to total protein was intentionally not performed, as the aim was to assess differences in total protein yield in the supernatant, which inherently reflected both protein expression and cell growth under each condition. Therefore, these results were intended to compare overall recovery between conditions, rather than normalised expression per cell or per protein content.

Large-scale expression of full-length KLHL12 in *pHTBV* vector after transient transfection in Expi293 cells (1 L) under the optimised conditions identified above (1:6 DNA/PEI 30 °C for 72 hours) was performed. FLAG-KLHL12 was captured from the clarified lysate on an anti-FLAG resin. The resin was washed, and the bound proteins were eluted with 3xFLAG peptide as described in methods section 2.1.3. After SDS-PAGE analysis of all samples, two observable bands of the expected molecular weight of KLHL12 were detected on SDS-PAGE for the first time[Fig. 6.3 (a)]. To confirm protein identity and distinguish between the two bands, each one was excised from the gel and analysed with

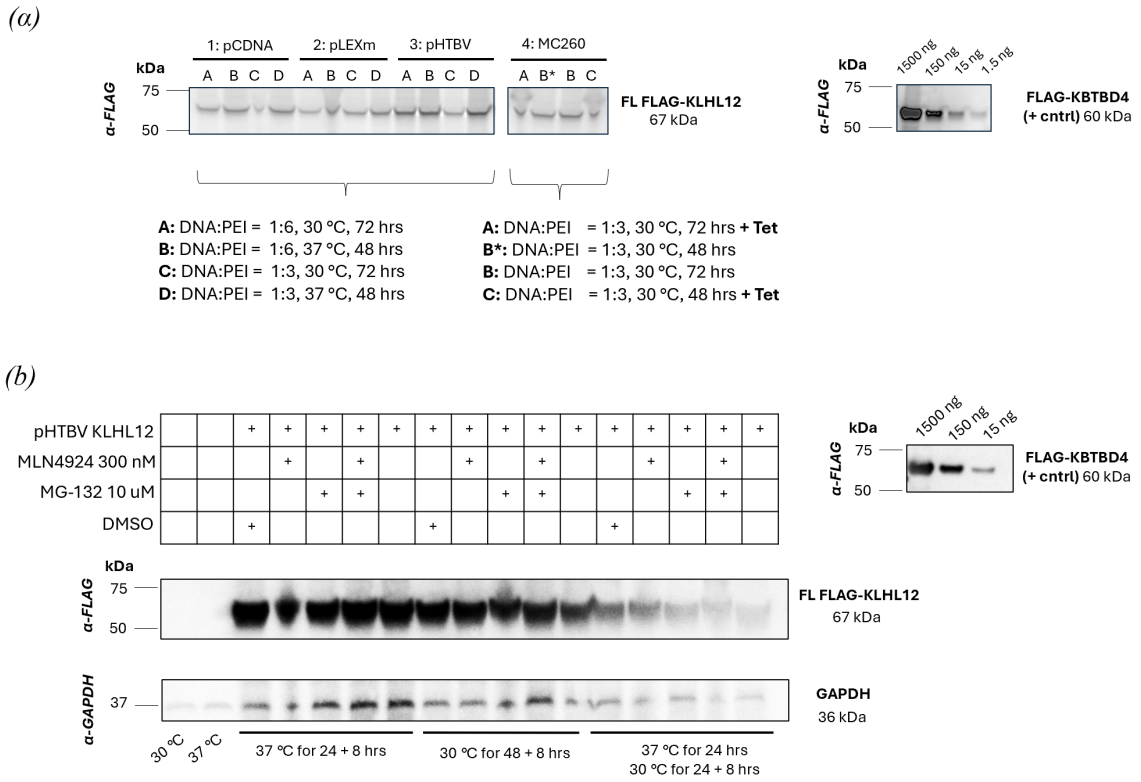


Figure 6.2: **Full length KLHL12 expression optimisation** Western blot analysis of full-length FLAG-tagged KLHL12 transiently transfected in Expi293 cells, probed with α-FLAG antibody. (a) Assessing expression levels of different vectors and transfection conditions in 1 mL small-scale cultures. A qualitative comparison of band intensity suggests that KLHL12 expressed from the *pHTBV* vector (highlighted in red) with a **1:6 DNA:PEI transfection ratio** and incubated at **30 °C for 72 hours** resulted in the highest recovery of protein in the soluble fraction. Additionally, a comparison with the KBTBD4-FLAG positive control indicates that the KLHL12 band corresponds to an approximate protein quantity of **15 ng** from 10 µL of clarified supernatant. (b) Assessing expression levels upon proteasomal and neddylation inhibition of 3 mL small-scale cultures. Cells were transfected with *pHTBV-KLHL12* and treated with **MLN4924** (300 nM, neddylation inhibitor) and/or **MG-132** (10 µM, proteasomal inhibitor) for the indicated time points. α-GAPDH was used as a loading control reflecting cell-growth from each expression condition. Additionally, a comparison with the KBTBD4-FLAG positive control indicates that the KLHL12 band corresponds to an approximate protein quantity of **1500 ng** from 10 µL of clarified supernatant. The observed higher expression in the 3 mL cultures compared to the smaller 1 mL cultures could be explained by improved aeration/shaking in larger volumes that enhances protein expression efficiency compared to miniature cultures. *Note on notation:* "37°C for 24 h + 8 h" indicates that cells were first incubated at 37°C for 24 hours, after which the neddylation or proteasomal inhibitor was added, and incubation was continued for an additional 8 hours before harvesting.

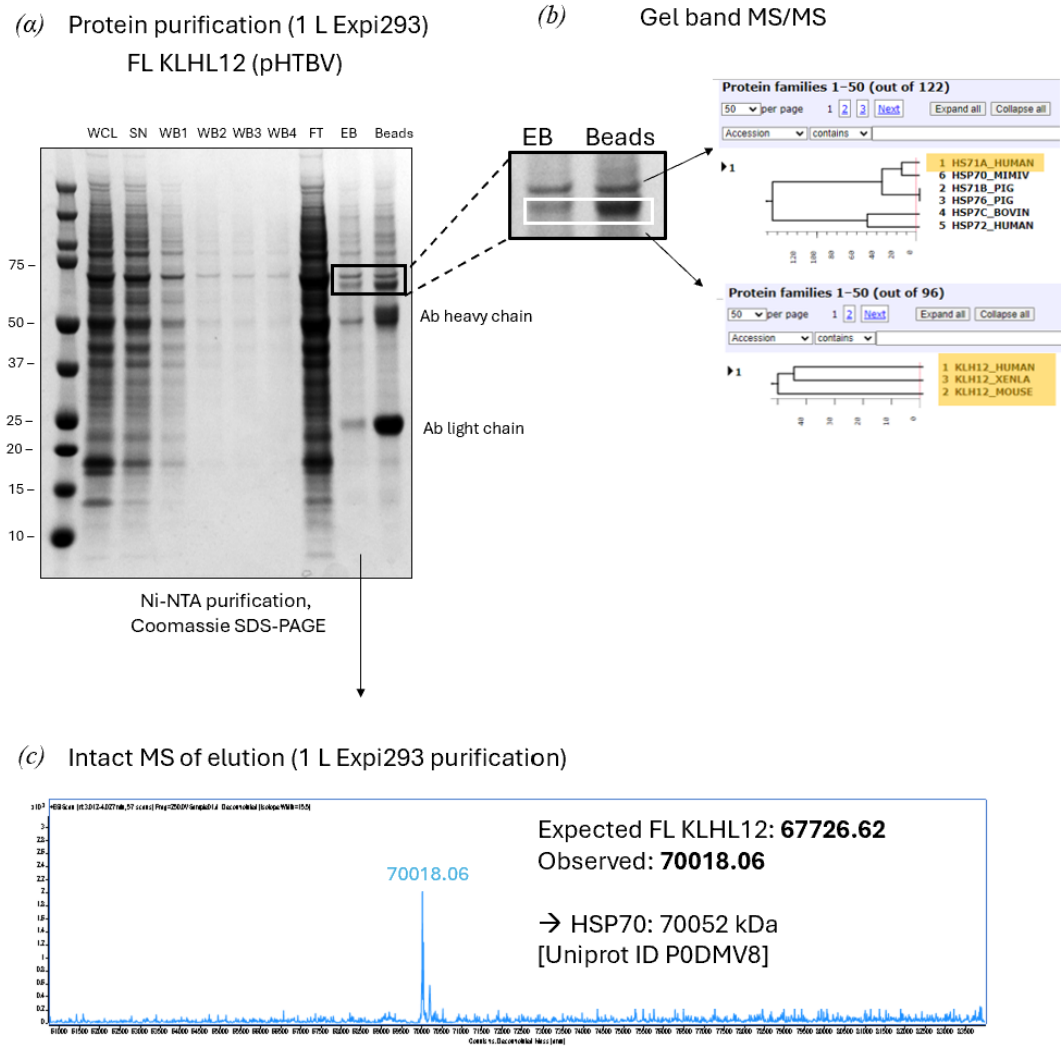


Figure 6.3: **Large scale protein expression of full-length KLHL12** (a) 1 L Expi293 FLAG purification analysed on SDS-PAGE gels. WCL = whole cell lysate; FT = flow-through; SN = supernatant; WB = wash buffer; EB = elution buffer, (b) MS/MS analysis of SDS-PAGE gel bands performed by Dr. Rod Chalk and Dr. Kavya Clement, (c) Intact MS of eluted fraction

tandem mass spectrometry MS/MS by Dr. Rod Chalk and Dr. Kavya Clement [method described in 2.6.2]. The lower band was confirmed to be full-length KLHL12, while the higher molecular weight band accompanying it was detected as heat shock protein 70 (HSP70)[Fig. 6.3 (b)]. Additionally, intact mass spectrometry under denaturing conditions of the eluted sample identified a peak likely to be only molecular chaperone HSP70 and not KLHL12 [Fig. 6.3 (c)]. These results taken together with heavy aggregation of the sample upon concentration, even with detergent supplementation (NP-40 0.25%), suggesting that the sample is unsuitable for further use.

Since it has been previously shown in cellular assays that engineered dimeric constructs with two KLHL12 Kelch domains rather than full-length KLHL12 was sufficient for USP21 interaction, protein expression trials of the engineered HiBit-LgBit KLHL12 were performed. Since Kelch domains and LgBIT-HiBIT complex individually express in BL21 *E. coli* in soluble lysate, bacterial expression of the fused HiBIT-LgBIT-KLHL12 Kelch domains was attempted. Mammalian expression constructs of ΔBTB-BACK HiBit and LgBit fusions to the KLHL12 Kelch domain were a kind gift from Dr. Dalia Barsyte Lovejoy and Dr. Peter Loppnau [Fig. 6.1]. These were subcloned and expressed following established protocols [Methods 2.1.1]. Briefly, The HiBit and LgBit inserts were initially amplified from the DNA templates of these mammalian expression vectors which were previously used in cellular NanoBRET assays. These inserts were then subcloned into the N-terminally His-tagged bacterial expression vector pNIC-Bsa4 and transformed into *E. coli* Mach1 cells [Fig. 6.4 (a), (b)]. Following colony PCR screening and sequence validation, large-scale expression and Ni-NTA purification of each KLHL12 fusion construct were performed in BL21(DE3)-RR Rosetta cells.

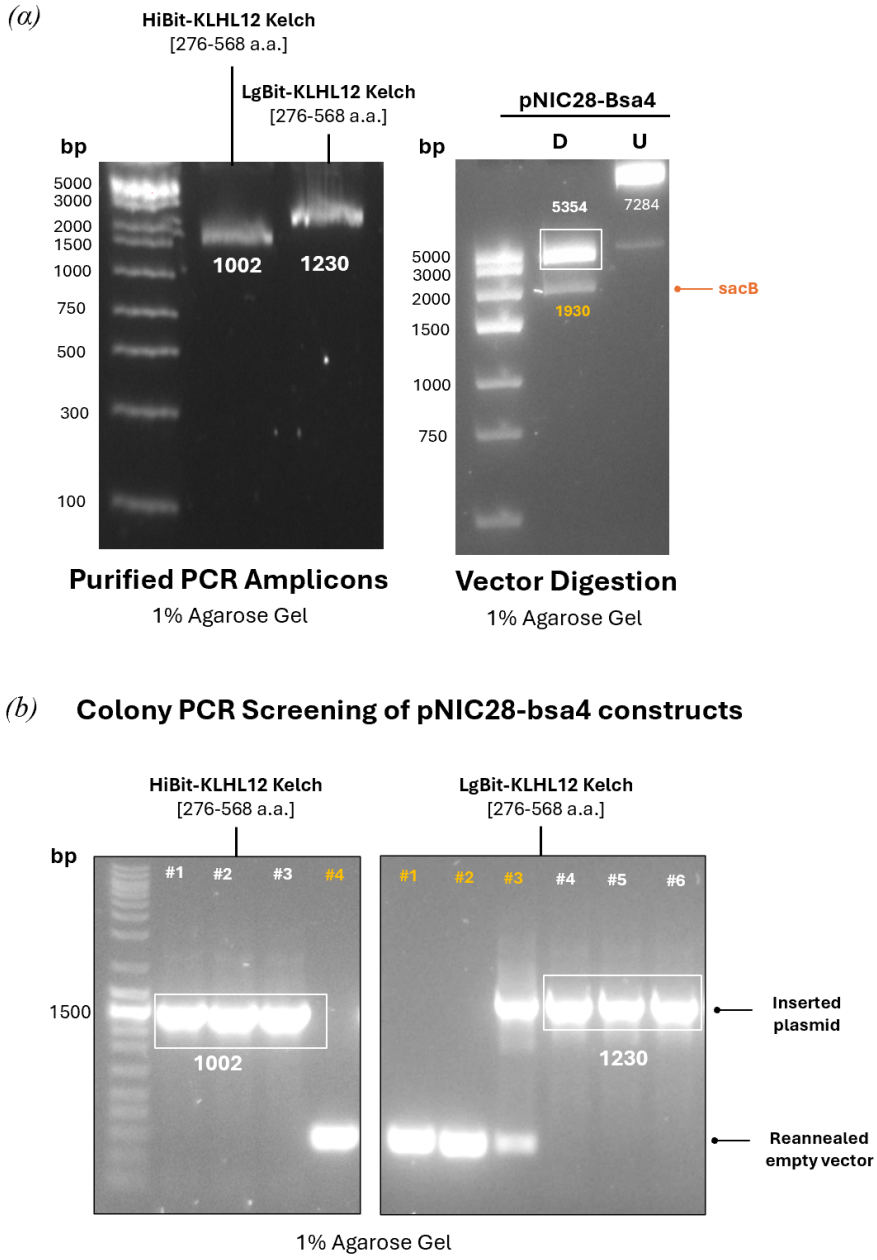
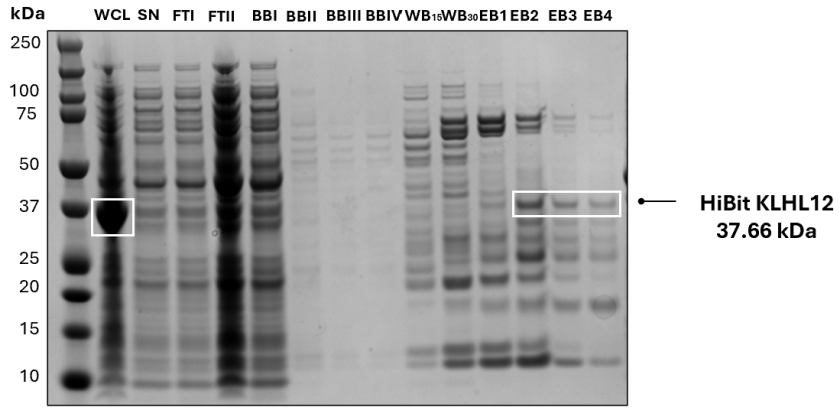


Figure 6.4: **Cloning HiBIT and LgBIT KLHL12 into bacterial expression vectors** (a) PCR amplification and vector digestion for cloning. *Left*: Agarose gel electrophoresis of purified PCR amplicons encoding HiBiT-KLHL12 Kelch (expected size: 1002 bp) and LgBiT-KLHL12 Kelch (expected size: 1230 bp). *Right*: Digestion of the pNIC28-Bsa4 vector (U: uncut, D: digested) with BsaI. The digestion results in two fragments: a larger 5354 bp fragment corresponding to the linearized vector backbone and a 1930 bp fragment containing the *sacB* counterselection marker. Successful removal of *sacB* allows for positive selection of recombinant clones. (b) Colony PCR screening of recombinant pNIC28-Bsa4 constructs. *Left*: Screening of HiBiT-KLHL12 Kelch clones. Positive colonies exhibit a PCR product of 1002 bp, confirming correct insertion. *Right*: Screening of LgBiT-KLHL12 Kelch clones. Positive colonies yield a 1230 bp band. In both panels, the presence of a lower band (corresponding to the reannealed empty vector) in negative clones confirms failed insertion and is coloured in orange.

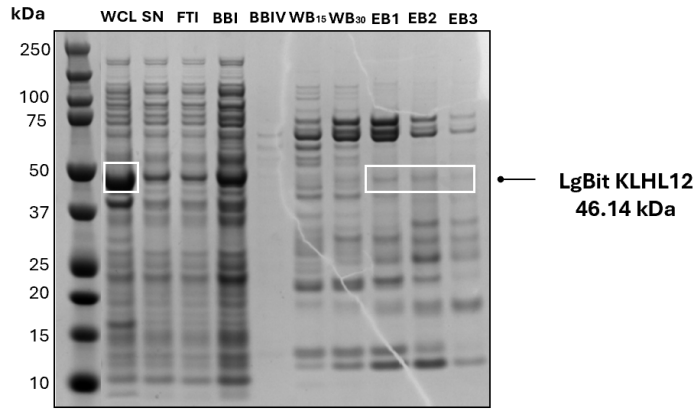
Large-scale expression (4 L culture) under standard KLHL12 purification conditions [2.1.4] resulted in only trace amounts of KLHL12 protein in the elution fractions [Fig. 6.5]. However, protein was readily detected in the whole-cell lysate and significantly depleted in the supernatant after centrifugation, indicating sufficient expression but poor solubility, potentially due to protein misfolding or aggregation. Bacterial co-expression of these constructs with NUDCD3 or TaKaRa chaperones did not yield properly folded protein [Supplementary Fig. 6.6].

(a) **HiBit-KLHL12 Kelch** [276-568 a.a., pNIC-Bsa4]



Ni-NTA purification, Coomassie SDS-PAGE

(b) **LgBit-KLHL12 Kelch** [276-568 a.a., pNIC-Bsa4]



Ni-NTA purification, Coomassie SDS-PAGE

Figure 6.5: **Large scale bacterial expression of HiBIT and LgBIT KLHL12 monomers** Even though protein was expressed (expected band present in whole cell lysate), the protein was not soluble (absence of expected band in soluble fractions). Ni-NTA purification analysed on SDS-PAGE gels. WCL = whole cell lysate; SN = supernatant; FT = flow through; BB = binding buffer, WB_{15} and WB_{30} = wash buffer containing 15 mM and 30 mM imidazole respectively; EB1, EB2 and EB3 = elution buffers containing 100, 150 and 200 mM imidazole.

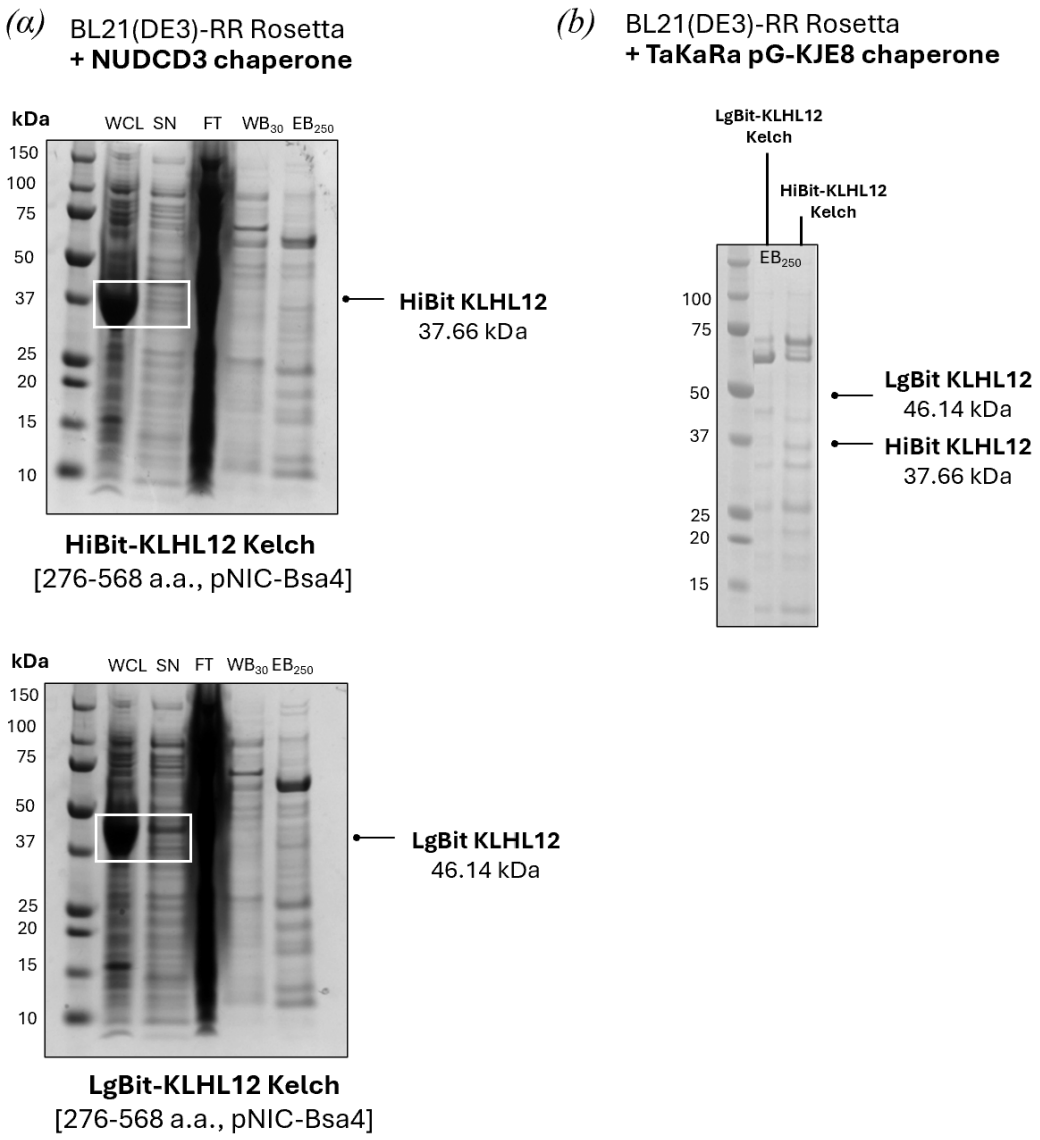


Figure 6.6: **Bacterial co-expression of HiBIT and LgBIT-KLHL12 monomer fusions with folding chaperones.** Small scale test expression of HiBIT and LgBit fusion constructs separately expressed in BL21(DE3)-RR-Rosetta *E.coli* cultures that express (a) NUDCD3 chaperone, and (b) TaKaRa chaperone panel (purification by Beatriz Montes Tan). The 50 mL cultures were harvested with centrifugation, and the clarified supernatant expressing His-tagged HiBit or LgBit KLHL12 was applied to a nickel column. The bound proteins were eluted with increasing imidazole and all samples were analysed in SDS-PAGE gel and visualised with Coomassie stain. Even though protein was expressed (expected band present in whole cell lysate), the protein was not soluble (absence of expected band in soluble fractions). WCL = whole cell lysate; SN = supernatant; FTI = flow through; BB = binding buffer, WB_{30} = wash buffer containing 30 mM imidazole; EB_{250} = elution buffer containing 250 mM imidazole.

To explore whether co-expression of the HiBit and LgBit monomers could enhance folding, the same ΔBTB-BACK HiBit and LgBit KLHL12-fusion constructs were cloned into pCDF-LIC bacterial vectors, as summarised in Fig. 6.7 (a). The pCDF vector was specifically chosen for its compatibility with pNIC due to its distinct origin of replication and antibiotic resistance while sharing a compatible promoter system, thereby facilitating co-expression. However, bacterial co-expression did not result in properly folded protein either [Fig. 6.7 (b)]. Finally, similar co-expression attempts were made using pBacMam2-DiEx-LIC vectors in insect Sf9 cells, by Dr. Peter Loppnau through our established collaboration with SGC at the University of Toronto. This approach successfully rescued expression, yielding a stable, engineered dimeric HiBit-LgBit KLHL12 Kelch domain species, as confirmed by Ni-NTA purification elutions [Fig. 6.8].

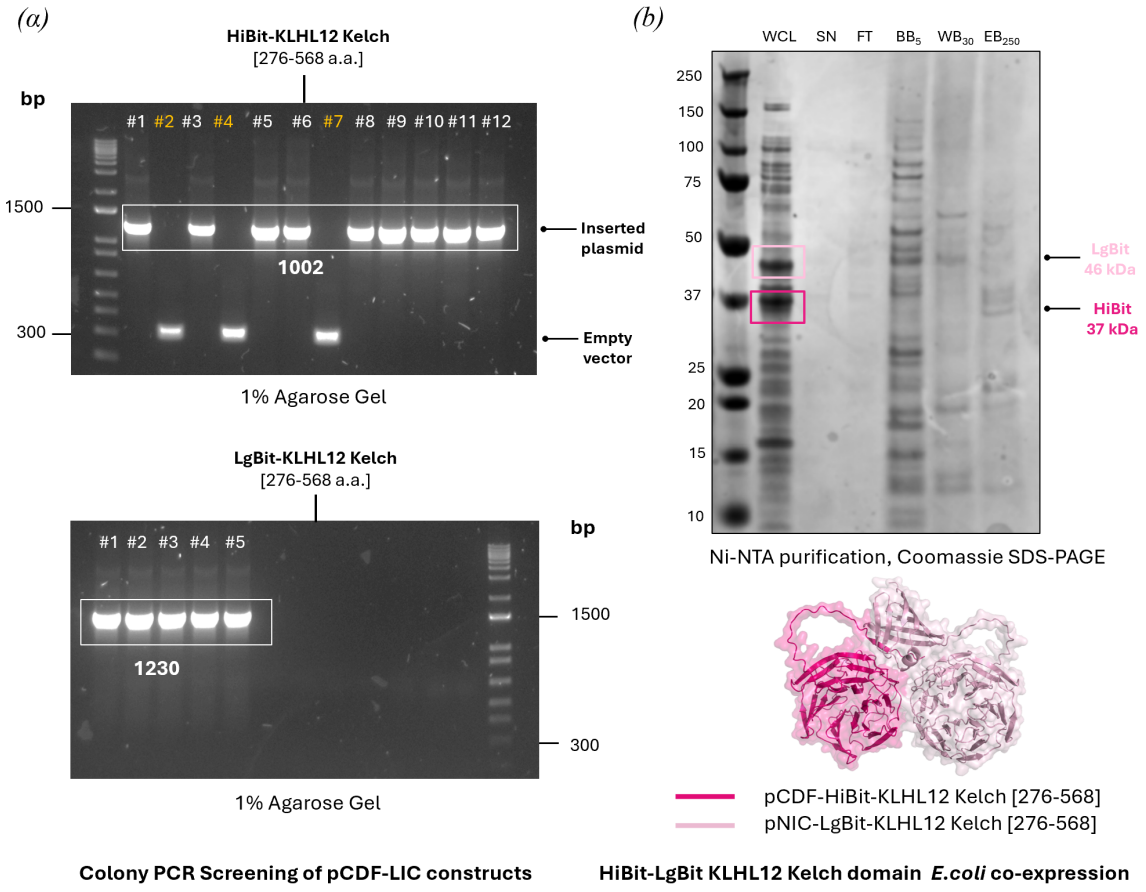


Figure 6.7: **Bacterial co-expression of HiBIT-LgBIT KLHL12 dimer.** (a) Colony PCR screening of recombinant *pCDF-LIC* constructs. *Top*: Screening of HiBiT-KLHL12 Kelch clones. Positive colonies showed a PCR product of 1002 bp, confirming correct insertion of the HiBiT-tagged KLHL12 Kelch domain. Negative colonies that had not uptaken the insert are coloured in orange and showed a lower band corresponding to the empty reannealed vector. *Bottom*: Screening of LgBiT-KLHL12 Kelch clones. Positive colonies exhibit a 1230 bp PCR band, confirming successful ligation. (b) Small scale test Ni-NTA test purification of *E. coli* co-expression of pCDF-LIC-HiBIT-KLHL12 and pNIC-Bsa4-LgBit-KLHL12 Kelch domain constructs. *Top*: Coomassie-stained SDS-PAGE gel of different purification fractions. HiBiT-KLHL12 (expected size: 37 kDa) and LgBiT-KLHL12 (expected size: 46 kDa) are highlighted in pink boxes. Even though proteins were expressed (expected bands present in whole cell lysate), the protein is not soluble (absence of expected band in soluble fractions). *Bottom*: Model of ΔBTB-BACK HiBIT-LgBIT KLHL12 reconstituted dimer by AFv3. Lane labels: WCL = whole cell lysate; SN = supernatant; FTI = flow through; BB = binding buffer, WB_{30} = wash buffer containing 30 mM imidazole respectively; EB_{250} = elution buffer containing 250 mM imidazole.

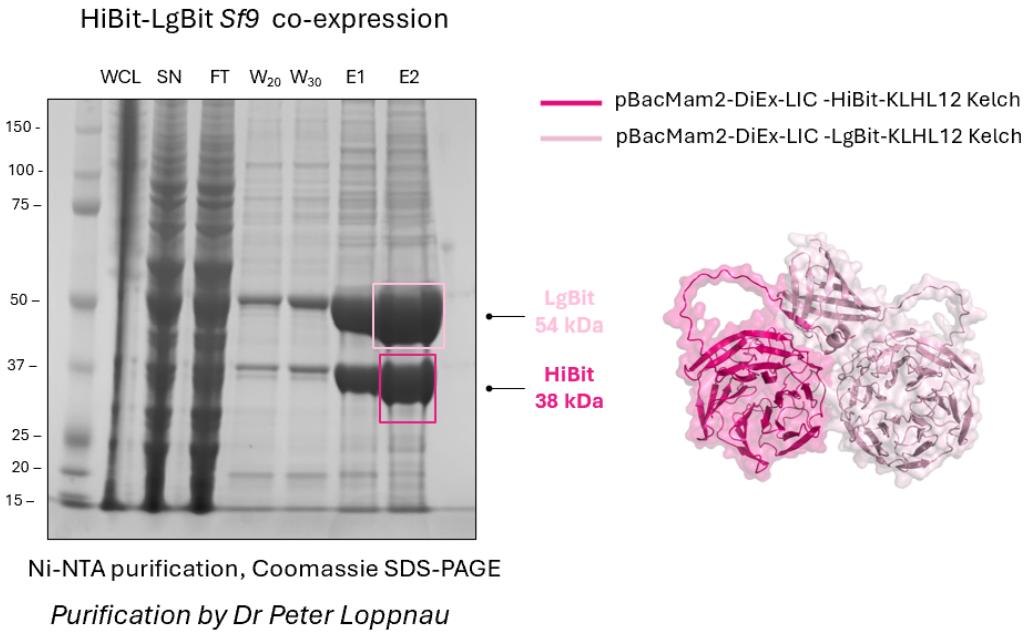


Figure 6.8: **High-yielding HiBit-LgBit KLHL12 Kelch domain dimer from Sf9 expression.** Δ-BTB-BACK HiBit-LgBit KLHL12 Kelch domain constructs co-expressed in Sf9 cells show high expression of stable KLHL12 Kelch domain complex. Ni-NTA purification elutions assessed in a Coomassie blue-stained SDS-PAGE gel. Construct cloning and purification by Dr Peter Loppnau, University of Toronto. Model of ΔBTB-BACK HiBIT-LgBIT KLHL12 reconstituted dimer by AFv3.

6.2.2 Indications of BAY-805 mediated USP21-HiBIT-LgBIT KLHL12 Kelch domain assembly *in vitro*

HiBIT-LgBIT KLHL12 Kelch domain-mediated interaction with USP21 in cellular nanoBRET assays has been demonstrated using BAY-805, however, no *in vitro* experiments had been performed previously due to the challenge of obtaining dimeric KLHL12 constructs required for the interaction. Having overcome this challenge, the *in vitro* interaction between purified recombinant USP21 catalytic domain [Uniprot Q9UK80, 210-561 a.a.] and the ΔBTB-KLHL12-HiBit-LgBit complex [Uniprot Q53G59, 268-567 a.a.] was assessed using MicroScale Thermophoresis (MST). MicroScale Thermophoresis measures the movement of molecules in a temperature gradient, which changes upon binding due to shifts in size, charge, or hydration, and relies on proteins being labeled with a fluorescent dye.

Preliminary trials after both molecules were covalently labelled by their lysines (RED-NHS dye) were performed to assess whether labeling abrogates known binding interactions. Preliminary trials with NHS-labeled HiBIT-LgBIT KLHL12 showed that the protein was stable under experimental conditions in standard capillaries [Fig. 6.9 (a)] and optimal protein concentration for adequate fluorescent signal was ascertained at 50 nM [Fig. 6.9 (b)]. The binding event of known interactor DVL1, tested across a range of concentrations up to 150 μM peptide, was not observed as expected [Fig. 6.9 (d)]. This suggests that lysine labeling occludes the KLHL12 binding site, making it unsuitable for interaction studies. Subsequent trials with NHS-labeled USP21 confirmed protein stability in premium capillaries [Fig. 6.10 (a)]. USP21, at 100 nM (as determined by preliminary concentration testing) [Fig. 6.10 (b)], exhibited an affinity for BAY-805 with a dissociation constant (K_d) of 29.3 nM, consistent with the low-nanomolar values reported in the original Bayer study [261]. That suggested that labelling USP21 cysteines does not affect BAY-805 binding, positing this molecule suitable for interaction studies.

BAY-805 at 1 μM (2% DMSO final concentration) was pre-incubated with varying concentrations

of C-terminally tagged HiBIT-LgBIT KLHL12 protein (16-point curve 0-106 μM) for 15 minutes at 0°C. Labeled USP21 (100 nM) was then added to the mixture, and the samples were incubated further for 15 minutes at room temperature before being transferred into premium MST capillaries. Measurements were conducted at 25 °C with 20% excitation power and 40% (medium) MST power. Data showed that, in the absence of molecular glue, the C-terminally tagged HiBIT-LgBIT KLHL12 protein and the catalytic domain of USP21 have a moderate affinity, with a K_d of 2.42 μM, which is potentiated by the addition of BAY-805 to a K_d of 36 nM [Fig. 6.11]. Additionally, in the binding experiment conducted with DMSO solvent only (without BAY-805), an inversion of the data was observed, where the expected binding curve appeared flipped [Fig. 6.11 (b)].

Such an effect on thermophoretic behavior can arise from changes in hydration shell, charge, or size, which may indicate true binding or result from aggregation or multivalent binding, leading to secondary effects on diffusion. Finally, it may also indicate fluorescent inhomogeneity caused by aggregation or a true binding event if binding occurs near the fluorescent label. These indications of affinity should be interpreted with caution as the protein quantity limitations did not allow for technical and biological repeats to ascertain robustness of the observed interaction. Further repeats and orthogonal assays to characterise the interaction should be performed in the future. However, these preliminary indications of binding were sufficient for attempting structural characterization of the complex using these protein reagents.

6.2.3 Structural characterisation attempts for the KLHL12-USP21 complex mediated by BAY-805

Purified reconstituted dimer of HiBit-LgBit KLHL12 Kelch domain fusions (N- or C-terminally tagged) was incubated with an excess of purified catalytic domain of USP21 [210-561 a.a.] at 4 °C for 2 hours at molar ratios of 2-fold and 5-fold excess of USP21 over KLHL12, in the presence of 100 μM BAY-805 (1% DMSO). The mixture was diluted to a total protein concentration of 1.6-2.0 mg/mL, applied to Quantifoil Cu R1.2/1.3 300-mesh grids, and vitrified. Fig. 6.12 (b) summarizes

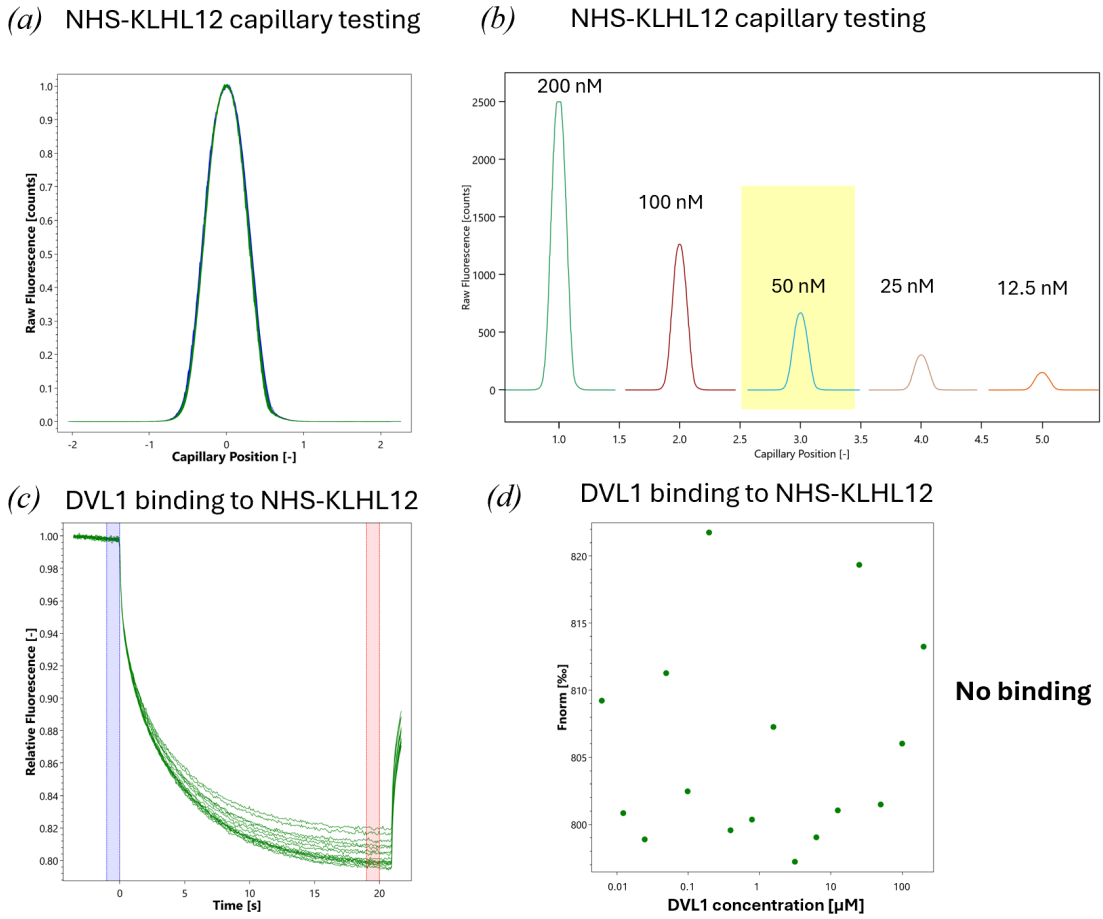


Figure 6.9: **Preliminary MST tests of NHS-labelled HiBIT-LgBIT KLHL12 Kelch domain.** Preliminary testing of NHS-labeled HiBIT-LgBIT KLHL12 Kelch domain for stability and binding assays prior to interaction experiments. (a) Fluorescence profile of NHS-labeled HiBIT-LgBIT KLHL12 Kelch domain in MST capillaries, confirming stability under experimental conditions. (b) Serial dilution of NHS-labeled KLHL12, indicating a reliable fluorescence signal at 50 nM (highlighted). (c) MST traces of NHS-labeled HiBIT-LgBIT KLHL12 Kelch domain with DVL1 20mer peptide across a concentration range up to 150 µM (positive control), showing expected fluorescence decay. (d) Binding analysis of NHS-labeled KLHL12 with DVL1 showing no detectable binding, suggesting lysine labeling may occlude the binding site.

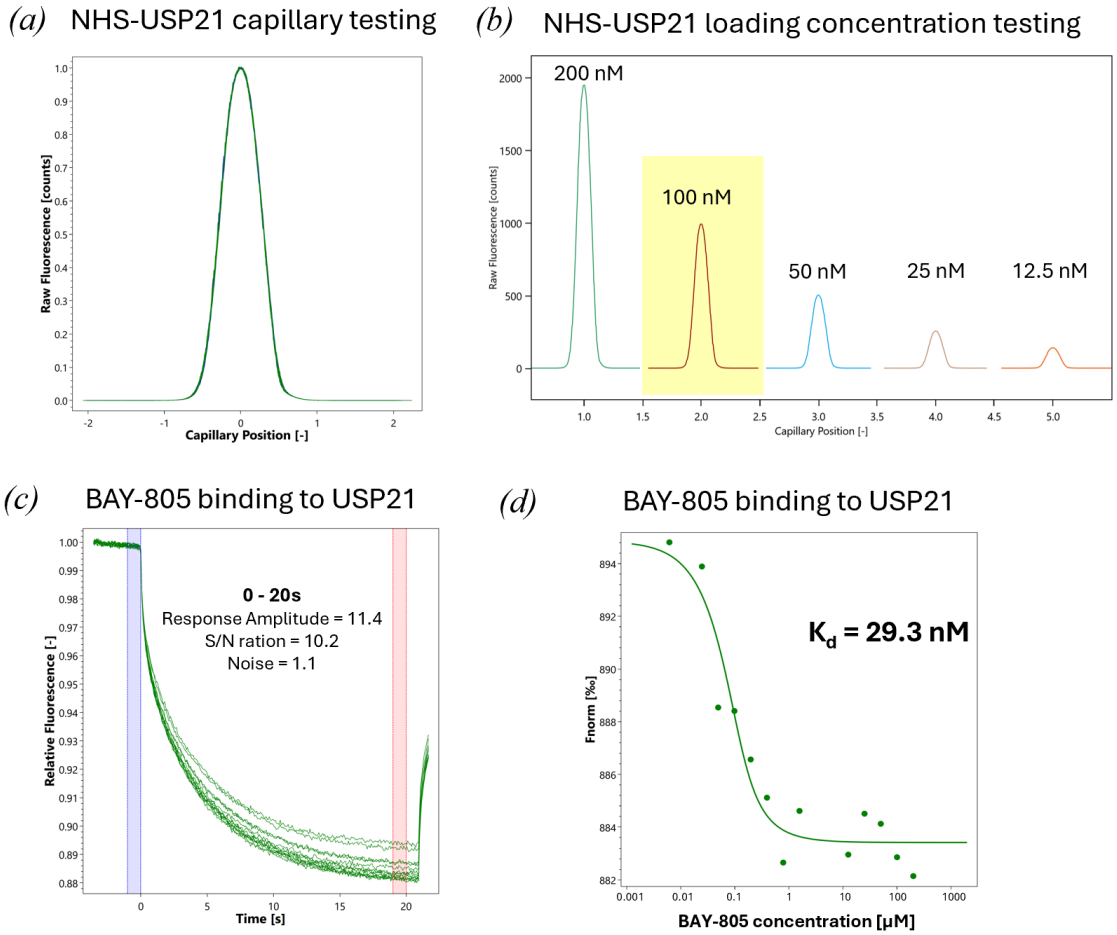


Figure 6.10: **Preliminary MST tests of NHS-labelled USP21 catalytic domain.** Preliminary testing of NHS-labeled USP21 for stability and binding to BAY-805 prior to interaction experiments. (a) Fluorescence profile of NHS-labeled USP21 in MST premium capillaries, confirming stability under experimental conditions. (b) Serial dilution of NHS-labeled USP21, indicating a reliable fluorescence signal at 100 nM (highlighted). (c) MST traces of NHS-labeled USP21 with BAY-805, showing a strong binding response with a high signal-to-noise ratio. (d) Binding analysis of BAY-805 to NHS-labeled USP21, yielding a dissociation constant (K_d) of 29.3 nM, consistent with previously reported values. Data were analyzed using the Monolith Evaluation software using the K_d model (1:1 Langmuir binding model).

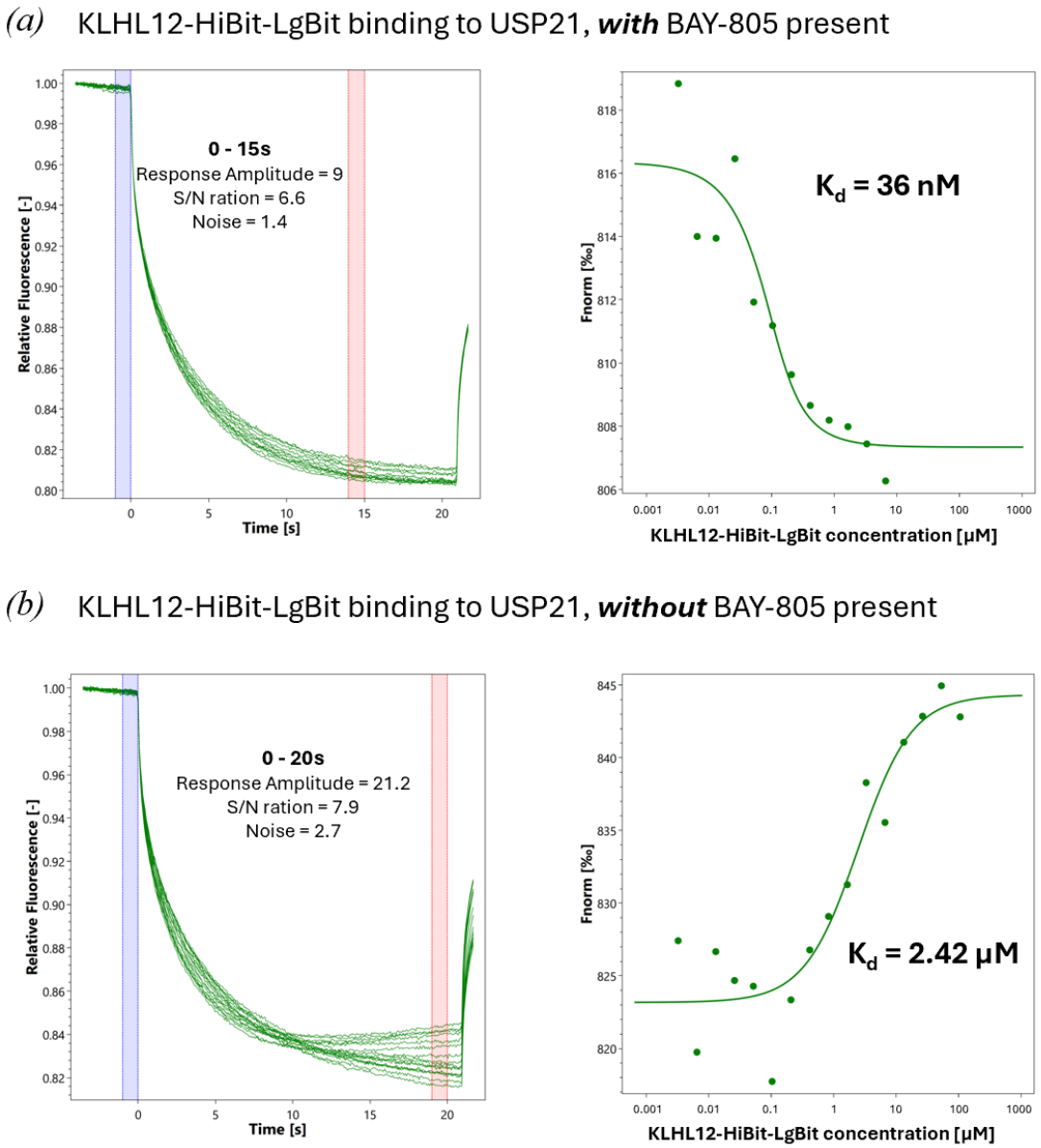


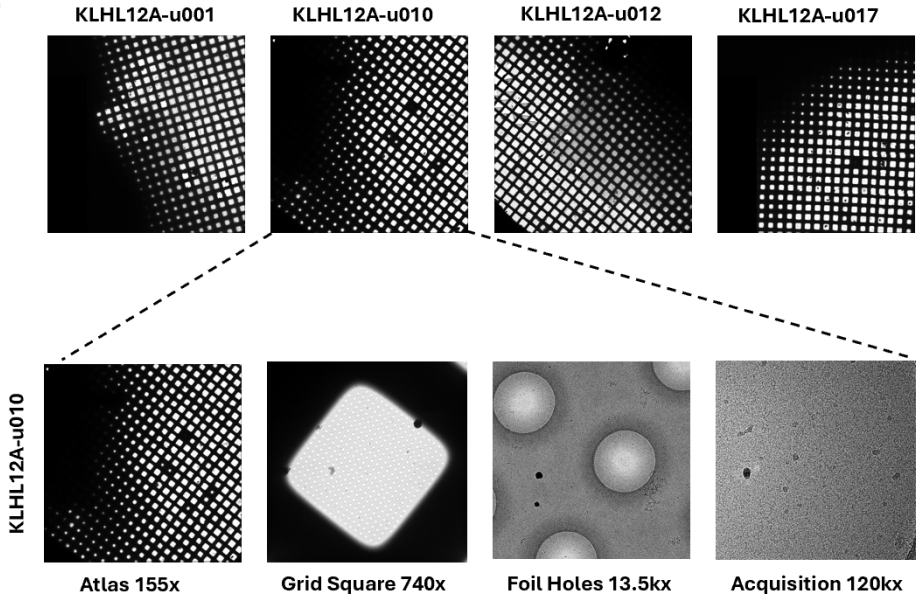
Figure 6.11: **BAY-805 may potentiate KLHL12-HiBiT-LgBiT binding to USP21 .** MST analysis of HiBIT-LgBIT KLHL12 Kelch domain (16-point curve 0-106 µM) binding to USP21 (100 nM) in the presence and absence of BAY-805 (1 µM). (a) MST traces (left) and binding curve (right) showing HiBIT-LgBIT KLHL12 Kelch domain interaction with USP21 in the presence of BAY-805. The interaction yielded a dissociation constant (K_d) of 36 nM. (b) MST traces (left) and binding curve (right) for KLHL12-HiBiT-LgBiT binding to USP21 in the presence of 2% DMSO, showing a weaker interaction with a K_d of 2.42 µM. Data were analyzed using the Monolith Evaluation software using the K_d model (1:1 Langmuir binding model).

the molecular ratios and blotting times used for grid preparation.

Grids were screened using an OPIC Glacios transmission electron microscope (TEM) operating at 200 kV, first at 155x magnification (Atlas images) to assess overall quality [Fig. 6.12 (a). Grids of sufficient quality underwent limited data collection at 120,000x magnification. However, no particles were observed, suggesting that either the complex had not formed or had dissociated upon vitrification; particles of the complex should be visible at 135 or 176 kDa depending on complex stoichiometry, but individual components at 93 kDa for dimeric KLHL12 and 42 kDa for USP21 would not be visible under these experimental conditions on a Glacios instrument operating at 200 kV.

Future studies should focus on molecular characterisation of the complex prior to vitrification to determine optimal conditions. Further optimisation of grid type, ice thickness, protein concentration, and vitrification parameters could significantly influence particle distribution. However, due to time and practical constraints, these aspects could be the focus of a future study.

(a)



(b)

Grid Identifier	Conc. (mg/ml)	Protein(s)	Blot time (s)
KLHL12A-u001	1.6	HiBit-LgBit KLHL12 (N-term)	4
KLHL12A-u002	1.6	HiBit-LgBit KLHL12 (N-term)	5
KLHL12A-u003	2.0	HiBit-LgBit KLHL12 (C-term)	5
KLHL12A-u005	1.9	HiBit-LgBit KLHL12 (N-term):USP21 1:2	4
KLHL12A-u008	1.9	HiBit-LgBit KLHL12 (N-term):USP21 1:2	7
KLHL12A-u010	1.7	HiBit-LgBit KLHL12 (N-term):USP21 1:5	5
KLHL12A-u011	1.7	HiBit-LgBit KLHL12 (N-term):USP21 1:5	6
KLHL12A-u012	1.7	HiBit-LgBit KLHL12 (N-term):USP21 1:5	7
KLHL12A-u017	1.7	HiBit-LgBit KLHL12 (C-term):USP21 1:5	4
KLHL12A-u019	1.7	HiBit-LgBit KLHL12 (C-term):USP21 1:5	6
KLHL12A-u020	1.7	HiBit-LgBit KLHL12 (C-term):USP21 1:5	7

Figure 6.12: **cryo-EM grid preparation overview for KLHL12-USP21 complexes.** (a) Representative grid atlases of screened HiBIT-LgBIT KLHL12 dimers mixed with USP21 and BAY-508. No clear particles were observed in acquisition magnification. Images acquired in a Glacios 200 kV instrument. (b) Summary of quantifoil Cu 1.2/1.3 grids with a 300 mesh, prepared for screening with varying molecular ratio compositions and blotting times.

6.3 Discussion

This chapter investigated small-molecule modulation of the full $CUL3^{KLHL12}$ E3 ligase complex, extending beyond the Kelch domain focus of **Chapters 4** and **5**. Conducted in collaboration with SGC Toronto and Dr. Dalia Barsyte-Lovejoy's group, this work was pursued in parallel with the chemical probe development in **Chapter 5**, as another avenue of exploring new chemical matter targeting KLHL12. Instead of focusing on the development of a conventional bifunctional degrader as previously, the focus was on the potential to recruit KLHL12 via molecular glue mechanisms, stabilising neo-interactions within the native ligase complex rather than relying on substrate-binding domain targeting. Unlike chimeric molecules, which typically function by tethering a substrate to the E3 ligase to induce degradation, molecular glues act by stabilising protein-protein interactions within the ligase assembly itself. This distinction is particularly significant, as modulating the assembly rather than just the substrate-binding interface could provide a broader and more tunable approach to regulating ubiquitination pathways.

In this chapter, I investigated the structural basis of a putative molecular glue in the KLHL12-USP21 assembly, which has not been reported before. I aimed to validate previously observed cellular interactions of a homodimeric KLHL12 E3 ligase assembly with deubiquitinase USP21 induced by inhibitor BAY-805 and reconstitute them *in vitro* [261]. In collaboration with SGC Toronto, a high-yielding HiBit-LgBit ΔBTB-Kelch construct *in lieu* of low expressing full-length dimeric KLHL12, that still retained the ability to bind with USP21 in the presence of BAY-805 in cells was developed.

The putative KLHL12-USP21 interaction proved challenging to probe, primarily due to difficulties in obtaining soluble and high-yielding protein constructs of full-length KLHL12 and USP21. As a result, several biophysical techniques - such as isothermal titration calorimetry (ITC) - were not feasible. Full-length surface plasmon resonance (SPR) studies also suffered from poor protein stability (data not shown). To circumvent these limitations, microscale thermophoresis (MST), a method with lower protein requirements, was employed. Using fluorescently labeled USP21 catalytic domain,

MST successfully recapitulated the previously reported interaction with BAY-805 [261], exhibiting a dissociation constant (K_d) of 30 nM. To investigate KLHL12-USP21 binding, an engineered dimeric KLHL12 Kelch domain (HiBit-LgBit ΔBTB-Kelch) was used in lieu of dimeric full-length KLHL12 *in vitro*. The results suggested a moderate constitutive affinity between USP21 and KLHL12 (K_d = 2.42 μM), which was significantly potentiated upon addition of BAY-805, reaching K_d = 36 nM [Fig. 6.11].

Interestingly, initial BioID experiments reported KLHL12 enrichment only in the presence of BAY-805, implying an inducible interaction. The observed moderate affinity by MST ($\sim$ 2 μM) in the absence of BAY-805 suggests a relatively weak yet possibly transient or dynamic interaction in cells, influenced by co-localization factors. However, these findings should be interpreted with caution, as protein yield limitations precluded technical and biological replicates for MST experiments. Further optimisation of full-length constructs will be necessary for additional *in vitro* validation of this interaction with orthogonal assays, such as SPR or analytical size exclusion chromatography.

Structural characterisation of the KLHL12-USP21 complex in the presence of BAY-805 was attempted using cryo-electron microscopy (cryo-EM). Despite obtaining high-quality grids suitable for data collection, no particles were observed at 120,000x magnification. This finding suggested that either the complex had not formed or had dissociated upon vitrification. Alternatively, the complex may have exhibited a higher affinity for the carbon-coated film rather than remaining suspended in the aqueous ice layer. Future studies could explore the use of chemical crosslinking reagents to stabilise these potentially weak-affinity complexes, as successfully demonstrated in the resolution of another molecular glue-mediated E3 ligase complex, dimeric KBTBD4-HDAC2 [270]. Additionally, alternative grid types, such as gold or graphene, may improve particle retention and distribution. Further optimisation of grid type, ice thickness, protein concentration, and vitrification conditions could significantly impact particle visualization. However, due to time and practical constraints, these optimisations will be addressed in future investigations.

Sequence analysis of USP21 revealed two regions of USP21 containing similar sequences to the KLHL12 recognition PGXPP motif [72]; (i) 91-PGSPP-95 in the N-terminal disordered region, and (ii) 334-PSPPR-338 in the catalytic region. Notably, the latter sequence, near the ubiquitin binding pocket, also corresponds to the predicted BAY-805 binding region [unpublished LiP-MS data, Viviane Reber, Matthias Gstaiger, Paolo Picotti Group]. Additionally, six sequences in the USP21 N-terminal domain follow a modified PXPP motif, raising the question of whether BAY-805 potentiates a weak constitutive interaction rather than mediating a novel one [Fig. 6.13].

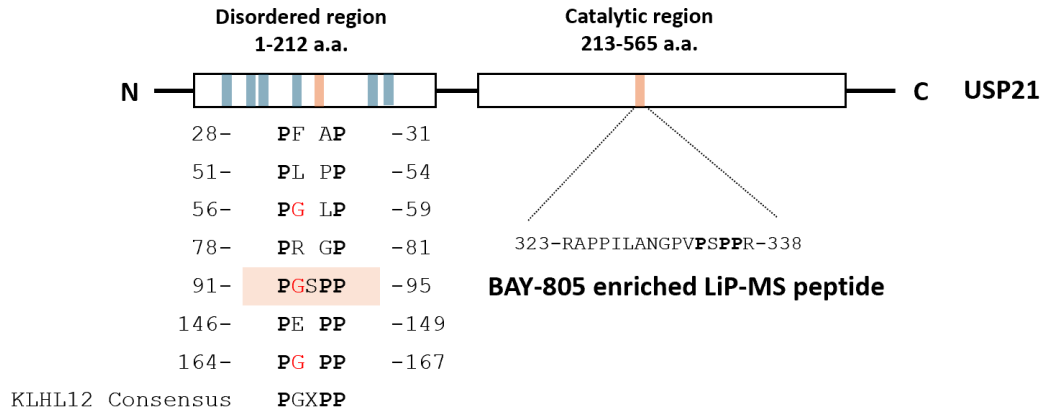


Figure 6.13: **USP21 has multiple KLHL12 PGXPP-like recognition sequences.** USP21 has one PGXPP KLHL12 recognition motifs (boxed and shaded orange), and six variants mostly located in the N-terminal disordered region and catalytic region.

Unfortunately, AFv3 is not enabled for including drug molecules such as BAY-805. Nonetheless, AlphaFold v3 [191] modelling of the CUL3^{KLHL12}-RBX1-USP21 complex predicted the expected CUL3-RBX1-KLHL12 architecture, whereby CUL3 (yellow) serves as the central structural scaffold, adopting an elongated arc-like conformation that bridges substrate recognition and ubiquitin transfer [Fig. 6.14 *(a)*]. Its C-terminal domain (CTD) recruits the RBX1 RING Finger Domain (red), which facilitates interaction with the ubiquitin-conjugating enzyme E2 (not modelled), while the N-terminal region (NTD) directly interacts with the BTB-domains (cyan and blue) of the KLHL12 substrate adaptor. The model also predicts two copies of USP21 (orange and green) fitting between the Kelch domains (magenta) [Fig. 6.14 *(a)*]. Despite the low confidence in the predicted USP21-KLHL12 interaction loop (pLDDT score < 50, Fig. 6.14 *(b)*), it is noteworthy that the putative interaction loop

of USP21 contains two PGXPP-like proline-rich variants.

The model docks a disordered N-terminal region of USP21 (76-PR**PRGP** LRADHGV **PLPGS** -93) onto the Kelch domain of KLHL12 [Fig. 6.15 *(a)*]. The USP21 N-terminal region is modelled to interact with KLHL12 somewhat similarly to DVL1 adopting a comparable U-shaped conformation. In this model USP21 residues R79 and D85 form hydrogen bonds with KLHL12 residues Q293 and Y512/R320 respectively, while USP21 P91 forms a salt bridge with KLHL12 R342 [Fig. 6.15 *(a)*]. Notably, the USP21 cryo-EM construct [210-561 a.a.] lacked the N-terminal disordered region potentially explaining the lack of visible complex particles in cryo-EM grids as that could potentially prevent complex formation.

Interestingly the potential binding site of BAY-805 does not lie in the predicted KLHL12-USP21 interface but on a flexible loop near the RBX1 CUL3 interface instead [shaded yellow and pink respectively, Fig. 6.14 *(b)*]. While great caution is needed in interpreting this speculative model, it does raise the possibility that the BAY-805 probe does not act as a bivalent molecular glue contacting both USP21 and KLHL12 directly, but could instead act allosterically to potentially introduce additional contacts between USP21 and KLHL12 as well as with RBX1-CUL3. The lateral could stabilise peripheral minor contacts of the assembly, potentiating the interaction by an avidity effect. Without an experimental structure, the exact nature of how this compound stabilises the complex remains unclear and possibilities like enhancing USP21 dimerisation or inducing conformational change should be explored with full-length USP21 constructs that include the N-terminal PGXPP and PGXPP-like rich domain.

In summary, many questions remain regarding USP21-KLHL12 assembly. Although a symmetrical interaction model is presented, the molecular stoichiometry could differ, for example with a KLHL12 dimer interacting asymmetrically with a USP21 dimer, or with one USP21 monomer, as seen with the UM171 molecular glue-mediated interaction of dimeric KBTBD4 and HDAC2 [270, 271]. NanoBIT interaction assays could be employed to test mutated USP21 constructs at the predicted PRGP and

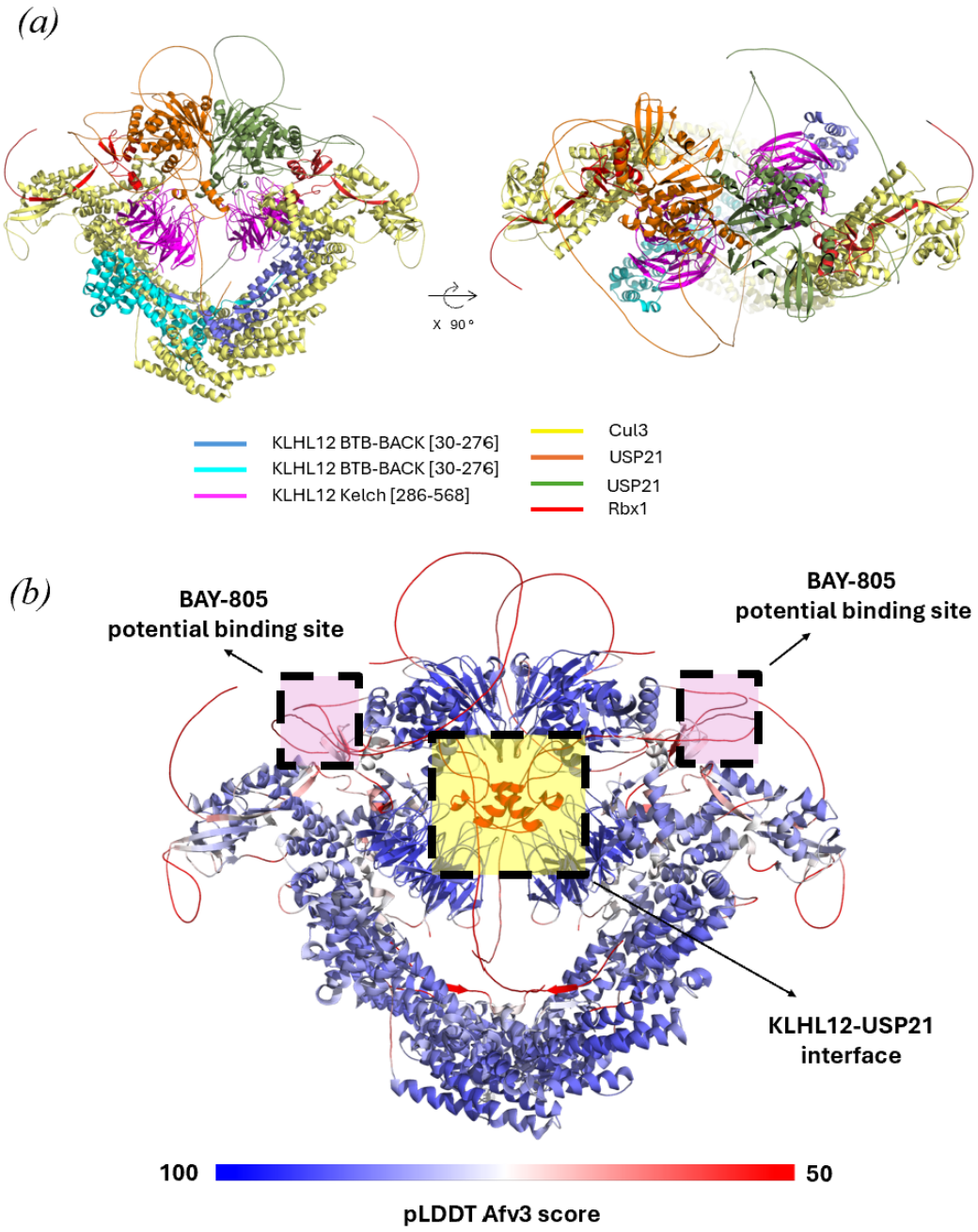


Figure 6.14: **Model of CUL3-RBX1-KLHL12 in complex with USP21.** (a) Model of CUL3-RBX1-KLHL12-USP21 assembly by AFv3. KLHL12 homoderises through its BTB domain and assembles with CUL3-RBX1 as expected. (b) Complex coloured by the predicted local distance difference test (pLDDT) score which is a per-residue measure of local confidence. Low scores denote low confidence (red), while high scores denote high confidence (blue). KLHL12-USP21 interface coloured in yellow, BAY-805 potential binding site coloured in pink.

PLPG motifs in the N-terminal disordered region and assess their impact on KLHL12 binding. Further cryo-EM experiments with full-length USP21, KLHL12, CUL3, and Rbx1 are necessary to investigate compound-mediated complex formation. However, purification of full-length KLHL12 remains an obstacle for attempting this.

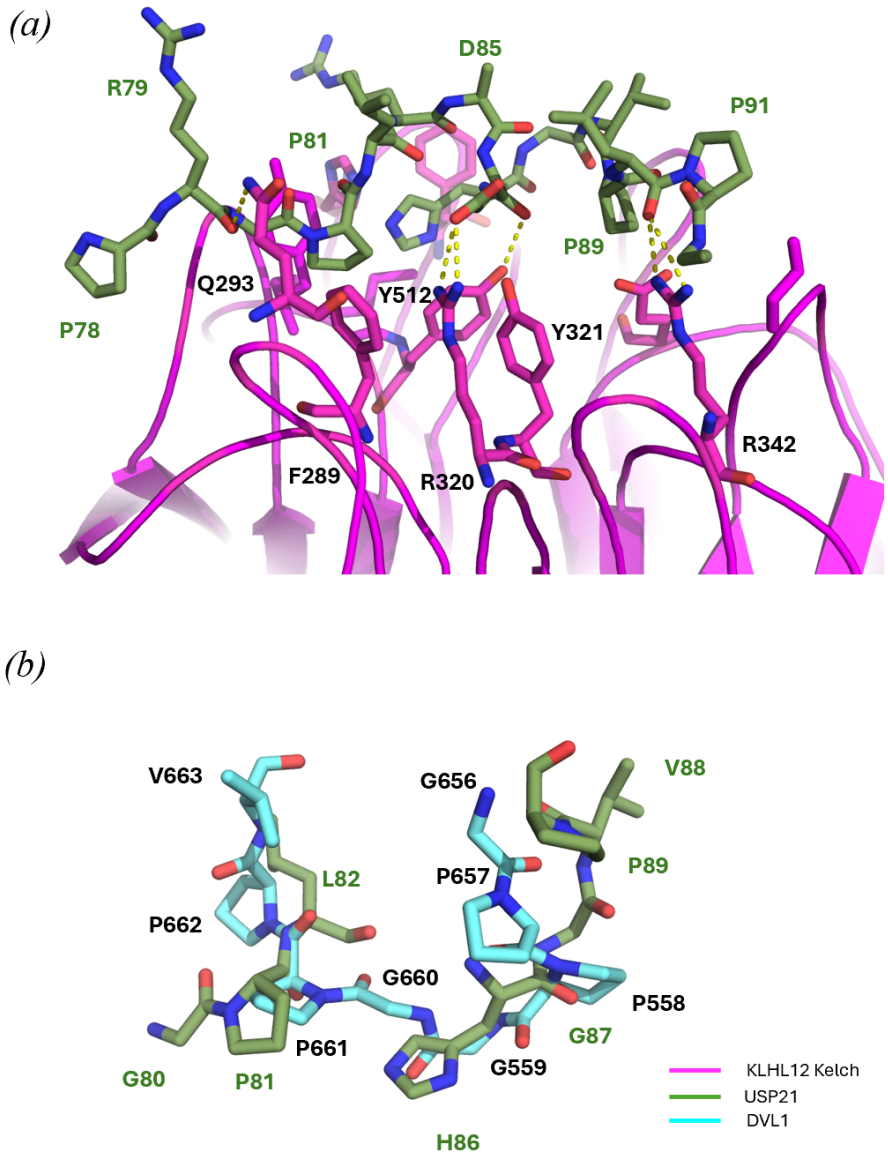


Figure 6.15: **Predicted interactions between USP21 and KLHL12.** (a) Predicted interactions between USP21 N-terminal disordered region 78-**PRGP** LRADHGV **PLPGS**-93 and KLHL12 Kelch domain, model by AlphaFold v3 [191]. Hydrogen bonds and salt bridges shown by yellow dashed lines. KLHL12 coloured in magenta, selected residues of USP21 in green. (b) Superposition of the putative USP21 binding model with the structurally resolved DVL1 peptide (PDB: 6TTK). DVL1 coloured in cyan.

Interestingly, the AlphaFold-predicted model appears to align with preliminary MST data, which suggest a constitutive USP21-KLHL12 interaction that is significantly enhanced by BAY-805 allosterically. Constitutive interactions between E3 ligases and deubiquitinases (DUBs) have been reported in the literature. For example, USP7, a well-characterized DUB, binds to the E3 ligase MDM2, stabilising it and preventing its degradation, facilitating the essential MDM2 role in regulating p53 levels in cells [272, 273]. The opposite regulatory mechanism can also hold true, as is the case ligase Ro52 that targets DUB USP4 for degradation [274]. Whether KLHL12 shares a similar co-regulatory mechanism with USP21 remains an open question for future investigation. Although no direct evidence has been reported, the constitutive association of some E3 ligases with deubiquitinases (DUBs) raises the possibility that co-recruitment of DUBs by PROTACs could attenuate target degradation. By opposing ubiquitination, DUBs may reduce the efficiency of PROTAC-mediated proteolysis.

In this thesis, two classes of small-molecule proximity-inducing agents were explored, bifunctional PROTACs and molecular glues. These modalities are attractive due to their synthetic tractability, tunability of physicochemical properties (e.g., pKa, solubility, metabolic stability), and demonstrated efficacy in both *in vitro* and *in vivo* systems [159, 169]. However, emerging proximity strategies include bioconjugate-based PROTACs, in which biomolecules such as antibodies or their oligonucleotide analogues, aptamers, are conjugated to small-molecule ligase recruiters [275–279]. In most cases, the antibody or aptamer confers tissue or cell-type specificity (e.g., HER2-positive cancer cells), facilitating targeted internalisation and degradation of the PROTAC. However, some aptamer-based PROTACs replace the small-molecule target-binding moiety altogether, using the aptamer itself to engage the target protein with high affinity [277, 278].

Another innovative platform leverages smaller alternatives to antibodies with lower immunogenicity such as nanobodies to target the neo-substrate proteins, wherein synthetic nanobodies are engineered via yeast or phase display methods, and are fused to minimal E3 ligase adaptors and cell-penetrating motifs [280]. Recent work has also demonstrated that the substrate-targeting ligand in PROTACs can

be entirely replaced with a polypeptide binder. These binders range from small peptides or synthetic macrocycles [184, 281, 282], to engineered mini-proteins or entire protein folded domains directly fused to an E3 ubiquitin ligase, effectively decoupling PROTAC development from the need for small-molecule ligandability [283].

With the advent of technologies such as phage display and CRISPR/Cas9-based screening, these strategies are beginning to move into the realm of synthetic biology, broadening the toolkit for selective and context-specific protein degradation beyond the constraints of conventional medicinal chemistry. Such approaches could be further explored in the context of KLHL12, for example by investigating tissue-specific modulation strategies using antibody/nanobody conjugation to address the ubiquitous expression of KLHL12. Alternatively, by engineering a synthetic peptide or aptamer ligand, based on the known PGXPP degron motif as an alternative recruitment module for KLHL12-driven degradation [**Chapter 1.1**]. These non-small-molecule approaches may be combined with the small-molecule inhibitor development explored here to give more modular and flexible proximity-inducing modalities in future studies.

CHAPTER 7

Exploring structural assemblies of the NRBP1 E3 ligase

7.1 Introduction

Although this is the final chapter of the thesis, it describes the first project I undertook: investigating the assembly and mechanism of a heterotypic E3 ligase complex and identifying a potential binding site on NRBP1, a pseudokinase proposed to function as a substrate adaptor [Fig. 1.13]. Given that both NRBP1- and KLHL12-based E3 ligases rely on dimeric substrate adaptors, insights from this study contribute to a broader mechanistic understanding of adaptor-mediated complex formation.

While Cullin-RING ligases (CRLs) typically assemble around a single Cullin scaffold, an unconventional heterotypic assembly involving both CUL2 and CUL4, bridged by an NRBP1 homodimer, has been proposed as a novel E3 ligase [Fig. 1.10 (a)]. This atypical complex, which challenges the established paradigm of Cullin recruitment, was first suggested by Yasukawa *et al*., who proposed NRBP1, a structurally uncharacterised pseudokinase, as the substrate adaptor for CUL2/CUL4^{NRBP1} E3 ligase targeting amyloidogenic substrates [126]. Such a heterotypic complex would represent a significant departure from the canonical architecture of CRLs, where each adaptor protein is thought to be typically organised around a single Cullin scaffold [Fig. 1.4]. In contrast, this model suggests that the same substrate adaptor would be able to assemble with multiple Cullins (CUL2 and CUL4), implying the potential for combinatorial CRL architectures with expanded substrate specificity than currently described. While the existence of this E3 ligase is supported by compelling co-immunoprecipitation and sequence alignment data, as previously discussed and summarised in Table 1.2 and Fig. 1.10 (b),

its precise assembly remains unclear.

Beyond its novel heterotypic assembly of mechanistic interest, NRBP1 provides an opportunity to explore the non-catalytic functions of pseudokinases, a growing class of regulatory proteins with increasing pharmacological relevance [284]. It is a possible this is an E3 ligase in which the pseudokinase domain of NRBP1 functions as a substrate recognition module, suggesting that kinase domains could be explored as ligandable sites in E3s. By elucidating the structural basis of this multi-subunit ligase, this work aimed to provide insights into its substrate specificity and potential therapeutic applications for amyloidogenic protein degradation by targeting the kinase's substrate-binding domain.

This chapter describes efforts to validate this proposed ligase assembly *in vitro* using purified proteins for the first time, with a particular focus on defining its structural architecture and substrate engagement mechanism. To investigate this assembly, full-length NRBP1 and its interacting partners were cloned using high-throughput ligation-independent cloning, enabling rapid generation of constructs with small truncations. Expression was assessed in both *E. coli* and insect Sf9 cells, and purification protocols were optimized for novel targets. The proposed protein interactions [Table 1.2] were assessed using native mass spectrometry and pull-down assays to confirm direct binding, while biolayer interferometry was used to measure binding affinities. Purified complexes were further used for structural studies, including crystallization trials and cryo-EM grid preparation.

7.2 *In vitro* exploration of putative heterodimeric CUL2/CUL4A^{NRBP1} ligase

7.2.1 ElonginBC co-expression rescues NRBP1 *E.coli* constructs

All previous studies on NRBP1 function were conducted *in cellulo*, and no protein constructs that yielded purified protein for *in vitro* studies had been reported in the literature. In order to structurally enable the proposed mixed assembly, initial cloning of different constructs was performed bearing small truncations around the predicted globular domains.

Modelling of NRBP1 with AlphaFoldv1 predicts a main globular fold that consists of the pseudokinase domain [68-327 a.a.] and two helices where the putative BC-/H- [331-341 a.a.] and CUL2-

boxes [342-357 a.a.] lie, connected with a disordered flexible loop to a second C-terminal globular fold [465-535 a.a.]. The N-terminus is predicted to be disordered [1-50 a.a.] [Fig. 7.1]. Initial attempts in our lab, prior to the start of my DPhil studies, to express the biggest predicted globular domain of NRBP1 pseudokinase domain [Uniprot Q9UHY1, 41-346 a.a.] were only successful in Sf9 cells. Following from that work, I purified the same construct by tandem Ni-NTA followed by size exclusion chromatography. Intact mass-spectrometry analysis of the denatured sample revealed that the main protein species was mono-phosphorylated by the eukaryotic insect cells kinases [Fig. 7.2 (a)]. After overnight incubation with λ-phosphatase, the protein was further mass analysed to reveal successful de-phosphorylation and significant degradation suggesting protein instability [Fig. 7.2 (b)].

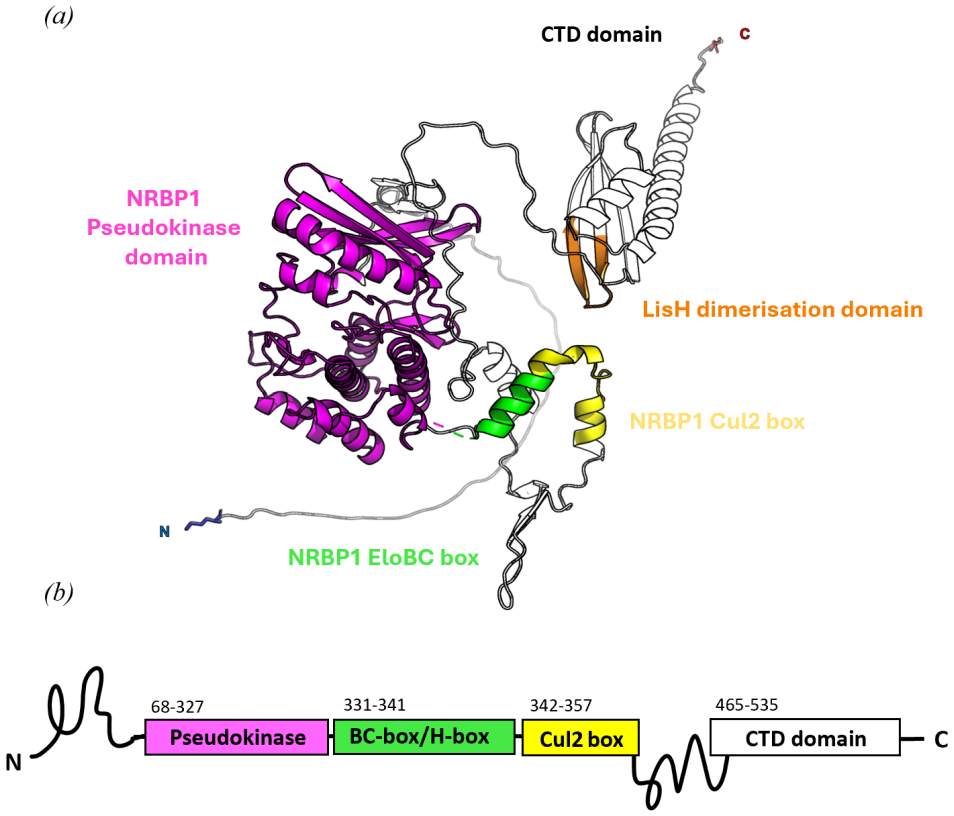
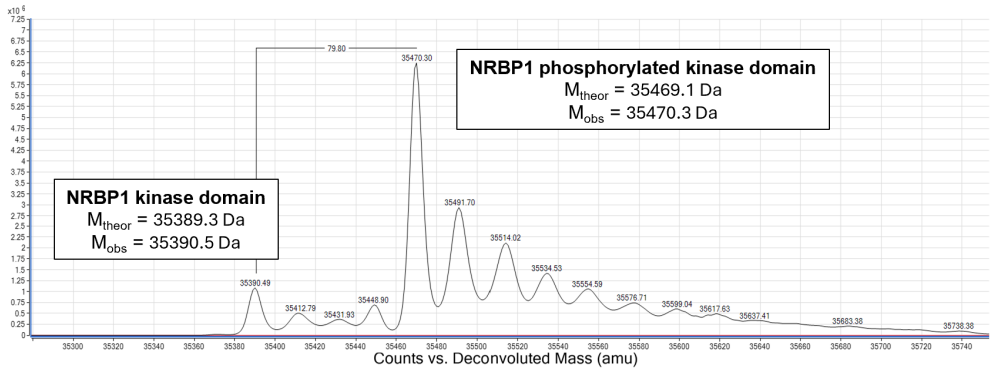


Figure 7.1: **NRBP1 AFv3 modeling and domain architecture.** (a) AFv1 model of NRBP1 protein. (b) Domain architecture of NRBP1. Predicted globular domains are boxed, while predicted disordered regions are drawn as flexible loops

(a) NRBP1 kinase domain expressed in Sf9 cells is mono-phosphorylated



(b) NRBP1 kinase domain after overnight incubation with λ-phosphatase

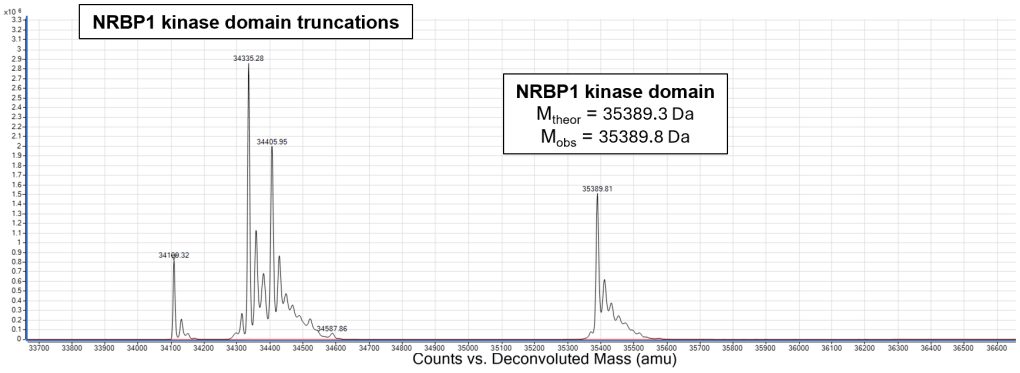


Figure 7.2: **Mass Spectrometry analysis of NRBP1 pseudokinase domain.** (a) Mass spectrometry analysis of NRBP1 pseudokinase domain (41-346 a.a.) from Sf9 cells. An adduct of + 79.8 denoted mono-phosphorylation HO(3)P. (b) Mass analysis of the same sample after overnight incubation with λ-phosphatase, to confirm dephosphorylation. Evidence of protein truncation are also present at lower molecular weights.

In order to address the limited protein stability and to rescue bacterial expression, 55 and 32 NRBP1 constructs were designed for Sf9 and *E. coli* expression respectively. Construct design was guided by the aforementioned modelling and construct boundaries around globular domains were sampled excluding disordered regions. A summary of constructs including protein sequence boundaries and vectors used are summarised in Tables 7.1 for bacterial NRBP1 constructs, 7.2 for insect NRBP1 constructs and 7.3 for other interactors. Full-length and truncated constructs of NRBP1 (Uniprot Q9UHY1) and its interacting partners DDB1 (Q16531), TSC22D2 (O75157), TSC22D3 (Q99576), TSC22D4 (Q9Y3Q8), BRI2 (Q9Y287), BRI3 (Q9NQX7), Elongin B (Q15370) and Elongin C (Q15369) were cloned using ligation independent cloning (LIC) and expressed in small scale, following established methods, as described in detail in Methods **Chapter 2.1.1** and **2.1.2**. The cloning of bacterial NRBP1 constructs summarised in Table 7.2 was kindly performed by colleague Ms Dong Wang.

Small-scale baculoviral expression and nickel purification of constructs in Sf9 cells, analysed via SDS-PAGE gel [Fig. 7.3], showed strong recovery of all NRBP1 constructs, including full-length NRBP1, with the exception of the isolated C-terminal domain alone. This marked the first successful production of the full-length NRBP1 protein. Additionally, the high molecular weight CUL4 adaptor DDB1 (150 kDa) could also be expressed in Sf9 cells allowing for co-transfection studies of the putative complex.

In contrast, small scale *E. coli* and nickel-purification trials, performed by Ms Dong Wang, failed to yield soluble protein for any construct. However, when NRBP1 was cloned into the pNIC-BsaI-ElonginBC vector, all constructs, except for the C-terminal domain, demonstrated robust expression, surpassing the levels seen in Sf9 cells [Fig. 7.4]. The pNIC-BsaI-ElonginBC vector incorporated ElonginB [Uniprot Q15370, 1-118 a.a.] and ElonginC [Uniprot Q15369, 17-112 a.a.] into the coding sequence region (CDS) for enhanced expression of SOCS-box-containing proteins. ElonginB and ElonginC co-expression was therefore essential to achieve *E. coli* expression and high-yielding constructs suitable for the sample concentrations needed for crystallography.

Construct	Vector	Residues	Forward Primer	Reverse Primer
NRBP1-c056	pNIC28-Bsa4	A437:S535	NRBP1-f016	NRBP1-r014
NRBP1-c057	pNIC28-Bsa4	T441:A530	NRBP1-f017	NRBP1-r015
NRBP1-c058	pNIC28-Bsa4	A437:A530	NRBP1-f016	NRBP1-r015
NRBP1-c059	pNIC28-Bsa4	T441:S535	NRBP1-f017	NRBP1-r014
NRBP1-c060	pNIC-BsaI-ElonginBC	S2:S535	NRBP1-f011	NRBP1-r014
NRBP1-c061	pNIC-BsaI-ElonginBC	S2:A530	NRBP1-f011	NRBP1-r015
NRBP1-c062	pNIC-BsaI-ElonginBC	A41:S535	NRBP1-f012	NRBP1-r014
NRBP1-c063	pNIC-BsaI-ElonginBC	A41:A530	NRBP1-f012	NRBP1-r015
NRBP1-c064	pNIC-BsaI-ElonginBC	S59:S535	NRBP1-f013	NRBP1-r014
NRBP1-c065	pNIC-BsaI-ElonginBC	S59:A530	NRBP1-f013	NRBP1-r015
NRBP1-c066	pNIC-BsaI-ElonginBC	R68:S535	NRBP1-f014	NRBP1-r014
NRBP1-c067	pNIC-BsaI-ElonginBC	R68:A530	NRBP1-f014	NRBP1-r015
NRBP1-c068	pNIC-BsaI-ElonginBC	V330:S535	NRBP1-f015	NRBP1-r014
NRBP1-c069	pNIC-BsaI-ElonginBC	V330:A530	NRBP1-f015	NRBP1-r015
NRBP1-c070	pNIC-BsaI-ElonginBC	S2:L409	NRBP1-f011	NRBP1-r017
NRBP1-c071	pNIC-BsaI-ElonginBC	A41:L409	NRBP1-f012	NRBP1-r017
NRBP1-c072	pNIC-BsaI-ElonginBC	S59:L409	NRBP1-f013	NRBP1-r017
NRBP1-c073	pNIC-BsaI-ElonginBC	R68:L409	NRBP1-f014	NRBP1-r017
NRBP1-c074	pNIC-BsaI-ElonginBC	V330:L409	NRBP1-f015	NRBP1-r017
NRBP1-c075	pNIC-BsaI-ElonginBC	V330:T441	NRBP1-f015	NRBP1-r016
NRBP1-c076	pNIC-BsaI-ElonginBC	S2:L388	NRBP1-f011	NRBP1-r018
NRBP1-c077	pNIC-BsaI-ElonginBC	A41:L388	NRBP1-f012	NRBP1-r018
NRBP1-c078	pNIC-BsaI-ElonginBC	S59:L388	NRBP1-f013	NRBP1-r018
NRBP1-c079	pNIC-BsaI-ElonginBC	R68:L388	NRBP1-f014	NRBP1-r018
NRBP1-c080	pNIC-BsaI-ElonginBC	S2:T362	NRBP1-f011	NRBP1-r019
NRBP1-c081	pNIC-BsaI-ElonginBC	A41:T362	NRBP1-f012	NRBP1-r019
NRBP1-c082	pNIC-BsaI-ElonginBC	S59:T362	NRBP1-f013	NRBP1-r019
NRBP1-c083	pNIC-BsaI-ElonginBC	R68:T362	NRBP1-f014	NRBP1-r019
NRBP1-c084	pNIC-BsaI-ElonginBC	S2:E350	NRBP1-f011	NRBP1-r020
NRBP1-c085	pNIC-BsaI-ElonginBC	A41:E350	NRBP1-f012	NRBP1-r020
NRBP1-c086	pNIC-BsaI-ElonginBC	S59:E350	NRBP1-f013	NRBP1-r020
NRBP1-c087	pNIC-BsaI-ElonginBC	R68:E350	NRBP1-f014	NRBP1-Fr020
TSC22D4-c001	pGTVL2	G56:W90	TSC22D4-f001	TSC22D4-r001
TSC22D2-c001	pGTVL2	P146:W180	TSC22D2-f001	TSC22D2-r002
TSC22D4-c002	pHGT-Bio	G56:W90	TSC22D4-f001	TSC22D4-r002
TSC22D2-c002	pHGT-Bio	P146:W180	TSC22D2-f001	TSC22D2-r001

Table 7.1: Construct design for NRBP1 and interactors for *E.coli* expression

Construct	Vector	Residues	Forward Primer	Reverse Primer
NRBP1-c020	pFB-LIC-Bse	A437:S535	NRBP1-f06	NRBP1-r01
NRBP1-c021	pFB-LIC-Bse	T441:A530	NRBP1-f07	NRBP1-r02
NRBP1-c022	pFB-LIC-Bse	A437:A530	NRBP1-f06	NRBP1-r02
NRBP1-c023	pFB-LIC-Bse	T441:S535	NRBP1-f07	NRBP1-r01
NRBP1-c024	pFB-NHStIIT-LIC	S2:S535	NRBP1-f01	NRBP1-r01
NRBP1-c025	pFB-NHStIIT-LIC	S2:A530	NRBP1-f01	NRBP1-r02
NRBP1-c026	pFB-NHStIIT-LIC	A41:S535	NRBP1-f02	NRBP1-r01
NRBP1-c027	pFB-NHStIIT-LIC	A41:A530	NRBP1-f02	NRBP1-r02
NRBP1-c028	pFB-NHStIIT-LIC	S59:S535	NRBP1-f03	NRBP1-r01
NRBP1-c029	pFB-NHStIIT-LIC	S59:A530	NRBP1-f03	NRBP1-r02
NRBP1-c030	pFB-NHStIIT-LIC	R68:S535	NRBP1-f04	NRBP1-r01
NRBP1-c031	pFB-NHStIIT-LIC	R68:A530	NRBP1-f04	NRBP1-r02
NRBP1-c032	pFB-LIC-Bse	S2:V330	NRBP1-f01	NRBP1-r08
NRBP1-c033	pFB-LIC-Bse	A41:V330	NRBP1-f02	NRBP1-r08
NRBP1-c034	pFB-LIC-Bse	S59:V330	NRBP1-f03	NRBP1-r08
NRBP1-c035	pFB-LIC-Bse	R68:V330	NRBP1-f04	NRBP1-r08
NRBP1-c036	pFB-LIC-Bse	V330:S535	NRBP1-f05	NRBP1-r01
NRBP1-c037	pFB-LIC-Bse	V330:A530	NRBP1-f05	NRBP1-r02
NRBP1-c038	pFB-LIC-Bse	S2:L409	NRBP1-f01	NRBP1-r04
NRBP1-c039	pFB-LIC-Bse	A41:L409	NRBP1-f02	NRBP1-r04
NRBP1-c040	pFB-LIC-Bse	S59:L409	NRBP1-f03	NRBP1-r04
NRBP1-c041	pFB-LIC-Bse	R68:L409	NRBP1-f04	NRBP1-r04
NRBP1-c042	pFB-LIC-Bse	V330:L409	NRBP1-f05	NRBP1-r04
NRBP1-c043	pFB-LIC-Bse	V330:T441	NRBP1-f05	NRBP1-r03
NRBP1-c044	pFB-LIC-Bse	S2:L388	NRBP1-f01	NRBP1-r05
NRBP1-c045	pFB-LIC-Bse	A41:L388	NRBP1-f02	NRBP1-r05
NRBP1-c046	pFB-LIC-Bse	S59:L388	NRBP1-f03	NRBP1-r05
NRBP1-c047	pFB-LIC-Bse	R68:L388	NRBP1-f04	NRBP1-r05
NRBP1-c048	pFB-LIC-Bse	S2:T362	NRBP1-f01	NRBP1-r06
NRBP1-c049	pFB-LIC-Bse	A41:T362	NRBP1-f02	NRBP1-r06
NRBP1-c050	pFB-LIC-Bse	S59:T362	NRBP1-f03	NRBP1-r06
NRBP1-c051	pFB-LIC-Bse	R68:T362	NRBP1-f04	NRBP1-r06
NRBP1-c052	pFB-LIC-Bse	S2:E350	NRBP1-f01	NRBP1-r07
NRBP1-c053	pFB-LIC-Bse	A41:E350	NRBP1-f02	NRBP1-r07
NRBP1-c054	pFB-LIC-Bse	S59:E350	NRBP1-f03	NRBP1-r07
NRBP1-c055	pFB-LIC-Bse	R68:E350	NRBP1-f04	NRBP1-r07
ElonginB-c002	pUCDM-CTF-LIC	D2:Q118	EloB-f01	EloB-r01
ElonginC-c004	pKL-CTF-LIC	M1:C96	EloC-f01	EloC-r01
DDB1-c006	pFB-GST3C-LIC	S2:H1140	DDB1-f02	DDB1-r02
DDB1-c007	pFB-LIC-Bse	S2:H1140	DDB1-f01	DDB1-r02
DDB1-c008	pFB-CT6H-LIC	S2:H1140	DDB1-f03	DDB1-r01
DDB1-c009	pFB-NHStIIT-LIC	S2:H1140	DDB1-f01	DDB1-r02

Table 7.2: Construct design for NRBP1 and interactors for Sf9 expression

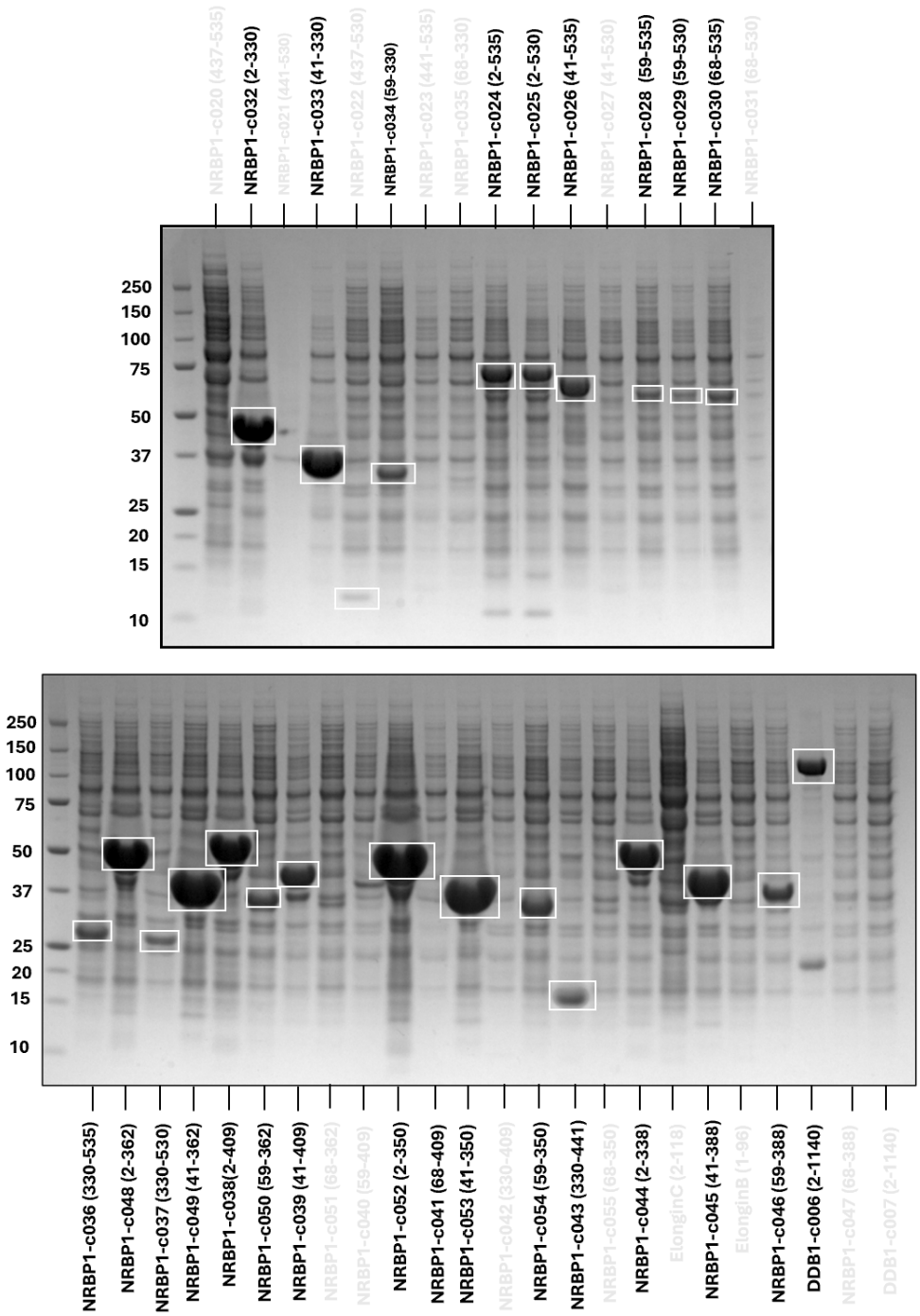


Figure 7.3: **NRBP1 construct Sf9 expression result summary.** Overview of small-scale test expression and IMAC purification results (SDS-PAGE analysis) of NRBP1 and interacting partners (as described in Methods 2.1.4 Constructs with medium to high protein expression are annotated in black and expressed protein bands are boxed in white. Construct with low and no protein expression are annotated in grey.

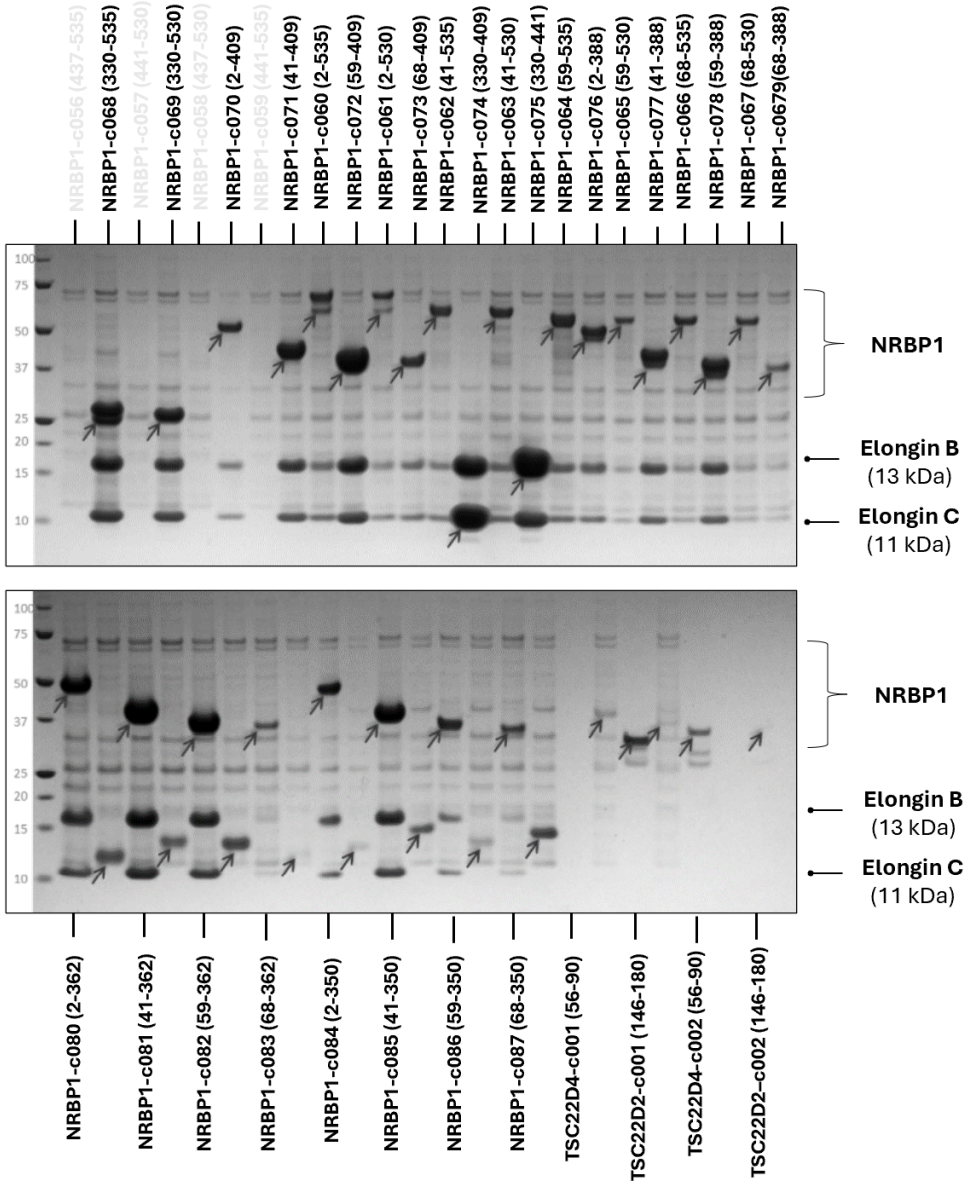
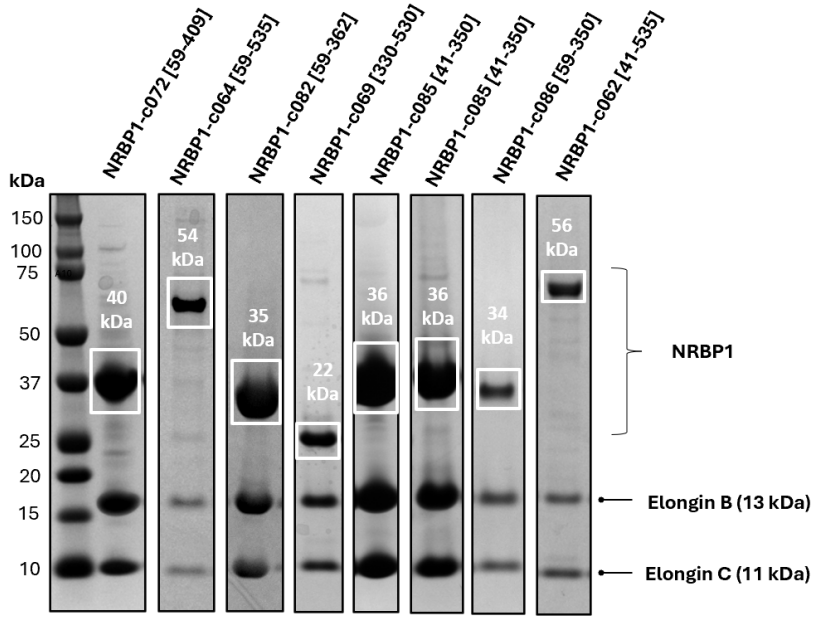


Figure 7.4: **NRBP1 construct *E. coli* expression result summary.** Overview of test expression and purification results (SDS-PAGE analysis) of NRBP1 and interactors. Constructs with medium to high protein expression are annotated in black and expressed protein bands are boxed in white. Construct with low and no protein expression are annotated in gray. Arrows denote the band corresponding to the expressed NRBP1 protein subunit. High-throughput cloning and test expression were performed by Ms Dong Wang.

(a)



(b)

		Purification				
Construct	**Boundaries**	**Step 1**	**Step 2**	**Step 3**	**Step 4**	**Step 5**
NRBP1-c072	59-409	Ni-NTA	Superdex 200 pg	Ni-NTA	PD-10	
NRBP1-c064	59-535	Ni-NTA	Superdex 200 pg			
NRBP1-c082	59-362	Ni-NTA	Superdex 200 pg	Mono Q	Ni-NTA	PD-10
NRBP1-c069	330-530	Ni-NTA	Superdex 200 pg	HiTrap Q SP		
NRBP1-c085	41-350	Ni-NTA	Superdex 200 pg	Mono Q		
NRBP1-c085	41-350	Ni-NTA	Superdex 200 pg	Mono Q	Ni-NTA	
NRBP1-c086	59-350	Ni-NTA	Superdex 200 pg	Mono Q	Ni-NTA	
NRBP1-c062	41-535	Ni-NTA	Superdex 200 pg			

Figure 7.5: **Summary of NRBP1-ElonginBC samples for crystallisation trials.** (a) SDS-PAGE analysis of NRBP1-ElonginBC protein complexes applied to crystal plates. (b) Summary of purification steps of each complex to achieve crystallography quality pure protein.

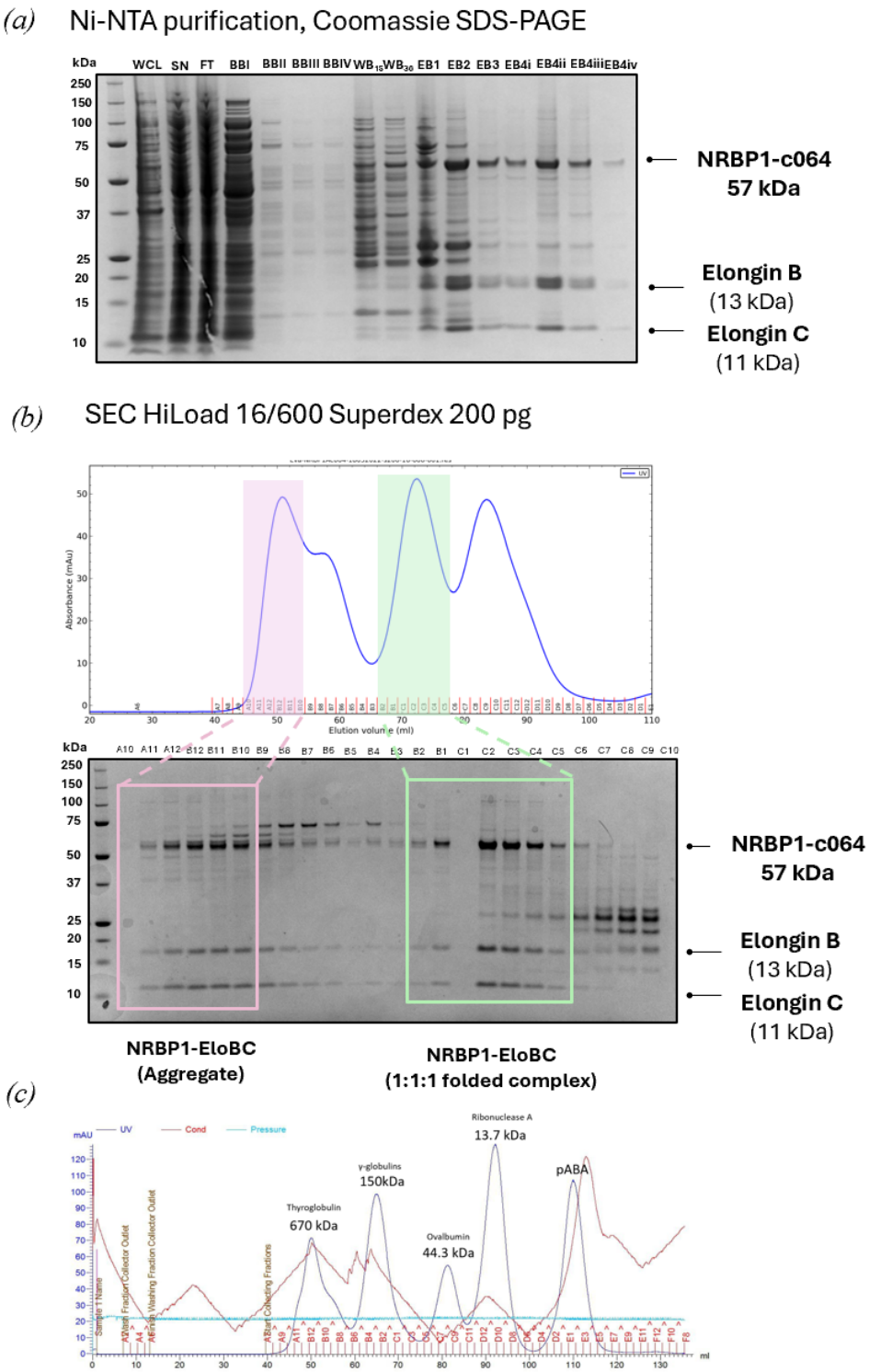


Figure 7.6: **Representative purification of NRBP1-ElonginBC complex.** a) Samples of IMAC purified NRBP1-ElonginBC analysed in SDS-PAGE and proteins were visualised with Coomassie stain. (b) Samples were subsequently further purified on a HiLoad 16/600 Superdex column 200 pg and the protein elution monitored by UV. The UV absorbance is plotted over elution volume (mL), and fractions were analysed in SDS-PAGE gel below. Shaded pink is aggregated complex, shaded green the folded NRBP1-ElonginBC complex, (c) protein standards run in the same column HiLoad 16/600 Superdex 200 pg for molecular weight comparisons. WCL = whole cell lysate, SN = supernatant, FT = flow-through, BB = binding buffer, WB = wash buffer, EB = elution buffer.

Construct	Vector	Residues	Forward Primer	Reverse Primer
ITM2C-c001	pHGT-Bio	A89:V267	ITM2C-f001	ITM2C-r001
ITM2C-c002	pNIC-Bio3	A89:V267	ITM2C-f001	ITM2C-r001
ITM2B-c001	pHGT-Bio	P83:S266	ITM2B-f001	ITM2B-r001
ITM2B-c002	pNIC-Bio3	P83:S266	ITM2B-f001	ITM2B-r001
CUL5-c001	pACYC-28cr	M1:K406	CUL5-f001	CUL5-r001
CUL2-c001	pACYC-28cr	M1:K406	CUL2-f001	CUL2-r001
TSC22D3-c001	pCDF-LIC	M1:V134	TSC22D3-f001	TSC22D3-r001
TSC22D3-c002	pCDF-LIC	M1:E117	TSC22D3-f001	TSC22D3-r002
TSC22D3-c003	pCDF-LIC	M5:V134	TSC22D3-f002	TSC22D3-r001
TSC22D3-c004	pCDF-LIC	M5:E117	TSC22D3-f002	TSC22D3-r002
TSC22D4-c001	pCDF-LIC	M1:V395	TSC22D4-f001	TSC22D4-r001
TSC22D4-c002	pCDF-LIC	M1:V382	TSC22D4-f001	TSC22D4-r002
TSC22D4-c003	pCDF-LIC	M240:V395	TSC22D4-f002	TSC22D4-r001
TSC22D4-c004	pCDF-LIC	M240:V382	TSC22D4-f002	TSC22D4-r002
TSC22D4-c005	pCDF-LIC	D281:V395	TSC22D4-f003	TSC22D4-r001
TSC22D4-c006	pCDF-LIC	D281:V382	TSC22D4-f003	TSC22D4-r002

Table 7.3: Construct design for NRBP1 interactors for *E.coli* co-expression

Finally, interacting partners of NRBP1 including TSC22D2 [Uniprot O75157, 146-180 a.a.], TSC22D4 [Uniprot Q9Y3Q8, 56-90 a.a.], CUL2 [Uniprot Q13617, 1-384 a.a.] and DDB1 [Uniprot Q16531, 2-1140 a.a.] were also cloned and expressed successfully in either *E. coli* or Sf9 systems, allowing for interaction studies.

7.2.2 NRBP1 associates with ElonginBC stoichiometrically *in vitro*

The dimerisation of NRBP1, possibly through the LisH motif (465-476 a.a.) in the C-terminal domain (CTD, 465-535 a.a.) [Fig. 1.10 (b)], has been reported by multiple sources in the literature through cellular studies but was never tested using purified protein *in vitro*. NRBP1 constructs with three different C-terminal truncations, 59-362 a.a., 59-409 a.a., and 59-535 a.a., were coexpressed in *E. coli* with ElonginBC, and the complexes captured by NRBP1's hexahistidine tag using Ni^{2+}-sepharose resin. The N-terminal hexahistidine tag was cleaved using TEV protease and the protein was further purified by size exclusion chromatography and ion exchange chromatography, as summarised in Fig. 7.5. Based on the gel filtration chromatography UV profile [representative sample profile, Fig 7.6 (b)], there is no indication of NRBP1 homodimerisation. NRBP1-ElonginBC complex with a mass ~80 kDa is eluted in a single peak at 70 mL of a SEC HiLoad 16/600 Superdex 200 pg column, and not at 60-65 mL where a hexameric 160 kDa complex (dimer of NRBP1-ElonginBC trimer) would be expected as protein standards suggest [Fig 7.6 (d)].

The samples were further analysed by mass spectrometry under denaturing conditions (0.1% formic acid) and the identities of all NRBP1 expected proteins were confirmed, alongside ElonginB and ElonginC [Fig. 7.7]. The same protein complexes were also investigated using native mass spectrometry under non-denaturing conditions. Analysis revealed that the primary species of all complexes excluding or including the C-terminal domain, was a NRBP1:ElonginB:ElonginC complex in 1:1:1 stoichiometry [Fig. 7.8]. No hexameric complex of NRBP1:ElonginB:ElonginC in a 2:2:2 stoichiometry could be discerned in the expected region [Fig. 7.8, marked with asterisk].

The data suggests that under these experimental conditions, the LisH motif in the C-terminal domain is not sufficient for NRBP1 homodimerisation.

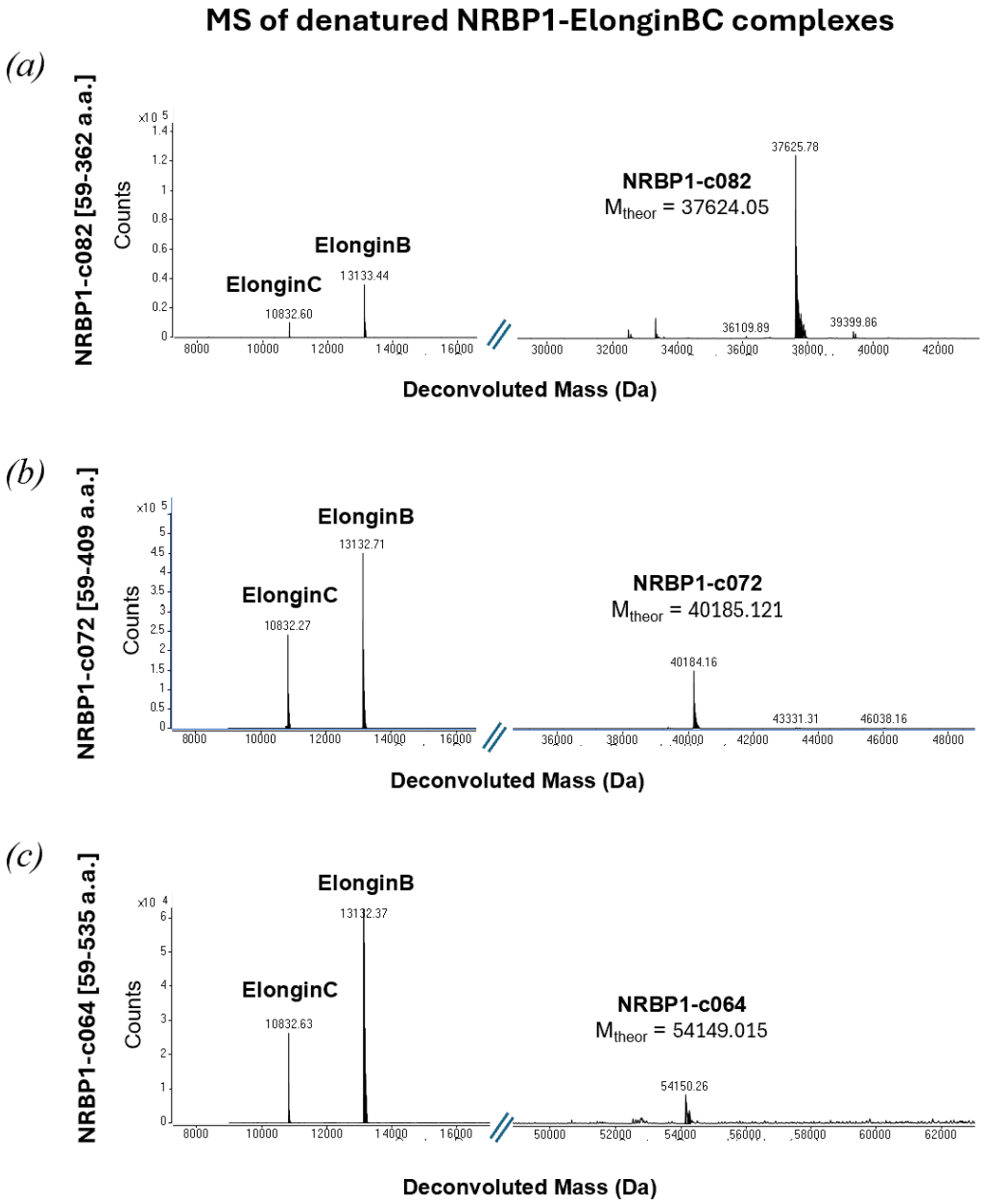


Figure 7.7: **Mass Spectrometry analysis of NRBP1-ElonginBC purified complexes.** Counts (y axis) plotted against deconvoluted mass (x axis) for LC-MS analysis under denaturing conditions confirmed the identity of NRBP1-ElonginBC co-expressed complexes of (a) NRBP1-c082 [59-362 a.a.] (b) NRBP1-c072 [59-409 a.a.], (c) NRBP1-c064 [59-535 a.a.]. The ppm error was calculated using the formula ppm error $= \left(\frac{M_{theo} - M_{.obs}}{M_{theor}}\right) \times 10^6$, where M_{theor} is the theoretical mass and M_{Obs} is the observed mass. Protein identity was confirmed as ppm error was below 50 ppm in all cases.

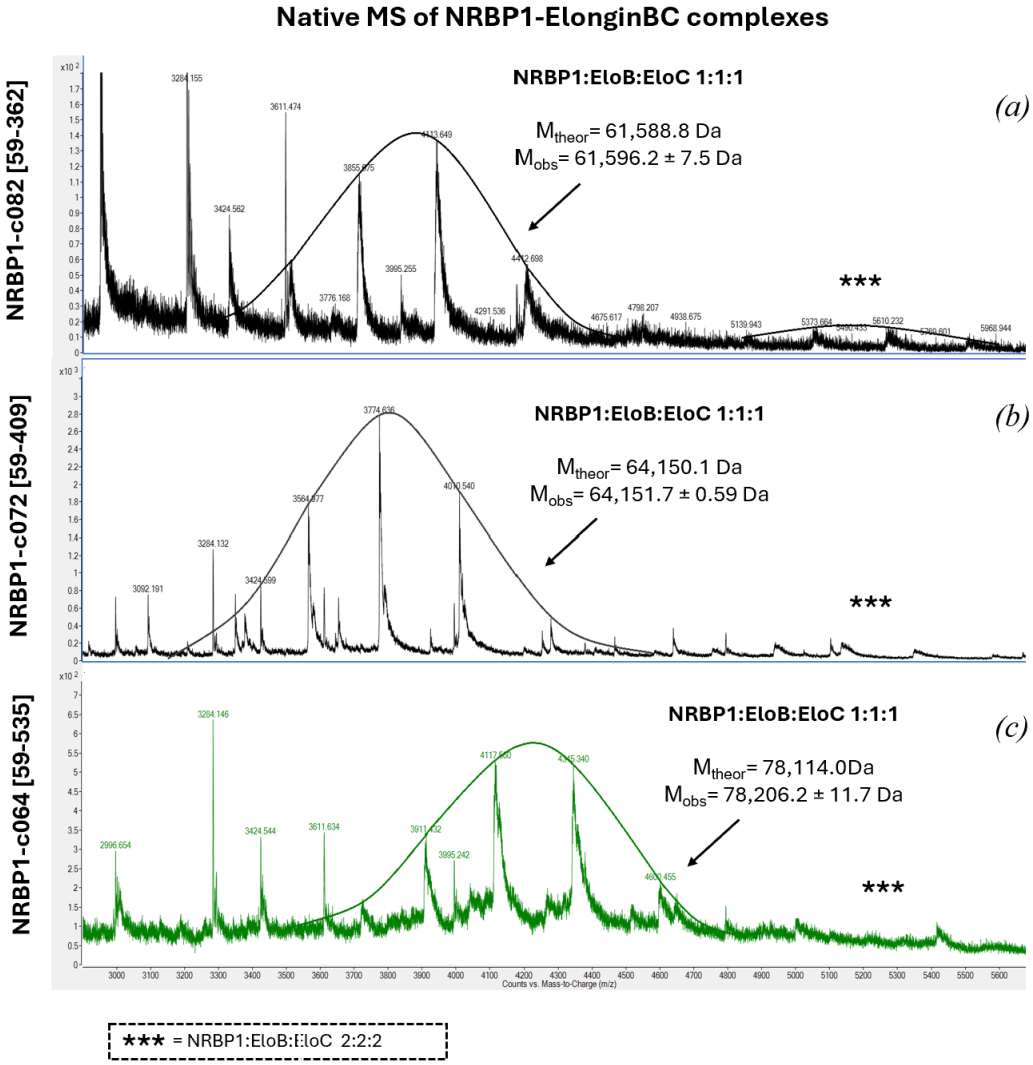


Figure 7.8: **NRBP1:ElonginB:ElonginC is a 1:1:1 complex**. Native mass spectrometry analysis of counts plotted over mass to charge ration (m/z). Purified NRBP1-ElonginBC samples were dialysed overnight from the purification buffer (50 mM HEPES, 300 mM NaCl, 0.5 mM TCEP) to ammonium acetate 50 mM solution, adjusted at pH 7.5, at 4 °C with stirring. The charged states of the protein complexes were used to calculate the de-convoluted mass of the NRBP1-ElonginBC complex using ESIprot [285]. Counts over Mass-to-Charge (m/z) were plotted for NRBP1 constructs (a) NRBP1-c082 [59-362 a.a.], (b) NRBP1-c072 [59-409 a.a.], (c) NRBP1-c064 [59-535 a.a.] co-expressed with ElonginBC. The data revealed that the primary species was one NRBP1 subunit binding one molecule of ElonginB and one molecule of ElonginC. The three asterisks denote where the dimeric complex species was expected to be observed.

7.2.3 Structural characterisation attempts for the NRBP1-ElonginBC complex

Initially, coarse screen crystallisation trials using NRBP1 pseudokinase domain (Q9UHY1, 41-346 a.a.) expressed from Sf9 cells in both its phosphorylated and de-phosphorylated state as previously described, were set up using the vapour diffusion method with sitting drops (Table 7.4). A subset of the samples were supplemented with 4 mM adenine in an effort to stabilise the pseudokinase domain [286], and rescue the limited protein stability as observed by MS [Fig. 7.2]. No crystal were observed in any of the tested conditions.

Table 7.4: NRBP1 Kinase domain crystallisation trials

Boundaries	**Protein Conc.** (mg/mL)	**Coarse Screens** 4 °C	**Coarse Screens** 20 °C	**Comments**
41-346	10.0	HCS3, HIN3	LFS4, JCSG7	supplemented with 4 mM adenine
41-346	6	JCSG, HIN3, HCS, LFS	LFS, HIN3, JCSG, HCS	dephosphorylated
41-346	6	JCSG, HIN3, HCS, LFS	HIN3, LFS, JCSG, HCS	supplemented with 4 mM adenine dephosphorylated
41-346	6	HCS, HIN3, Morpheus	JCSG7, LFS, Morpheus	supplemented with 4mM adenine, phosphorylated

As a next step, the NRBP1-ElonginBC *E. coli* expression constructs that showed high levels of expression in small scale (1 mL of culture media) [Fig. 7.4], were selected for large scale purification (multiple litres of 2xYT media). All constructs were captured by their N-terminal hexahistidine tag, and purified by nickel-affinity chromatography, followed by tag cleavage and gel filtration using standard techniques as outlined previously in **Chapter 3.2.1**. When necessary, ion-exchange chromatography or a second nickel affinity column followed. For ion exchange chromatography, a positively charged cationic column was used to capture negatively charged proteins in low salt buffer (50 mM Tris pH = 7.4, 2.5% glycerol). The proteins were eluted with increasing salt gradient. The construct boundaries of the expressed proteins and the purification steps are summarised in Fig. 7.5 (b). In contrast with the previous crystal sample of NRBP1 alone, both SDS-PAGE analysis of purified samples

[Fig. 7.5 (a)] and mass spectrometry analysis [Fig. 7.7] did not show any signs of protein degradation and verified the stability of the complex. This observation highlighted the importance of Elongin BC in NRBP1 stability, suggesting that its stability depends on co-folding with Elongin BC as an obligate partner. Coarse screen crystallisation trials were thus set up as summarised in Table 7.5. Despite the sample purity and homogeneity no crystals were observed in any of the tested conditions.

Table 7.5: NRBP1-ElonginBC complex crystallisation trials

Construct	Conc. (mg/mL)	Coarse Screens 4 °C	Coarse Screens 20 °C
-c062 41-535 a.a.	11.5	JCSG+ECO, HCS	N/A
-c064 59-535 a.a	11.5	N/A	JCSG, LFS
-c069 330-530 a.a	8.6	JCSG, HCS HIN, LFS	Morpheus Midas+, LFS
-c072 59-409 a.a	13.3	JCSG7, HCS HIN, LFS Morpheus, Midas+	JCSG7, HIN HCS, Morpheus Midas+, LFS
-c082 59-362 a.a	13.1	JCSG, BCS HIN, LFS Morpheus	JCSG, BCS HIN, LFS Morpheus
-c085 41-350 a.a	10.5	HCS, LFS, HIN, JCSG, BCS, Midas+ Morpheus	HCS, LFS, HIN, JCSG, BCS, Midas+ Morpheus
-c085 41-350 a.a	12.0	Stura Footprint, PGA Screen	Stura Footprint, PGA Screen
-c086 59-350 a.a	10.2	N/A	LFS, BCS, JCSG

Since no crystals were obtained in the trials above, structure determination of purified reconstituted NRBP1-ElonginBC complexes was attempted by cryo-EM. Near full-length NRBP1-ElonginBC 59-535 a.a. was purified as described above in 50 mM HEPES, 150 mM NaCl, 0.5 mM TCEP [Fig. 7.6] yielding a complex at 80 kDa that is at the cusp of the resolution limit of the technique for single-particle analysis. Initial grids were prepared and vitrified with sample applied to quantifoil Cu 1.2/1.3 µM 300 mesh. Two different concentrations 0.8 and 3 mg/mL and a range of blotting times were sampled to ascertain optimal conditions and protein behaviour, since this would be the first time NRBP1-ElonginBC was applied to grids. Grid preparation conditions are summarised in Table 7.6. These initial grids were screened for particle distribution and ice thickness on a Glacios

microscope (Thermo Fisher Scientific) operating at 200keV, on a Falcon 3 detector at 120kv nominal magnification. Grids at lower concentration 0.8 mg/mL appeared mostly dry, while those at higher conc 3 mg/ml and with longer blotting times (5, 5.5 s) showed a gradient of ice and potential for further screening.

Table 7.6: NRBP1-c064[59:535]-ElonginBC complex grid screening conditions

Grid Identifier	Grid Type	Conc. (mg/mL)	Blot time (s)
NRBPA-u007	Qfoil Cu 1.2/1.3 µm 300 mesh	0.8	3
NRBPA-u008	Qfoil Cu 1.2/1.3 µm 300 mesh	0.8	4
NRBPA-u009	Qfoil Cu 1.2/1.3 µm 300 mesh	0.8	5
NRBPA-u010	Qfoil Cu 1.2/1.3 µm 300 mesh	**3**	**5.5**
NRBPA-u011	Qfoil Cu 1.2/1.3 µm 300 mesh	3	7

Further grids were prepared and screened near the optimised conditions; at 2.6 mg/mL and sampling longer blotting times as summarised in Table 7.7. In addition to quantifoil Cu 1.2/1.3 µM 300 mesh grids, Au 1.2/1.3 µM 200 mesh grids affording the possibility of thinner ice and potentially enhancing signal to noise ratio [Fig. 7.9 (a)]. However, the lower density gold mesh grids (NRBP1-u018) exhibited thicker, uneven ice, whereas the higher density copper grids (NRBP1-u013) produced an optimal ice gradient [Fig. 7.9 (b)]. A large data set (11,573 micrographs) was acquired from grid NRBP1-u014 (Quantifoil Cu 1.2/1.3 µM 300 mesh, 2.6 mg/mL, 7.5 s blotting time) on a Krios IV microscope (Diamond, eBIC) at 300 keV, using a K3 detector at 105 kx nominal magnification.

After initial processing using the eBIC's auto-processing pipeline, for motion correction and CTF estimation, 1,021,837 particles were picked by CRYOLO. The processing continued in cryoSPARC, with 2D classification using an inner mask diameter of 130 Å, consistent with particle diameter as predicted by AFv3 [Fig. 7.9 (d)]. After iterative rounds of 2D classification and selection, the 2D classes that showed sufficient alignment were selected (576,737 particles) but failed to give meaningful alignment in ab initio reconstruction. Alternative approaches were employed where selected 2D classes with well-aligned features were used as templates for particle picking or as a training set for the machine-learning-based particle-picking tool Topaz, to repick particles from the dataset. In all

attempts, 2D classification yielded similar 2D classes to the CRYOLO picks and did not improve 3D reconstruction. A summary of the workflow of data refinement can be found in Fig 7.10.

Table 7.7: NRBP1-c062[41:535]-ElonginBC complex grids for collection

Grid Identifier	Grid Type	Conc. (mg/mL)	Blot time (s)
NRBP1-u012	Qfoil Cu 1.2/1.3 µm 300 mesh	2.6	5.5
NRBPA-u013	**Qfoil Cu 1.2/1.3 µm 300 mesh**	**2.6**	**6.5**
NRBP1-u014	Qfoil Cu 1.2/1.3 µm 300 mesh	2.6	7.5
NRBP1-u015	Qfoil Cu 1.2/1.3 µm 300 mesh	2.6	8.5
NRBP1-u016	Qfoil Au 1.2/1.3 µm 200 mesh	2.6	5.5
NRBP1-u017	Qfoil Au 1.2/1.3 µm 200 mesh	2.6	6.5
NRBP1-u018	Qfoil Au 1.2/1.3 µm 200 mesh	2.6	7.5
NRBP1-u019	Qfoil Au 1.2/1.3 µm 200 mesh	2.6	8.5

The protein particles were clearly visible in the micrographs and readily picked by both eBIC's CRYOLO auto-processing and cryoSPARC's blob picker [Fig. 7.9 (c)]. The particle diameter was approximately 100 nm, consistent with the expected particle size based on the AFv3 model of the NRBP1-ElonginBC complex (100x70 nm) [Fig. 7.9 (d)]. Since, the particle selection is consistent with protein particles, the poor volume reconstruction betrays an inherent limitation of the system leading to the discontinuation of further attempts at processing. Thus, the focus shifted to validating other NRBP1 cellular interactors *in vitro* to increase the molecular weight of the complex and assist any future 3D reconstruction attempts by increasing the signal to noise ratio.

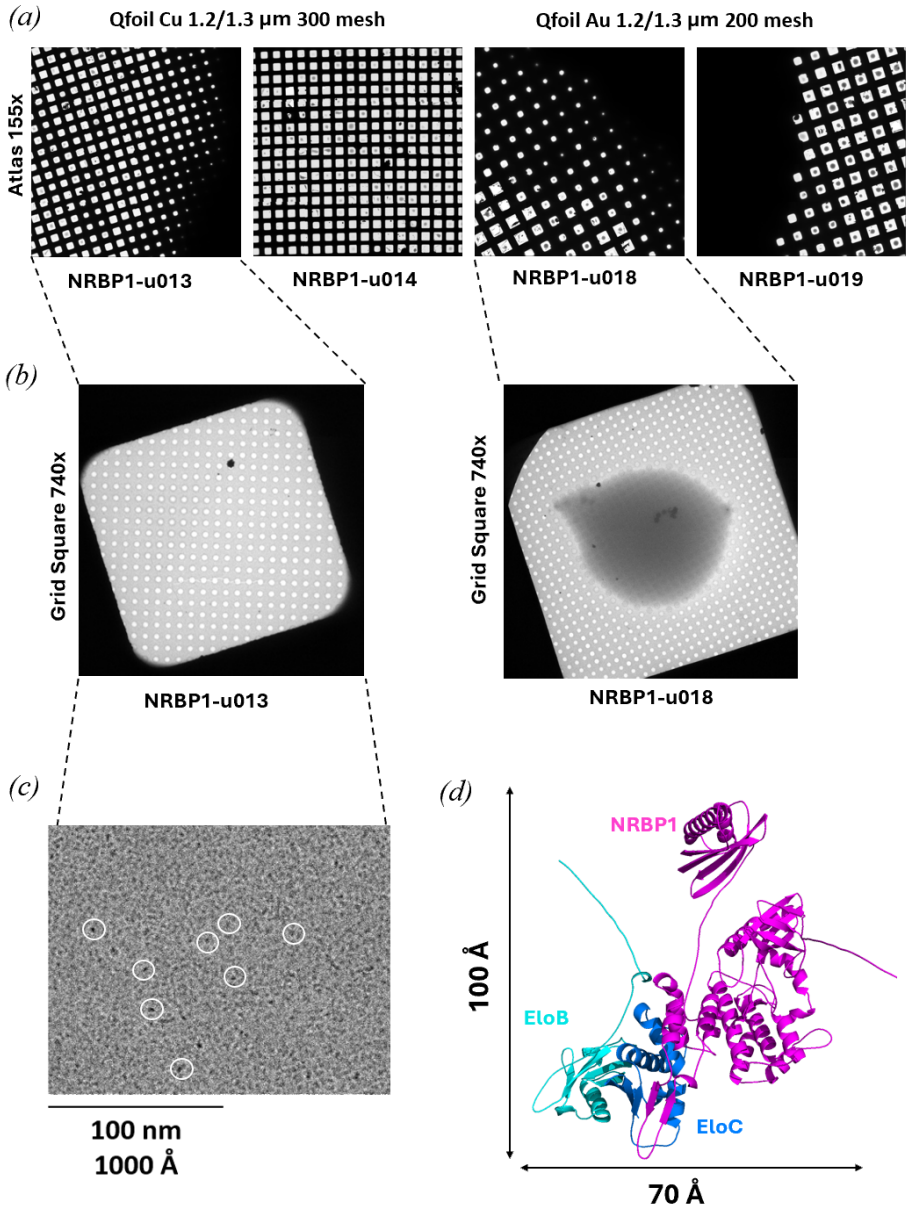


Figure 7.9: **Cryo-EM grid screening overview for NRBP1-ElonginBC complex.** (a) Representative grid atlases of copper Qfoil 1.2/1.3 µm 300 mesh and gold Au 1.2/1.3 µm 200 mesh grids. (b) Representative images at the square level, showing uneven ice gradient of the gold grid compared with even blotting of the copper grid, (c) Representative micrograph of copper grid NRBP1-u013. Protein particles were clearly visible; for illustrative purposes a few particles are circled in white. Images acquired in a Glacios 200 kV instrument, (d) NRBP1-ElonginBC complex modelled with AFv3. Approximate particle diameter measured in Å in PyMol. Of note, due to the likely flexibility of C-term CCT-like domain the visualized size might be smaller than predicted.

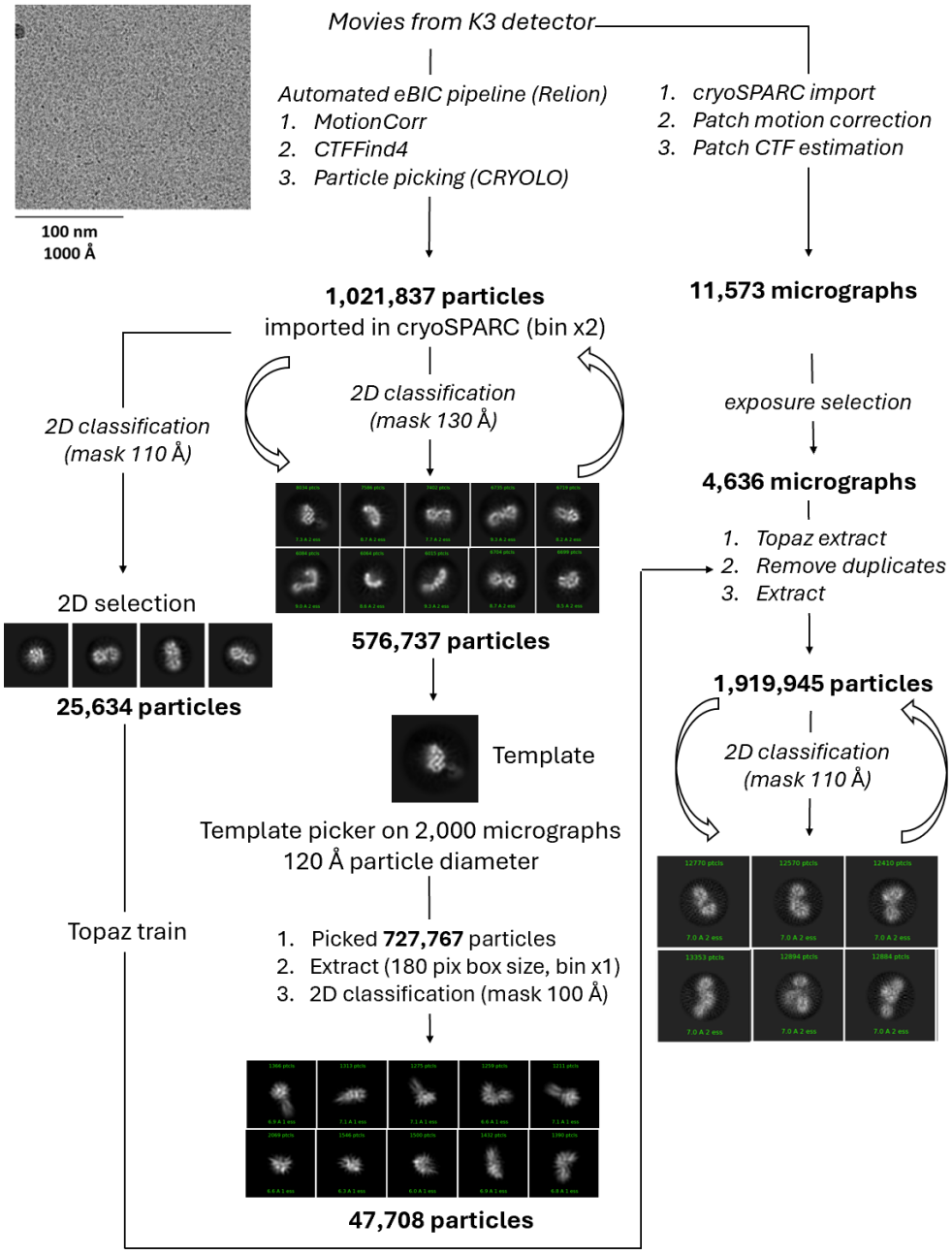


Figure 7.10: **Cryo-EM data refinement workflow of NRBP1-ElonginBC complex.** Analysis performed in cryoSPARC v3.3.3 and eBIC automated pipeline based on Relion.

7.2.4 Validating NRBP1-ElonginBC complex interactions

7.2.4.1 ElonginBC is not sufficient for CUL2 recruitment

The NRBP1-ElonginBC complex was previously shown to be stable and reproducible *in vitro* [Fig. 7.5 (a)]. Since ElonginBC is a known adaptor for CUL2 and CUL5, I sought to verify whether NRBP1 also associates with CUL2 *in vitro*, as previously reported in cellular co-immunoprecipitation experiments [126]. FLAG-tagged human CUL2 N-terminal domain [Uniprot Q13617, 1-384 a.a.] was cloned, purified and mass confirmed as described in **Chapter 2.1.4**. Untagged NRBP1-c072 [59-490 a.a.] construct that included both the SOCS box and putative CUL2 box was also purified and confirmed for correct mass as previously described [Fig. 7.5]. The purified NRBP1 [59-409 a.a.] in complex with ElonginBC was mixed with N-terminal CUL2 [1-384 a.a.] in equimolar quantities and incubated for 3 hours at 4 °C with end-to-end mixing before being applied to buffer-equilibrated FLAG resin. The resin was subsequently washed with wash buffer and the samples analysed in SDS-PAGE as seen in Fig. 7.11 and described in Methods 2.2. No interaction between NRBP1-ElonginBC and CUL2 was detected in the pull-down assay under the experimental conditions.

7.2.4.2 Indications of DDB1 recruitment in the absence of ElonginBC

Based on sequence alignment of consensus motifs, Yasukawa *et al*. have suggested that the H-box for CUL4A recruitment overlaps with the BC-box for CUL2 recruitment in the same α-helix 331-341 a.a., making DDB1 and EloBC binding mutually exclusive [126]. In order to investigate that hypothesis, His-tagged NRBP1 construct that included the H-box responsible for DDB1 recruitment [NRBP1-c026, 41-535 a.a.] and full-length GST-tagged DDB1 were co-expressed in Sf9 insect cells and were pulled-down by their respective tags from the cell-lysate. The samples were analysed by SDS-PAGE gel and probed by α-His and α-GST antibodies. His-tagged NRBP1 was co-eluted with GST-tagged DDB1 from the GST-sepharose beads by 20 mM glutathione, while DDB1 was co-eluted with NRBP1 from the nickel beads with imidazole [Fig. 7.12 (a) and (b) respectively]. The relative

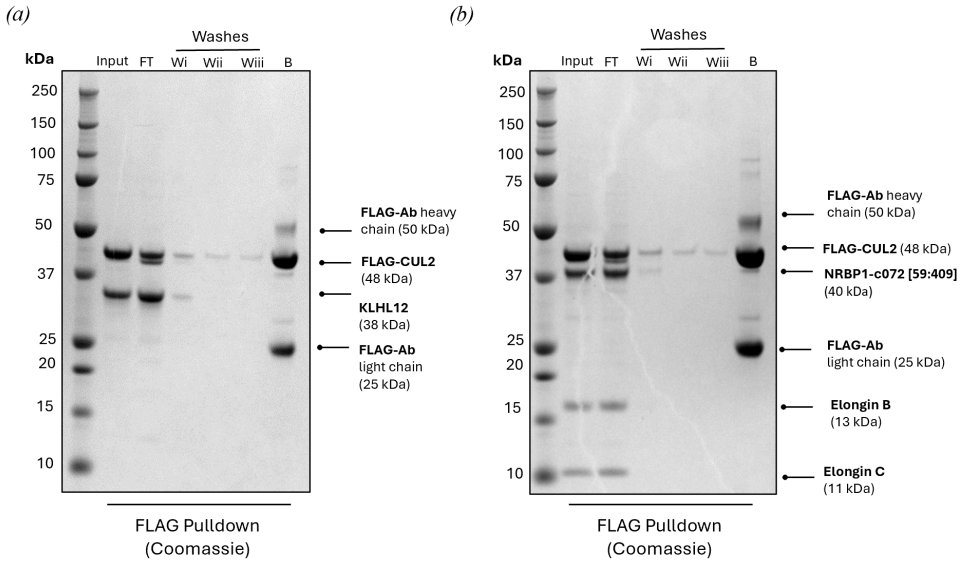


Figure 7.11: **NRBP1-ElonginBC CUL2-box is not sufficient for CUL2 recruitment.** (a) FLAG-pulldown of FLAG-tagged CUL2 and untagged Kelch domain of KLHL12, that does not associate with CUL2 (negative control). (b) FLAG-pulldown of FLAG-tagged CUL2 and untagged NRBP1-c072 [59-409 a.a.] in complex with ElonginBC. In both cases, the two purified proteins were mixed (input sample) in equimolar quantities and incubated for 3 hours before being applied to equilibrated FLAG resin (FT = flow-through). The resin was washed (W = wash). All samples including beads (B = beads) were analysed using SDS-PAGE. Neither KLHL12 (negative control), nor NRBP1 were not captured the FLAG-CUL2 immobilised on the beads, in contrast to what was expected.

intensities of NRBP1 and DDB1 for the same fractions suggested the formation of a NRBP1-DDB1 complex. Interestingly, His-tag pulldown of the His-tagged NRBP1-ElonginBC complex [NRBP1-c072, 59-409 a.a.] failed to detect interaction with GST-DDB1 from *E. coli* lysate, despite the presence of the DDB1-binding H-box motif in NRBP1 [Fig. 7.12 (c)]. Taken together, these data support the hypothesis that NRBP1 alone is sufficient for association with DDB1, an interaction that is not possible in the presence of ElonginBC. This observation suggests that ElonginBC and DDB1 might compete for the same binding region on NRBP1.

7.2.4.3 TSC22D4 transcription factor binds to NRBP1

Although the interaction between the TSC22D protein family and NRBP1 has been documented, the specific binding site within the protein remains unknown. An initial homology model of NRBP1 using SWISS-MODEL [192, 287], that I generated before the advent of AlphaFold, modelled the NRBP1 pseudokinase domain 66-332 residues based on the crystal structure of WNK3 kinase domain

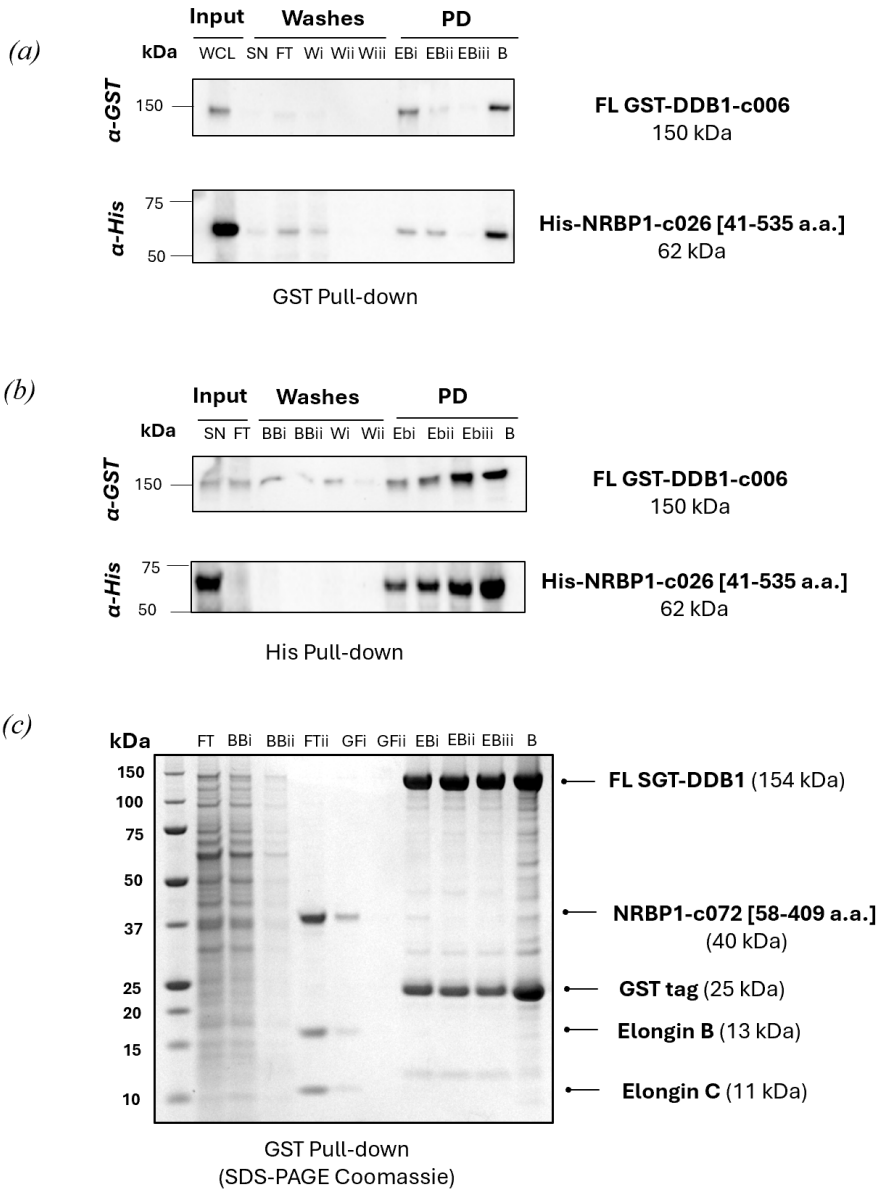


Figure 7.12: **NRBP1-ElonginBC binding precludes DDB1 association**. His-tagged NRBP1-c026 42-535 a.a. and full-length GST-tagged DDB1 were co-expressed in Sf9 insect cells. (a) GST tagged DDB1 was bound to sepharose resin, and eluted with glutathione, alongside His-tagged NRBP1. (b) His-tagged NRBP1 was bound to Ni^{2+} resin and eluted with imidazole, alongside GST-tagged DDB1. Pull down samples were analysed on SDS-PAGE and Western blot. Blots were probed with anti-GST and anti-His antibodies. (c) His-pulldown of *E. coli* cell lysate expressing GST-tagged full-length DDB1 incubated with purified NRBP1-ElonginBC complex [NRBP1-c072, 59-409 a.a.] does not show any DDB1-NRBP1 association. WCL: whole-cell lysate, SN: supernatant, FT: flow-through, BB: binding buffer, EB: elution buffer, B: beads.

in its monophosphorylated apo state (A-loop swapped) (PDB 5O1V [288], Seq Id 32.82% Average Model Confidence = 0.62 $\pm$ 0.05). This finding informed the hypothesis that NRBP1 C-terminal domain [430-530 a.a.] may adopt the fold of a conserved carboxy-terminal CCT-like RFXV-motif binding domain, as found in the homologous WNK1-4 atypical kinases [289, 290]. Interestingly, the sequence motif that CCT-like domain recognises, "RFXV," was present in binding partners TSC22D2 [O75157, 71-RFRV-74 a.a.] and TSC22D4 [Q9Y3Q8, 161-RFRV-164 a.a.]. To test this hypothesis, purified NRBP1 protein constructs that include the C-terminal domain *NRBP1 (+CTD)* [NRBP1-c069, 330-530 a.a.] and exclude it *NRBP1 (-CTD)* [NRBP1-c072, 59-409 a.a.] were tested for RFXV binding. Biolayer Interferometry (BLI) was chosen as a label-free, low-volume, high-throughput method suitable for determining binding affinities and kinetics, in contrast to label-dependent MST or the high sample demands and lower throughput of ITC and SPR.

The predicted disordered TSC22D2[1] and TSC22D4[2] peptides, with 10-15 residues on either side of central "RFXV" motif, were expressed in *E. coli* with biotin supplementation in the expression media to enable biotinylation of their C-terminal Avi-tag, as detailed in Methods 2.1.3. Purification by their N-terminal GST-tag followed, as detailed in Methods 2.1.4. The well-characterized interaction between the "RFTV" motif of KCC2[3] and the CCT-like domain of the kinase SPAK[4] [291, 292] was used as a positive control in the BLI assay. This served to confirm that RFXV-motif binding to the CCT-like domain could be reliably detected under the experimental conditions, which involve immobilised (rather than freely diffusing) proteins. For this purpose, a biotinylated KCC2 peptide [Q9H2X9, 1-MSRRFTVTSLPPAGPA-17 a.a.] was immobilised on the streptavidin biosensor, while the C-terminal domain of SPAK kinase [Q9UEW8, 448-547 a.a.] was used as the analyte.

The biotinylated peptides TSC22D2, TSC22D4 and KCC2 were immobilised on super-streptavidin sensors which were dipped in wells containing the protein analytes. A summary of immobilised

[1] 146- PVTAAPSQPPTTCSS**RFRV**IKLDHGSGEPYRRGRW-180
[2] 56-GSPPPGAPSS**RFRV**VKLPHGLGEPYRRGRW-90
[3] Potassium-chloride transporter member 5
[4] SPS1-related proline/alanine-rich kinase

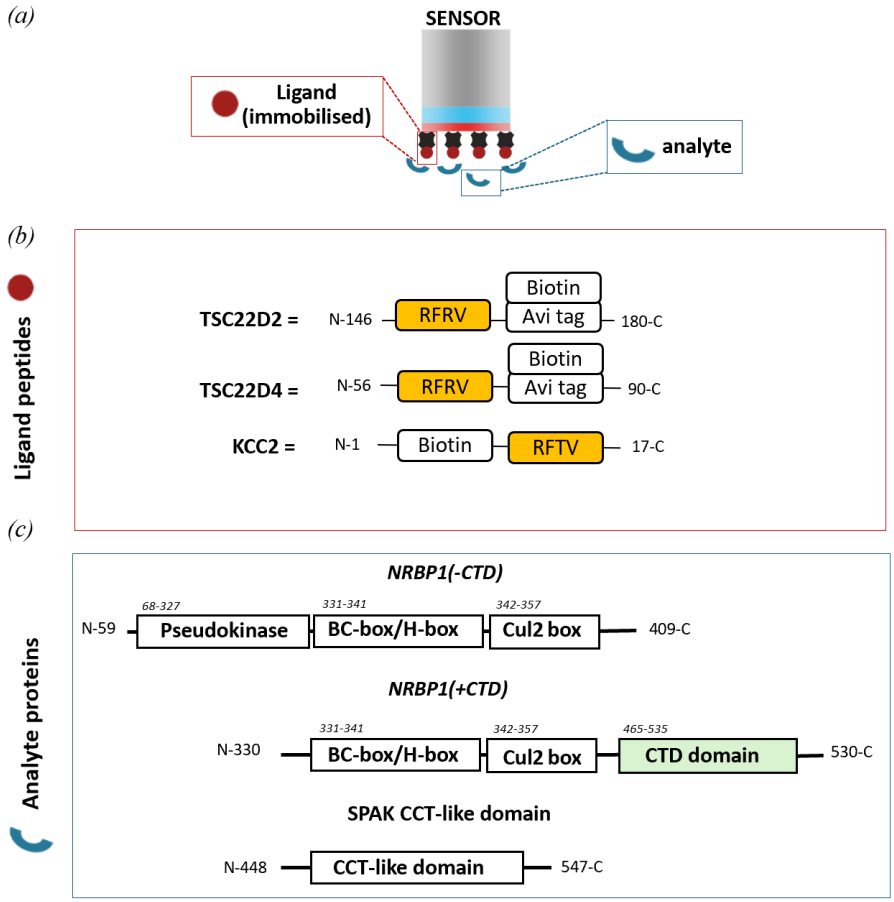


Figure 7.13: **BLI experimental setup to probe NRBP1-TSC22D interaction**. (a) Schematic representation of the biolayer interferometry (BLI) experimental setup, where biotinylated ligand peptides (red) are immobilized on a streptavidin-coated sensor (black), allowing interaction with analyte proteins in solution (blue). (b) Sequences and constructs of ligand peptides used in BLI assays, including TSC22D2, TSC22D4, and KCC2, highlighting the presence of conserved RFXV motifs and biotinylation tags. (c) Domain architecture of analyte proteins tested for interaction, including NRBP1 variants (CTD-truncated and full-length) and the CCT-like domain of SPAK, with key structural features such as pseudokinase, BC-box/H-box, Cul2 box, and CTD domain indicated.

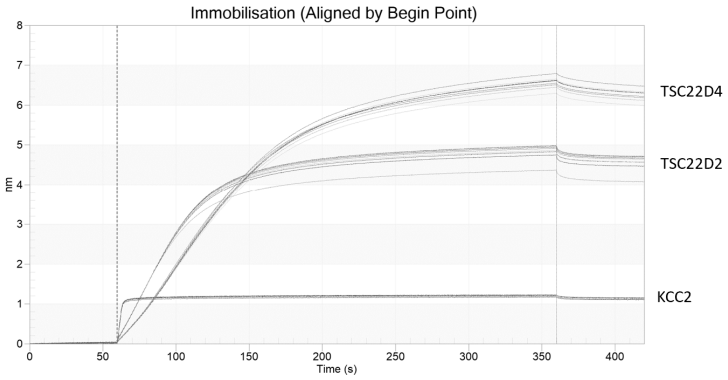


Figure 7.14: **Immobilisation of RFXV-containing biotinylated peptides on super-streptavidin BLI Sensors.** Peptides TSC22DD2 [146-180 a.a.] and TSC22D4 [56-90 a.a.] were immobilised on SSA biosensors to test their predicted interaction with NRBP1's putative CCT-like domain. Peptide KCC2 [1-17 a.a.] bearing an RFTV motif was also immobilised as a positive control for SPAK.

peptides containing RFXV motifs and protein constructs with CCT-like domains can be found in Fig. 7.13. Contrary to expectations, *NRBP1(+CTD)* in concentrations up to 90 μM did not show binding to any RFXV motif, while *NRBP1(-CTD)* showed concentration-dependent affinity for TSC22D4 [Fig. 7.15]. Global fit analysis, assuming one-site binding, measured the TSC22D4 affinity to K_d = 10 ± 2.7 μM [Fig. 7.16 (a)]. The measured K_d value should be interpreted with caution, as the binding curve did not reach saturation, potentially leading to an underestimation or overestimation of the actual affinity. Of note, the difference in magnitude of response (nm) between TSC22D peptides and the positive control KCC2 peptide reflected the lower immobilisation levels of KCC2 [Fig. 7.14].

Taken together, the data suggested that the "RFRV" motif contributed to the interaction with NRBP1 but not through the C-terminal domain as originally hypothesised. In addition, since no NRBP1 constructs had affinity for the positive control KCC2, the "RFXV" motif is not the sole mediator of NRBP1-TSC22D4 association. Finally, since the CCT-like domain of SPAK recognised and showed measurable binding to the RFRV motif of TSC22D2/TSC22D4, those motifs could be recognised by proteins bearing CCT-like motifs, like WNK kinases.

The interaction between NRBP1 and the RFRV motifs of TSC22D2/TSC22D4 was further investigated through a pull-down experiment using purified proteins. GST-tagged TSC22D2 and TSC22D4 peptides, as described above, were purified and immobilized on GST-Sepharose resin. The resin was then incubated with purified NRBP1-ElonginBC [59-409 a.a.] at 4°C with end-over-end mixing for 2 hours. After washing, GST-tagged proteins were eluted using a buffer containing 20 mM reduced glutathione. In contrast to the BLI finding, no interaction between NRBP1 and either TSC22D2 or TSC22D4 was observed [Fig. 7.17]. This lack of binding may result from steric hindrance imposed by the large GST tag, in comparison to the smaller disordered peptide, or from interactions with the resin that interfere with binding. The weak nature of the interaction (K_d > 10 μM) may be also lost due to dissociation during washing steps.

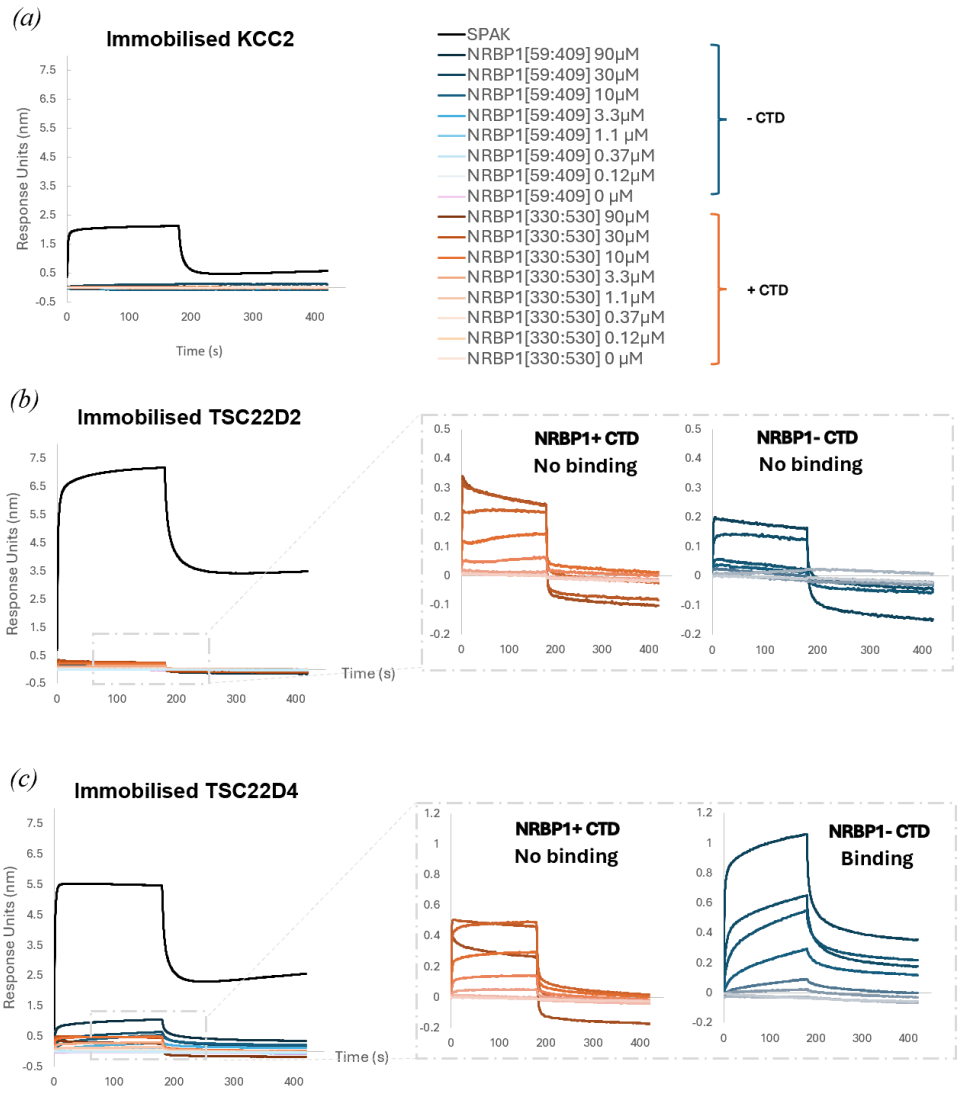
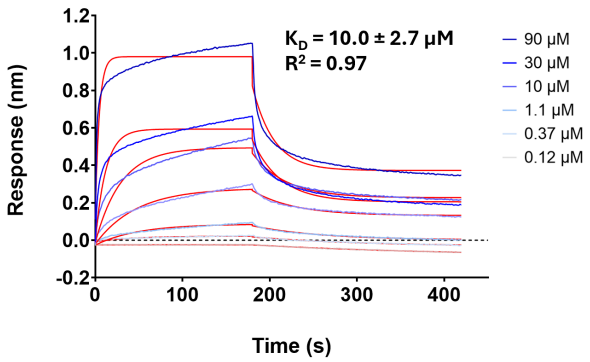


Figure 7.15: **NRBP1 binds to the TSC22D4 RFTV motif but not to TSC22D2**. BLI experiments where used to test the hypothesis that NRBP1 the putative CCT-like domain of NRBP1 binds to TSC22DF via their RFXV motif. Response Units (nm) plotted over time (s). (a) Results of analyte binding to immobilised KCC2 peptide [1-17 a.a.]. Only SPAK kinase binding was detected (positive control), (b) Results of analyte binding to immobilised TSC22D2 peptide [146-180 a.a.]. SPAK, but not NRBP1, binding was detected (dashed boxed, shown right, denotes close up view of the NRBP1 samples). (c) Results of analyte binding to immobilised TSC22D4 [56-90 a.a.]. *NRBP1(-CTD)* was bound to TSC22D4, but not TSC22D2. SPAK bound both TSC22D RFRV motifs, without fully disassociating when washed. SPAK is coloured black, *NRBP1(-CTD)* [59-409 a.a.] blue and *NRBP1(+CTD)* [330-530 a.a.] orange. Colour intensity increases with protein concentration [0-90 μM]. Data obtained from BLI Octet were reference-subtracted, and Savitzky-Golay filtered.

(a)



(b)

Conc. (µM)	Response	K_D (µM)	k_{on} (1/µM·s)	K_{dis} (1/s)	Assoc R^2
90	0.970	20.9	1960	0.041	0.704
30	0.629	18.1	1720	0.0311	0.769
10	0.528	33.8	776.0	0.0263	0.887
1.11	0.097	1.8	7120	0.0128	0.909
0.37	0.027	0.203	38,400	0.0078	0.847
0.12	0	-	-	-	-
0.04	0	-	-	-	-

Figure 7.16: **RFXV-containing biotinylated peptides are immobilised on SSA BLI Sensors**. Peptides TSC22DD2 [146-180 a.a.] and TSC22D4 [56-90 a.a.] are immobilised on SSA biosensors to test hypothesised interaction with NRBP1 putative CCT-like domain. Peptide KCC2 [1-17 a.a.] bearing an RFTV motif was also immobilised as a positive control. (a) Binding curve derived from BLI (Biolayer Interferometry) data, showing the sensor response (nm) over time (s). The experimental response of different concentrations of *NRBP1(-CTD)* is coloured in blue. Colour intensity increases with protein concentration [0-90 μM]. The red curve represents model fit. The dissociation constant K_D was determined by fitting the response to a 1:1 Langmuir binding model $R = R_{\text{max}} \cdot \frac{\text{Conc.}}{K_D + \text{Conc.}}$. The data fit well to the model, with an R^2 of 0.965, indicating a reliable fit. The estimated global K_D value is 10.0 μM, with a standard error of ± 2.7 μM. (b) Kinetic parameters from BLI binding experiment. The table presents association (k_{on}), dissociation (k_{dis}), and dissociation constant (K_D) values at different concentrations of *NRBP1(-CTD)* in μM. The association R^2 values indicate the goodness of fit for the binding model. The data indicate a non-saturating binding response, and should be interpreted with caution.

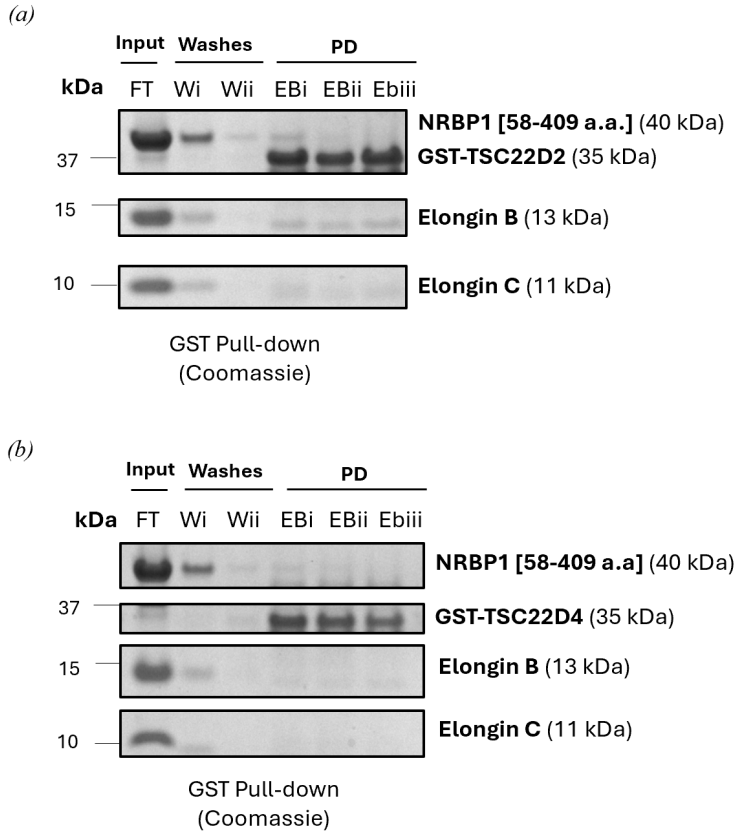


Figure 7.17: **NRBP1-TSC22DF pull-down validation experiment** Purified GST-tagged TSC22DF peptides were immobilised on GST-Sepharose resin. The resin was then incubated with purified NRBP1-ElonginBC [59-409 a.a.]. After washing, GST-tagged proteins were eluted using a buffer containing 20 mM reduced glutathione. Pull-down samples were analysed on SDS-PAGE. (a) Immobilised TSC22D2 [146-180 a.a.], and (b) immobilised TSC22D4 [56-90 a.a.]. FT: flow-through, W: wash buffer, EB: elution buffer. No binding was detected.

7.2.4.4 Modeling studies of complex assembly

Due to time constraints, a full experimental characterization of this complex assembly was not feasible. Instead, structural modeling was performed using AlphaFold v3, providing computationally predicted structures that may guide future experimental validation. Initially, the experimentally validated interaction of NRBP1 with ElonginBC was modeled. AlphaFold v3 modeling predicted the canonical interaction between ElonginBC and the NRBP1 EloBC box [331-341 a.a.] as expected, along with a second, potentially novel interaction between ElonginC [40-51 a.a.] and the antiparallel β-strands of NRBP1 [362-382 a.a.] [Fig. 7.18 (b)]. This bipartite binding is not observed in the canonical VHL-ElonginBC interaction [Fig. 7.18 (b)] and may influence the structural basis of CUL2 recruitment. Notably, this β-strand region could introduce steric clashes with CUL2, potentially necessitating the involvement of the co-adaptor TSC22D2 to stabilize the complex and establish additional contacts with CUL2.

To test this hypothesis, the NRBP1-ElonginBC-TSC22D4 complex was modeled. The resulting prediction suggested that TSC22D4 dimerises via its coiled-coil domain, forming a leucine zipper that acts as a scaffold to facilitate NRBP1-ElonginBC protomer dimerisation [Fig. 7.19 (a)]. Notably, the predicted interaction between TSC22D4 and NRBP1 was bipartite [Fig. 7.19 (c)]. The CCT-like C-terminal domain of NRBP1 (465-535 a.a.) interacted with an "RwTC" motif located within an antiparallel β-strand of TSC22D4 (85-97 a.a.), rather than the hypothesized "RFRV" motif. Additionally, the downstream "RFRV" motif of TSC22D4 was predicted to engage with the N-terminal lobe of NRBP1's pseudokinase domain, specifically within antiparallel strands spanning NRBP1 residues 55-95 [Fig. 7.19 (c)]. This prediction aligns with the biolayer interferometry (BLI) data presented in this study [Fig. 7.15], which indicated that the "RFRV"-containing TSC22D4 residues 56-90 contributed to NRBP1 binding via interaction with NRBP1 residues 59-409. The "RWTC" motif, located in residues 90-97, fell outside the tested residue range 56-90 in BLI and pull-down assays, potentially explaining why this interaction was not observed.

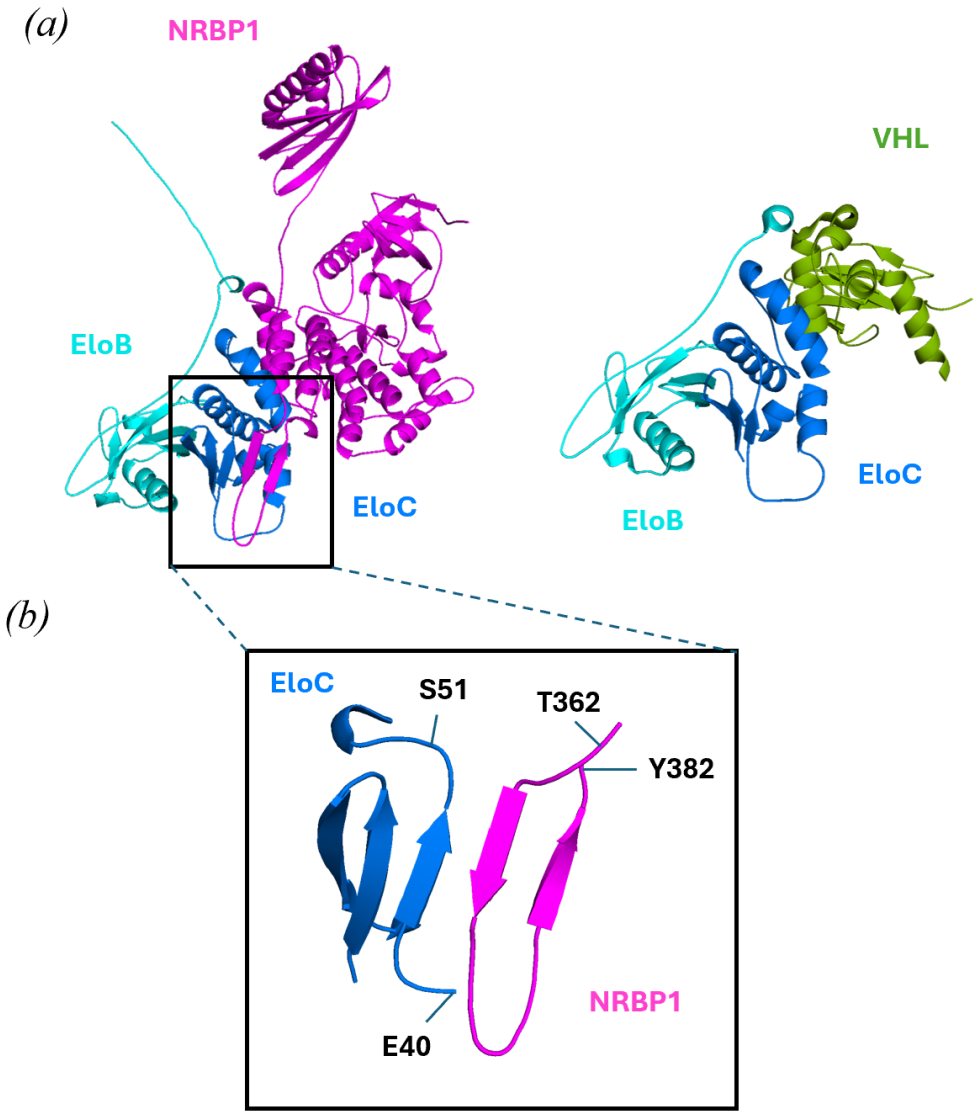


Figure 7.18: **NRBP1 associates with ElonginBC by bipartite binding** (a) *Left*: NRBP1-ElonginBC model by AlphaFold v3, *Right*: structure of VHL-ElonginBC, PDB = 4WQO [293]. (b) AlphaFoldv3 model of NRBP1-ElonginBC model predicts an additional interaction not observed in canonical VHL binding. Specifically, NRBP1 362-382 a.a. pack against Elongin C 40-51 a.a., forming antiparallel β-strands.

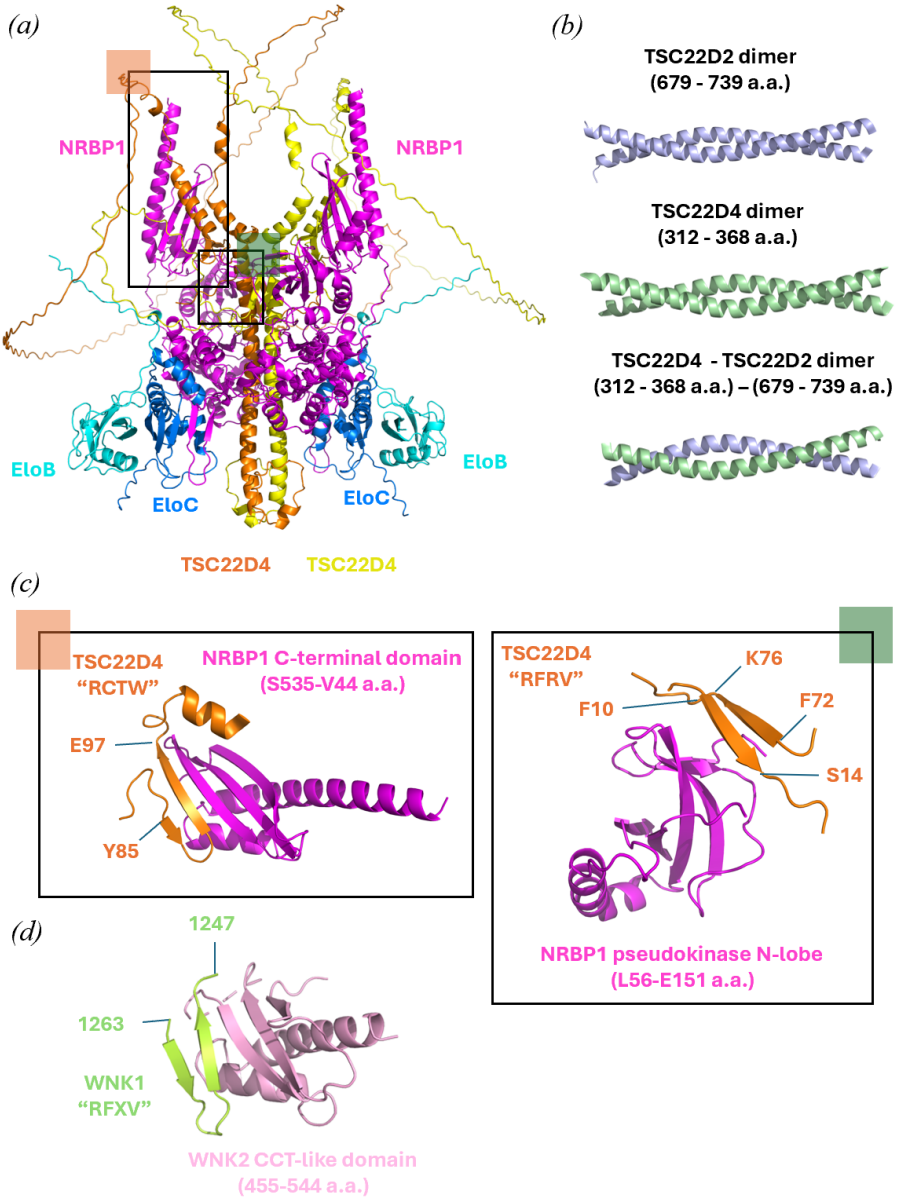


Figure 7.19: **Putative model of TSC22DF mediated NRBP1 dimerisation** (a) AlphaFold model of the NRBP1-ElonginBC-TSC22D4 assembly in 2:2:2:2 stoichiometry. TSC22D4 mediates NRBP1 dimerization through homodimeric leucine zipper formation. (b) AlphaFold models of TSC22DF family members forming a leucine zipper via their central coiled-coil region. (c) Close-up view of the bipartite TSC22D4-NRBP1 interaction. The TSC22D4 "RWTC" motif is predicted to interact with the CCT-like domain of NRBP1, while the "RFRV" motif is predicted to engage the NRBP1 pseudokinase domain. (d) Crystal structure of the human WNK2 CCT-like 1 domain in complex with a WNK1 1257-RFIV-1260 peptide, PDB = 6FBK.

To test the initial hypothesis, whether the β-strand region of NRBP1 introduces steric clashes with CUL2, potentially necessitating the involvement of the co-adaptor TSC22D2 for complex stabilization and additional CUL2 contacts, the next step was to incorporate the N-terminal region of CUL2 into the TSC22D complex. Interestingly, the predicted structure suggests that TSC22D4 indeed adopts a scaffolding role, facilitating CUL2 binding while preserving the canonical CUL2-ElonginBC interaction [Fig. 7.20]. This is only a putative model, however it would explain the inability to observe CUL2 recruitment by an NRBP1-EloBC complex.

Similar attempts to model the NRBP1-DDB1 interaction, with or without TSC22D4, using AlphaFold v3 did not yield meaningful results. However, previous evidence in this work suggested that NRBP1, in the absence of ElonginBC or TSC22D4, could still associate with DDB1 [Fig. 7.12]. To explore this possibility, a putative model was generated by superimposing the predicted H-box of NRBP1 - α-helix spanning residues 331-341 - onto the known H-box of DNA damage-binding protein 2 (DDB2) [Fig. 7.21]. The model revealed that the conserved hydrophobic φ residues of the H-box motif align well at the interface between the NRBP1 BC-box and DDB1, closely resembling the canonical DDB2-DDB1 interaction. This finding suggests that NRBP1-DDB1 binding might be feasible even in the absence of the scaffolding effect of TSC22D4.

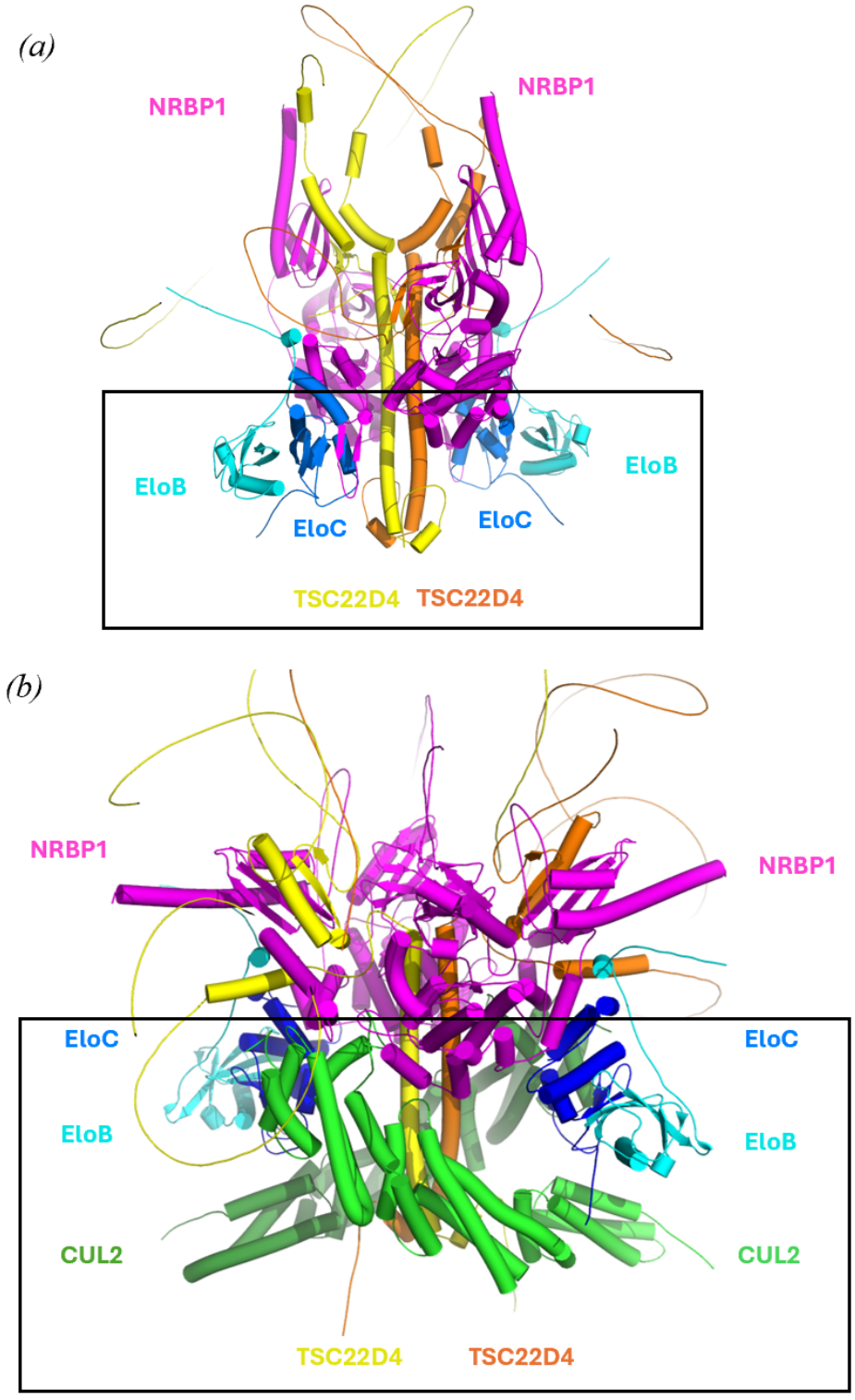


Figure 7.20: **TSC22D4 may act as a scaffold to facilitate CUL2 binding with NRBP1**. AlphaFold v3 structural modeling of the NRBP1-ElonginBC-TSC22D4 complex, with and without CUL2. (a) The predicted NRBP1-ElonginBC-TSC22D4 complex shows TSC22D4 (yellow and orange) as a potential dimeric scaffold bridging NRBP1 (magenta) and ElonginBC (cyan and blue). (b) The addition of CUL2 (green) suggests that TSC22D4 may facilitate CUL2 recruitment by stabilizing the interaction interface. While these models provide structural hypotheses, experimental validation is required to confirm the role of TSC22D4 as a scaffold in CUL2 binding.

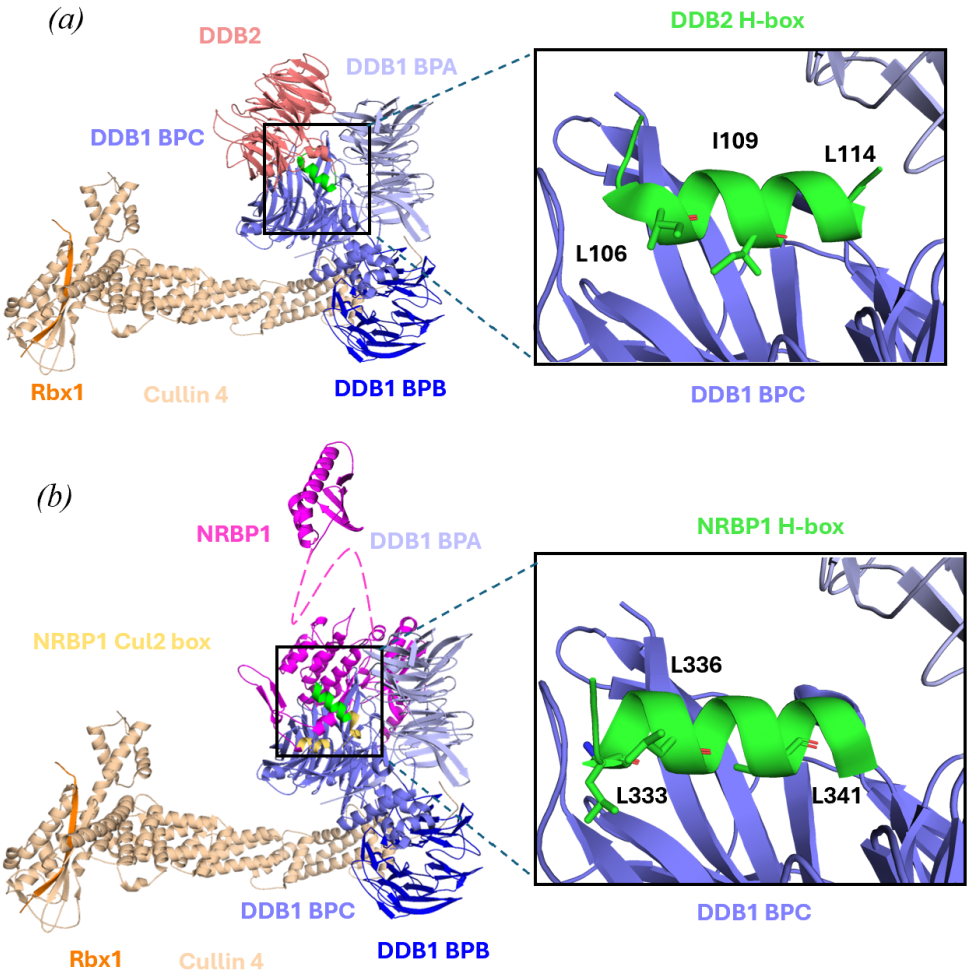


Figure 7.21: **Putative model of CUL4^{NRBP1}** (a) Crystal structure of DDB1-DDB2-CUL4A-Rbx1 assembly PDB = 4A0K [294]. (b) Putative model of NRBP1-CUL4A-Rbx1 generated by superposing the predicted H-box of NRBP1 [331-341 a.a.] onto the H-box of DDB2. The NRBP1 structure was modeled using AlphaFold v3. The model predicts the region between the C-terminal and pseudokinase domains of NRBP1 to be disordered, represented as a dashed line. The H-boxes of both DDB2 and NRBP1 are highlighted in green. The conserved φ hydrophobic residues that are key residues for DDB1 interaction are shown as sticks in both structures.

7.3 Discussion

In this exploratory chapter, I aimed to elucidate a proposed unique mechanism of assembly for a mixed CUL2/CUL4^{NRBP1} E3 ligase complex. The interest laid not only in the unusual assembly, which challenges the established paradigm that a single substrate adaptor associates with only one type of cullin, but also in the opportunity to solve its structure and identify potential ligandable sites on the pseudokinase fold. This chapter focused on generating protein reagents through high-throughput cloning and attempting to reconstitute these complexes *in vitro* for biophysical and structural studies.

Investigating the extensive reported interactome of NRBP1 proved to be a multi-parametric challenge. Association with ElonginBC remained stable and consistent across multiple NRBP1 constructs. Since both NRBP1 [330-441 a.a.] and NRBP1 [59-388 a.a.] retained ElonginBC binding [**Chapter. 7.2.1**], the minimal recruitment region can be narrowed to residues 330-388. This aligns with the proposed ElonginBC box [331-340 a.a.] identified through sequence alignment of ElonginBC-associated proteins [120, 126]. Despite ElonginBC association and the presence of a putative CUL2 interaction site, NRBP1 did not co-purify with CUL2 in a pull-down assay [Fig. 7.11]. This observation is consistent with affinity mass spectrometry studies by Mahrour *et al.* [120] and Yasukawa *et al.* [126], which reported CUL2 co-immunoprecipitation only upon TSC22D overexpression. On the other hand, DDB1 association could be observed after NRBP1-DDB1 coexpression in Sf9 cells in the absence of ElonginBC [Fig. 7.12]. The interaction was abolished upon ElonginBC binding, supporting further the argument that CUL2 or CUL4 recruitment are mutually exclusive and mediated through overlapping ElonginBC/H-boxes in NRBP1 residues 331-341 [126]. Predictive modeling using AlphaFold v3 revealed a potential mechanism for NRBP1's ability to recruit DDB1 [Fig. 7.21]. In contrast, CUL2 binding may require the scaffolding role of TSC22D4 to mediate interaction [Fig. 7.20], possibly due to an unusual bipartite binding of EloC, which might disrupt the association of the NRBP1 CUL2 box with CUL2.

It is widely reported that NRBP1 homodimerises [118, 126], similarly to the CUL4A substrate recep-

tor DCAF1 [85], via a LisH motif in its C-terminal domain [118, 126]. However, both size exclusion chromatography [Fig. 7.5] and native mass spectrometry [Fig. 7.8] detected no NRBP1-ElonginBC protomer dimerisation, regardless of C-terminal domain presence. Given that transcription factors TSC22D3 and TSC22D4 have been shown to enhance NRBP1 dimerisation by binding its C-terminal domain [126], these results suggest that their presence is not merely an enhancer but a prerequisite for NRBP1 dimerisation *in vitro*.

Given the structural homology between the NRBP1 pseudokinase domain and WNK atypical kinases - which recognize RΦXV motifs (where Φ denotes a hydrophobic amino acid) via a C-terminal CCT-like domain [289, 295] - it was hypothesized that the NRBP1 C-terminal domain might similarly bind TSC22DF leucine zippers through their RFRV motifs in disordered regions. It should be noted that despite the high structural homology, the putative CCT-like domain of NRBP1 and of WNK kinases share low sequence identity. To investigate TSC22DF binding, BLI affinity measurements revealed weak direct NRBP1 binding (10 μM after kinetic global fit) to TSC22D4 "RFRV"-containing peptide [56-90 a.a.], whereas no interaction was observed with TSC22D2 "RFRV"-containing peptide [146-180 a.a.] [Fig. 7.15]. Surprisingly, TSC22D4 association did not involve the NRBP1 CCT-like domain, but upstream NRBP1 regions, such as the kinase fold or the adjacent post-kinase domain, instead.

A recent study published at the time of writing this dissertation confirmed, through mutational studies and co-immunoprecipitation in cells, that TSC22D4 associates with the CCT-like domain of kinase WNK1 via its 68-RFRV-71 motif and with the CCT-like domain of NRBP1 via its 86-RWTC-89 motif [296]. These findings corroborate the NRBP1-TSC22D4 computational model described above [Fig. 7.19]. The same study reports an equivalent binding mode for TSC22D2, where residues 158-RFRV-161 interact with the CCT-like domain of WNK1, while residues 176-RWTC-179 bind NRBP1's CCT-like domain. Additionally, the study demonstrated that WNK1 monophosphorylates the T-loop of NRBP1's pseudokinase domain at Thr-232. This observation is consistent with the findings of this

study, which detected monophosphorylated NRBP1 when expressed in Sf9 cells [Fig. 7.2].

It can thus be hypothesized that TSC22D family members function as molecular scaffolds, either facilitating the proximity of WNK1 and NRBP1 for NRBP1 phosphorylation [296] or, promoting NRBP1 dimerisation to assemble the putative mixed CUL2/CUL4 E3 ligase complex [126]. Given that multiple TSC22D family members have been reported to interact with NRBP1 [126, 297] and possess the capacity to form both homo- and heterodimers via leucine zipper motifs as modeled by AlphaFold [Fig. 7.19 (b)], it is plausible that any homo- or hetero-dimer can fulfill a similar scaffolding role to the one modeled for TSC22D4.

The role of substrate receptor dimerisation remains unclear. It may function similarly to KLHL12, where co-factor concentration modulates substrate ubiquitination [**Chapter 3.3**]. Alternatively, it could allow for different cullins to catalyse proximal and distal ubiquitination, or it may be required for high-avidity binding or optimal spatial presentation of suboptimal degrons [298, 299]. The NRBP1 homodimer may simultaneously recognise two distinct substrate degrons, similar to the hinge and latch mechanism of KEAP1, where the ligase ubiquitylates the substrate only when both degrons are engaged [300]. This hypothesis is further supported by findings that two distinct regions of the substrate BRI3 [121-140 a.a. and 191-210 a.a.] are important for its recruitment [126].

In this chapter multiple biochemical and biophysical approaches were used to probe NRBP1 interactions, and even though reconstitution of CUL2/CUL4^{NRBP1} *in vitro* proved challenging, significant progress was made in exploring this assembly with recombinant proteins for the first time. Despite cryo-EM not yielding a successful volume reconstruction due to poor alignment of low signal-to-noise particles, the protein demonstrated stability and was amenable to the technique. To assess whether multivalent interactions among binding partners are required to achieve stoichiometric assemblies (e.g. NRBP1-DDB1, NRBP1-TSC22D), future work could employ co-expression of multiple full-length partners. This strategy may facilitate the formation of higher molecular weight complexes amenable to cryo-EM analysis. The recombinant protein material generated in this study provides

a valuable foundation for future structural and mechanistic investigations. In particular, it opens the door to exploring how this novel E3 ligase employs a kinase fold in the context of ubiquitination as an exciting new avenue for understanding its function and regulation.

CHAPTER 8

Conclusions and Future Directions

The ubiquitin-proteasome system maintains protein homeostasis and regulates the half-life of most soluble proteins through precise spatiotemporal control. Given its ability to accommodate much of the human proteome as substrates, this system requires sophisticated regulation. Current research has only begun to uncover its mechanisms, which include the extraordinary modularity of multi-protein complexes, higher-order assemblies for targeting complex substrates, and post-translational modifications such as phosphorylation and neddylation that modulate E3 ligase assembly, activity, and substrate recognition. E3 ligases, positioned at the substrate-facing end of these complexes, determine substrate specificity through direct protein binding. Despite the critical role of E3 ligases in substrate selectivity, their diverse functions and assembly mechanisms remain poorly understood. The human genome encodes more than 700 E3 ligases, many of which lack chemical tools to interrogate their function. This work tried to address this gap by developing structures, assays, and selective chemical probes to better understand and modulate previously unexplored $CUL3^{KLHL12}$ and $CUL2/CUL4^{NRBP1}$ that are implicated in neurodegenerative disease pathology, different types of cancer and developmental disorders.

KLHL12, a substrate-specific adaptor of E3 ubiquitin ligases, engages target-specific co-adaptors through its N-terminal BTB domain to ubiquitinylate different substrates bound to its C-terminal Kelch domain. Substrate, co-factor and antagonist binding to KLHL12 had been mapped to a "PGXPP" motif, necessary for both recruitment and degradation. Since that posited interesting questions on

ligase assembly, there was an incentive to identify the binding mode of all these interactors and characterise the Kelch domain binding site. The 1.9 Å crystal structure of KLHL12 Kelch domain in complex with the antagonist peptide PLEKHA4, compared with co-factor PEF1 and substrate DVL1 co-crystals revealed a conserved binding mode between all known interactors defining the binding site interactions [**Chapter 3**)].

PLEKHA4 binding to the same site as the physiological substrate DVL1 suggests a competitive mode of antagonism for substrate recognition by the Kelch domain of the $CRL3^{KLHL12}$ E3 ligase. The C-terminal Kelch domain of KLHL12 exhibits the highest binding affinity for DVL1 among all tested interactors, whereas PLEKHA4 binds with approximately half the affinity of the substrate. Consequently, PLEKHA4 acts as a weaker competitive inhibitor, only effectively outcompeting DVL1 when present at a significantly higher concentration than the substrate. Notably, PLEKHA4 oligomerizes into clusters at the membrane, sequestering KLHL12 from the cytoplasmic pool of DVL1. Thus, this sequestration could not only spatially exclude KLHL12 from its substrate but also could partially compensate for PLEKHA4 weaker binding affinity, by increasing its local concentration.

Competitive inhibition of E3 ligases as a regulatory mechanism within the ubiquitin-proteasome system has been previously documented. An example of such a regulatory mechanism is β-Transducin repeats-containing protein (β-TrCP), the substrate recognition subunit of the SCF (SKP1, CUL1, and F-box protein)-type E3 ubiquitin ligase complex. β-TrCP recognizes phosphorylated substrates, promoting their subsequent proteasomal degradation. The physiological inhibitor, Nuclear factor NFκB p105 subunit [301], competes with substrates such as IκBα and IκBβ [302], as well as the DEP domain-containing mTOR-interacting protein (DEPTOR) [303], for binding at the substrate recognition site of $SCF^{\beta\text{-}TrCP}$. This competition limits the degradation of these inhibitors and affects NF-κB and mTOR signaling pathways.

Similarly, multiple pathways regulate the recognition of NF-E2-related factor 2 (NRF2) by the $CRL3^{KEAP1}$ E3 ligase, including competitive inhibition by intracellular proteins such as P62/Sqstm1 [304], PGAM5

[305], and CIP/WFAI [306]. Analogous to how PLEKHA4 clusters sequester KLHL12 at the membrane, P62 has been shown to recruit KEAP1 to autophagosomes, thereby preventing its ability to ubiquitinate NRF2 [304]. These findings suggest a broader regulatory mechanism in which competitive inhibitors alter the subcellular localization of E3 ligases and their substrates, effectively preventing their interaction and modulating ubiquitination-dependent degradation.

Furthermore, the observation that the binding site is shared not only between substrates and the competitive inhibitor PLEKHA4 but also with co-factors such as PEF1 has intriguing implications for the putative assembly of $CUL3^{KLHL12}$-substrate complexes. The reported simultaneous binding of both the PEF1 co-factor and the SEC31 substrate [95] would necessitate a homodimeric $CUL3^{KLHL12}$ E3 ligase, in which KLHL12 substrate receptors dimerize through their BTB domains [Fig. 1.7].

This mechanism may enable fine-tuning of KLHL12 E3 ligase activity through the relative abundance of its binding partners. A similar regulatory paradigm has been observed in the Keap1[1]-Nrf2[2] system, where E3 ligase activity is modulated by the fluctuating levels of a substrate/co-adaptor pair [84]. Future cryo-electron microscopy (cryoEM) studies of the reconstituted dimeric $CUL3^{KLHL12}$ E3 ligase, in complex with the PEF1 co-adaptor and SEC31 substrate, could provide critical structural insights into this assembly and clarify why PEF1 binding is essential for SEC31 recognition by KLHL12.

The cognate "PGXPP"-containing peptides facilitated competitive displacement assays, enabling the screening of in-house fragment collections for molecules that could target the KLHL12 Kelch domain binding site. Non-covalent fragments were also further validated with orthogonal biophysical assays, and covalent ones with mass spectrometry to assess adduct formation efficiency. One fragment, **EVD-1**, was identified as a selective covalent modifier of KLHL12, over other Kelch family members [**Chapter 4**]. Co-crystallisation of **EVD-1** with the Kelch domain of KLHL12 yielded a 2.6 Å co-crystal structure of mono-conjugated KLHL12, which revealed that **EVD-1** selectively bound to C511 as a thioimidate in the peptide-binding site, with no evidence of adduct formation at the neighboring

[1] Kelch ECH-associating protein 1
[2] nuclear factor erythroid 2-related factor 2

cysteine C558 within the same pocket.

Further exploration of the structure-activity relationships of this chemotype using a combination of high-throughput docking and ligand-based pharmacophore screening of commercial libraries to perform SAR by catalogue, identified key pharmacophoric features of **EVD-1**. The importance of electronic configuration of the aromatic ring bearing the reactive nitrile to fine-tune reactivity to selectivity ratios of these reactive covalent warheads was highlighted. **EVD-1** showed remarkable stability and low chemical reactivity in a glutathione assay, exhibiting 55 hours half-life, paired with high selectivity for the KLHL12 in a selectivity assay tested against other Kelch domains of E3 ligases with accessible cysteines. Low intrinsic reactivity combined with the sub-micromolar apparent IC_{50} = 0.46 μM after 3-hour incubation time suggests a potent compound whose affinity for KLHL12 is not driven my chemical reactivity alone [**Chapter 5**]. Given that highly potent compounds are typically dosed at low cellular concentrations, polypharmacology effects are unlikely if **EVD-1** or its future derivatives are dosed appropriately as probes.

This study marked only the beginning of covalent probe development for KLHL12. **EVD-1** exhibited several characteristics of a high-quality covalent probe, including chemical stability, low intrinsic reactivity, a well-defined binding site, and inhibition of the native protein binder. It also demonstrated great selectivity over other Kelch domain-containing E3 ligases of the same family, and an available inactive negative control was identified (fragments **17**, **78**, and **30**) by covalent SAR studies.

However, further studies would be desirable to optimise and refine the scaffold core into a fully developed high potency small-molecule probe capable of effectively interrogating KLHL12 function.

First, although IC_{50} measurements were a sufficient approach to compare activity at a specific incubation time, a more comprehensive characterisation should include determination of K_{inact}/K_i values. This would allow for quantification of both the binding affinity (K_i) and the maximum rate of KLHL12 inactivation (K_{inact}) [308]. Second, while **EVD-1** demonstrated high selectivity for KLHL12 over other family members tested in our selectivity assay, unbiased proteomic studies are necessary

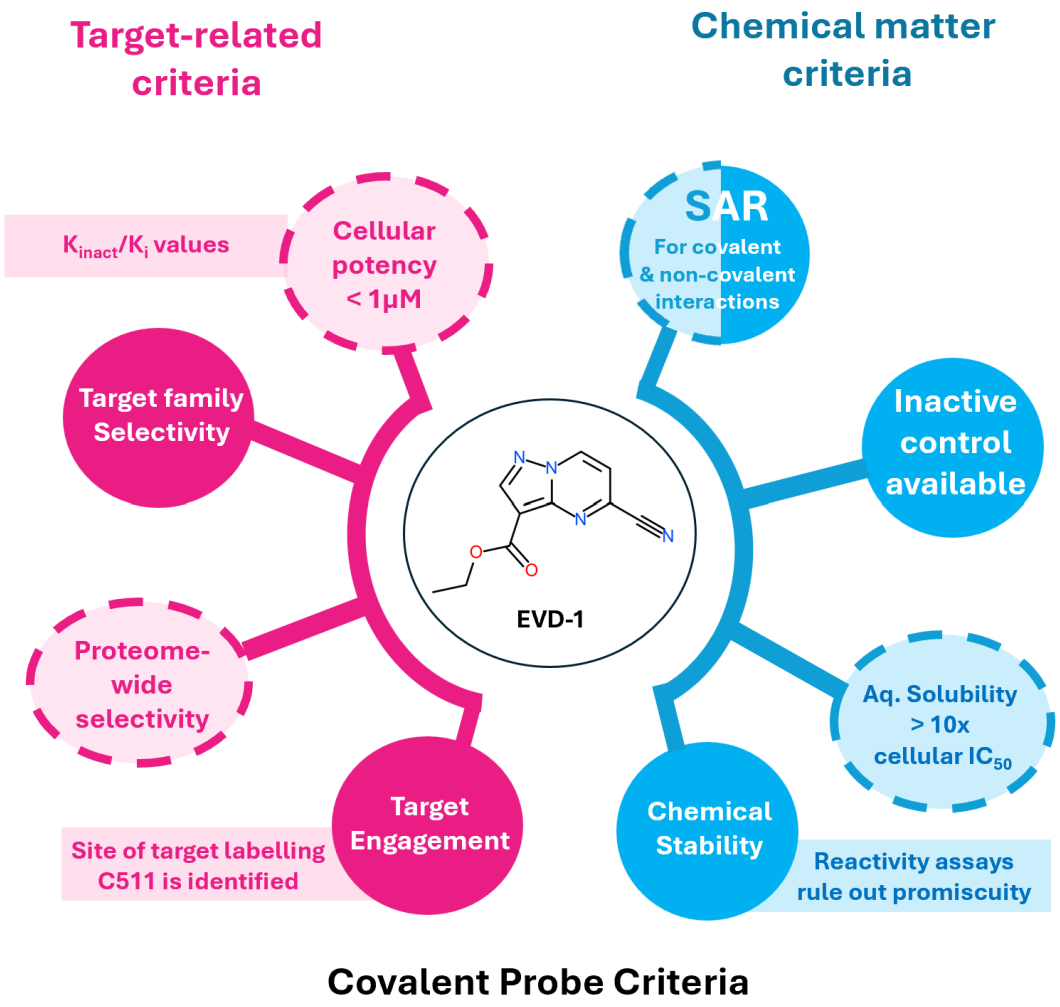


Figure 8.1: **EVD-1 meets many covalent chemical probe criteria.** Covalent probe criteria that **EVD-1** already meets as demonstrated are circled with solid colour line. Those that need further study are shaded in dashed colour line. Criteria outlined in study from Hartung *et al.* [307].

to confirm whether this selectivity extends to the entire proteome. Finally, in-cell target engagement studies should be conducted to verify that an **EVD-1 derivative** is cell-permeable and capable of engaging KLHL12 in a cellular context. A summary of the characterisation criteria for covalent inhibitors, as proposed by Hartung *et al.* [307], highlighting the already fulfilled aspects and those requiring further study, is presented in Fig. 8.1.

Importantly, a tractable synthetic route has been proposed in this study for accessing extended analogues of **EVD-1**, with the potential to enhance selectivity and potency by engaging a currently unexploited pocket area in KLHL12. Future studies should focus on overcoming the low solubility limiting the last cyanation step by heteroatom containing thioether and ether extensions, ultimately yielding a library of extended analogues ready for testing.

An alternative approach would be to design extended compounds incorporating two covalent warheads to engage both KLHL12 binding-site cysteines C511 and C558, enabling a bipartite binding mode as proposed in Fig. 8.2. Although crystallization experiments provided no evidence of C558 targeting, it is not uncommon for binding to occur sequentially in a hinge-like mechanism, where the more reactive C511 engages the **EVD-1** covalent warhead first, followed by slow reactivity with C558 driven by an avidity effect. These structural extensions could yield lead compounds with higher molecular weight and potentially greater selectivity, as the introduced steric clashes would restrict off-target binding by imposing additional spatial constraints on the targeted pockets.

Beyond developing a selective KLHL12 probe to investigate its role in collagen export, α-synuclein degradation, and osteogenesis, **EVD-1** could also serve as a molecular handle to harness KLHL12 as a novel E3 ligase for targeted protein degradation. Crystal structures of both **EVD-1** and the more promiscuous fragment **23** revealed synthetically accessible exit vectors suitable for linker functionalisation, enabling the development of proximity-inducing modalities such as proteolysis-targeting chimeras (PROTACs).

PROTAC development has largely focused on E3 ligases with well-characterised ligands, namely

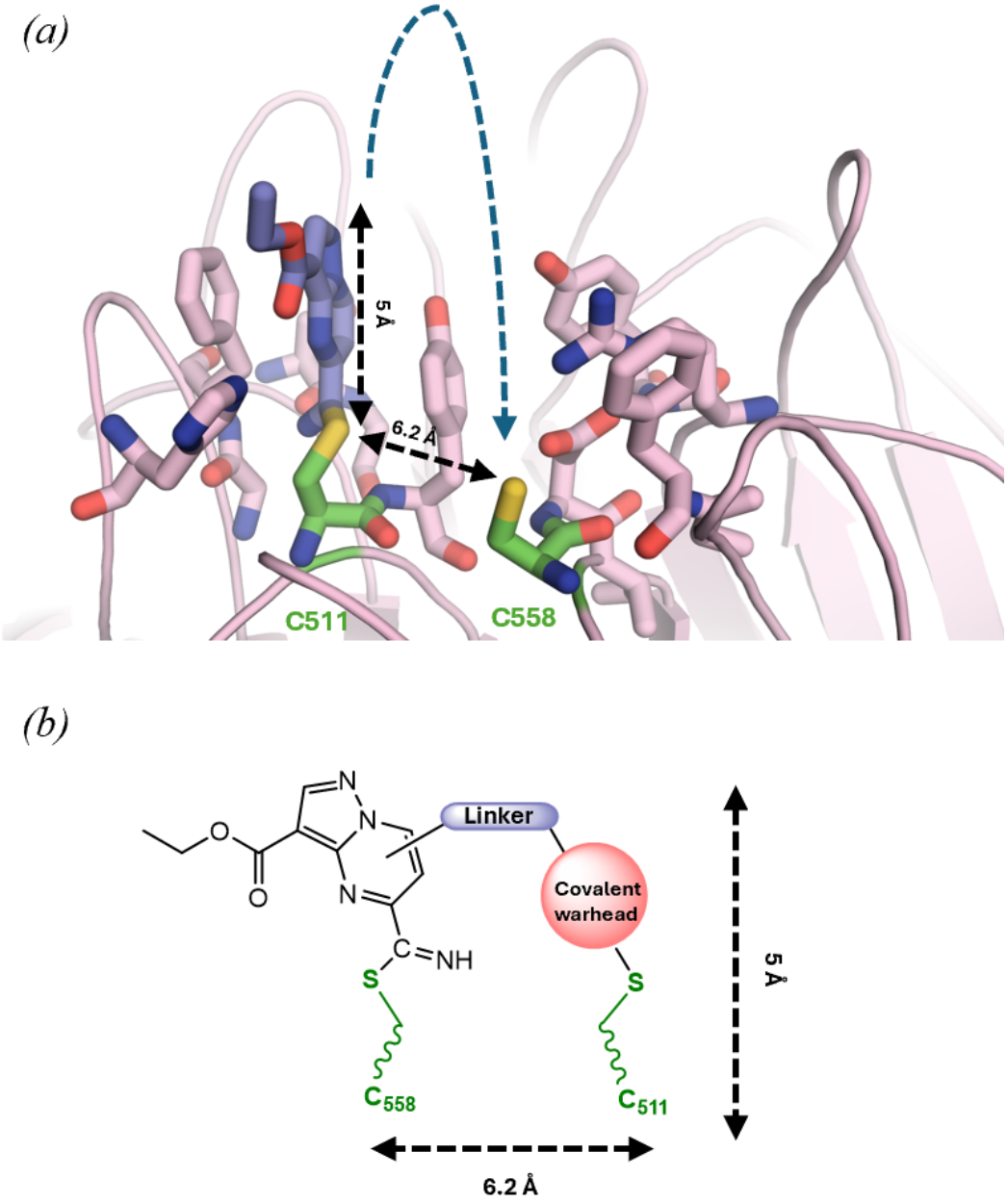


Figure 8.2: **Proposed bipartite binding by targeting both KLHL12 binding cysteines.** Proposed design of extended compounds incorporating dual covalent warheads to engage KLHL12 Cysteines C511 and C558. (a) KLHL12 residues within 5 Å of both C558 and C511 are shown as pink sticks, highlighting an unexplored proximal pocket near C558, (b) Proposed design of an **EVD-1**-based fusion compound incorporating a second covalent warhead to target C558, with a linker spanning at least 6.2 Å to enable dual engagement. KLHL12 coloured in pink, **EVD-1** coloured in purple, cysteines coloured in green.

CRBN and VHL. Although some Kelch domain-containing E3 ubiquitin ligases - KLHL20 [184], KEAP1 [185, 186], and KLHDC2 [182] - have been explored in this context, no reported covalent PROTACs specifically targeting Kelch domain proteins currently exist. However, studies have shown that both reversible and irreversible covalent modifications of the neo-substrate and the E3 ligases, KEAP1 [187] and CRBN [309], are compatible with PROTAC-mediated degradation, even enhancing degradation efficiency. This precedent supports the feasibility of developing a KLHL12-based PROTAC.

The development of covalent inhibitors can achieve high target engagement without a corresponding increase in molecular mass, thereby surpassing the conventional "ligand efficiency" ceiling of 0.3 kcal/mol per heavy atom observed with non-covalent ligands [310, 311]. In addition, covalent modifiers often exhibit prolonged residence time and sustained target occupancy, which can translate into lower effective doses. Covalent inhibition is also advantageous when targeting proteins with shallow or otherwise intractable binding sites that are poorly suited to conventional ligand design, such as the broad KLHL12 Kelch domain presented here, thus enabling modulation of proteins that lack discrete ligandable pockets. Such a case are the covalent inhibitors of the $KRAS^{G12C}$ mutant [312], that gave rise to covalent PROTACs of this oncogenic target [313].

Due to these favourable properties, covalent warheads have recently been incorporated into proximity-inducing modalities such as PROTACs. While degradation has been achieved by covalently recruiting neo-substrates [314–317], covalent binding to the POI can impair the catalytic cycle of PROTACs, inhibiting substrate turnover despite confirmed target engagement [318]. Covalent targeting of the E3 ligase on the other hand, as explored in this study, avoids this limitation, as the catalytic nature of PROTAC-mediated degradation is preserved, since the covalently bound KLHL12-based PROTAC can be recycled, maintaining stoichiometric efficiency. By covalently engaging the ligase while allowing for substrate release, the degradation mechanism transitions from classical ternary-complex kinetics to a "pseudo-binary" model, potentially simplifying linker length and geometry optimisation

[319]. Additionally, covalent ligase targeting may enhance selectivity and reduce off-target toxicity [320]. Only recently have such covalent E3 ligase recruiters been reported; terpene-derived molecules that covalently bind cysteines of RNF114 and KEAP1, as well as CCW16, which targets a zinc-coordinating cysteine in RNF4 [187, 320–323].

In addition, similar to VHL and CRBN, KLHL12 has a shallow binding pocket that could be advantageous for PROTAC-mediated substrate recruitment, as it would not limit potential neo-substrate trajectories and would not impact selective and/or cooperative $CUL3^{KLHL12}$-PROTAC-substrate interactions. KLHL12 also appears to be expressed across nearly all tissues providing wide utility. However, it differs from VHL and CRBN in its subcellular localisation to vesicles, making it particularly suitable for targeting neo-substrates that co-localize with these vesicular structures and potentially avoid off-target effects.

A proof-of-concept study for a KLHL12-based degrader could be conducted by fusing **EVD-1** with a pan-kinase inhibitor and assessing kinase degradation, as reported for KLHDC2 [188]. Alternatively, a physiologically relevant target that potentially shares the same subcellular localization as KLHL12 in vesicles, such as SMARCA2[3] could be considered (Human Protein Atlas, [88]). A study by Farnaby *et al.* reported a potent and cooperative degrader of SMARCA2, by using a potent ligand of SMARCA2 bromodomain [324] to recruit E3 ubiquitin ligase VHL [325]. A similar approach could exploit KLHL12 instead. Future work should focus on synthesising candidate molecules with varied linker types and lengths (as outlined in Fig. 8.3), followed by in-cell assessment of target engagement, SMARCA2 degradation, UPS dependency, and half-maximal degradation concentration (DC_{50}) to establish proof of concept.

Building on the investigation of KLHL12 binders, this study also explored two promising new avenues for small molecule development: the KLHL12-USP21 glue-mediated complex [**Chapter 6**] and the novel $CUL4/CUL2^{NRBP1}$ E3 ligase assembly [**Chapter 7**]. In addition to their potential for

[3] a subunit of a chromatin-remodeling complex

ligand development, both complexes utilize co-adaptors in their assembly and mechanisms. KLHL12 employs co-adaptors at its substrate-binding site to regulate substrate recruitment, while NRBP1 facilitates Cullin binding in a heterotypic assembly previously unreported. These complexes were studied not only for their potential to be chemistry-enabled but also for the valuable insights they provide into novel E3 ligase regulatory mechanisms of assembly.

Initially, BAY-805, a potent inhibitor of the deubiquitinase USP21, was identified in a proteomic study as a compound that induces KLHL12-USP21 binding. NanoBRET experiments further revealed that the dimeric KLHL12 Kelch domain is essential for USP21 recruitment via the BAY-805 glue. However, reconstituting the putative complex *in vitro* proved challenging due to the limited stability of full-length KLHL12. Attempts to recover an engineered HiBit-LgBit KLHL12 Kelch domain complex, which could assemble with USP21, in cryoEM studies yielded no clear particles. This may be attributed to either complex dissociation or weak affinity interactions between the components. Thus, future cryoEM studies could incorporate crosslinking reagents to stabilize these potentially weak-affinity complexes, an approach that was successfully employed in the resolution of another glue mediated E3-ligase complex of dimeric KBTBD4 and HDAC2 [270]. Furthermore, additional studies to identify more stable full-length KLHL12 constructs, possibly co-expressed with Cullin3 and Rbx1 to stabilize the dimeric conformation, may be explored. This approach would offer a twofold advantage: enhanced stability and the potential for larger complexes that could facilitate cryoEM processing, similarly to the recently reported cryoEM structure of the E3 ligase $CRL3^{KLHL22}$ in complex with a mitochondrial ligase substrate by cryoEM [83]. Although a full structure was not resolved in this study, the molecular glue finding provides a promising validation of KLHL12 as a viable host for PROTAC degraders or molecular glues.

In addition to the well-established KLHL12 paradigm, which involves the recruitment of adaptors as co-factors to divert substrate selectivity, I also explored the potential heterotypic assembly of a novel E3 ligase. This model proposed the recruitment of adaptors that facilitate Cullin binding instead [126].

While this concept was intriguing on its own, it was further compounded by the suggestion that the NRBP1 substrate adaptor might recruit both Cullin 4 via DDB1 and Cullin 2/5 through ElonginBC, in addition to a novel co-adaptor, which was hypothesized to be a transcription factor. My *in vitro* attempts to validate this proposed assembly corroborated a key observation: contrary to the traditional paradigm of ElonginBC and SOCS-box-containing proteins, ElonginBC alone was insufficient for Cullin 2/5 recruitment. To the best of my knowledge, this is the first reported case in which a substrate adaptor does not associate with the Cullin 2/5, despite the presence of ElonginBC, highlighting an exception to canonical E3 ligase assembly mechanisms and raising intriguing questions about the regulatory requirements governing this interaction.

Beyond its mechanistic significance, this ligase would represent a novel entity, as it would be the first example of a protein containing a kinase fold acting as the substrate adaptor of an E3 ligase. Despite challenges in resolving the NRBP1 structure in combination with various binding partners using crystallography, cryo-electron microscopy (cryo-EM) yielded promising results. Preliminary data revealed well-distributed particles that survived vitrification and could be readily identified by the software. Although volume reconstruction was unsuccessful with the small molecular-weight complex (NRBP1-ElonginBC) applied to the grids, these findings lay the groundwork for further structural characterization. Future studies exploring potential co-adaptors, such as TSC22D2/D4, DDB1, or other Cullin-binding partners, may have a higher likelihood of yielding a more stable and resolvable asymmetric complex.

This investigation into the unusual E3 ligase assembly mechanism holds promise not only for identifying novel assembly processes but also for evaluating the role of the pseudokinase domain in NRBP1 as a potential ligand-binding site. This could have important implications for ligand discovery, as small-molecule kinase inhibitors are well-studied and could be leveraged in PROTAC-based functionalization for targeted protein degradation. A recent large-scale chemoproteomic study screening 400 fragments against 2,500 proteins identified a fragment binding to NRBP1, fragment **C338** [326].

Given this finding, it is reasonable to hypothesize that fragment **C338** could interact with the NRBP1 pseudokinase fold [118, 119, 327], either at the ATP-binding site, similar to conventional kinase inhibitors, or potentially at a downstream substrate-binding site. Future studies on NRBP1-compound interactions could incorporate this fragment to further investigate its binding mode and potential functional implications.

In summary the findings of this thesis expand the repertoire of chemically tractable E3 ligases, positioning KLHL12, as a promising targets for the development of small-molecules or degraders.

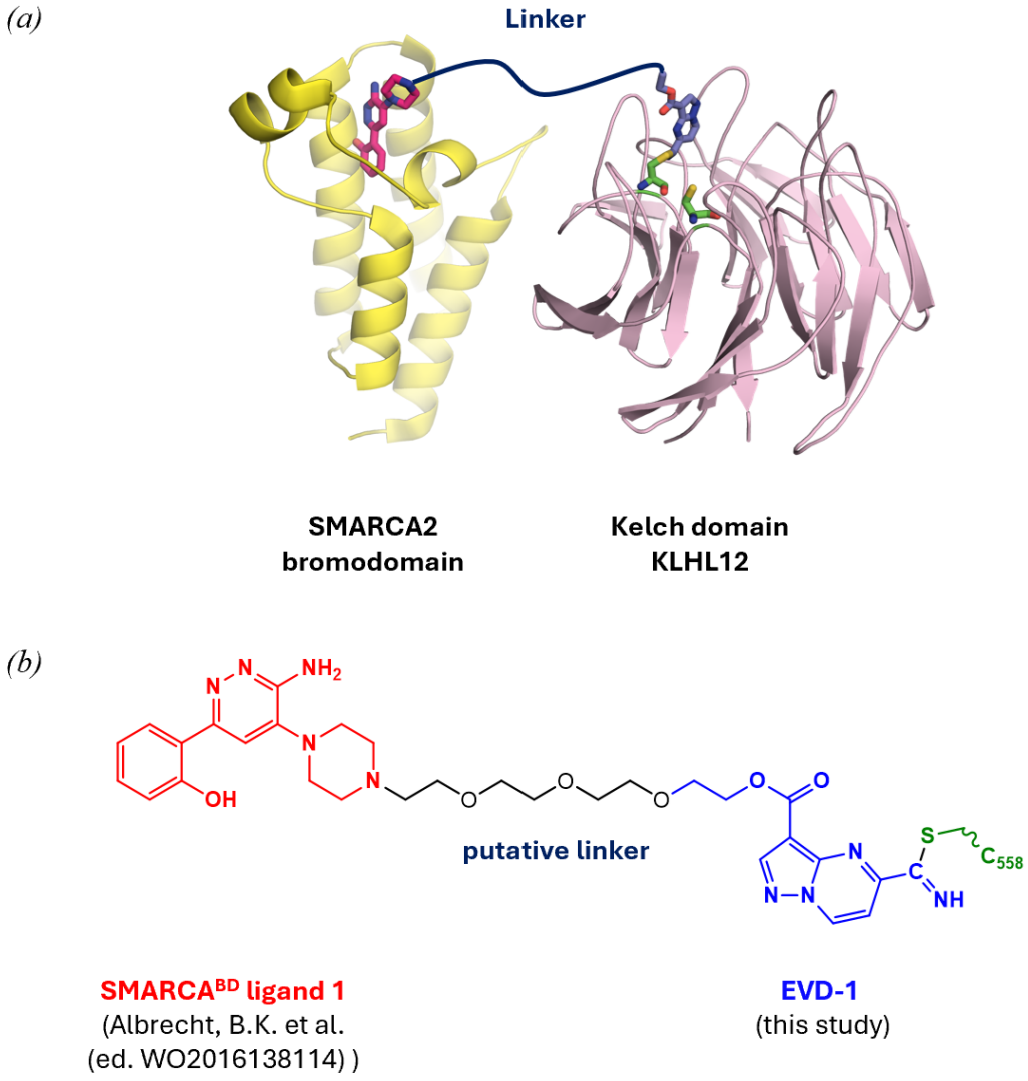


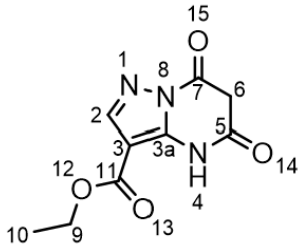
Figure 8.3: **Design of EVD-1 as a handle for KLHL12-based degraders.** Proposed strategy for developing KLHL12-based PROTACs by leveraging the **EVD-1** exit vector. (a) Conceptual design of a bivalent degrader, recruiting KLHL12 via **EVD-1** to mediate degradation of SMARCA2 using a potent ligand binding to its bromodomain [324]. (b) Structural feasibility of a SMARCA2-KLHL12 PROTAC, demonstrating synthetic accessibility through ester or amide coupling chemistry.

CHAPTER 9

Experimental Procedures

9.0.1 Experimental Procedures

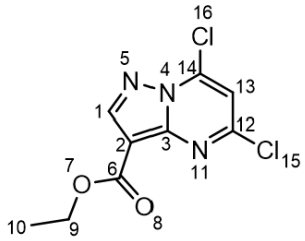
ethyl 5,7-dioxo-4,5,6,7-tetrahydropyrazolo[1,5-a]pyrimidine-3-carboxylate, **EVD-3**



Ethyl 3-imino-2,3-dihydro-1H-pyrazole-4-carboxylate (10 g, 64.45 mmol) and diethyl malonate (14.70 mL, 96.76 mmol) were dissolved in dry EtOH (72 mL). 21% w/w EtONa (36 mL, 96.76 mmol) was added and the solution was stirred at 90 ° C for 24 hrs. The reaction was cooled to room temperature and was acidified to pH = 3 with 1N HCl aqueous solution. The resulting precipitate was filtered and oven-dried to give **EVD-3** compound as an off-white solid (9.667 mg, 43.31 mmol, 67.2%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 9.06 (s, 1H, H4), 7.78 (s, 1H, H2), 4.25 (s, 2H, H9), 4.23 (d, *J* = 8.3 Hz, 2H, H6), 1.27 (t, *J* = 7.0 Hz, 3H, H10); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 162.80 (C5), 162.51 (C7), 157.88 (C11), 142.59 (C3a), 140.05 (C2), 93.75 (C3), 77.25 (C9), 59.30 (C6), 14.47 (C10); **ESI-MS** $C_9H_9N_3O_4$ $[M-H]^-$ *m/z* found 222.24 cald 222.06.

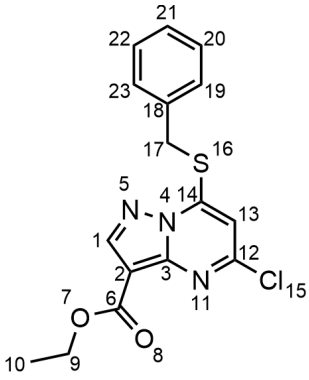
ethyl 5,7-dichloropyrazolo[1,5-a]pyrimidine-3-carboxylate, EVD-4



EVD-3 (7.445 g, 33.36 mmol) was suspended in acetonitrile (31.80 mL), and heated to 50 °C, before addition of $POCl_3$ (12.4 mL, 133.4 mmol) followed by pyridine (2.7 mL, 33.36 mmol). The reaction mixture was stirred at 100 °C for 5 hours, cooled to room temperature, and quenched by pouring into an ice-water mixture. The solution was neutralized with saturated aqueous sodium bicarbonate. The resulting precipitate was collected by filtration, washed with cold water, and obtained as a light yellow paste. The wet paste was dissolved in ethyl acetate, dried over sodium sulfate, and filtered. Removal of the solvent under reduced pressure yielded **EVD-4** as a yellow solid (4.825 g, 18.55 mmol, 55.6%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 8.74 (s, 1H, H1), 7.93 (s, 1H, H13), 4.32 (d, *J* = 7.1 Hz, 2H, H9), 1.32 (t, *J* = 7.1 Hz, 3H, H10); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 161.47 (C6), 152.93 (C12), 148.16 (C1), 147.37 (C14), 141.57 (C3), 111.74 (C13), 103.79 (C2), 60.48 (C9), 14.81 (C10); **ESI-MS** $C_9H_7Cl_2N_3O_2$, $[M\text{-}C_2H_5O]^+$ *m/z* found 214.03 cald 214.97.

ethyl 7-(benzylthio)-5-chloropyrazolo[1,5-a]pyrimidine-3-carboxylate, EVD-6

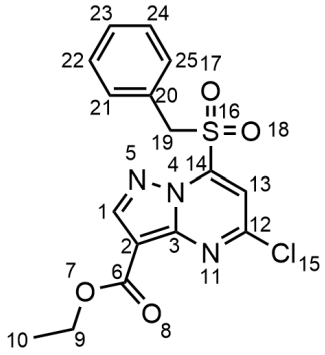


To a solution of **EVD-4** (500 mg, 1.923 mmol) in acetonitrile (6.2 mL) cooled down to 0 °C, triethylamine (268 µL, 1.923 mmol) was added followed by slow addition of benzyl mercaptan (226 µL, 1.923 mmol). The reaction stirred for 2 hours and diluted with water, filtered, and dried under vacuum overnight to afford the product **EVD-6** as a yellow solid (596 mg, 1.714 mmol, 89%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.54 (s, 1H, H1), 7.74 - 7.20 (m, 5H, H19-23), 6.82 (s, 1H, H13), 4.45 - 4.38 (m, 2H, H9), 4.39 (s, 1H, H17), 1.41 (t, *J* = 7.1 Hz, 3H, H10); **^{13}C NMR** (101 MHz, Chloroform-*d*) δ 162.11 (C6), 153.12(C12), 152.58 (C14), 147.69

(C1), 146.63 (C3), 132.90 (C18), 129.38 (C19, C23), 129.06 (C22, C20), 128.80 (C21), 105.64 (C13), 103.44 (C2), 60.76 (C9), 35.84 (C17), 14.59 (C10); **ESI-MS** $C_{16}H_{14}ClN_3O_2S$ $[M+H]^+$ *m/z* found 348.13 cald 348.05.

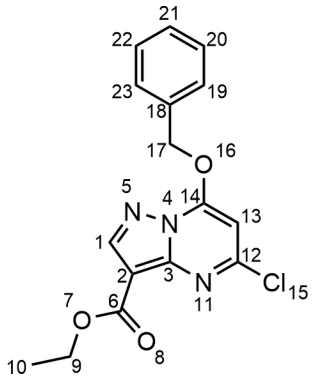
ethyl 7-(benzylsulfonyl)-5-chloropyrazolo[1,5-a]pyrimidine-3-carboxylate, **EVD-7**



EVD-6 (309 mg, 0.888 mmol) and 3- chlorobenzoperoxoic acid (70%, 987 mg, 4 mmol) were suspended in DCM (8.9 mL) at 0°C. The reaction was stirred and left to reach ambient temperature for 16 hours. The reaction was treated with saturated aqueous sodium bicarbonate, and extracted between DCM and water (x3). The combined organic phases were washed with brine (x3) and treated with sodium sulfate, before being filtered, and evaporated under *vacuo* to afford brown oil. The crude product was purified with silica gel chromatography (gradient elution EtOAc 5%-40% in cHex) to yield **EVD-7** as a white solid (270 mg, 0.711 mmol, 80%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.79 (s, 1H, H1), 7.32 (s, 1H, H13), 7.34 - 7.28 (m, 1H, H23), 7.24 - 7.19 (m, 2H, H22, H24), 7.10 - 7.01 (m, 2H, H21, H25), 5.09 (s, 2H, H19), 4.46 (d, *J* = 7.1 Hz, 2H, H9), 1.45 (t, *J* = 7.1 Hz, 3H, H10); **^{13}C NMR** (101 MHz, Chloroform-*d*) δ 161.32 (C6), 153.86 (C12), 148.75 (C1), 147.42 (C14), 143.43(C3), 130.26(C21, C25), 129.97 (C20), 129.31 (C22, C24), 125.29 (C23), 112.57 (C13), 104.93 (C2), 61.20 (C9), 60.22 (C19), 14.41 (C10); **ESI-MS** $C_{16}H_{14}ClN_3O_4S$ $[M+H]^+$ *m/z* found 380.12 cald 380.04.

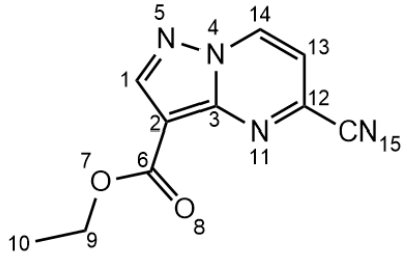
ethyl 7-(benzyloxy)-5-chloropyrazolo[1,5-a]pyrimidine-3-carboxylate, EVD-11



Benzyl alcohol (131 µL, 1.269 mmol) was dissolved in THF (2.1 ml) and the solution cooled to 0 °C. Anhydrous sodium hydride (60%, 76.1 mg, 1.904 mmol) was added portion-wise and the reaction was left stirring for 30 min at 15 °C. The suspension was slowly added to a stirring solution of **EVD-4** (300 mg, 1.154 mmol) in THF (1.2 mL) at 0 °C. The resulting mixture was left to stir for 3 hrs at 0 °C. The reaction mixture was quenched with saturated aqueous NH_4C1 (3 mL) and extracted with EtOAc (x3). The combined organic phases were washed with brine (x3), dried over sodium sulphate and evaporated to dryness. The crude residue was purified by silica gel chromatography (gradient elution EtOAc 5%-40% in cHex) to yield **EVD-11** (29.8 mg, 0.090 mmol, 8%) as a white solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.48 (s, 1H, H1), 7.54 (dt, *J* = 6.1, 1.6 Hz, 2H, H23, H19), 7.44 - 7.31 (m, 3H, H20-22), 6.66 (s, 1H, H13), 5.57 (s, 2H, H17), 4.41 (q, *J* = 7.1 Hz, 2H, H9), 1.43 (t, *J* = 7.1 Hz, 3H, H10); **ESI-MS** $C_{16}H_{14}ClN_3O_3$ $[M+H]^+$ *m/z* found 332.13 cald 332.07.

ethyl 5-cyanopyrazolo[1,5-a]pyrimidine-3-carboxylate, EVD-1

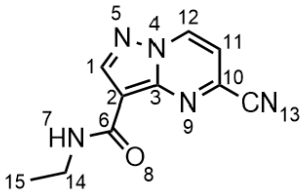


Ethyl 5-chloropyrazolo[1,5-a]pyrimidine-3-carboxylate (1.5 g, 6.648 mmol) **EVD-13**, [(2-Di-tert-butylphosphino-2',4',6'-triisopropyl-1,1'-biphenyl)-2- (2'-amino- 1,1'-biphenyl)] palladium(II) methanesulfonate (164 mg, 0.239 mmol) and $Zn(CN)_2$ (515 mg, 4.388 mmol) were all added as solids to a reaction tube which was then sealed. De-gassed deionised water (16.6 mL) and THF (3.3 mL) were added to the reaction tube with a syringe and the mixture was stirred rapidly such that the two phases appear homogenous at 50 °C for 24 hours. The reaction was treated with saturated aqueous sodium bicarbonate, and extracted with ethyl acetate

(x3). The combined organic layers were dried over magnesium sulphate, filtered, and concentrated *in vacuo*. The crude residue was purified by silica gel chromatography (gradient elution EtOAc 5%-40% in cHex) to afford **EVD-1** as a pale yellow solid (830 mg, 3.837 mmol, 57.7%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.92 (d, *J* = 7.1 Hz, 1H, H14), 8.73 (s, 1H, H1), 7.29 (d, *J* = 7.2 Hz, 1H, H13), 4.45 (q, *J* = 7.1 Hz, 2H, H9), 1.43 (t, *J* = 7.1 Hz, 3H, H10); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 161.59 (C6), 149.82 (C1), 146.69 (C3), 137.22 (C14), 134.89 (C12), 115.50 (C15), 111.52 (C13), 105.96 (C2), 61.19 (C9), 14.54 (C10); **ESI-MS** $C_{10}H_8N_4O_2$ $[M-H]^-$ *m/z* found 215.10 cald 215.06.

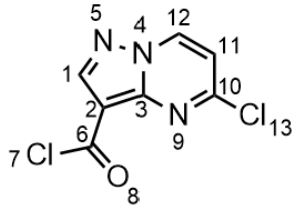
5-cyano-N-ethylpyrazolo[1,5-a]pyrimidine-3-carboxamide, EVD-15



[(2-Di-tert-butylphosphino-2',4',6'-triisopropyl- 1,1'- biphenyl) -2-(2'-amino-1,1'-biphenyl)] palladium(II) methanesulfonate (29 mg, 0.042 mmol) and $Zn(CN)_2$ (81.8 mg, 0.696 mmol) were added as solids to a reaction tube which was then sealed. Degassed deionised water (5.3 mL) and THF (1 mL) were added to the reaction tube with a syringe and the mixture was stirred rapidly such that the two phases appear homogenous at 50 °C for 24 hours. The reaction was treated with saturated aqueous sodium bicarbonate, and extracted with ethyl acetate (x3). The combined organic layers were dried over magnesium sulphate, filtered, and concentrated *in vacuo*. The crude residue was purified by silica gel chromatography (gradient elution EtOAc 5%-40% in cHex) to afford a mixture of **EVD-15** and **EVD-18** as a pale brown solid (199 mg, 0.925 mmol, 87.6%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.94 (d, *J* = 7.1 Hz, 1H, H12), 8.86 (s, 1H, H1), 8.69 - 8.62 (m, 1H, H11), 7.46 (s, 1H, H7), 3.74 - 3.32 (m, 2H, H14), 1.31 (t, *J* = 7.4 Hz, 3H, H15); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 160.56 (C6), 149.59 (C1), 148.16 (C3), 137.49 (C12), 133.00 (C10), 110.56 (C11), 110.12 (C13), 108.85 (C2), 34.43 (C14), 15.23 (C15); **ESI-MS** $C_{10}H_9N_5O$, $[M\text{-}C_2H_6N]^+$ *m/z* found 171.08 cald 171.03; **HR-ESI-MS**, **EVD-15** $C_{10}H_9N_5O$ $[M+H]^+$ *m/z* found 216.0859 cald 216.0879, alongside starting material **EVD-18** $C_9H_9ClN_4O$ $[M+H]^+$ *m/z* found 225.0518 cald 225.0538.

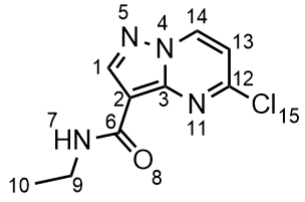
5-chloropyrazolo[1,5-a]pyrimidine-3-carbonyl chloride, EVD-12



Thionyl chloride (10 ml, 137 mmol) was added dropwise to 5-chloropyrazolo[1,5-a] pyrimidine -3-carboxylic acid (500 mg, 2.531 mmol) dissolved in DCM (3 mL). After the addition, the reaction was heated to 80 °C and left stirring for 3 hours. The thionyl chloride was removed under reduced pressure to yield crude residue of reactive intermediate **EVD-12** (546.7 mg, 2.531 mmol, 100%).

Reactive intermediate, **ESI-MS** of morpholine adduct, (5-chloropyrazolo[1,5-a] pyrimidin-3-yl (morpholino)-methanone, $C_{11}H_{11}Cl_2N_4O_2$, $[M-H]^-$ *m/z* found 265.11 cald 265.06.

5-chloro-N-ethylpyrazolo[1,5-a]pyrimidine-3-carboxamide, EVD-18



Ethylamine hydrochloride solid (309.6 mg, 3.797 mmol) was dissolved in anhydrous DCM (6 mL) and the solution was cooled down 0 °C, before addition of triethylamine (768.4 mg, 7.593 mmol). Addition of crude residue **EVD-17** (546.7 mg, 2.531 mmol) was done dropwise. The mixture was allowed to react to room temperature for 24 hours. Potassium carbonate 2 M was added to neutralise the mixture, and the mixture was extracted with DCM (x3). The combined organic phases were dried over sodium sulphate, filtered, and the crude residue was purified by gel chromatography (gradient elution EtOAc 5%-40% in cHex) to give EVD-13 as an orange solid (236 mg, 1.051 mmol, 41.5 %).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.65 (d, *J* = 7.1 Hz, 2H, H1, H14), 7.46 (s, 1H, H7), 6.95 (d, *J* = 7.2 Hz, 1H, H13), 3.54 (qd, *J* = 7.2, 5.7 Hz, 2H, H9), 1.29 (t,*J* = 7.3 Hz, 3H, H10); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 161.28 (C6), 152.84 (C12), 148.15 (C1), 144.66 (C3), 137.46 (C14), 110.12 (C13), 106.54 (C2), 34.24 (C9), 15.25 (C10); **ESI-MS** $C_9H_9ClN_4O$, $[M-C_2H_6N]^+$ *m/z* found 180.08 cald 180.00.

APPENDIX A

Cloning &Construct Design

Primer Name	Primer Sequence	Amplification Position	Tm (°C)	% GC
NRBP1-f011	TACTTCCAATCCATGTCGGAGGGGGAGTCCCAG	Ser2	54	57.6
NRBP1-f012	TACTTCCAATCCATGGCTTCCCCAGAGGAAGAA	Ala41	47.3	48.5
NRBP1-f013	TACTTCCAATCCATGTCGCCCTGTGGGCGCTGG	Ser59	59.3	60.6
NRBP1-f014	TACTTCCAATCCATGCGAGAAGAGGTGAATCAACGG	Arg68	49.4	47.2
NRBP1-f015	TACTTCCAATCCATGGTGCCCTCGCTCAAACTC	Val330	49.5	51.5
NRBP1-f016	TACTTCCAATCCATGGCTGAGGTGGAGACTCGC	Ala437	50.9	54.5
NRBP1-f017	TACTTCCAATCCATGACTCGCAAGGTGGTGCTG	Thr441	51.6	51.5
TSC22D2-f001	TACTTCCAATCCATGCCAGTGACTGCAGCCCCA	Pro146	53.6	54.5
TSC22D4-f001	TACTTCCAATCCATGGGCACCCCCCGGAATGGC	Gly56	57.6	60.6
NRBP1-r014	TATCCACCTTTACTGTCAAGAGGAGACGGTGACAGC	Ser535	49.7	50
NRBP1-r015	TATCCACCTTTACTGTCAAGCGGCTGAGTTGAGGGT	Ala530	52.6	50
NRBP1-r016	TATCCACCTTTACTGTCAAGTCTCCACCTCAGCTGG	Thr441	49.5	50
NRBP1-r017	TATCCACCTTTACTGTCACAGCCCAAAGGCTGTCAG	Leu409	49.6	50
NRBP1-r018	TATCCACCTTTACTGTCACAGAGCTGGTGACTGAGA	Leu388	46.9	47.2
NRBP1-r019	TATCCACCTTTACTGTCAAGTATCCATGTTTTTGGT GATCTCCT	Thr362	51.2	38.6
NRBP1-r020	TATCCACCTTTACTGTCACTCTGGGATCATGTGTTGG	Glu350	46.4	45.9
TSC22D2-r001	TATCCACCTTTACTGCTCCATCGGCCGCGTCTATA	Trp180	50.2	51.4
TSC22D2-r002	TATCCACCTTTACTGTCACCATCGGCCGCGTCTATA	Trp180	50.2	50
TSC22D4-r001	TATCCACCTTTACTGTCACCAGCGACCGCGGCGATA	Trp90	56.3	55.6
TSC22D4-r002	TATCCACCTTTACTGCTCCAGCGACCGCGGCGATA	Trp90	56.3	57.1

Figure A.1: **Oligonucleotide design for NRBP1 interactors for *E.coli* expression** Bases coloured in red anneal to the vector, bases coloured in black anneal to the entry clone.

Primer Name	Primer Sequence	Amplification Position	Tm (°C)	% GC
DDB1A-f001	TACTTCCAATCCATGTCGTACAACTACGTGGTAACGG	Ser2	50.7	45.9
DDB1A-f002	CCCACTCTCTATGATGTCGTACAACTACGTGGTAACGG	Ser2	50.7	50
DDB1A-f003	TTAAGAAGGAGATATACTATGTCGTACAACTACGTGGTAACGGCC	Ser2	55.3	42.2
NRBP1-f001	TACTTCCAATCCATGTCGGAGGGGGAGTCCCAG	Ser2	54	57.6
NRBP1-f002	TACTTCCAATCCATGGCTTCCCCAGAGGAAGAA	Ala41	47.3	48.5
NRBP1-f003	TACTTCCAATCCATGTCGCCCTGTGGGCGCTGG	Ser59	59.3	60.6
NRBP1-f004	TACTTCCAATCCATGCGAGAAGAGGTGAATCAACGG	Arg68	49.4	47.2
NRBP1-f005	TACTTCCAATCCATGGTGCCCTCGCTCAAACTC	Val330	49.5	51.5
NRBP1-f006	TACTTCCAATCCATGGCTGAGGTGGAGACTCGC	Ala437	50.9	54.5
NRBP1-f007	TACTTCCAATCCATGACTCGCAAGGTGGTGCTG	Thr441	51.6	51.5
TCEB1A-f001	TTAAGAAGGAGATATACTATGTATGTCAAATTGATATCATCTGATG GCCATGAATTTATTGTAAAAAGAGA	Met1	58.3	28.2
TCEB2A-f001	TTAAGAAGGAGATATACTATGGACGTGTTCCTCATGATCCGGC	Asp2	54.3	44.2
DDB1A-r001	GATTGGAAGTAGAGGTTCTCTGCATGGATCCGAGTTAGCTCC	His1140	47.3	50
DDB1A-r002	TATCCACCTTTACTGTCAATGGATCCGAGTTAGCTCCT	His1140	49.1	44.7
NRBP1-r001	TATCCACCTTTACTGTCAAGAGGAGACGGTGACAGC	Ser535	49.7	50
NRBP1-r002	TATCCACCTTTACTGTCAAGCGGCTGAGTTGAGGGT	Ala530	52.6	50
NRBP1-r003	TATCCACCTTTACTGTCAAGTCTCCACCTCAGCTGG	Thr441	49.5	50
NRBP1-r004	TATCCACCTTTACTGTCACAGCCCAAAGGCTGTCAG	Leu409	49.6	50
NRBP1-r005	TATCCACCTTTACTGTCACAGAGCTGGTGACTGAGA	Leu388	46.9	47.2
NRBP1-r006	TATCCACCTTTACTGTCAAGTATCCATGTTTTTGGTGATCTCCT	Thr362	51.2	38.6
NRBP1-r007	TATCCACCTTTACTGTCACTCTGGGATCATGTGTTGG	Glu350	46.4	45.9
NRBP1-r008	TATCCACCTTTACTGTCACACTTCAAACAATGCTGGGT	Val330	47.9	42.1
TCEB1A-r001	GATTGGAAGTAGAGGTTCTCTGCTCAACAATCTAAGAAGTTCGCAGC	Cys96	47.4	44.7
TCEB2A-r002	GATTGGAAGTAGAGGTTCTCTGCTCACTGCACGGCTTGTTCATT	Gln118	47.4	47.7

Figure A.2: **Oligonucleotide design for NRBP1 interactors for** Sf9 **expression.** Bases coloured in red anneal to the vector, bases coloured in black anneal to the entry clone.

Primer Name	Primer Sequence	Amplification Position	Tm (°C)	% GC
CUL2-f001	TACTTCCAATCCATGAGCCTGAAACCGCGTGT	Met1	53.9	50
CUL5-f001	TACTTCCAATCCATGGCGACGTCTAATCTGTTAAAGAATAAAGGT	Met1	55.1	37.8
ITM2B-f001	TACTTCCAATCCATGCCAGATGACGTGTACTACTGTG	Pro83	48.9	45.9
ITM2C-f001	TACTTCCAATCCATGGCCCGAGATAACTTCTTCCG	Ala89	49	48.6
TSC22D3-f001	TACTTCCAATCCATGAACACCGAAATGTATCAGACCCC	Met1	53.8	44.7
TSC22D3-f002	TACTTCCAATCCATGTATCAGACCCCCATGGAGG	Met5	52.5	50
TSC22D4-f001	TACTTCCAATCCATGAGCGGGGGCAAGAAGAA	Met1	53.3	50
TSC22D4-f002	TACTTCCAATCCATGCGGCTGCGGATGGAGTT	Met240	56.4	53.1
TSC22D4-f003	TACTTCCAATCCATGGACCCCTTCGGAGCAGTA	Asp281	49.5	51.5
CUL2-r001	TATCCACCTTTACTGTCACTTATCGTCGTCATCCTTGTAATCATG	Lys406	50.4	40
CUL5-r001	TATCCACCTTTACTGTCACTTATCGTCGTCATCCTTGTAATCATG	Lys406	50.4	40
ITM2B-r001	TATCCACCTTTACTGCTAGAACAAATTAAAGTTTCCACGGCA	Ser266	50.9	38.1
ITM2C-r001	TATCCACCTTTACTGCTCACCACCCCGCAGATGAG	Val267	51.4	54.3
TSC22D3-r001	TATCCACCTTTACTGTCACACCGCAGAACCACCAGG	Val134	51.9	52.8
TSC22D3-r002	TATCCACCTTTACTGTCACTCTTCAGGGCTCAGACA	Glu117	46.9	47.2
TSC22D4-r001	TATCCACCTTTACTGTCAGACGGAGGGCCCATTGGG	Val395	54.1	55.6
TSC22D4-r002	TATCCACCTTTACTGTCAGACCCCCGAGGAGGGCAG	Val382	55.6	58.3

Figure A.3: **Oligonucleotide design for NRBP1 interactors for *E.coli* co-expression**. Bases coloured in red anneal to the vector, bases coloured in black anneal to the entry clone.

Step	PCR Temp. (°C)	Time (min)	Cycles
Initial Denaturation	95	10	1x
1. Denaturation	95	0.5	5x
2. Annealing	68	0.5	
3. Extension	68	1.5	
4. Denaturation	95	0.5	5x
5. Annealing	60	1.5	
6. Extension	68	1.5	
7. Denaturation	95	0.5	5x
8. Annealing	55	1.5	
9. Extension	68	1.5	
10. Denaturation	95	0.5	20x
11. Annealing	50	1.5	
12. Extension	68	1.5	
Final Extension	68	10	1x
Storage	4	∞	-

Components	Volume (µL)	Final Quantity/Concentration
Template DNA (2.5 ng/µL)	2.5	50 ng
dNTPs (10 mM)	0.75	300 µM
Forward Primer (5 µM)	1.5	0.3 µM
Reverse Primer (5 µM)	1.5	0.3 µM
5x Buffer	5	1x
Herculase II Fusion DNA Pol	0.25	0.02 U/µL
Nuclease-Free H_20	15	-
Total Volume	25 µL	

Figure A.4: **Thermocycling conditions and PCR reaction for PCR amplification** using Herculase DNA Polymerase II.

Step	PCR Temp. (°C)	Time (s)	Cycles
Initial Denaturation	98	30	1x
1. Denaturation	98	30	5x
2. Annealing	72	30	
3. Extension	72	60	
4. Denaturation	98	30	5x
5. Annealing	70	30	
6. Extension	72	60	
7. Denaturation	98	30	5x
8. Annealing	65	30	
9. Extension	72	60	
10. Denaturation	98	30	20x
11. Annealing	60	30	
12. Extension	72	60	
Final Extension	72	120	1x
Storage	4	∞	-

Components	Volume (µL)	Final Quantity/Concentration
Template DNA (2.5 ng/µL)	2.5	< 1,000 ng
dNTPs (10 mM)	0.5	200 µM
Forward Primer (10 µM)	1.25	0.5 µM
Reverse Primer (10 µM)	1.25	0.5 µM
5x Q5 Reaction Buffer	5	1x
Q5 High-Fidelity Pol	0.25	0.02 U/µL
(5x Q5 High GC Enhancer)	(5 µl)	1x
Nuclease-Free H20	q.s. 25	-
Total Volume	25 µL	

Figure A.5: Thermocycling Conditions and PCR Reaction for PCR Amplification using Q5 Polymerase

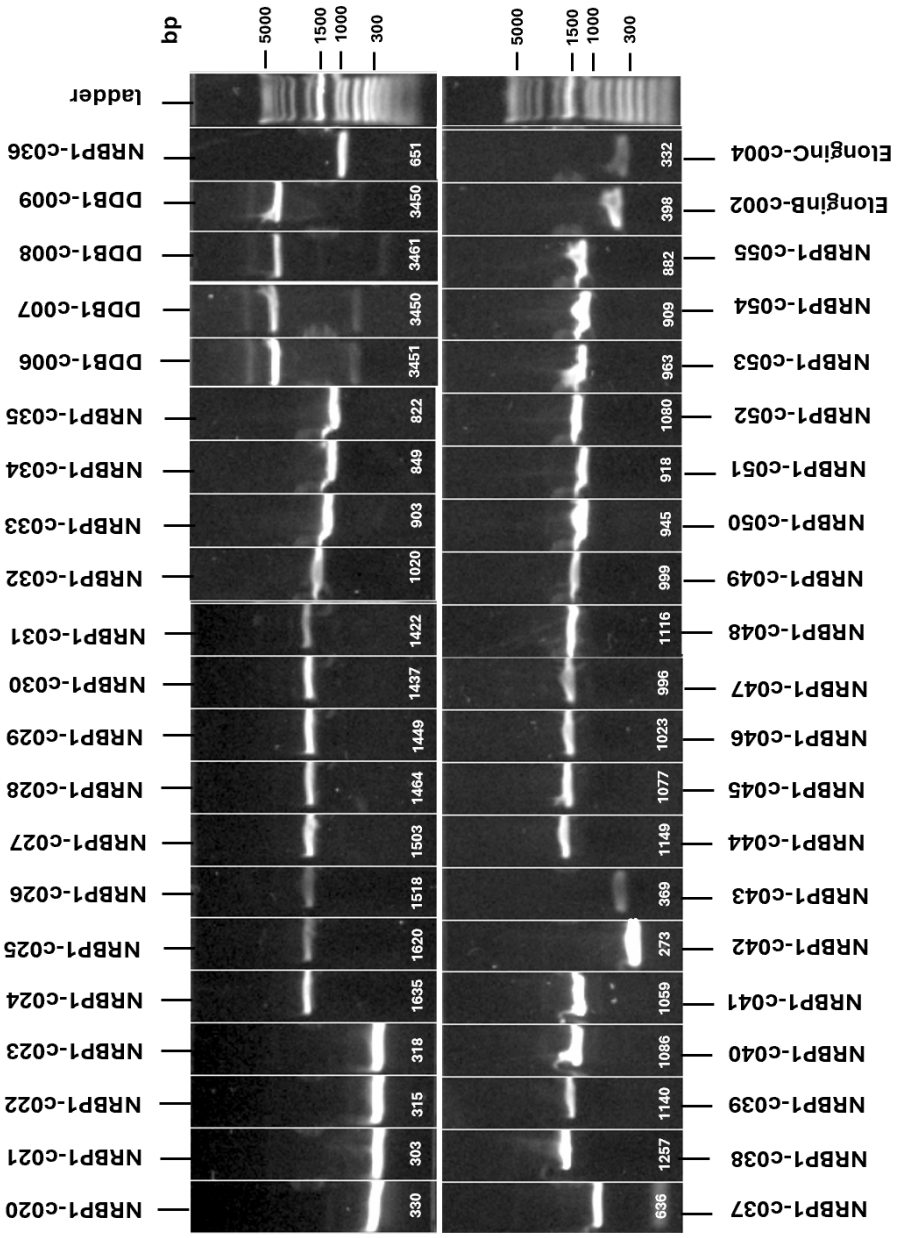


Figure A.6: **Purified PCR amplicons summary for baculo constructs: 96 successful amplifications**. Purified PCR amplicons were analyzed using the E-Gel®system and compared to E-Gel®1 Kb Plus DNA Ladder. Due to the low resolution of these gels the products are judged based on the sizing of neighboring bands.

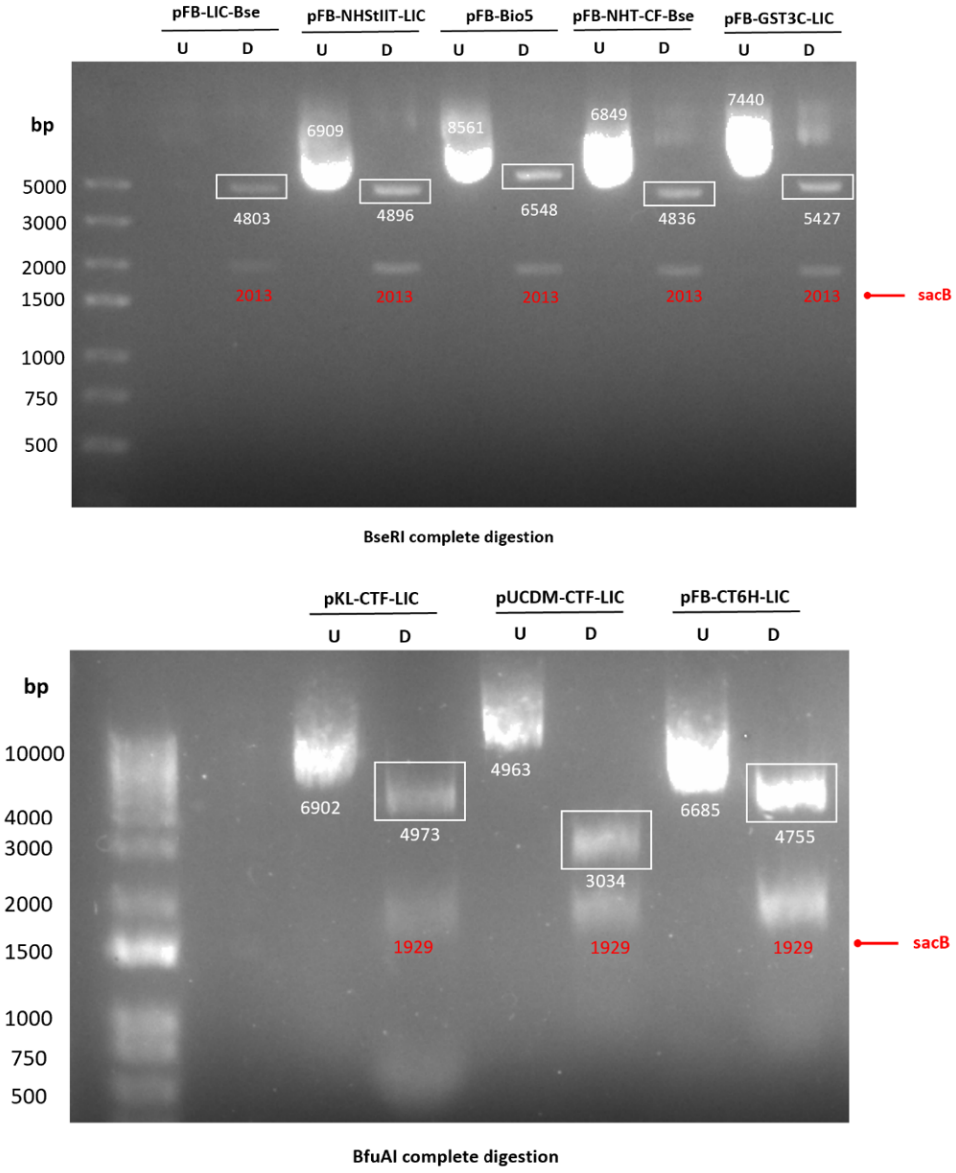


Figure A.7: **Vector digestion of vectors for Sf9 expression**. In-house LIC-adapted circular vectors containing the sacB gene were linearised by restriction digest. U = undigested vector, D = digested vector. The appearance of the sacB gene (in red) signifies successful linearisation.

Step	PCR Temp. (°C)	Time (min)	Cycles
Initial Denaturation	95	10	1x
1. Denaturation	95	0.5	30x
2. Annealing	50	0.5	
3. Extension	72	3	
Final Extension	72	5	1x
Storage	15	∞	-

Components	Volume (µL)	Final Quantity/Concentration
Resuspended Colony in H_20	5	-
2xMyTaq Reaction Buffer	10	1x
FBAC1 Primer (100 µM)	0.1	0.5 µM
FBAC2 Primer (100 µM)	0.1	0.5 µM
Nuclease-Free H_20	4.8	-
Total Volume	20 µL	

Figure A.8: Representative Thermocycling conditions and PCR reaction for PCR bacmid screening

Step	PCR Temp. (°C)	Time (min)	Cycles
Initial Denaturation	95	5	1x
1. Denaturation	95	0.75	25x
2. Annealing	50	0.75	
3. Extension	72	5	
Final Extension	72	7	1x
Storage	15	∞	-

Components	Volume (µL)	Final Quantity/Concentration
Bacmid DNA	1.5	variable
Bioline 5x Buffer	4	1x
FBAC1 Primer (10 µM)	0.5	0.25 µM
M13bac_rev (10 µM)	0.5	0.25 µM
MyTaq™ Red DNA Pol (5 unit/µL, Bioline)	0.05	0.0125 units
Nuclease-Free H20	q.s 20	-
Total Volume	20 µL	

Figure A.9: Thermocycling conditions and PCR reaction for PCR bacmid screening

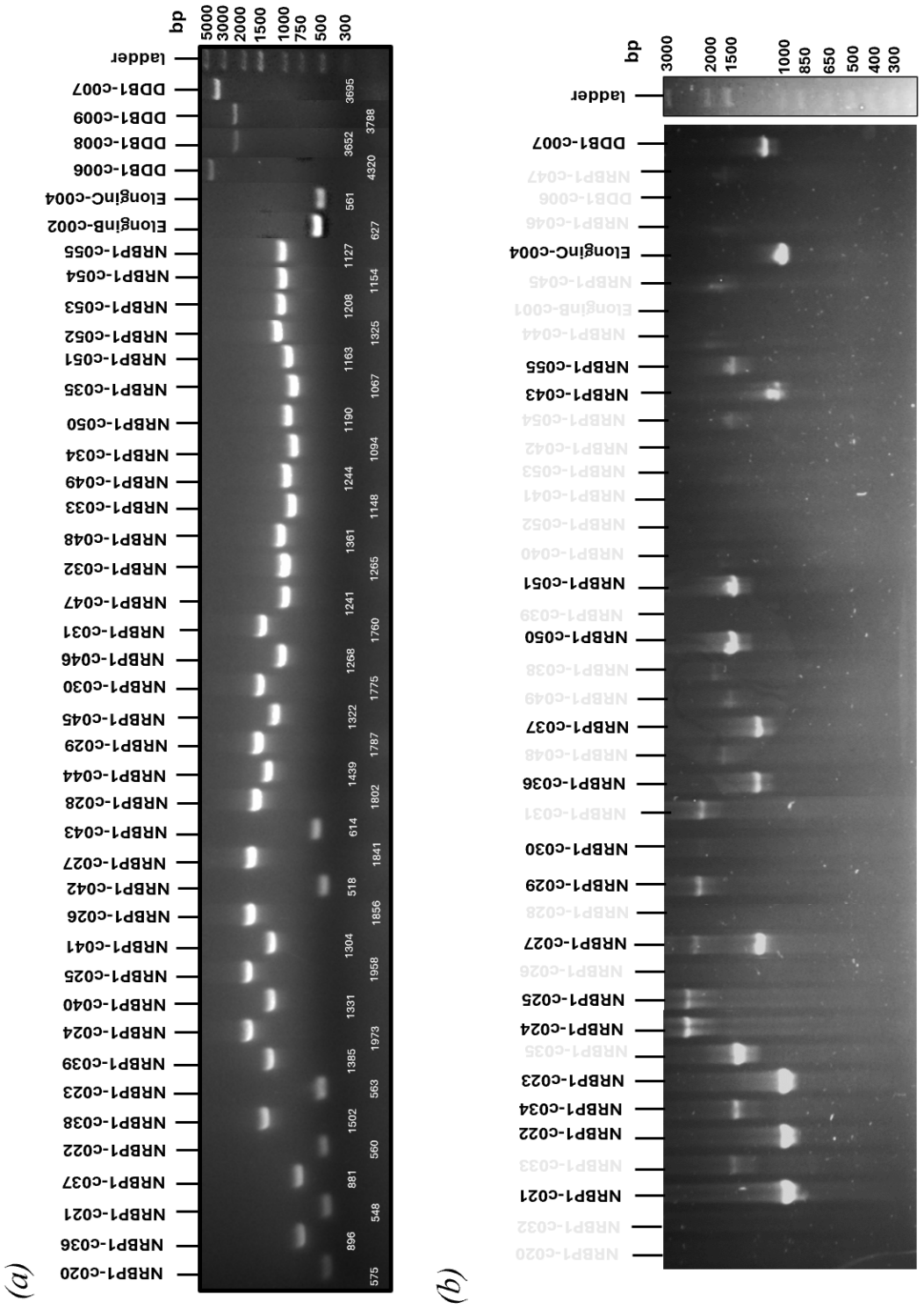


Figure A.10: **Colony PCR and bacmid screening of Sf9 constructs**. (a) Colony PCR Screening performed in a in a 96-well format, analyzed on a 1.5% TAE agarose gel. Monocolonies are picked and amplified in the presence of FBAC1 and FBAC2 oligonucleotides that anneal to a common sequence in all Baculo vectors' backbone external to their LIC sites. An amplification product of the correct size indicates a successful transformation. Wells in blue failed to yield any colonies. (b) Bacmid Screening performed in a in a 96-well format, analyzed on a 1.5% TAE agarose gel. White monocolonies were picked and amplified in the presence of FBAC1 and M13bac oligonucleotides that anneal to a common sequence in all Baculo vectors' backbone external to their LIC sites and to the lacZ gene respectively. An amplification product of the correct size indicates successful bacmid recombination. Wells coloured blue had failed to yield any colonies. Wells coloured in red failed to amplify.

REFERENCES

[1] G. Goldstein, M. Scheid, U. Hammerling, D. H. Schlesinger, H. D. Niall, and E. A. Boyse. Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proceedings of the National Academy of Sciences of the United States of America*, 72(1):11–15, 1975.

[2] J. D. Etlinger and A. L. Goldberg. A soluble ATP-dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes. *Proceedings of the National Academy of Sciences of the United States of America*, 74(1):54–58, January 1977.

[3] A. Ciechanover, Y. Hod, and A. Hershko. A heat-stable polypeptide component of an atp-dependent proteolytic system from reticulocytes. *Biochemical and Biophysical Research Communications*, 81(4):1100–1105, April 1978.

[4] A. Hershko, A. Ciechanover, and I. A. Rose. Resolution of the ATP-dependent proteolytic system from reticulocytes: a component that interacts with ATP. *Proceedings of the National Academy of Sciences of the United States of America*, 76(7):3107–3110, July 1979.

[5] A. Ciechanover, H. Heller, S. Elias, A. L. Haas, and A. Hershko. ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. *Proceedings of the National Academy of Sciences of the United States of America*, 77(3):1365–1368, March 1980.

[6] A. Hershko, A. Ciechanover, H. Heller, A. L. Haas, and I. A. Rose. Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. *Proceedings of the National Academy of Sciences of the United States of America*, 77(4):1783–1786, April 1980.

[7] A. Ciechanover, S. Elias, H. Heller, S. Ferber, and A. Hershko. Characterization of the heat-stable polypeptide of the ATP-dependent proteolytic system from reticulocytes. *Journal of Biological Chemistry*, 255(16):7525–7528, August 1980.

[8] K. D. Wilkinson, M. K. Urban, and A. L. Haas. Ubiquitin is the ATP-dependent proteolysis factor I of rabbit reticulocytes. *Journal of Biological Chemistry*, 255(16):7529–7532, 1980.

[9] A. Ciechanover, H. Heller, R. Katz-Etzion, and A. Hershko. Activation of the heat-stable polypeptide of the ATP-dependent proteolytic system. *Proceedings of the National Academy of Sciences of the United States of America*, 78(2):761–765, February 1981.

[10] A. Hershko, E. Eytan, A. Ciechanover, and A. L. Haas. Immunochemical analysis of the turnover of ubiquitin-protein conjugates in intact cells. Relationship to the breakdown of abnormal proteins. *Journal of Biological Chemistry*, 257(23):13964–13970, 1982.

[11] A. Hershko, H. Heller, S. Elias, and A. Ciechanover. Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. *Journal of Biological Chemistry*, 258(13):8206–8214, July 1983.

[12] Vincent Chau, John W. Tobias, Andreas Bachmair, David Marriott, David J. Ecker, David K. Gonda, and Alexander Varshavsky. A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science*, 243(4898):1576–1583, 1989.

[13] A. Ciechanover. Intracellular protein degradation: from a vague idea thru the lysosome and the ubiquitin-proteasome system and onto human diseases and drug targeting. *Cell Death & Differentiation*, 12(9):1178–1190, 2005.

[14] A. Hershko and A. Ciechanover. The ubiquitin system. *Annual Review of Biochemistry*, 67:425–479, 1998.

[15] M. Hochstrasser. Ubiquitin-dependent protein degradation. *Annual Review of Genetics*, 30:405–439, 1996.

[16] A. L. Goldberg. Degradation of abnormal proteins in *E. coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 69:422–426, 1972.

[17] Y. Klemes, J. D. Etlinger, and A. L. Goldberg. Properties of proteins degraded rapidly in reticulocytes: intracellular aggregation of the globin molecules prior to hydrolysis. *Journal of Biological Chemistry*, 256(16):8436–8444, 1981.

[18] W. F. Prouty and A. L. Goldberg. Fate of abnormal proteins in *E. coli*: accumulation in intracellular granules before catabolism. *Nature New Biology*, 240:147–150, 1972.

[19] Deanna M. Koepp, J. Wade Harper, and Stephen J. Elledge. How the cyclin became a cyclin: Regulated proteolysis in the cell cycle. *Cell*, 97(4):431–434, 1999.

[20] E. Latres, D. S. Chiaur, and M. Pagano. The human F box protein β-Trcp associates with the Cul1/Skp1 complex and regulates the stability of β-catenin. *Oncogene*, 18(4):849–854, 1999.

[21] R. W. King, J. M. Peters, S. Tugendreich, M. Rolfe, P. Hieter, and M. W. Kirschner. A 20S complex containing CDC27 and CDC16 catalyzes the mitosis-specific conjugation of ubiquitin to cyclin B. *Cell*, 81(2):279–288, April 1995.

[22] Miki Tsukada and Yoshinori Ohsumi. Isolation and characterization of autophagy-defective mutants of *Saccharomyces cerevisiae*. *FEBS Letters*, 333(1-2):169–174, 1993.

[23] K. Takeshige, M. Baba, S. Tsuboi, T. Noda, and Y. Ohsumi. Autophagy in yeast demonstrated with proteinase-deficient mutants and conditions for its induction. *The Journal of Cell Biology*, 119(2):301–311, Oct 1992. Research Support, Non-U.S. Gov't.

[24] M. Thumm, R. Egner, B. Koch, M. Schlumpberger, M. Straub, M. Veenhuis, and D.H. Wolf. Isolation of autophagocytosis mutants of *Saccharomyces cerevisiae*. *FEBS Letters*, 349(2):275–280, 1994.

[25] Akiko Kuma, Masahiko Hatano, Makoto Matsui, Akitsugu Yamamoto, Haruaki Nakaya, Tamotsu Yoshimori, Yoshinori Ohsumi, Takeshi Tokuhisa, and Noboru Mizushima. The role of autophagy during the early neonatal starvation period. *Nature*, 432(7020):1032–1036, 2004.

[26] Noboru Mizushima, Hisao Sugita, Tamotsu Yoshimori, and Yoshinori Ohsumi. A New Protein Conjugation System in Human: THE COUNTERPART OF THE YEAST Apg12p CONJUGATION SYSTEM ESSENTIAL FOR AUTOPHAGY. *Journal of Biological Chemistry*, 273(51):33889–33892, 1998.

[27] R. S. Marshall, F. Li, D. C. Gemperline, A. J. Book, and R. D. Vierstra. Autophagic degradation of the 26S proteasome is mediated by the dual ATG8/ubiquitin receptor RPN10 in Arabidopsis. *Molecular Cell*, 58(6):1053–1066, 2015.

[28] R. S. Marshall, F. McLoughlin, and R. D. Vierstra. Autophagic turnover of inactive 26S proteasomes in yeast is directed by the ubiquitin receptor Cue5 and the Hsp42 chaperone. *Cell Reports*, 16(7):1717–1732, 2016.

[29] K. A. Waite, A. De La Mota-Peynado, G. Vontz, and J. Roelofs. Starvation induces proteasome autophagy with different pathways for core and regulatory particles. *Journal of Biological Chemistry*, 291(7):3239–3253, 2016.

[30] Shaoning Jiang, Dae Won Park, Yong Gao, Saranya Ravi, Victor Darley-Usmar, Edward Abraham, and Jaroslaw W. Zmijewski. Participation of proteasome-ubiquitin protein degradation in autophagy and the activation of AMP-activated protein kinase. *Cellular Signalling*, 27(6):1186–1197, 2015.

[31] Manuela Antonioli, Federica Albiero, Francesca Nazio, Tiziana Vescovo, Ariel Basulto Perdomo, Marco Corazzari, Claudia Marsella, Pierluca Piselli, Christine Gretzmeier, Jörn Dengjel, Francesco Cecconi, Mauro Piacentini, and Gian Maria Fimia. AMBRA1 Interplay with Cullin E3 Ubiquitin Ligases Regulates Autophagy Dynamics. *Developmental Cell*, 31(6):734–746, 2014.

[32] Chin-Chih Liu, Yu-Ching Lin, Yu-Hsuan Chen, Chun-Ming Chen, Liang-Yu Pang, Hsuan-An Chen, Pei-Rung Wu, Mei-Yao Lin, Si-Tse Jiang, Ting-Fen Tsai, and Ruey-Hwa Chen. Cul3-KLHL20 Ubiquitin Ligase Governs the Turnover of ULK1 and VPS34 Complexes to Control Autophagy Termination. *Molecular Cell*, 61(1):84–97, 2016.

[33] Zhonghua Gao, Noor Gammoh, Pui-Mun Wong, Hediye Erdjument-Bromage, Paul Tempst, and Xuejun Jiang and. Processing of autophagic protein LC3 by the 20S proteasome. *Autophagy*, 6(1):126–137, 2010. PMID: 20061800.

[34] Xiao J Wang, Jun Yu, Sunny H Wong, Alfred SL Cheng, Francis KL Chan, Simon SM Ng, Chi H Cho, Joseph JY Sung, and William KK Wu and. A novel crosstalk between two major protein degradation systems. *Autophagy*, 9(10):1500–1508, 2013. PMID: 23934082.

[35] Michael Lazarou, Danielle A. Sliter, Lesley A. Kane, Shireen A. Sarraf, Chunxin Wang, Jonathon L. Burman, Dionisia P. Sideris, Adam I. Fogel, and Richard J. Youle. The ubiquitin kinase PINK1 recruits autophagy receptors to induce mitophagy. *Nature*, 524(7565):309–314, 2015.

[36] Kenneth L. Rock, Colette Gramm, Lisa Rothstein, Karen Clark, Ross Stein, Lawrence Dick, Daniel Hwang, and Alfred L. Goldberg. Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. *Cell*, 78(5):761–771, 1994.

[37] Thomas Sommer and Dieter H. Wolf. Endoplasmic reticulum degradation: reverse protein flow of no return. *The FASEB Journal*, 11(14):1227–1233, 1997.

[38] L Gregori, M S Poosch, G Cousins, and V Chau. A uniform isopeptide-linked multiubiquitin chain is sufficient to target substrate for degradation in ubiquitin-mediated proteolysis. *Journal of Biological Chemistry*, 265(15):8354–8357, 1990.

[39] C M Pickart and A T Vella. Ubiquitin carrier protein-catalyzed ubiquitin transfer to histones. Mechanism and specificity. *Journal of Biological Chemistry*, 263(29):15076–15082, 1988.

[40] Avram Hershko and Aaron Ciechanover. The ubiquitin system. *Annual Review of Biochemistry*, 67(Volume 67, 1998):425–479, 1998.

[41] David Komander and Michael Rape. The ubiquitin code. *Annual Review of Biochemistry*, 81:203–229, 2012.

[42] Liang Jin, Amanda Williamson, Suman Banerjee, Ines Philipp, and Marc Rape. Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex. *Cell*, 133(4):653–665, May 2008.

[43] Julia S. Thrower, Laura Hoffman, Martin Rechsteiner, and Cecile M. Pickart. Recognition of the polyubiquitin proteolytic signal. *The EMBO Journal*, 19(1):94–102, 2000.

[44] Linda Hicke. Protein regulation by monoubiquitin. *Nature Reviews Molecular Cell Biology*, 2(3):195–201, 2001.

[45] Dhritiman Mukhopadhyay and Henrik Riezman. Proteasome-independent functions of ubiquitin in endocytosis and signaling. *Science*, 315(5809):201–205, 2007.

[46] Elsa Lauwers, Christophe Jacob, and Bruno André. K63-linked ubiquitin chains as a specific signal for protein sorting into the multivesicular body pathway. *Journal of Cell Biology*, 185(3):493–502, 04 2009.

[47] J. M. Galan and R. Haguenauer-Tsapis. Ubiquitin lys63 is involved in ubiquitination of a yeast plasma membrane protein. *EMBO Journal*, 16:5847–5854, 1997.

[48] J.-Y. Springael, J.-M. Galan, R. Haguenauer-Tsapis, and B. André. NH4+-induced down-regulation of the Saccharomyces cerevisiae Gap1p permease involves its ubiquitination with lysine-63-linked chains. *Journal of Cell Science*, 112:1375–1383, 1999.

[49] Z. Erpapazoglou, M. Froissard, I. Nondier, E. Lesuisse, R. Haguenauer-Tsapis, and N. Belgareh-Touze. Substrate and ubiquitin-dependent trafficking of the yeast siderophore transporter Sit1. *Traffic*, 9:1372–1391, 2008.

[50] F. Huang, D. Kirkpatrick, X. Jiang, S. Gygi, and A. Sorkin. Differential regulation of EGF receptor internalization and degradation by multiubiquitination within the kinase domain. *Molecular Cell*, 21:737–748, 2006.

[51] L. M. Duncan, S. Piper, R. B. Dodd, M. K. Saville, C. M. Sanderson, J. P. Luzio, and P. J. Lehner. Lysine-63-linked ubiquitination is required for endolysosomal degradation of class I molecules. *EMBO Journal*, 25:1635–1645, 2006.

[52] Takayoshi Kirisako, Kiyoko Kamei, Shigeo Murata, Michiko Kato, Hiromi Fukumoto, Masato Kanie, Soichi Sano, Fuminori Tokunaga, Keiji Tanaka, and Kazuhiro Iwai. A ubiquitin ligase complex assembles linear polyubiquitin chains. *The EMBO Journal*, 25(20):4877–4887, 2006.

[53] Fuminori Tokunaga, Shin ichi Sakata, Yasushi Saeki, Yoshinori Satomi, Takayoshi Kirisako, Kiyoko Kamei, Tomoko Nakagawa, Michiko Kato, Shigeo Murata, Shoji Yamaoka, Masahiro Yamamoto, Shizuo Akira, Toshifumi Takao, Keiji Tanaka, and Kazuhiro Iwai. Involvement of linear polyubiquitylation of NEMO in NF-κB activation. *Nature Cell Biology*, 11(2):123–132, 2009.

[54] A Hershko, H Heller, S Elias, and A Ciechanover. Components of ubiquitin-protein ligase system. resolution, affinity purification, and role in protein breakdown. *Journal of Biological Chemistry*, 258(13):8206–8214, 1983.

[55] Avram Hershko and Aaron Ciechanover. THE UBIQUITIN SYSTEM. *https://doi.org/10.1146/annurev.biochem.67.1.425*, 67:425–479, nov 2003.

[56] Karin Breitschopf, Eyal Bengal, Tamar Ziv, Arie Admon, and Aaron Ciechanover. A novel site for ubiquitination: the N-terminal residue, and not internal lysines of MyoD, is essential for conjugation and degradation of the protein. *The EMBO Journal*, 17(20):5964–5973, 1998.

[57] Lionel A. Tintignac, Julie Lagirand, Sabrina Batonnet, Valentina Sirri, Marie Pierre Leibovitch, and Serge A. Leibovitch. Degradation of MyoD Mediated by the SCF (MAFbx) Ubiquitin Ligase. *Journal of Biological Chemistry*, 280(4):2847–2856, 2005.

[58] Peng Gao, Wei Wang, Wenjing Zhao, Feng Yu, Tao Liu, Yujie Du, Xin Chen, Jiang Sun, Xiaojun Wang, Lianfeng Zhang, and Zhigang Jiang. E3 ligase Nedd4l promotes antiviral innate immunity by catalyzing K29-linked cysteine ubiquitination of TRAF3. *Nature Communications*, 12:1194, 2021.

[59] Valentina G. Anania, Daniel J. Bustos, Jennifer R. Lill, Derek S. Kirkpatrick, and Laurent Coscoy. A novel peptide-based SILAC method to identify the posttranslational modifications provides evidence for unconventional ubiquitination in the ER-associated degradation pathway. *International Journal of Proteomics*, page 857918, 2013.

[60] Yuki Shimizu, Yuko Okuda-Shimizu, and Linda M. Hendershot. Ubiquitylation of an ERAD substrate occurs on multiple types of amino acids. *Molecular Cell*, 40:917–926, 2010.

[61] Satoshi Ishikura, Alan M. Weissman, and Juan S. Bonifacino. Serine residues in the cytosolic tail of the T-cell antigen receptor alpha-chain mediate ubiquitination and endoplasmic reticulum-associated degradation of the unassembled protein. *The Journal of Biological Chemistry*, 285:23916–23924, 2010.

[62] Michael L. Burr et al. MHC class I molecules are preferentially ubiquitinated on endoplasmic reticulum luminal residues during HRD1 ubiquitin E3 ligase-mediated dislocation. *Proceedings of the National Academy of Sciences of the United States of America*, 110:14290–14295, 2013.

[63] Kevin H. Wang, Giselle Roman-Hernandez, Robert A. Grant, Robert T. Sauer, and Tania A. Baker. The Molecular Basis of N-end Rule Recognition. *Molecular cell*, 32(3):406, nov 2008.

[64] Kuan-Chuan Pao, Mathew Stanley, Cong Han, Yu-Chiang Lai, Paul Murphy, Kristin Balk, Nicola T Wood, Olga Corti, Jean-Christophe Corvol, Miratul M K Muqit, and Satpal Virdee. Probes of ubiquitin E3 ligases enable systematic dissection of parkin activation. *Nature Chemical Biology*, 2016.

[65] I. R. Kelsall, J. Zhang, A. Knebel, J. S. C. Arthur, and P. Cohen. The E3 ligase HOIL-1 catalyses ester bond formation between ubiquitin and components of the Myddosome in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, 116(27):13293–13298, 2019.

[66] Sven M. Lange, Lee A. Armstrong, and Yogesh Kulathu. Deubiquitinases: From mechanisms to their inhibition by small molecules. *Molecular Cell*, 82(1):15–29, 2022.

[67] Michael J. Clague, Sylvie Urbé, and David Komander. Breaking the chains: deubiquitylating enzyme specificity begets function. *Nature Reviews Molecular Cell Biology*, 20(6):338–352, 2019.

[68] Yu Ye, Masato Akutsu, Francisca Reyes-Turcu, Radoslav I Enchev, Keith D Wilkinson, and David Komander. Polyubiquitin binding and cross-reactivity in the USP domain deubiquitinase USP21. *EMBO reports*, 12(4):350–357, 2011.

[69] Berthe Katrine Fiil, Rune Busk Damgaard, Sebastian Alexander Wagner, Kirstin Keusekotten, Melanie Fritsch, Simon Bekker-Jensen, Niels Mailand, Chunaram Choudhary, David Komander, and Mads Gyrd-Hansen. OTULIN restricts Met1-linked ubiquitination to control innate immune signaling. *Molecular Cell*, 50(6):818–830, June 2013.

[70] Rati Verma, L. Aravind, Robert Oania, W. Hayes McDonald, John R. Yates III, Eugene V. Koonin, and Raymond J. Deshaies. Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. *Science*, 298(5593):611–615, October 2002.

[71] Matthieu Schapira, Matthew F. Calabrese, Alex N. Bullock, and Craig M. Crews. Targeted protein degradation: expanding the toolbox. *Nature Reviews Drug Discovery 2019 18:12*, 18(12):949–963, oct 2019.

[72] Zhuoyao Chen, Gregory A. Wasney, Sarah Picaud, Panagis Filippakopoulos, Masoud Vedadi, Vincenzo D'Angiolella, and Alex N. Bullock. Identification of a PGXPP degron motif in dishevelled and structural basis for its binding to the E3 ligase KLHL12. *Open Biology*, 10(6):200041, jun 2020.

[73] Recognition of the Diglycine C-End Degron by $CRL2^{KLHDC2}$ Ubiquitin Ligase.

[74] Alexander Varshavsky. N-degron and C-degron pathways of protein degradation. *Proceedings of the National Academy of Sciences of the United States of America*, 116(2):358–366, 2019.

[75] Geng Wu, Guozhou Xu, Brenda A. Schulman, Philip D. Jeffrey, J. Wade Harper, and Nikola P. Pavletich. Structure of a β-TrCP1-Skp1-β-Catenin Complex: Destruction Motif Binding and Lysine Specificity of the SCFβ-TrCP1 Ubiquitin Ligase. *Molecular Cell*, 11(6):1445–1456, 2003.

[76] Shiyun Cao, Sheena Faye Garcia, Huigang Shi, Ellie I. James, Yuki Kito, Hui Shi, Haibin Mao, Sharon Kaisari, Gergely Rona, Sophia Deng, Hailey V. Goldberg, Jackeline Ponce, Beatrix Ueberheide, Luca Lignitto, Miklos Guttman, Michele Pagano, and Ning Zheng. Recognition of BACH1 quaternary structure degrons by two F-box proteins under oxidative stress. *Cell*, 187(26):7568–7584.e22, 2024.

[77] Laura B. Sheard, Xu Tan, Haibin Mao, John Withers, Gili Ben-Nissan, Thomas R. Hinds, Yuichi Kobayashi, Fong-Fu Hsu, Michal Sharon, John Browse, Sheng Yang He, Josep Rizo, Gregg A. Howe, and Ning Zheng. Jasmonate perception by inositol-phosphate-potentiated COI1-JAZ co-receptor. *Nature*, 468(7322):400–405, 2010.

[78] Xu Tan, Luz Irina A. Calderon-Villalobos, Michal Sharon, Changxue Zheng, Carol V. Robinson, Mark Estelle, and Ning Zheng. Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature*, 446(7136):640–645, 2007.

[79] Elijah L. Mena, Predrag JevtiÄ, Basil J. Greber, Christine L. Gee, Brandon G. Lew, David Akopian, Eva Nogales, John Kuriyan, and Michael Rape. Structural basis for dimerization quality control. *Nature*, 586(7829):452–456, 2020.

[80] Edna Matta-Camacho, Guennadi Kozlov, Flora F. Li, and Kalle Gehring. Structural basis of substrate recognition and specificity in the N-end rule pathway. *Nature Structural & Molecular Biology*, 17(10):1182–1187, 2010.

[81] L. Schrödinger. Pymol: Molecular visualization system. Version 1.8, 2015.

[82] C. W. Liew, H. Sun, T. Hunter, and C. L. Day. RING domain dimerization is essential for RNF4 function. *Biochemical Journal*, 431(1):23–29, 2010.

[83] Fei Teng, Yang Wang, Ming Liu, Shuyun Tian, Goran Stjepanovic, and Ming-Yuan Su. Cryo-EM structure of the KLHL22 E3 ligase bound to an oligomeric metabolic enzyme. *Structure*, 31(11):1431–1440.e5, 2023.

[84] Peter Canning, Fiona J. Sorrell, and Alex N. Bullock. Structural basis of Keap1 interactions with Nrf2. *Free Radical Biology & Medicine*, 88(Pt B):101, 2015.

[85] Jinwoo Ahn, Zach Novince, Jason Concel, Chang-Hyeock Byeon, Alexander M. Makhov, In-Ja L. Byeon, Peijun Zhang, and Angela M. Gronenborn. The Cullin-RING E3 ubiquitin ligase CRL4-DCAF1 complex dimerizes via a short helical region in DCAF1. *Biochemistry*, 50(8):1359, mar 2011.

[86] Weaam I Mohamed, Andreas D Schenk, Georg Kempf, Simone Cavadini, Anja Basters, Alessandro Potenza, Wassim Abdul Rahman, Julius Rabl, Kurt Reichermeier, and Nicolas H Thomä. The $CRL4^{DCAF1}$ cullin-RING ubiquitin ligase is activated following a switch in oligomerization state. *The EMBO Journal*, 40(22):e108008, 2021.

[87] D. C. Scott, M. T. King, K. Baek, C. T. Gee, R. Kalathur, J. Li, N. Purser, A. Nourse, S. C. Chai, S. Vaithiyalingam, T. Chen, R. E. Lee, S. J. Elledge, G. Kleiger, and B. A. Schulman. E3 ligase autoinhibition by C-degron mimicry maintains C-degron substrate fidelity. *Molecular Cell*, 83(5):770–786.e9, 2023.

[88] Human Protein Atlas. KLHL12 protein expression profile, https://www.proteinatlas.org/ensg00000117153-klhl12, 2025. Accessed: 2025-02-23.

[89] Stephane Angers, Chris J. Thorpe, Travis L. Biechele, Seth J. Goldenberg, Ning Zheng, Michael J. MacCoss, and Randall T. Moon. The KLHL12-Cullin-3 ubiquitin ligase negatively regulates the Wnt-β-catenin pathway by targeting Dishevelled for degradation. *Nature Cell Biology 2006 8:4*, 8(4):348–357, mar 2006.

[90] Adnan Shami Shah, Alex G. Batrouni, Dongsung Kim, Amith Punyala, Wendy Cao, Chun Han, Michael L. Goldberg, Marcus B. Smolka, and Jeremy M. Baskin. PLEKHA4/kramer Attenuates Dishevelled Ubiquitination to Modulate Wnt and Planar Cell Polarity Signaling. *Cell Reports*, 27(7):2157–2170.e8, may 2019.

[91] Marek Mlodzik. The Dishevelled Protein Family: Still Rather a Mystery After Over 20 Years of Molecular Studies. *Current topics in developmental biology*, 117:75–91, dec 2016.

[92] Hans Clevers and Roel Nusse. Wnt/β-catenin signaling and disease. *Cell*, 149(6):1192–1205, jun 2012.

[93] Chutikarn Butkinaree, Liang Guo, Bhama Ramkhelawon, Amarylis Wanschel, Jeffrey L. Brodsky, Kathryn J. Moore, and Edward A. Fisher. A regulator of secretory vesicle size, Kelch-like protein 12, facilitates the secretion of apolipoprotein B100 and very-low-density lipoproteins–brief report. *Arteriosclerosis, thrombosis, and vascular biology*, 34(2):251–254, feb 2014.

[94] Lingyan Jin, Kanika Bajaj Pahuja, Katherine E. Wickliffe, Amita Gorur, Christine Baumgärtel, Randy Schekman, and Michael Rape. Ubiquitin-dependent regulation of COPII coat size and function. *Nature 2012 482:7386*, 482(7386):495–500, feb 2012.

[95] Colleen A. McGourty, David Akopian, Carolyn Walsh, Amita Gorur, Achim Werner, Randy Schekman, Diana Bautista, and Michael Rape. Regulation of the CUL3 Ubiquitin Ligase by a Calcium-Dependent Co-adaptor. *Cell*, 167(2):525–538.e14, oct 2016.

[96] Malin Sund, Timo Väisänen, Sami Kaukinen, Mika Ilves, Hongmin Tu, Helena Autio-Harmainen, Heikki Rauvala, and Taina Pihlajaniemi. Distinct expression of type XIII collagen in neuronal structures and other tissues during mouse development. *Matrix Biology*, 20(4):215–231, 2001.

[97] J Su, K Gorse, F Ramirez, and MA Fox. Collagen XIX is expressed by interneurons and contributes to the formation of hippocampal synapses. *Journal of Comparative Neurology*, 518(2):229–253, 2010.

[98] Michael A. Fox, Joshua R. Sanes, Dorin-Bogdan Borza, Veraragavan P. Eswarakumar, Reinhard Fässler, Billy G. Hudson, Simon W.M. John, Yoshifumi Ninomiya, Vadim Pedchenko, Samuel L. Pfaff, Michelle N. Rheault, Yoshikazu Sado, Yoav Segal, Michael J. Werle, and Hisashi Umemori. Distinct target-derived signals organize formation, maturation, and maintenance of motor nerve terminals. *Cell*, 129(1):179–193, 2007.

[99] I. Gregorio, P. Braghetta, P. Bonaldo, and M. Cescon. Collagen VI in healthy and diseased nervous system. *Disease Models Mechanisms*, 11:dmm032946, 2018.

[100] S. A. Boyadjiev et al. Cranio-lenticulo-sutural dysplasia is caused by a SEC23A mutation leading to abnormal endoplasmic-reticulum-to-Golgi trafficking. *Nature Genetics*, 38:1192–1197, 2006.

[101] Audrey M. G. Ragagnin, Vinod Sundaramoorthy, Fabiha Farzana, Shashi Gautam, Sayanthooran Saravanabavan, Zeinab Takalloo, Prachi Mehta, Dzung Do-Ha, Sonam Parakh, Sina Shadfar, Julie Hunter, Marta Vidal, Cyril J. Jagaraj, Mariana Brocardo, Anna Konopka, Shu Yang, Stephanie L. Rayner, Kelly L. Williams, Ian P. Blair, Roger S. Chung, Albert Lee, Lezanne Ooi, and Julie D. Atkin. ALS/FTD-associated mutation in cyclin F inhibits ER-Golgi trafficking, inducing ER stress, ERAD and Golgi fragmentation. *Scientific Reports*, 13(1):20467, November 2023.

[102] Laurensia Yuniati, Angela Lauriola, Manouk Gerritsen, Susana Abreu, Eric Ni, Chiara Tesoriero, Jacob O. Onireti, Teck Yew Low, Albert J.R. Heck, Andrea Vettori, Timothy Cardozo, and Daniele Guardavaccaro. Ubiquitylation of the ER-Shaping Protein Lunapark via the CRL3KLHL12 Ubiquitin Ligase Complex. *Cell Reports*, 31(7):107664, may 2020.

[103] Pieter Rondou, Guy Haegeman, Peter Vanhoenacker, and Kathleen Van Craenenbroeck. BTB Protein KLHL12 Targets the Dopamine D4 Receptor for Ubiquitination by a Cul3-based E3 Ligase. *Journal of Biological Chemistry*, 283(17):11083–11096, apr 2008.

[104] Pieter Rondou, Kamila Skieterska, Ann Packeu, Béatrice Lintermans, Peter Vanhoenacker, Georges Vauquelin, Guy Haegeman, and Kathleen Van Craenenbroeck. KLHL12-mediated ubiquitination of the dopamine D4 receptor does not target the receptor for degradation. *Cellular Signalling*, 22(6):900–913, jun 2010.

[105] Kamila Skieterska, Pieter Rondou, Béatrice Lintermans, and Kathleen Van Craenenbroeck. KLHL12 Promotes Non-Lysine Ubiquitination of the Dopamine Receptors D4.2 and D4.4, but Not of the ADHD-Associated D4.7 Variant. *PLOS ONE*, 10(12):e0145654, dec 2015.

[106] Peter Canning, Christopher D. O. Cooper, Tobias Krojer, James W. Murray, Ashley C. W. Pike, Apirat Chaikuad, Tracy Keates, Chancievan Thangaratnarajah, Viktorija Hojzan, Brian D. Marsden, Opher Gileadi, Stefan Knapp, Frank von Delft, and Alex N. Bullock. Structural Basis for Cul3 Protein Assembly with the BTB-Kelch Family of E3 Ubiquitin Ligases. *The Journal of Biological Chemistry*, 288(11):7803, mar 2013.

[107] Alan X. Ji and Gilbert G. Privé. Crystal Structure of KLHL3 in Complex with Cullin3. *PLOS ONE*, 8(4):e60445, apr 2013.

[108] Min Zhuang, Matthew F. Calabrese, Jiang Liu, M. Brett Waddell, Amanda Nourse, Michal Hammel, Darcie J. Miller, Helen Walden, David M. Duda, Steven N. Seyedin, Timothy Hoggard, J. Wade Harper, Kevin P. White, and Brenda A. Schulman. Structures of SPOP-Substrate Complexes: Insights into Molecular Architectures of BTB-Cul3 Ubiquitin Ligases. *Molecular Cell*, 36(1):39–50, oct 2009.

[109] Peter J. Stogios and Gilbert G. Privé. The BACK domain in BTB-kelch proteins. *Trends in Biochemical Sciences*, 29(12):634–637, 12 2004.

[110] K. Farid Ahmad, Christian K. Engel, and Gilbert G. Privé. Crystal structure of the BTB domain from PLZF. *Proceedings of the National Academy of Sciences*, 95(21):12123–12128, 10 1998.

[111] Lionel Pintard, Andrew Willems, and Matthias Peter. Cullin-based ubiquitin ligases: Cul3-BTB complexes join the family. *The EMBO Journal*, 23(8):1681, 4 2004.

[112] Elena L. Mena, Rune A. S. Kjolby, Rachel A. Saxton, Amanda Werner, Ben G. Lew, James M. Boyle, Richard Harland, and Michael Rape. Dimerization quality control ensures neuronal development and survival. *Science*, 362(6411):eaap8236, 2018.

[113] Daniel C. Scott, Vladislav O. Sviderskiy, JulieK. Monda, JohnR. Lydeard, Shein Ei Cho, J.Wade Harper, and Brenda A. Schulman. Structure of a RING E3 Trapped in Action Reveals Ligation Mechanism for the Ubiquitin-like Protein NEDD8. *Cell*, 157(7):1671–1684, 2014.

[114] Bin Zhao, William G. Payne, Jiqing Sai, Zhenwei Lu, Edward T. Olejniczak, and Stephen W. Fesik. Structural Elucidation of Peptide Binding to KLHL-12, a Substrate Specific Adapter Protein in a Cul3-Ring E3 Ligase Complex. *Biochemistry*, 59(8):964–969, mar 2020.

[115] Eugene Krissinel and Kim Henrick. Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology*, 372(3):774–797, 2007.

[116] Jerome Boudeau, Diego Miranda-Saavedra, Geoffrey J. Barton, and Dario R. Alessi. Emerging roles of pseudokinases. *Trends in Cell Biology*, 16(9):443–452, 2006.

[117] Jason S. Kerr and Catherine H. Wilson. Nuclear receptor-binding protein 1: a novel tumour suppressor and pseudokinase. *Biochemical Society Transactions*, 41(4):1055–1060, aug 2013.

[118] Raelene Lim, Louise N Winteringham, James H Williams, Ross K McCulloch, Evan Ingley, Jim Y-H. Tiao, Jean-Philippe Lalonde, Schickwann Tsai, Peta A Tilbrook, Yi Sun, Xiaohua Wu, Stephan W Morris, and S Peter Klinken. MADM, a Novel Adaptor Protein That Mediates Phosphorylation of the 14-3-3 Binding Site of Myeloid Leukemia Factor 1. *Journal of Biological Chemistry*, 277(43):40997–41008, oct 2002.

[119] Wang H, Sun X, Luo Y, Lin Z, and Wu J. Adapter protein NRBP associates with Jab1 and negatively regulates AP-1 activity. *FEBS letters*, 580(25):6015–6021, oct 2006.

[120] Nawel Mahrour, William B. Redwine, Laurence Florens, Selene K. Swanson, Skylar Martin-Brown, William D. Bradford, Karen Staehling-Hampton, Michael P. Washburn, Ronald C. Conaway, and Joan W. Conaway. Characterization of Cullin-box Sequences That Direct Recruitment of Cul2-Rbx1 and Cul5-Rbx2 Modules to Elongin BC-based Ubiquitin Ligases. *Journal of Biological Chemistry*, 283(12):8005–8013, mar 2008.

[121] Silvia Gluderer, Erich Brunner, Markus Germann, Virginija Jovaisaite, Changqing Li, Cyrill A. Rentsch, Ernst Hafen, and Hugo Stocker. Madm (Mlf1 adapter molecule) cooperates with Bunched A to promote growth in *Drosophila*. *Journal of Biology*, 9(1):9, 2010.

[122] Catherine H Wilson, Catriona Crombie, Louise van der Weyden, George Poulogiannis, Alistair G Rust, Mercedes Pardo, Tannia Gracia, Lu Yu, Jyoti Choudhary, Gino B Poulin, Rebecca E McIntyre, Douglas J Winton, H Nikki March, Mark J Arends, Andrew G Fraser, and David J Adams. Nuclear receptor binding protein 1 regulates intestinal progenitor cell homeostasis and tumour formation. *The EMBO Journal*, 31(11):2486–2497, may 2012.

[123] Madhuri Warren, Wei Wang, Sarah Spiden, Dongrong Chen-Murchie, David Tannahill, Karen P. Steel, and Allan Bradley. A Sall4 mutant mouse model useful for studying the role of Sall4 in early embryonic development and organogenesis. *genesis*, 45(1):51–58, 2007.

[124] T. Aso, W. S. Lane, J. W. Conaway, and R. C. Conaway. Elongin (SIII): a multisubunit regulator of elongation by RNA polymerase II. *Science*, 269(5229):1439–1443, 1995.

[125] T. Kamura, S. Sato, D. Haque, L. Liu, W. G. Jr. Kaelin, R. C. Conaway, and J. W. Conaway. The Elongin BC complex interacts with the conserved SOCS-box motif present in members of the SOCS, ras, WD-40 repeat, and ankyrin repeat families. *Genes Development*, 12(24):3872–3881, 1998.

[126] Takashi Yasukawa, Aya Tsutsui, Chieri Tomomori-Sato, Shigeo Sato, Anita Saraf, Michael P. Washburn, Laurence Florens, Tohru Terada, Kentaro Shimizu, Ronald C. Conaway, Joan W. Conaway, and Teijiro Aso. NRBP1-Containing CRL2/CRL4A Regulates Amyloid β Production by Targeting BRI2 and BRI3 for Degradation. *Cell Reports*, 30(10):3478–3491.e6, mar 2020.

[127] T. Li, E. I. Robert, P. C. van Breugel, M. Strubin, and N. Zheng. A promiscuous alpha-helical motif anchors viral hijackers and substrate receptors to the CUL4-DDB1 ubiquitin ligase machinery. *Nature Structural & Molecular Biology*, 17(1):105–111, 2010. Epub 2009 Dec 6.

[128] Huang MW, Lin YJ, Chang CW, Lei FJ, Ho EP, Liu RS, Shyu WC, and Hsieh CH. RGS4 deficit in prefrontal cortex contributes to the behaviors related to schizophrenia via system x c–mediated glutamatergic dysfunction in mice. *Theranostics*, 8(17):4781–4794, 2018.

[129] Myung Hee Kim, David R. Cooper, Arkadiusz Oleksy, Yancho Devedjiev, Urszula Derewenda, Orly Reiner, Jacek Otlewski, and Zygmunt S. Derewenda. The Structure of the N-Terminal Domain of the Product of the Lissencephaly Gene Lis1 and Its Functional Implications. *Structure*, 12(6):987–998, 2004.

[130] Robert Eferl and Erwin F. Wagner. AP-1: a double-edged sword in tumorigenesis. *Nature Reviews Cancer*, 3(11):859–868, 2003.

[131] Ee Hong Tan and Owen James Sansom. A new tumour suppressor enters the network of intestinal progenitor cell homeostasis. *The EMBO Journal*, 31(11):2444–2445, 2012.

[132] Yasuhito Okano, Hiroyuki Suzuki, Yukihide Watanabe, Mohammed Abdelaziz, Lev Manevich, Kunio Kawanishi, Haruka Ozaki, Ryota Ishii, Shin Matsumoto, Nohara Goto, Ling Zheng, Yukari Okita, Jongchan Hwang, Masahiro Nakayama, Yoshihide Shima, Noriaki Sakamoto, Masayuki Noguchi, Keiji Tabuchi, and Mitsuyasu Kato. THG-1/TSC22D4 promotes interleukin-1 signaling through stabilization of traf6 in squamous cell carcinoma. *Molecular Cancer Research*, 01 2025.

[133] Shuji Matsuda, Luca Giliberto, Yukiko Matsuda, Eileen M. McGowan, and Luciano D'Adamio. BRI2 Inhibits Amyloid β-Peptide Precursor Protein Processing by Interfering with the Docking of Secretases to the Substrate. *Journal of Neuroscience*, 28(35):8668–8676, aug 2008.

[134] Daniel B. Rosoff, Josephin Wagner, Andrew S. Bell, Lucas A. Mavromatis, Jeesun Jung, and Falk W. Lohoff. A multi-omics mendelian randomization study identifies new therapeutic targets for alcohol use disorder and problem drinking. *Nature Human Behaviour*, 9(1):188–207, January 2025.

[135] Alice Shi Ming Li, Serah Kimani, Brian Wilson, Mahmoud Noureldin, Héctor González-Álvarez, Ahmed Mamai, Laurent Hoffer, John P. Guilinger, Ying Zhang, Moritz von Rechenberg, Jeremy S. Disch, Christopher J. Mulhern, Belinda L. Slakman, John W. Cuozzo, Aiping Dong, Gennady Poda, Mohammed Mohammed, Punit Saraon, Manish Mittal, Pratik Modh, Vaibhavi Rathod, Bhashant Patel, Suzanne Ackloo, Vijayaratnam Santhakumar, Magdalena M. Szewczyk, Dalia Barsyte-Lovejoy, Cheryl H. Arrowsmith, Richard Marcellus, Marie-Aude Guie, Anthony D. Keefe, Peter J. Brown, Levon Halabelian, Rima Al-awar, and Masoud

Vedadi. Discovery of Nanomolar DCAF1 Small Molecule Ligands. *Journal of Medicinal Chemistry*, 66(7):5041–5060, 2023. Published online: 2023-04-13.

[136] M. Li, R. Ding, X. Yang, and D. Ran. Study on Biomarkers Related to the Treatment of Post-Stroke Depression and Alternative Medical Treatment Methods. *Neuropsychiatric Disease and Treatment*, 18:1861–1873, Aug 26 2022.

[137] L. Y. Zhang, Y. H. Chu, Y. F. You, M. H. Dong, X. W. Pang, L. Chen, L. F. Zhu, S. Yang, L. Q. Zhou, K. Shang, G. Deng, J. Xiao, W. Wang, C. Qin, and D. S. Tian. Systematic druggable genome-wide mendelian randomization identifies therapeutic targets for functional outcome after ischemic stroke. *Journal of the American Heart Association*, 13(16):e034749, Aug 20 2024. Epub 2024 Aug 9.

[138] Yu-Xi Xiao, Seon Yong Lee, Magali Aguilera-Uribe, Reuben Samson, Aaron Au, Yukti Khanna, Zetao Liu, Ran Cheng, Kamaldeep Aulakh, Jiarun Wei, Adrian Granda Farias, Taylor Reilly, Saba Birkadze, Andrea Habsid, Kevin R. Brown, Katherine Chan, Patricia Mero, Jie Qi Huang, Maximilian Billmann, Mahfuzur Rahman, Chad Myers, Brenda J. Andrews, Ji-Young Youn, Christopher M. Yip, Daniela Rotin, W. Brent Derry, Julie D. Forman-Kay, Alan M. Moses, Iva Pritišanac, Anne-Claude Gingras, and Jason Moffat. The TSC22D, WNK, and NRBP gene families exhibit functional buffering and evolved with metazoa for cell volume regulation. *Cell Reports*, 43(7):114417, 2024.

[139] Mingyuan Fan, Zhangjun Yun, Jiushu Yuan, Sai Zhang, Hongyan Xie, Dingyi Lu, Haipo Yuan, and Hong Gao. Genetic insights into therapeutic targets for gout: evidence from a multi-omics Mendelian randomization study. *Hereditas*, 161(1):56, 2024.

[140] H. Li, W. Li, D. Li, L. Yuan, Y. Xu, P. Su, L. Wu, and Z. Zhang. Based on systematic druggable genome-wide Mendelian randomization identifies therapeutic targets for diabetes. *Frontiers in Endocrinology*, 15:1366290, 2024.

[141] Alexander Varshavsky. The N-end rule pathway and regulation by proteolysis. *Protein science : a publication of the Protein Society*, 20(8):1298–1345, aug 2011.

[142] Michael Rape. Ubiquitylation at the crossroads of development and disease. *Nature Reviews Molecular Cell Biology*, 19(1):59–70, 2018.

[143] Cheryl H Arrowsmith, James E Audia, Christopher Austin, Jonathan Baell, Jonathan Bennett, Julian Blagg, Chas Bountra, Paul E Brennan, Peter J Brown, Mark E Bunnage, Carolyn Buser-Doepner, Robert M Campbell, Adrian J Carter, Philip Cohen, Robert A Copeland, Ben Cravatt, Jayme L Dahlin, Dashyant Dhanak, Aled M Edwards, Mathias Frederiksen, Stephen V Frye, Nathanael Gray, Charles E Grimshaw, David Hepworth, Trevor Howe, Kilian V M Huber, Jian Jin, Stefan Knapp, Joanne D Kotz, Ryan G Kruger, Derek Lowe, Mary M Mader, Brian Marsden, Anke Mueller-Fahrnow, Susanne Mueller, Ronan C O'Hagan, John P Overington, Dafydd R Owen, Saul H Rosenberg, Ruth Ross, Bryan Roth, Matthieu Schapira, Stuart L Schreiber, Brian Shoichet, Michael Sundstrom, Giulio Superti-Furga, Jack Taunton, Leticia Toledo-Sherman, Chris Walpole, Michael A Walters, Timothy M Willson, Paul Workman, Robert N Young, and William J Zuercher. The promise and peril of chemical probes. *Nature Chemical Biology*, 11(8):536–541, 2015.

[144] William A. Weiss, Susan S. Taylor, and Kevan M. Shokat. Recognizing and exploiting differences between RNAi and small-molecule inhibitors. *Nature Chemical Biology*, 3(12):739–744, 2007.

[145] Gregory M. Simon, Micah J. Niphakis, and Benjamin F. Cravatt. Determining target engagement in living systems. *Nature Chemical Biology*, 9(4):200–205, 2013.

[146] Paul Workman and Ian Collins. Probing the probes: Fitness factors for small molecule tools. *Chemistry Biology*, 17(6):561–577, 2010.

[147] Jonathan Baell and Michael A Walters. Chemistry: Chemical con artists foil drug discovery. *Nature*, 513(7519):481–483, 2014.

[148] John Seidler, Sarah L. McGovern, Thomas N. Doman, and Brian K. Shoichet. Identification and prediction of promiscuous aggregating inhibitors among known drugs. *Journal of Medicinal Chemistry*, 46(21):4477–4486, 2003.

[149] Teresa A.F. Cardote, Morgan S. Gadd, and Alessio Ciulli. Crystal Structure of the Cul2-Rbx1-EloBC-VHL Ubiquitin Ligase Complex. *Structure*, 25(6):901–911.e3, 2017.

[150] T. Kamura, D. M. Koepp, M. N. Conrad, D. Skowyra, R. J. Moreland, O. Iliopoulos, W. S. Lane, W. G. Kaelin, S. J. Elledge, R. C. Conaway, J. W. Harper, and J. W. Conaway. Rbx1, a Component of the VHL Tumor Suppressor Complex and SCF Ubiquitin Ligase. *Science*, 284(5414):657–661, 1999.

[151] John S Schneekloth, Fabiana N Fonseca, Michael Koldobskiy, Amit Mandal, Raymond Deshaies, Kathleen Sakamoto, and Craig M Crews. Chemical Genetic Control of Protein Levels: Selective in Vivo Targeted Degradation. 2004.

[152] Dennis L. Buckley, Jeffrey L. Gustafson, Inge Van-Molle, Anke G. Roth, Hyun Seop Tae, Peter C. Gareiss, William L. Jorgensen, Alessio Ciulli, and Craig M. Crews. Small-molecule inhibitors of the interaction between the E3 ligase VHL and HIF1α. *Angewandte Chemie - International Edition*, 51:11463–11467, 11 2012.

[153] Inge Van Molle, Andreas Thomann, Dennis L. Buckley, Ernest C. So, Steffen Lang, Craig M. Crews, and Alessio Ciulli. Dissecting fragment-based lead discovery at the von Hippel-Lindau protein:Hypoxia Inducible Factor 1α protein-protein interface. *Chemistry Biology*, 19(10):1300–1312, 2012.

[154] D. R. Lynch, S. Perlman, and K. Schadt. Omaveloxolone for the treatment of Friedreich ataxia: clinical trial results and practical considerations. *Expert Review of Neurotherapeutics*, 24(3):251–258, 2024. Epub 2024 Jan 30.

[155] Michael J. Niphakis and Benjamin F. Cravatt. Enzyme inhibitor discovery by activity-based protein profiling. *Annual Review of Biochemistry*, 83:341–377, 2014.

[156] K. C. Pao, N. T. Wood, A. Knebel, K. Rafie, M. Stanley, P. D. Mabbitt, R. Sundaramoorthy, K. Hofmann, D. M. van Aalten, and S. Virdee. Activity-based E3 ligase profiling uncovers an E3 ligase with esterification activity. *Nature*, 556:381–385, 2018.

[157] Juraj Ahel, Adam Fletcher, Daniel B. Grabarczyk, Elisabeth Roitinger, Luiza Deszcz, Anita Lehner, Satpal Virdee, and Tim Clausen. E3 ubiquitin ligase RNF213 employs a non-canonical zinc finger active site and is allosterically regulated by ATP. *bioRxiv*, 2021.

[158] Sunil Mathur, Adam J. Fletcher, Emma Branigan, Ronald T. Hay, and Satpal Virdee. Photocrosslinking Activity-Based Probes for Ubiquitin RING E3 Ligases. *Cell Chemical Biology*, 27(1):74–82.e6, 2020.

[159] Kathleen M. Sakamoto, Kyung B. Kim, Akiko Kumagai, Frank Mercurio, Craig M. Crews, and Raymond J. Deshaies. Protacs: Chimeric molecules that target proteins to the Skp1-Cullin-F box complex for ubiquitination and degradation. *Proceedings of the National Academy of Sciences*, 98(15):8554–8559, 2001.

[160] Georg E. Winter, Dennis L. Buckley, Joshiawa Paulk, Justin M. Roberts, Amanda Souza, Sirano Dhe-Paganon, and James E. Bradner. Phthalimide conjugation as a strategy for *in vivo* target protein degradation. *Science*, 348(6241):1376–1381, 2015.

[161] Dennis L. Buckley, Inge Van Molle, Peter C. Gareiss, Hyun Seop Tae, Julien Michel, Devin J. Noblin, William L. Jorgensen, Alessio Ciulli, and Craig M. Crews. Targeting the von Hippel-Lindau E3 ubiquitin ligase using small molecules to disrupt the VHL/HIF1α interaction. *Journal of the American Chemical Society*, 134:4465–4468, 3 2012.

[162] Marc Zengerle, Kai-Hung Chan, and Alessio Ciulli. Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. *ACS Chemical Biology*, 10(8):1770–1777, August 2015.

[163] Hsiu-Ting Huang, Dmitry Dobrovolsky, Jonathan Paulk, Guanglin Yang, Ellen L. Weisberg, Zachary M. Doctor, David L. Buckley, Jun Hyung Cho, Eunice Ko, Jae Jang, Ke Shi, Hyun Gyu Choi, James D. Griffin, Ying Li, Steven P. Treon, Eric S. Fischer, James E. Bradner, Liang Tan, and Nathanael S. Gray. A chemoproteomic approach to query the degradable kinome using a multi-kinase degrader. *Cell Chemical Biology*, 25(1):88–99.e6, January 18 2018.

[164] Yonghui Sun, Xingwang Zhao, Ning Ding, Hongying Gao, Yue Wu, Yiqing Yang, Meng Zhao, Jinseok Hwang, Yuqin Song, Wanli Liu, and Yu Rao. PROTAC-induced BTK degradation as a novel therapy for mutated BTK C481S induced ibrutinib-resistant B-cell malignancies. *Cell Research*, 28(7):779–781, July 2018.

[165] Graham M. Burslem, Benjamin E. Smith, Adrian C. Lai, Sebastian Jaime-Figueroa, Daniel C. McQuaid, Daniel P. Bondeson, Momar Toure, Haining Dong, Yifeng Qian, Jian Wang, Andrew P. Crew, John Hines, and Craig M. Crews. The Advantages of Targeted Protein Degradation Over Inhibition: An RTK Case Study. *Cell Chemical Biology*, 25(1):67–77.e3, January 2018.

[166] Jessie Hao-Ru Hsu, Timothy Rasmusson, James Robinson, Fiona Pachl, Jon Read, Sameer Kawatkar, Daniel H. O'Donovan, Sharan Bagal, Erin Code, Philip Rawlins, Argyrides Argyrou, Ronald Tomlinson, Ning Gao, Xiahui Zhu, Elisabetta Chiarparin, Kelly Jacques, Minhui Shen, Haley Woods, Emma Bednarski, David M. Wilson, Lisa Drew, M. Paola Castaldi, Stephen Fawell, and Andrew Bloecher. EED-Targeted PROTACs Degrade EED, EZH2, and SUZ12 in the PRC2 Complex. *Cell Chemical Biology*, 27(1):41–46.e17, 2020.

[167] M Catarina Silva, Fleur M Ferguson, Quan Cai, Katherine A Donovan, Ghata Nandi, Debasis Patnaik, Tinghu Zhang, Hai-Tsang Huang, Diane E Lucente, Bradford C Dickerson, Timothy J Mitchison, Eric S Fischer, Nathanael S Gray, and Stephen J Haggarty. Targeted degradation of aberrant tau in frontotemporal dementia patient-derived neuronal cell models. *eLife*, 8:e45457, mar 2019.

[168] Alexander Hanzl, Ryan Casement, Hana Imrichova, Scott J. Hughes, Eleonora Barone, Andrea Testa, Sophie Bauer, Jane Wright, Matthias Brand, Alessio Ciulli, and Georg E. Winter. Functional E3 ligase hotspots and resistance mechanisms to small-molecule degraders. *Nature Chemical Biology*, 19(3):323–333, 2023.

[169] Takumi Ito, Hideki Ando, Takayuki Suzuki, Toshihiko Ogura, Kentaro Hotta, Yoshimasa Imamura, Yuki Yamaguchi, and Hiroshi Handa. Identification of a primary target of thalidomide teratogenicity. *Science*, 327(5971):1345–1350, 2010.

[170] William Griffith McBride. Thalidomide and congenital abnormalities. *Lancet*, 2(1358):90927–8, 1961.

[171] S. Singhal, J. Mehta, R. Desikan, D. Ayers, P. Roberson, P. Eddlemon, N. Munshi, E. Anaissie, C. Wilson, M. Dhodapkar, J. Zeddis, and B. Barlogie. Antitumor activity of thalidomide in refractory multiple myeloma. *N Engl J Med*, 341(21):1565–1571, 1999.

[172] Georg Petzold, Eric S. Fischer, and Nicolas H. Thoma. Structural basis of lenalidomide-induced CK1α degradation by the CRL4$^{\mathrm{CRBN}}$ ubiquitin ligase. *Nature*, 532(7597):127–130, 2016.

[173] Jan Krönke, Namrata D. Udeshi, Anupama Narla, Peter Grauman, Slater N. Hurst, Marie McConkey, Tanya Svinkina, Dirk Heckl, Eamon Comer, Xiaoyu Li, Christie Ciarlo, Emily Hartman, Nikhil Munshi, Monica Schenone, Stuart L. Schreiber, Steven A. Carr, and Benjamin L. Ebert. Lenalidomide Causes Selective Degradation of IKZF1 and IKZF3 in Multiple Myeloma Cells. *Science*, 343(6168):301–305, 2014.

[174] Philipp M. Cromm and Craig M. Crews. Targeted protein degradation: from chemical biology to drug discovery. *Cell Chemical Biology*, 24(9):1181–1190, 2017.

[175] David Norton, William G Bonnette, James F Callahan, Maria G Carr, Charlotte M Griffiths-Jones, Tom D Heightman, Jeffrey K Kerns, Hong Nie, Sharna J Rich, Caroline Richardson, William Rumsey, Yolanda Sanchez, Marcel L Verdonk, Henriëtte M G Willems, William E Wixted, Lawrence Wolfe III, Alison J-A Woolford, Zining Wu, and Thomas G Davies. Fragment-Guided Discovery of Pyrazole Carboxylic Acid Inhibitors of the Kelch-like ECH-Associated Protein 1: Nuclear Factor Erythroid 2 Related Factor 2 (KEAP1:NRF2) Protein-Protein Interaction. *J. Med. Chem*, 64:15949–15972, 2021.

[176] Jakob S. Pallesen, Dilip Narayanan, Kim T. Tran, Sara M. Ø. Solbak, Giuseppe Marseglia, Louis M. E. Sørensen, Lars J. Høj, Federico Munafò, Rosa M. C. Carmona, Anthony D. Garcia, Haritha L. Desu, Roberta Brambilla, Tommy N. Johansen, Grzegorz M. Popowicz, Michael Sattler, Michael Gajhede, and Anders Bach. Deconstructing Noncovalent Kelch-like ECH-Associated Protein

1 (Keap1) Inhibitors into Fragments to Reconstruct New Potent Compounds. *Journal of Medicinal Chemistry*, 64(8):4623–4661, 2021. PMID: 33818106.

[177] Albena T. Dinkova-Kostova, Karen T. Liby, Katherine K. Stephenson, W. David Holtzclaw, Xiangqun Gao, Nanjoo Suh, Charlotte Williams, Renee Risingsong, Tadashi Honda, Gordon W. Gribble, Michael B. Sporn, and Paul Talalay. Extremely potent triterpenoid inducers of the phase 2 response: Correlations of protection against oxidant and inflammatory stress. *Proceedings of the National Academy of Sciences*, 102(12):4584–4589, 2005.

[178] Melinda S. Yates, Masafumi Tauchi, Fumiki Katsuoka, Kathleen C. Flanders, Karen T. Liby, Tadashi Honda, Gordon W. Gribble, Delinda A. Johnson, Jeffrey A. Johnson, Neal C. Burton, TomaÌs R. Guilarte, Masayuki Yamamoto, Michael B. Sporn, and Thomas W. Kensler. Pharmacodynamic characterization of chemopreventive triterpenoids as exceptionally potent inducers of Nrf2-regulated genes. *Molecular Cancer Therapeutics*, 6(1):154–162, 01 2007.

[179] Imre Fejes, Piroska Markacz, Janos Tatai, Monika Rudas, Petra Dunkel, Mario Gyuris, Miklos Nyerges, Nicolas Provost, Valérie Duvivier, Philippe Delerive, Virginie Martiny, Alexandra Bristiel, Brice Vidal, William Richardson, Elisabeth M Rothweiler, Jeppe Tranberg-Jensen, Charlotte E Manning, Melissa N Sweeney, Rod Chalk, Kilian V M Huber, Alex N Bullock, Andras Herner, Klaus Seedorf, Cedric Vinson, Csaba Weber, and Andras Kotschy. Covalent Inhibitors of KEAP1 with Exquisite Selectivity. *Journal of Medicinal Chemistry*, 67:21208–21222, 12 2024. doi: 10.1021/acs.jmedchem.4c02019.

[180] Thomas G. Davies, William E. Wixted, Joseph E. Coyle, Charlotte Griffiths-Jones, Keisha Hearn, Rachel McMenamin, David Norton, Sharna J. Rich, Caroline Richardson, Gordon Saxty, Henriëtte M.G. Willems, Alison J.A. Woolford, Joshua E. Cottom, Jen Pyng Kou, John G. Yonchuk, Heidi G. Feldser, Yolanda Sanchez, Joseph P. Foley, Brian J. Bolognese, Gregory Logan, Patricia L. Podolin, Hongxing Yan, James F. Callahan, Tom D. Heightman, and Jeffrey K. Kerns. Monoacidic Inhibitors of the Kelch-like ECH-Associated Protein 1: Nuclear Factor Erythroid 2-Related Factor 2 (KEAP1:NRF2) Protein-Protein Interaction with High Cell Potency Identified by Fragment-Based Discovery. *Journal of Medicinal Chemistry*, 59:3991–4006, 4 2016.

[181] Tom D. Heightman, James F. Callahan, Elisabetta Chiarparin, Joseph E. Coyle, Charlotte Griffiths-Jones, Ami S. Lakdawala, Rachel McMenamin, Paul N. Mortenson, David Norton, Torren M. Peakman, Sharna J. Rich, Caroline Richardson, William L. Rumsey, Yolanda Sanchez, Gordon Saxty, Henriëtte M. G. Willems, Lawrence Wolfe, Alison J.-A. Woolford, Zining Wu, Hongxing Yan, Jeffrey K. Kerns, and Thomas G. Davies. Structure-Activity and Structure-Conformation Relationships of Aryl Propionic Acid Inhibitors of the Kelch-like ECH-Associated Protein 1/Nuclear Factor Erythroid 2-Related Factor 2 (KEAP1/NRF2) Protein-Protein Interaction, journal = Journal of Medicinal Chemistry. 62(9):4683–4702, 2019.

[182] Christopher M. Hickey, Katherine M. Digianantonio, Kurt Zimmermann, Alicia Harbin, Connor Quinn, Avani Patel, Peter Gareiss, Amanda Chapman, Bernadette Tiberi, Jennifer Dobrodziej, John Corradi, Angela M. Cacace, David R. Langley, and Miklós Békés. Co-opting the E3 ligase KLHDC2 for targeted protein degradation by small molecules. *Nature Structural and Molecular Biology*, 31:311–322, 2 2024.

[183] Daniel C Scott, Suresh Dharuman, Elizabeth Griffith, Sergio C Chai, Jarrid Ronnebaum, Moeko T King, Rajendra Tangallapally, Chan Lee, Clifford T Gee, Lei Yang, Yong Li, Victoria C Loudon, Ha Won Lee, Jason Ochoada, Darcie J Miller, Thilina Jayasinghe, Joao A Paulo, Stephen J Elledge, J Wade Harper, Taosheng Chen, Richard E Lee, and Brenda A Schulman. Principles of paralog-specific targeted protein degradation engaging the C-degron E3 KLHDC2. *Nature Communications*, 15:8829, 2024.

[184] Brian M. Farrell, Fabian Gerth, Cheng Hao R. Yang, and Johannes T.H. Yeh. A synthetic KLHL20 ligand to validate $CUL3^{KLHL20}$ as a potent E3 ligase for targeted protein degradation. *Genes and Development*, 36:1031–1042, 9 2022.

[185] Guangyan Du, Jie Jiang, Nathaniel J. Henning, Nozhat Safaee, Eriko Koide, RadosÅaw P. Nowak, Katherine A. Donovan, Hojong Yoon, Inchul You, Hong Yue, Nicholas A. Eleuteri, Zhixiang He, Zhengnian Li, Hubert T. Huang, Jianwei Che, Behnam Nabet, Tinghu Zhang, Eric S. Fischer, and Nathanael S. Gray. Exploring the target scope of KEAP1 E3 ligase-based PROTACs. *Cell Chemical Biology*, 29(10):1470–1481.e31, October 2022. Epub 2022 Sep 6.

[186] Jieli Wei, Fanye Meng, Kwang Su Park, Hyerin Yim, Julia Velez, Prashasti Kumar, Li Wang, Ling Xie, He Chen, Yudao Shen, Emily Teichman, Dongxu Li, Gang Greg Wang, Xian Chen, H. Ümit Kaniskan, and Jian Jin. Harnessing the E3 Ligase KEAP1 for Targeted Protein Degradation. *Journal of the American Chemical Society*, 143:15073–15083, 9 2021.

[187] Bingqi Tong, Mai Luo, Yi Xie, Jessica N. Spradlin, John A. Tallarico, Jeffrey M. McKenna, Markus Schirle, Thomas J. Maimone, and Daniel K. Nomura. Bardoxolone conjugation enables targeted protein degradation of BRD4. *Scientific Reports*, 10(1):15543, September 2020.

[188] Younghoon Kim, Pooreum Seo, Eunhye Jeon, Inchul You, Kyubin Hwang, Namkyoung Kim, Jason Tse, Juhyeon Bae, Ha-Soon Choi, Stephen M. Hinshaw, Nathanael S. Gray, and Taebo Sim. Targeted kinase degradation via the KLHDC2 ubiquitin E3 ligase. *Cell Chemical Biology*, 30(11):1414–1420.e5, 2023.

[189] Rune Linding, Robert B. Russell, Victor Neduva, and Toby J. Gibson. Globplot: Exploring protein sequences for globularity and disorder. *Nucleic Acids Research*, 31(13):3701–3708, 2003.

[190] David T. Jones. Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology*, 292(2):195–202, 1999.

[191] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon Kohl, Andrew J Ballard, Andrew Cowie, Benito Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, Martin Steinegger, Michalina Pacholska, Tamas Berghammer, Sebastian Bodenstein, David Silver, Oriol Vinyals, Andrew W Senior, Koray Kavukcuoglu, Pushmeet Kohli, and Demis Hassabis. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596:583–589, 2021.

[192] Andrew Waterhouse, Marco Bertoni, Stefan Bienert, Gabriel Studer, Gabriele Tauriello, Roman Gumienny, Florian T. Heer, Tim A. P. de Beer, Christian Rempfer, Lorenza Bordoli, Rosalba Lepore, and Torsten Schwede. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, 46(W1):W296–W303, 2018.

[193] Claire Strain-Damerell, Pravin Mahajan, Alejandra Fernandez-Cid, Opher Gileadi, and Nicola A. Burgess-Brown. *Screening and Production of Recombinant Human Proteins: Ligation-Independent Cloning*, pages 23–43. Springer US, New York, NY, 2021.

[194] Douglas Hanahan, Joel Jessee, and Fredric R. Bloom. Plasmid transformation of *Escherichia coli* and other bacteria. In *Bacterial Genetic Systems*, volume 204 of *Methods in Enzymology*, pages 63–113. Academic Press, 1991.

[195] Pravin Mahajan, Claire Strain-Damerell, Shubhashish Mukhopadhyay, Alejandra Fernandez-Cid, Opher Gileadi, and Nicola A. Burgess-Brown. *Screening and Production of Recombinant Human Proteins: Protein Production in Insect Cells*, pages 67–94. Springer US, New York, NY, 2021.

[196] Dorothy Beckett, Elena Kovaleva, and Philip J. Schatz. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Science*, 8(4):921–929, 1999.

[197] Matthew Fairhead and Mark Howarth. Site-specific biotinylation of purified proteins using BirA. *Methods in Molecular Biology*, 1266:171–184, 2015.

[198] Ryne C. Johnston, Kun Yao, Zachary Kaplan, Monica Chelliah, Karl Leswing, Sean Seekins, Shawn Watts, David Calkins, Jackson Chief Elk, Steven V. Jerome, Matthew P. Repasky, and John C. Shelley. Epik: pKa and Protonation State Prediction through Machine Learning. *Journal of Chemical Theory and Computation*, 19(8):2380–2388, 2023. PMID: 37023332.

[199] Richard A. Friesner, Robert B. Murphy, Matthew P. Repasky, Leah L. Frye, Jeremy R. Greenwood, Thomas A. Halgren, Paul C. Sanschagrin, and Daniel T. Mainz. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *Journal of Medicinal Chemistry*, 49(21):6177–6196, 2006. PMID: 17034125.

[200] Ruben Abagyan, Maxim Totrov, and Dmitry Kuznetsov. ICM-A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *Journal of Computational Chemistry*, 15(5):488–506, 1994.

[201] V. Katritch, C. M. Byrd, V. Tseitin, D. Dai, E. Raush, M. Totrov, R. Abagyan, R. Jordan, and D. E. Hruby. Discovery of small molecule inhibitors of ubiquitin-like poxvirus proteinase I7L using homology modeling and covalent docking approaches. *Journal of Computer-Aided Molecular Design*, 21(10-11):549–558, Oct-Nov 2007.

[202] A. D'Arcy, T. Bergfors, S.W. Cowan-Jacob, and M. Marsh. Microseed matrix screening for optimization in protein crystallization: what have we learned? *Acta Crystallographica Section F: Structural Biology Communications*, 70(Pt 9):1117–1126, September 2014.

[203] T. Geoff G. Battye, Luke Kontogiannis, Owen Johnson, Harold R. Powell, and Andrew G.W. Leslie. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica Section D: Biological Crystallography*, 67(4):271–281, apr 2011.

[204] Philip R. Evans and Garib N. Murshudov. How good are my data and what is the resolution? *Acta Crystallographica Section D*, 69(7):1204–1214, Jul 2013.

[205] Airlie J. McCoy, Ralf W. Grosse-Kunstleve, Paul D. Adams, Martyn D. Winn, Laurent C. Storoni, and Randy J. Read. *Phaser* crystallographic software. *Journal of Applied Crystallography*, 40(4):658–674, Aug 2007.

[206] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan. Features and development of *Coot*. *Acta Crystallographica Section D*, 66(4):486–501, Apr 2010.

[207] Paul D. Adams, Pavel V. Afonine, Gábor Bunkóczi, Vincent B. Chen, Ian W. Davis, Nathaniel Echols, Jeffrey J. Headd, Li-Wei Hung, Gary J. Kapral, Ralf W. Grosse-Kunstleve, Airlie J. McCoy, Nigel W. Moriarty, Robert Oeffner, Randy J. Read, David C. Richardson, Jane S. Richardson, Thomas C. Terwilliger, and Peter H. Zwart. *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D*, 66(2):213–221, Feb 2010.

[208] E. Blanc, P. Roversi, C. Vonrhein, C. Flensburg, S. M. Lea, and G. Bricogne. Refinement of severely incomplete structures with maximum likelihood in *BUSTER–TNT*. *Acta Crystallographica Section D*, 60(12 Part 1):2210–2221, Dec 2004.

[209] Vincent B. Chen, W. Bryan Arendall, III, Jeffrey J. Headd, Daniel A. Keedy, Robert M. Immormino, Gary J. Kapral, Laura W. Murray, Jane S. Richardson, and David C. Richardson. *MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D*, 66(1):12–21, Jan 2010.

[210] Tristan Bepler, Andrew Morin, Micah Rapp, Julia Brasch, Lawrence Shapiro, Alex J. Noble, and Bonnie Berger. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nature Methods*, 16(11):1153–1160, 2019.

[211] Rod Chalk. *Mass Spectrometric Analysis of Proteins*, pages 373–395. Springer New York, New York, NY, 2017.

[212] Loredana Puca, Patricia Chastagner, Vannary Meas-Yedid, Alain Israël, and Christel Brou. α-arrestin 1 (ARRDC1) and β-arrestins cooperate to mediate Notch degradation in mammals. *Journal of Cell Science*, 126(19):4457–4468, 10 2013.

[213] N. Yamamoto, S. Yamamoto, F. Inagaki, M. Kawaichi, A. Fukamizu, N. Kishi, K. Matsuno, K. Nakamura, G. Weinmaster, H. Okano, and M. Nakafuku. Role of Deltex-1 as a transcriptional regulator downstream of the Notch receptor. *Journal of Biological Chemistry*, 276(48):45031–45040, November 2001. Epub 2001 Sep 19.

[214] K. Matsuno, D. Eastman, T. Mitsiades, A. M. Quinn, M. L. Carcanciu, P. Ordentlich, T. Kadesch, and S. Artavanis-Tsakonas. Human deltex is a conserved regulator of Notch signalling. *Nature Genetics*, 19(1):74–78, May 1998.

[215] S. Saxena, O. Foresti, A. Liu, S. Androulaki, M. Pena Rodriguez, I. Raote, M. Aridor, B. Cui, and V. Malhotra. Endoplasmic reticulum exit sites are segregated for secretion based on cargo size. *Developmental Cell*, 59(19):2593–2608.e6, October 2024. Epub 2024 Jul 10.

[216] K. Tani, T. Mizoguchi, A. Iwamatsu, K. Hatsuzawa, and M. Tagaya. p125 is a novel mammalian Sec23p-interacting protein with structural similarity to phospholipid-modifying proteins. *Journal of Biological Chemistry*, 274(29):20505–20512, July 1999.

[217] C. Y. Wu and L. A. Taneyhill. Cadherin-7 mediates proper neural crest cell-placodal neuron interactions during trigeminal ganglion assembly. *Genesis*, 57(1):e23264, January 2019. Epub 2018 Dec 24.

[218] C. A. Halmi, C. Y. Wu, and L. A. Taneyhill. Neural crest cell-placodal neuron interactions are mediated by Cadherin-7 and N-cadherin during early chick trigeminal ganglion assembly. *F1000Research*, 11:741, July 2022.

[219] H. Yamada, H. Osaka, N. Tatsumi, M. Araki, T. Abe, K. Kaihara, K. Takahashi, E. Takashima, T. Uchihashi, K. Naruse, and K. Takei. Direct Binding of Synaptopodin 2-Like Protein to Alpha-Actinin Contributes to Actin Bundle Formation in Cardiomyocytes. *Cells*, 13(16):1373, August 2024.

[220] Edouard De Castro, Christian J.A. Sigrist, Alexandre Gattiker, Valérie Bulliard, Petra S. Langendijk-Genevaux, Elisabeth Gasteiger, Amos Bairoch, and Nicolas Hulo. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Research*, 34(Web Server issue):W362–W365, July 2006.

[221] Kevin McCloskey, Eric A. Sigel, Steven Kearnes, Ling Xue, Xia Tian, Dennis Moccia, Diana Gikunju, Sana Bazzaz, Betty Chan, Matthew A. Clark, John W. Cuozzo, Marie-Aude Guié, John P. Guilinger, Christelle Huguet, Christopher D. Hupp, Anthony D. Keefe, Christopher J. Mulhern, Ying Zhang, and Patrick Riley. Machine Learning on DNA-Encoded Libraries: A New Paradigm for Hit Finding. *Journal of Medicinal Chemistry*, 63(16):8857–8866, 2020. PMID: 32525674.

[222] Zhuoyao Chen, Sarah Picaud, Panagis Filippakopoulos, Vincenzo D'Angiolella, and Alex N. Bullock. Structural Basis for Recruitment of DAPK1 to the KLHL20 E3 Ligase. *Structure*, 27:1395–1404.e4, 9 2019.

[223] Jennifer C. Pai, Jeffrey A. Culver, Jason E. Drury, Rakesh S. Motani, Raquel L. Lieberman, and Jennifer A. Maynard. Conversion of scfv peptide-binding specificity for crystal chaperone development. *Protein Engineering, Design and Selection*, 24(5):419–428, 01 2011.

[224] Jiangdong Huo, Audrey Le Bas, Reinis R. Ruza, Helen M.E. Duyvesteyn, Halina Mikolajek, Tomas Malinauskas, Tiong Kit Tan, Pramila Rijal, Maud Dumoux, Philip N. Ward, Jingshan Ren, Daming Zhou, Peter J. Harrison, Miriam Weckener, Daniel K. Clare, Vinod K. Vogirala, Julika Radecke, Lucile Moynié, Yuguang Zhao, Javier Gilbert-Jaramillo, Michael L. Knight, Julia A. Tree, Karen R. Buttigieg, Naomi Coombes, Michael J. Elmore, Miles W. Carroll, Loic Carrique, Pranav N.M. Shah, William James, Alain R. Townsend, David I. Stuart, Raymond J. Owens, and James H. Naismith. Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2. *Nature Structural and Molecular Biology*, 27(9):846–854, 9 2020.

[225] Shih-Ching Lo, Xuchu Li, Michael T Henzl, Lesa J Beamer, and Mark Hannink. Structure of the Keap1:Nrf2 interface provides mechanistic insight into Nrf2 signaling. *The EMBO Journal*, 25(15):3605–3617, 2006.

[226] Sruthi Unni, Prashant Deshmukh, Gopinatha Krishnappa, Padmini Kommu, and Balasundaram Padmanabhan. Structural insights into the multiple binding modes of Dimethyl Fumarate (DMF) and its analogs to the Kelch domain of Keap1. *The FEBS Journal*, 288(5):1599–1613, 2021.

[227] Kenji Schorpp, Ina Rothenaigner, Elena Salmina, Jeanette Reinshagen, Terence Low, Jara K. Brenke, Jay Gopalakrishnan, Igor V. Tetko, Sheraz Gul, and Kamyar Hadian. Identification of Small-Molecule Frequent Hitters from AlphaScreen High-Throughput Screens. *SLAS Discovery*, 19(5):715–726, 2014. Special Issue: Knowledge from Small-Molecule Screening and Profiling Data.

[228] Oleksandr O. Grygorenko, Dmytro S. Radchenko, Igor Dziuba, Alexander Chuprina, Kateryna E. Gubina, and Yurii S. Moroz. Generating multibillion chemical space of readily accessible screening compounds. *iScience*, 23(11):101681, 2020.

[229] Matthieu Schapira, Mike Tyers, Maricel Torrent, and Cheryl H Arrowsmith. WD40 repeat domain proteins: a novel target class? *Nature Reviews Drug Discovery*, 16:773–786, 2017.

[230] Suzanne Ackloo, Fengling Li, Magda Szewczyk, Almagul Seitova, Peter Loppnau, Hong Zeng, Jin Xu, Shabbir Ahmad, Yelena A Arnautova, A. J. Baghaie, Serap Beldar, Albina Bolotokova, Paolo A. Centrella, Irene Chau, Matthew A. Clark, John W. Cuozzo, Saba Dehghani-Tafti, Jeremy S. Disch, Aiping Dong, Antoine Dumas, Jianwen A. Feng, Pegah Ghiabi, Elisa Gibson, Justin Gilmer, Brian Goldman, Stuart R Green, Marie-Aude Guié, John P. Guilinger, Nathan Harms, Oleksandra Herasymenko, Scott Houliston, Ashley Hutchinson, Steven Kearnes, Anthony D. Keefe, Serah W. Kimani, Trevor Kramer, Maria Kutera, Haejin A. Kwak, Cristina Lento, Yanjun Li, Jenny Liu, Joachim Loup, Raquel A. C. Machado, Christopher J. Mulhern, Sumera Perveen, Germanna L. Righetto, Patrick Riley, Suman Shrestha, Eric A. Sigel, Madhushika Silva, Michael D. Sintchak, Belinda L. Slakman, Rhys D. Taylor, James Thompson, Wen Torng, Carl Underkoffler, Moritz von Rechenberg, Ryan T. Walsh, Ian Watson, Derek J. Wilson, Esther Wolf, Manisha Yadav, Aliakbar K. Yazdi, Junyi Zhang, Ying Zhang, Vijayaratnam Santhakumar, Aled M. Edwards, Dalia Barsyte-Lovejoy, Matthieu Schapira, Peter J. Brown, Levon Halabelian, and Cheryl H. Arrowsmith. A Target Class Ligandability Evaluation of WD40 Repeat-Containing Proteins. *Journal of Medicinal Chemistry*, 0(0):null, 0. PMID: 39495097.

[231] Upendra P. Dahal, Adam M. Gilbert, R. Scott Obach, Mark E. Flanagan, Jinshan M. Chen, Carmen Garcia-Irizarry, Jeremy T. Starr, Brandon Schuff, Daniel P. Uccello, and Jennifer A. Young. Intrinsic reactivity profile of electrophilic moieties to guide covalent drug design: N-α-acetyl-L-lysine as an amine nucleophile. *MedChemComm*, 7(5):864–872, 2016.

[232] Eranthie Weerapana, Chu Wang, Gabriel M Simon, Florian Richter, Sagar Khare, Myles B D Dillon, Daniel A Bachovchin, Kerri Mowen, David Baker, and Benjamin F Cravatt. Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature*, 468(7325):790–795, Dec 2010. Epub 2010 Nov 17.

[233] Keriann M. Backus, Bruno E. Correia, Kenneth M. Lum, Stefano Forli, Benjamin D. Horning, Gonzalo E. González-Páez, Sandip Chatterjee, Bryan R. Lanning, John R. Teijaro, Arthur J. Olson, Dennis W. Wolan, and Benjamin F. Cravatt. Proteome-wide covalent ligand discovery in native biological systems. *Nature*, 534(7608):570–574, 2016. Epub 2016 Jun 15.

[234] Ka Yang, Rebecca L. Whitehouse, Shane L. Dawson, Lu Zhang, Jeffrey G. Martin, Douglas S. Johnson, Joao A. Paulo, Steven P. Gygi, and Qing Yu. Accelerating multiplexed profiling of protein-ligand interactions: High-throughput plate-based reactive cysteine profiling with minimal input. *Cell Chemical Biology*, 31(3):565–576.e4, 2024.

[235] Yankun Wang and Chu Wang. Quantitative reactive cysteinome profiling reveals a functional link between ferroptosis and proteasome-mediated degradation. *Cell Death & Differentiation*, 30(1):125–136, 2023.

[236] Ekaterina V Vinogradova, Xiaoyu Zhang, David Remillard, Daniel C Lazar, Radu M Suciu, Yujia Wang, Giulia Bianco, Yu Yamashita, Vincent M Crowley, Michael A Schafroth, Minoru Yokoyama, David B Konrad, Kenneth M Lum, Gabriel M Simon, Esther K Kemper, Michael R Lazear, Sifei Yin, Megan M Blewett, Melissa M Dix, Nhan Nguyen, Maxim N Shokhirev, Emily N Chin, Luke L Lairson, Bruno Melillo, Stuart L Schreiber, Stefano Forli, John R Teijaro, and Benjamin F Cravatt. An activity-guided map of electrophile-cysteine interactions in primary human t cells. *Cell*, 182(4):1009–1026.e29, 2020. Epub 2020 Jul 29.

[237] Ling Fu, Zongmin Li, Keke Liu, Caiping Tian, Jixiang He, Jingyang He, Fuchu He, Ping Xu, and Jing Yang. A quantitative thiol reactivity profiling platform to analyze redox and electrophile reactive cysteine proteomes. *Nature Protocols*, 15(9):2891–2919, 2020.

[238] George S. Biggs, Emma E. Cawood, Aini Vuorinen, William J. McCarthy, Harry Wilders, Ioannis G. Riziotis, Antonie J. van der Zouwen, Jonathan Pettinger, Luke Nightingale, Peiling Chen, Andrew J. Powell, David House, Simon J. Boulton, J. Mark Skehel, Katrin Rittinger, and Jacob T. Bush. Robust proteome profiling of cysteine-reactive fragments using label-free chemoproteomics. *Nature Communications*, 16(1):73, 2025.

[239] Tin-Yan Koo, Hinyuk Lai, Daniel K. Nomura, and Clive Yik-Sham Chung. N-Acryloylindole-alkyne (NAIA) enables imaging and profiling new ligandable cysteines and oxidized thiols by chemoproteomics. *Nature Communications*, 14(1):3564, 2023.

[240] Nikolas R. Burton and Keriann M. Backus. Functionalizing tandem mass tags for streamlining click-based quantitative chemoproteomics. *Communications Chemistry*, 7(1):80, 2024.

[241] Miljan Kuljanin, Dylan C. Mitchell, Devin K. Schweppe, Ajami S. Gikandi, David P. Nusinow, Nathan J. Bulloch, Ekaterina V. Vinogradova, David L. Wilson, Eric T. Kool, Joseph D. Mancias, Benjamin F. Cravatt, and Steven P. Gygi. Reimagining high-throughput profiling of reactive cysteines for cell-based screening of large electrophile libraries. *Nature Biotechnology*, 39(5):630–641, 2021.

[242] Eliram Resnick, Andrew Bradley, Jian Gan, Anna Douangamath, Tobias Krojer, Rima Sethi, Paul P. Geurink, Audrey Aimon, Gil Amitai, Daniele Bellini, Jamie Bennett, Martin Fairhead, Oleg Fedorov, Ron Gabizon, Jian Gan, Jie Guo, Alex Plotnikov, Noa Reznik, Georg F. Ruda, Laura Díaz-Sáez, Volker M. Straub, Tamas Szommer, Sharmilee Velupillai, Daniel Zaidman, Yaru Zhang, Andrew R. Coker, Christopher G. Dowson, Hamish M. Barr, Chen Wang, Kilian V. M. Huber, Paul E. Brennan, Huib Ovaa, Frank von Delft, and Nir London. Rapid Covalent-Probe Discovery by Electrophile-Fragment Screening. *Journal of the American Chemical Society*, 141(22):8951–8968, June 2019. Epub 2019 May 22.

[243] Brandon P. Belcher, Christopher C. Ward, and Daniel K. Nomura. Ligandability of E3 Ligases for Targeted Protein Degradation Applications. *Biochemistry*, 62(3):588–600, 2023. Epub 2021 Sep 2.

[244] Lindsey M. Boatner, Maria F. Palafox, Devin K. Schweppe, and Keriann M. Backus. CysDB: a human cysteine database based on experimental quantitative chemoproteomics. *Cell Chemical Biology*, 30(6):683–698.e3, June 2023. Epub 2023 Apr 28.

[245] Imre Fejes, Piroska Markacz, Janos Tatai, Monika Rudas, Petra Dunkel, Mario Gyuris, Miklos Nyerges, Nicolas Provost, Valérie Duvivier, Philippe Delerive, Virginie Martiny, Alexandra Bristiel, Brice Vidal, William Richardson, Elisabeth M Rothweiler, Jeppe Tranberg-Jensen, Charlotte E Manning, Melissa N Sweeney, Rod Chalk, Kilian V M Huber, Alex N Bullock, Andras Herner, Klaus Seedorf, Cedric Vinson, Csaba Weber, and Andras Kotschy. Covalent Inhibitors of KEAP1 with Exquisite Selectivity. *Journal of Medicinal Chemistry*, 67(23):21208–21222, 12 2024.

[246] Gisbert Schneider, Wolfgang Neidhart, Thomas Giller, and Georg Schmid. "Scaffold-hopping" by topological pharmacophore search: A contribution to virtual screening. *Angewandte Chemie International Edition*, 38(19):2894–2896, 1999.

[247] Daniel P. Bondeson, Blake E. Smith, George M. Burslem, Alexandru D. Buhimschi, John Hines, Saul Jaime-Figueroa, Jing Wang, Brian D. Hamman, Alexey Ishchenko, and Craig M. Crews. Lessons in PROTAC Design from Selective Degradation with a Promiscuous Warhead. *Cell Chemical Biology*, 25(1):78–87.e5, 2018.

[248] M. S. Gadd, A. Testa, X. Lucas, K.-H. Chan, W. Chen, D. J. Lamont, M. Zengerle, and A. Ciulli. Structural basis of PROTAC cooperative recognition for selective protein degradation. *Nature Chemical Biology*, 13(5):514–521, 2017.

[249] Yuting Qin, Cecilie Poulsen, Dilip Narayanan, Camilla B. Chan, Xiangrong Chen, Beatriz Ralsi Montes, Kim T. Tran, Elina Mukminova, Chunyu Lin, Michael Gajhede, Alex N. Bullock, David Olagnier, and Anders Bach. Structure-Guided Conformational Restriction Leading to High-Affinity, Selective, and Cell-Active Tetrahydroisoquinoline-Based Noncovalent Keap1-Nrf2 Inhibitors. *Journal of Medicinal Chemistry*, 67(21):18828–18864, 2024. Published online: 2024-11-14.

[250] Roger Shen-Chu Chen, George W. Muller, and David I. Stirling. Method of reducing tnfα levels with amino substituted 2-(2,6-dioxopiperidin-3-yl)-1-oxo- and 1,3-dioxoisoindolines.

[251] Ian J. Tickle, Carsten Flensburg, Paul Keller, Wladek Paciorek, Andrew Sharff, Clemens Vonrhein, and Gerard Bricogne. Staraniso. `http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi`, 2016. Cambridge, United Kingdom: Global Phasing Ltd.

[252] O. S. Smart, T. O. Womack, A. Sharff, C. Flensburg, P. Keller, W. Paciorek, C. Vonrhein, and G. Bricogne. grade, version 1.2.20. https://www.globalphasing.com, 2021.

[253] Daniel T. Cohen and Stephen L. Buchwald. Mild Palladium-Catalyzed Cyanation of (Hetero)aryl Halides and Triflates in Aqueous Media. *Organic Letters*, 17(2):202–205, 2015. PMID: 25555140.

[254] Tsuyoshi Morita, Yoshiki Okamoto, and Hiroshi Sakurai. Novel method for dealkylation of esters, ethers, and acetals by chlorotrimethylsilane-sodium iodide. *J. Chem. Soc., Chem. Commun.*, pages 874–875, 1978.

[255] K C Nicolaou, Anthony A Estrada, Mark Zak, Sang Hyup Lee, Brian S Safina, A A Estrada, M Zak, S H Lee, and B S Safina. A Mild and Selective Method for the Hydrolysis of Esters with Trimethyltin Hydroxide. *Angew. Chem. Int. Ed*, 44:1378–1382, 2005.

[256] Sam P. De Visser, Matthew G. Quesne, and Richard A. Ward. Cysteine protease inhibition by nitrile-based inhibitors: a computational study. *Frontiers in Chemistry*, 1, 2013.

[257] Fábio Madeira, Nandana Madhusoodanan, Joonheung Lee, Alberto Eusebi, Ania Niewielska, Adrian R N Tivey, Rodrigo Lopez, and Sarah Butcher. The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic acids research*, 52(W1):W521–W525, July 2024.

[258] Andrew M. Waterhouse, James B. Procter, David M. A. Martin, Michèle Clamp, and Geoffrey J. Barton. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25(9):1189–1191, 01 2009.

[259] J. J. M. Rood, A. Jamalpoor, S. van Hoppe, M. J. van Haren, R. E. Wasmann, M. J. Janssen, A. H. Schinkel, R. Masereeuw, J. H. Beijnen, and R. W. Sparidans. Extrahepatic metabolism of ibrutinib. *Investigational New Drugs*, 39(1):1–14, Feb 2021. Epub 2020 Jul 4.

[260] Vinícius Bonatto, Rafael F. Lameiro, Fernanda R. Rocho, Jeronimo Lameira, Andrei Leitão, and Carlos A. Montanari. Nitriles: an attractive approach to the development of covalent inhibitors. *RSC Med. Chem.*, 14:201–217, 2023.

[261] Fabian Göricke, Victoria Vu, Leanna Smith, Ulrike Scheib, Raphael Böhm, Namik Akkilic, Gerd Wohlfahrt, Jörg Weiske, Ulf Bömer, Krzysztof Brzezinka, Niels Lindner, Philip Lienau, Stefan Gradl, Hartmut Beck, Peter J. Brown, Vijayaratnam Santhakumar, Masoud Vedadi, Dalia Barsyte-Lovejoy, Cheryl H. Arrowsmith, Norbert Schmees, and Kirstin Petersen. Discovery and Characterization of BAY-805, a Potent and Selective Inhibitor of Ubiquitin-Specific Protease USP21. *Cite This: J. Med. Chem*, 66:3431–3447, 2023.

[262] A. Radu Aricescu, Weixian Lu, and E. Yvonne Jones. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallographica Section D*, 62(10):1243–1250, Oct 2006.

[263] Jonathan Elegheert, Ester Behiels, Benjamin Bishop, Suzanne Scott, Rachel E Woolley, Samuel C Griffiths, Eamon F X Byrne, Veronica T Chang, David I Stuart, E Yvonne Jones, Christian Siebold, and A Radu Aricescu. Lentiviral transduction of mammalian cells for fast, scalable and high-level production of soluble and membrane proteins.

[264] Jean-Marc Galan and Matthias Peter. Ubiquitin-dependent degradation of multiple F-box proteins by an autocatalytic mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, 96(16):9124–9129, 1999.

[265] Jon M. Huibregtse, Martin Scheffner, Sylvie Beaudenon, and Peter M. Howley. A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. *Proceedings of the National Academy of Sciences of the United States of America*, 92(7):2563–2567, 1995.

[266] Guohong Hu and Eric R. Fearon. SIAH-1 N-terminal RING domain is required for proteolysis function, and C-terminal sequences regulate oligomerization and binding to target proteins. *Molecular and Cellular Biology*, 19(1):724–732, 1999.

[267] Qian Yin, Shu-Ching Lin, Benoit Lamothe, Min Lu, Yuen-Chun Lo, Greg Hura, Lei Zheng, Ronald L Rich, Alex D Campos, David G Myszka, Michael J Lenardo, Bruce G Darnay, and Hao Wu. E2 interaction and dimerization in the crystal structure of TRAF6. *Nature Structural & Molecular Biology*, 16(6):658–666, 2009.

[268] Teresa A Soucy, Peter G Smith, Michael A Milhollen, Allison J Berger, James M Gavin, Sharmila Adhikari, James E Brownell, Kristine E Burke, David P Cardin, Stephen Critchley, Courtney A Cullis, Amanda Doucette, James J Garnsey, Jeffrey L Gaulin, Rachel E Gershman, Anna R Lublinsky, Alice Mcdonald, Hirotake Mizutani, Usha Narayanan, Edward J Olhava, Stephane Peluso, Mansoureh Rezaei, Michael D Sintchak, Tina Talreja, Michael P Thomas, Tary Traore, Stepan Vyskocil, Gabriel S Weatherhead, Jie Yu, Julie Zhang, Lawrence R Dick, Christopher F Claiborne, Mark Rolfe, Joseph B Bolen, and Steven P Langston. An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. 458, 2009.

[269] Daniel H. Lee and Alfred L. Goldberg. Proteasome inhibitors: valuable new tools for cell biologists. *Trends in Cell Biology*, 8(10):397–403, 1998.

[270] Zhuoyao Chen, Gamma Chi, Timea Balo, Xiangrong Chen, Beatriz Montes, Steven C Clifford, Timea Szabo, Arpad Kiss, András Herner, András Kotschy, and Alex N Bullock. Structural mimicry of UM171 and neomorphic cancer mutants co-opts E3 ligase KBTBD4 for HDAC1/2 recruitment. *bioRxiv*, 2024.

[271] X Xie, O Zhang, MJR Yeo, C Lee, SA Harry, L Paul, Y Li, NC Payne, E Nam, HS Kwok, H Jiang, H Mao, JL Hadley, H Lin, M Batts, PM Gosavi, V D'Angiolella, PA Cole, R Mazitschek, PA Northcott, N Zheng, and BB Liau. KBTBD4 Cancer Hotspot Mutations Drive Neomorphic Degradation of HDAC1/2 Corepressor Complexes. *bioRxiv*, page 2024.05.14.593970, May 2024. [Preprint].

[272] Yi Sheng, Vivian Saridakis, Feroz Sarkari, Shili Duan, Tianne Wu, Cheryl H Arrowsmith, and Lori Frappier. Molecular recognition of p53 and MDM2 by USP7/HAUSP. *Nature Structural & Molecular Biology*, 13(3):285–291, March 2006.

[273] Hannah V. Clegg, Koji Itahana, and Yanping Zhang. Unlocking the Mdm2-p53 loop: ubiquitin is the key. *Cell Cycle*, 7(3):287–292, 2008.

[274] Kenji Wada and Takashi Kamitani. UnpEL/Usp4 is ubiquitinated by Ro52 and deubiquitinated by itself. *Biochemical and Biophysical Research Communications*, 342(1):253–258, 2006.

[275] Maria Maneiro, Nafsika Forte, Maria M. Shchepinova, Cyrille S. Kounde, Vijay Chudasama, James Richard Baker, and Edward W. Tate. Antibody-PROTAC Conjugates Enable HER2-Dependent Targeted Protein Degradation of BRD4. *ACS Chemical Biology*, 15(6):1306–1312, 2020. PMID: 32338867.

[276] Sheng He, Fei Gao, Jing Ma, Hao Ma, Guoying Dong, and Chunhua Sheng. Aptamer-PROTAC Conjugates (APCs) for Tumor-Specific Targeting in Breast Cancer. *Angewandte Chemie International Edition*, 60(43):23299–23305, 2021.

[277] Meng Chen, Ping Zhou, Yao Kong, Jing Li, Ying Li, Yujie Zhang, Jia Ran, Jun Zhou, Yu Chen, and Shitao Xie. Inducible Degradation of Oncogenic Nucleolin Using an Aptamer-Based PROTAC. *Journal of Medicinal Chemistry*, 66(2):1339–1348, 2023.

[278] Po-Chun Shih, Masaki Naganuma, Gaku Tsuji, Yuta Demizu, and Masatoshi Naito. Development of decoy oligonucleotide-warheaded chimeric molecules targeting STAT3. *Bioorganic & Medicinal Chemistry*, 95:117507, 2023.

[279] Lin Zhang, Lin Li, Xiaomin Wang, Haonan Liu, Yihan Zhang, Ting Xie, Heng Zhang, Xi Li, Ting Peng, Xiaoyuan Sun, Jun Dai, Jie Liu, Wei Wu, Minghui Ye, and Weihong Tan. Development of a novel PROTAC using the nucleic acid aptamer as a targeting ligand for tumor selective degradation of nucleolin. *Molecular Therapy - Nucleic Acids*, 30:66–79, 2022.

[280] Fangfang Shen, Ge Zheng, Mekedlawit Setegne, Karin Tenglin, Manizheh Izadi, Henry Xie, Liting Zhai, Stuart H. Orkin, and Laura M. K. Dassama. A Cell-Permeant Nanobody-Based Degrader That Induces Fetal Hemoglobin. *ACS Central Science*, 8(12):1695–1703, 2022.

[281] Kun Wang, Xiaotian Dai, Aoling Yu, Miao Xu, Chang Wang, Youyang Zhao, Wei Wang, Hong Zhou, Xiaowei Chen, Ke Yao, Xinghua Yang, and Zhiwei Wang. Peptide-based PROTAC degrader of FOXM1 suppresses cancer and decreases GLUT1 and PD-L1 expression. *Journal of Experimental & Clinical Cancer Research*, 41(1):289, 2022.

[282] Min-Yi Dai, Ying-Ying Shi, Ai-Jun Wang, Chao Yang, Yu-Ying Kong, Xiao-Lei Zhao, Jie Liu, Yuan Liu, Yi-Quan Wei, Yun-Yun Yu, Zhi-Hao Liu, Zhi-Liang Li, Yi Pan, Ping Wang, Zhen-Hua Zhang, and Yi Tang. High-potency PD-1/PD-L1 degradation induced by Peptide-PROTAC in human cancer cells. *Cell Death & Disease*, 13:924, 2022.

[283] Shuhui Lim, Regina Khoo, Khong Ming Peh, Jinkai Teo, Shih Chieh Chang, Simon Ng, Greg L. Beilhartz, Roman A. Melnyk, Charles W. Johannes, Christopher J. Brown, David P. Lane, Brian Henry, and Anthony W. Partridge. bioPROTACs as versatile modulators of intracellular therapeutic targets including proliferating cell nuclear antigen (PCNA). *Proceedings of the National Academy of Sciences*, 117(11):5791–5800, 2020.

[284] Jennifer E. Kung and Natalia Jura. Prospects for pharmacological targeting of pseudokinases. *Nature reviews. Drug discovery*, 18(7):501, jul 2019.

[285] Robert Winkler. ESIprot: A universal tool for charge state determination and molecular weight calculation of proteins from electrospray ionization mass spectrometry data. *Rapid Communications in Mass Spectrometry*, 24(3):285–294, 2010.

[286] Yibing Shan, Kavitha Gnanasambandan, Daniel Ungureanu, Eunhee T. Kim, Henrik Hammaren, Kazuhiro Yamashita, Olli Silvennoinen, Henrik Hammarén, Stevan R. Hubbard, Olli Silvennoinen, David E. Shaw, Thomas E. Smithgall, and Stevan R. Hubbard. ATP binding promotes conformational changes in the JAK2 pseudokinase domain that enhance pathogenic activation. *Proceedings of the National Academy of Sciences (PNAS)*, 111(50):E5363–E5372, 2014.

[287] Gabriel Studer, Christine Rempfer, Andrew M Waterhouse, Rafal Gumienny, Juergen Haas, and Torsten Schwede. QMEANDisCo-distance constraints applied on model quality estimation. *Bioinformatics*, 36(6):1765–1771, mar 2020.

[288] D.M. Pinkas, J.C. Bufton, K. Kupinska, D. Wang, M. Fairhead, J. Kopec, R. Sethi, S.E. Dixon-Clarke, R. Chalk, G. Berridge, N.A. Burgess-Brown, F. von Delft, C.H. Arrowsmith, A.M. Edwards, C. Bountra, A. Bullock, and Structural Genomics Consortium (SGC). Crystal structure of WNK3 kinase domain in a monophosphorylated apo state (A-loop swapped). *Protein Data Bank*, 2017. PDB entry 5O1V, deposited on 19 May 2017, released on 28 June 2017, latest revision on 20 November 2024.

[289] Fabrizio Villa, Jürgen Goebel, Fatema H Rafiqi, Maria Deak, Jacob Thastrup, Dario R Alessi, and Daan M F van Aalten. Structural insights into the recognition of substrates and activators by the OSR1 kinase. *EMBO reports*, 8(9):839–845, sep 2007.

[290] Jinwei Zhang, Keith Siew, Thomas Macartney, Kevin M O'Shaughnessy, and Dario R Alessi. Critical role of the SPAK protein kinase CCT domain in controlling blood pressure. *Human molecular genetics*, 24(16):4545–4558, aug 2015.

[291] Ciaran Richardson, Kei Sakamoto, Paola de los Heros, Maria Deak, David G. Campbell, Alan R. Prescott, and Dario R. Alessi. Regulation of the NKCC2 ion cotransporter by SPAK-OSR1-dependent and -independent pathways. *Journal of Cell Science*, 124(5):789–800, 03 2011.

[292] P. de Los Heros, D. R. Alessi, R. Gourlay, D. G. Campbell, M. Deak, T. J. Macartney, K. T. Kahle, and J. Zhang. The WNK-regulated SPAK/OSR1 kinases directly phosphorylate and inhibit the K^+-Cl^- co-transporters. *Biochemical Journal*, 458(3):559–573, March 2014.

[293] Henry C. Nguyen, Haitao Yang, Jennifer L. Fribourgh, Leslie S. Wolfe, and Yong Xiong. Insights into Cullin-RING E3 Ubiquitin Ligase Recruitment: Structure of the VHL-EloBC-Cul2 Complex. *Structure*, 23(3):441–449, 2015.

[294] Eric S. Fischer, Andrea Scrima, Kerstin Böhm, Syota Matsumoto, Gondichatnahalli M. Lingaraju, Mahamadou Faty, Takeshi Yasuda, Simone Cavadini, Mitsuo Wakasugi, Fumio Hanaoka, Shigenori Iwai, Heinz Gut, Kaoru Sugasawa, and Nicolas H. Thomä.

The Molecular Basis of $CRL4^{DDB2/CSA}$ Ubiquitin Ligase Architecture, Targeting, and Activation. *Cell*, 147(5):1024–1039, November 2011.

[295] Zhuoyao Chen, Jinwei Zhang, Adrián R. Murillo-de Ozores, María Castañeda Bueno, Francesca D'Amico, Raphael Heilig, Charlotte E. Manning, Fiona J. Sorrell, Vincenzo D'Angiolella, Roman Fischer, Monique P. C. Mulder, Gerardo Gamba, Dario R. Alessi, and Alex N. Bullock. Sequence and structural variations determining the recruitment of WNK kinases to the KLHL3 E3 ligase. *Biochemical Journal*, 479(5):661–675, 03 2022.

[296] Ramchandra V Amnekar, Toby Dite, Pawel Lis, Sebastian Bell, Fiona Brown, Clare Johnson, Stuart Wilkinson, Samantha Raggett, Mark Dorward, Mel Wightman, Thomas Macartney, Renata F. Soares, Frederic Lamoliatte, and Dario R Alessi. NRBP1 pseudokinase binds to and activates the WNK pathway in response to osmotic stress. *bioRxiv*, 2024.

[297] Germán Magaña-Ávila, Héctor Carbajal-Contreras, Ramchandra Amnekar, Toby Dite, Michelle Téllez-Sutterlin, Kevin García-Ávila, Brenda Marquina-Castillo, Alejandro Lopez-Saavedra, Norma Vazquez, Eréndira Rojas-Ortega, Eric Delpire, David H. Ellison, Dario R. Alessi, Gerardo Gamba, and María Castañeda-Bueno. NRBP1 and TSC22D proteins impact distal convoluted tubule physiology through modulation of the WNK pathway. *bioRxiv*, 2025.

[298] Markus Welcker and Bruce E Clurman. Fbw7/hCDC4 dimerization regulates its substrate interactions. *Cell Division*, 2:7, feb 2007.

[299] Welcker M and Clurman BE. FBW7 ubiquitin ligase: a tumour suppressor at the crossroads of cell division, growth and differentiation. *Nature reviews. Cancer*, 8(2):83–93, feb 2008.

[300] Erik S. Zimmerman, Brenda A. Schulman, and Ning Zheng. Structural assembly of cullin-RING ubiquitin ligase complexes. *Current Opinion in Structural Biology*, 20(6):714–721, dec 2010.

[301] Amir Orian, Hagit Gonen, Bracha Bercovich, Irit Fajerman, Esther Eytan, Alain Israël, F. Mercurio, Kazuhiro Iwai, Arthur L. Schwartz, and Aaron Ciechanover. SCF(β-TrCP) ubiquitin ligase-mediated processing of NF-κB p105 requires phosphorylation of its C-terminus by IκB kinase. *EMBO Journal*, 19(11):2580–2591, June 2000.

[302] Chun Wu and Sankar Ghosh. β-trcp mediates the signal-induced ubiquitination of iκbβ. *Journal of Biological Chemistry*, 274(42):29591–29594, 1999.

[303] Daming Gao, Hiroyuki Inuzuka, Meng-Kwang Marcus Tan, Hidefumi Fukushima, Jason W. Locasale, Pengda Liu, Lixin Wan, Bo Zhai, Y. Rebecca Chin, Shavali Shaik, Costas A. Lyssiotis, Steven P. Gygi, Alex Toker, Lewis C. Cantley, John M. Asara, J. Wade Harper, and Wenyi Wei. mTOR Drives Its Own Activation via $SCF^{\beta TrCP}$-Dependent Degradation of the mTOR Inhibitor DEPTOR. *Molecular Cell*, 44(2):290–303, October 2011.

[304] Masaaki Komatsu, Hirofumi Kurokawa, Satoshi Waguri, Keiko Taguchi, Akira Kobayashi, Yoshinobu Ichimura, Yu-Shin Sou, Izumi Ueno, Ayako Sakamoto, Kit I. Tong, Mihee Kim, Yasumasa Nishito, Shun-ichiro Iemura, Tohru Natsume, Takashi Ueno, Eiki Kominami, Hozumi Motohashi, Keiji Tanaka, and Masayuki Yamamoto. The selective autophagy substrate p62 activates the stress responsive transcription factor Nrf2 through inactivation of Keap1. *Nature Cell Biology*, 12(3):213–223, March 2010.

[305] Shih-Ching Lo and Mark Hannink. PGAM5, a Bcl-XL-interacting Protein, Is a Novel Substrate for the Redox-regulated Keap1-dependent Ubiquitin Ligase Complex. *Journal of Biological Chemistry*, 281(49):37893–37903, 2006.

[306] Wei Chen, Zhide Sun, Xuejun J. Wang, Tian Jiang, Zhen Huang, Dong Fang, and Donna D. Zhang. Direct interaction between Nrf2 and p21(Cip1/WAF1) upregulates the Nrf2-mediated antioxidant response. *Molecular Cell*, 34(6):663–673, June 2009.

[307] Ingo V Hartung, Joachim Rudolph, Mary M Mader, Monique P C Mulder, and Paul Workman. Expanding chemical probe space: Quality criteria for covalent and degrader probes. *Cite This: J. Med. Chem*, 66:9312, 2023.

[308] Elma Mons, Sander Roet, Robbert Q. Kim, and Monique P. C. Mulder. A comprehensive guide for assessing covalent inhibition in enzymatic assays illustrated with kinetic simulations. *Current Protocols*, 2(6):e419, 2022.

[309] Radoslaw P. Nowak, Leah Ragosta, Fidel Huerta, Hu Liu, Scott B. Ficarro, Justin T. Cruite, Rebecca J. Metivier, Katherine A. Donovan, Jarrod A. Marto, Eric S. Fischer, Breanna L. Zerfas, and Lyn H. Jones. Development of a covalent cereblon-based PROTAC employing a fluorosulfate warhead. *RSC Chem. Biol.*, 4:906–912, 2023.

[310] A. L. Hopkins, C. R. Groom, and A. Alex. Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today*, 9(10):430–431, 2004.

[311] Philip J. Hajduk. Fragment-based drug design: how big is too big? *Journal of Medicinal Chemistry*, 49(23):6972–6976, 2006.

[312] Jonathan M. Ostrem, Ulf Peters, Martin L. Sos, James A. Wells, and Kevan M. Shokat. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature*, 503(7477):548–551, 2013.

[313] Mei Zeng, Yuan Xiong, Nozhat Safaee, Radoslaw P. Nowak, Katherine A. Donovan, Christine J. Yuan, Behnam Nabet, Thomas W. Gero, Frederic Feru, Lianbo Li, Sudershan Gondi, Lincoln J. Ombelets, Chunshan Quan, Pasi A. Jänne, Milka Kostic, David A. Scott, Kenneth D. Westover, Eric S. Fischer, and Nathanael S. Gray. Exploring Targeted Degradation Strategy for Oncogenic $KRAS^{G12C}$. *Cell Chemical Biology*, 27(1):19–31.e6, 2020.

[314] James Schiemer, Andrew Maxwell, Reto Horst, Shenping Liu, Daniel P. Uccello, Kris Borzilleri, Nisha Rajamohan, Matthew F. Brown, and Matthew F. Calabrese. A covalent BTK ternary complex compatible with targeted protein degradation. *Nature Communications*, 14(1):1189, 2023.

[315] Shang Li, Liangliang Ma, Xinxin Li, Yuhan Jiang, Zhongwen Luo, Fucheng Yin, Yonglei Zhang, Yifan Chen, Siyuan Wan, Han Zhou, Lingyi Kong, and Xiaobing Wang. Discovery of Covalent MLKL PROTAC Degraders via Optimization of a Theophylline Derivative Ligand for Treating Necroptosis. *Journal of Medicinal Chemistry*, 67(17):15353–15372, September 2024.

[316] Ronen Gabizon, Amit Shraga, Paul Gehrtz, Ella Livnah, Yamit Shorer, Neta Gurwicz, Liat Avram, Tamar Unger, Hila Aharoni, Shira Albeck, Alexander Brandis, Ziv Shulman, Ben-Zion Katz, Yair Herishanu, and Nir London. Efficient Targeted Degradation via Reversible and Irreversible Covalent PROTACs. *Journal of the American Chemical Society*, 142(27):11734–11742, 2020. PMID: 32369353.

[317] Gang Xue, Jiahui Chen, Lihong Liu, Danli Zhou, Yingying Zuo, Tiancheng Fu, and Zhengying Pan. Protein degradation through covalent inhibitor-based PROTACs. *Chem. Commun.*, 56:1521–1524, 2020.

[318] Christopher P. Tinworth, Hannah Lithgow, Lars Dittus, Zuni I. Bassi, Sophie E. Hughes, Marcel Muelbaier, Han Dai, Ian E. D. Smith, William J. Kerr, Glenn A. Burley, Marcus Bantscheff, and John D. Harling. PROTAC-Mediated Degradation of Bruton's Tyrosine Kinase Is Inhibited by Covalent Binding. *ACS Chemical Biology*, 14(3):342–347, 2019. PMID: 30807093.

[319] Dandan Lu, Xiaobo Yu, Huan Lin, Ruzhou Cheng, Eric Y. Monroy, Xiaofei Qi, Meng C. Wang, and Jian Wang. Applications of covalent chemistry in targeted protein degradation. *Chemical Society Reviews*, 51(22):9243–9261, November 2022.

[320] Jennifer N. Spradlin, Xinning Hu, Christopher C. Ward, Sarah M. Brittain, Matthew D. Jones, Linda Ou, Michelle To, Anne Proudfoot, Elizabeth Ornelas, Mulu Woldegiorgis, James A. Olzmann, Derek E. Bussiere, John R. Thomas, John A. Tallarico, Jonathan M. McKenna, Markus Schirle, Thomas J. Maimone, and Daniel K. Nomura. Harnessing the anti-cancer natural product nimbolide for targeted protein degradation. *Nature Chemical Biology*, 15(7):747–755, July 2019. Epub 2019 Jun 17.

[321] Bing Tong, Jennifer N. Spradlin, Livia F. T. Novaes, Emily Zhang, Xinning Hu, Max Moeller, Sarah M. Brittain, Lauren M. McGregor, Jonathan M. McKenna, John A. Tallarico, Markus Schirle, Thomas J. Maimone, and Daniel K. Nomura. A Nimbolide-Based Kinase Degrader Preferentially Degrades Oncogenic BCR-ABL. *ACS Chemical Biology*, 15(7):1788–1794, July 2020. Epub 2020 Jun 25.

[322] Cameron C. Ward, Jacob I. Kleinman, Scott M. Brittain, Patrick S. Lee, Cindy Y. S. Chung, Kyung Kim, Yannik Petri, Jason R. Thomas, John A. Tallarico, James M. McKenna, Markus Schirle, and Daniel K. Nomura. Covalent Ligand Screening Uncovers a RNF4 E3 Ligase Recruiter for Targeted Protein Degradation Applications. *ACS Chemical Biology*, 14(11):2430–2440, Nov 2019. Epub 2019 May 13.

[323] Xiaoyu Zhang, Vincent M. Crowley, Thomas G. Wucherpfennig, Melissa M. Dix, and Benjamin F. Cravatt. Electrophilic PROTACs that degrade nuclear proteins by engaging DCAF16. *Nature Chemical Biology*, 15(8):737–746, 2019.

[324] Brian K. Albrecht et al. Bromodomain inhibitors and uses thereof. *Patent WO2016138114A1*, 2016. Genentech, Inc., Constellation Pharmaceuticals, Inc.

[325] William Farnaby, Manfred Koegl, Michael J. Roy, Claire Whitworth, Emelyne Diers, Nicole Trainor, David Zollman, Steffen Steurer, Jale Karolyi-Oezguer, Carina Riedmueller, Teresa Gmaschitz, Johannes Wachter, Christian Dank, Michael Galant, Bernadette Sharps, Klaus Rumpel, Elisabeth Traxler, Thomas Gerstberger, Renate Schnitzer, Oliver Petermann, Peter Greb, Harald Weinstabl, Gerd Bader, Andreas Zoephel, Alexander Weiss-Puxbaum, Katharina Ehrenhöfer-Wölfer, Simon Wöhrle, Guido Boehmelt, Joerg Rinnenthal, Heribert Arnhof, Nicola Wiechens, Meng-Ying Wu, Tom Owen-Hughes, Peter Ettmayer, Mark Pearson, Darryl B. McConnell, and Alessio Ciulli. BAF complex vulnerabilities in cancer demonstrated via structure-based PROTAC design. *Nature Chemical Biology*, 15(7):672–680, July 2019.

[326] Fabian Offensperger, Gary Tin, Miquel Duran-Frigola, Elisa Hahn, Sarah Dobner, Christopher W. am Ende, Joseph W. Strohbach, Andrea Rukavina, Vincenth Brennsteiner, Kevin Ogilvie, Nara Marella, Katharina Kladnik, Rodolfo Ciuffa, Jaimeen D. Majmudar, S. Denise Field, Ariel Bensimon, Luca Ferrari, Evandro Ferrada, Amanda Ng, Zhechun Zhang, Gianluca Degliesposti, Andras Boeszoermenyi, Sascha Martens, Robert Stanton, André C. Müller, J. Thomas Hannich, David Hepworth, Giulio Superti-Furga, Stefan Kubicek, Monica Schenone, and Georg E. Winter. Large-scale chemoproteomics expedites ligand discovery and predicts ligand behavior in cells. *Science*, 384(6694):eadk5864, 2024.

[327] Miller CJ and Turk BE. Homing in: Mechanisms of Substrate Targeting by Protein Kinases. *Trends in biochemical sciences*, 43(5):380–394, may 2018.