Chemical genetic inhibition of E2 ubiquitin-conjugating enzymes

Submitted to Cardiff University for the degree of Doctor of Philosophy by

Luke Anthony Spear

Supervisors: Dr. Yu-Hsuan Tsai, Dr. Louis Y. P. Luk

May 2023

PREFACE

The following publications have been authored or co-authored during the work carried out on this PhD thesis, with the data presented in Chapters 2 and 3 contributing to the first listed publication.

1. Spear, L. A., Huang, Y., Chen, J., Nödling, A. R., Virdee, S. & Tsai, Y.-H. Selective Inhibition of Cysteine-Dependent Enzymes by Bioorthogonal Tethering. *J Mol Biol* **434,** 167524 (2022).

2. Nödling, A. R., Spear, L. A., Williams, T. L., Luk, L. Y. P. & Tsai, Y.-H. Using genetically incorporated unnatural amino acids to control protein functions in mammalian cells. *Essays Biochem* **63**, 237–266 (2019).

ABSTRACT

The development of a general approach for the rapid and selective inhibition of enzymes in cells using a common tool holds great promise for research and therapeutic applications. In this study, we expanded on a previously reported chemogenetic strategy for kinase inhibition to address the challenge of inhibiting cysteine-dependent enzymes. We achieved selective inhibition of two E2 ubiquitin-conjugating enzymes, UBE2L3 and UBE2D1, through bioorthogonal tethering of electrophilic warheads. The successful inhibition was demonstrated in biochemical assays, with consistent inhibition of polyubiquitin chain formation observed using SDS-PAGE analysis. Mass spectrometry data confirmed the inhibitor tethering to the protein and supported the reaction between a chloroacetamide warhead in one of the proposed inhibitor conjugates and the E2 catalytic cysteine. Further optimization of the strategy is necessary to achieve complete inhibition without residual polyubiquitination.

The transferability of this chemogenetic approach was demonstrated by selective inhibition of a UBE2D1 variant using the same inhibitor complex that achieved success in UBE2L3. This indicates the potential application of the strategy to other enzymes within the same family and other cysteine-dependent enzyme classes.

Additionally, efforts were made to investigate the inhibition of UBE2L3 in living mammalian cells using the NF-κB signaling pathway as a model for the E2 enzyme's activity. Although successful UBE2L3 overexpression was achieved for wild-type protein, UBE2L3 variant expression requires further tuning.Cell viability assays revealed toxicity at higher concentrations of the successful inhibitor complex, highlighting the need for considering toxicity in future inhibition assays.

While the results presented here offer promising insights into selective enzyme inhibition and the potential use of this strategy in various cysteine-dependent enzymes, further investigation and optimization are required. The study provides a foundation for future research on the development of the proposed strategy and the investigation of cellular functions and validation of therapeutic targets.

ACKNOWLEDGEMENTS

I would like to first thank my supervisory team Dr Yu-Hsuan Tsai and Dr Louis Luk for their support and guidance throughout this project. The years spent at Cardiff University under their supervision has been enlightening and an amazing opportunity to establish myself as a researcher. It is through this project I have been given the chance to grow as a scientist and learn invaluable techniques and skills through the tutelage of my supervisory team and the support of the Chemical Biology research groups. I am very grateful for the support on the initial project development and materials provided by Dr Satpal Virdee, as well as the support provided by Dr Myles Lewis in bringing UBE2L3 inhibition to a mammalian cell platform. Importantly, this project could not have been pursued without the financial support from the Biotechnology and Biological Sciences Research Council, and the Engineering and Physical Sciences Research Council that sponsored this PhD position.

I would like to give special thanks to Dr. Alexander Nödling (especially for those inhibitor complexes too!), Dr. Thomas Williams, Dr. Nicolo Santi, Dr Emily Mills, Dr Simon Tang, Dr Patrick Baumann and Sanjay Patel for such a warm reception to the research group. Together you fostered such a nurturing and friendly environment I will be forever thankful for. This feeling is extended to all past and present members of the Luk, Tsai, Jin and Alleman research groups. I couldn't go without saying thank you to Dr Victoria Barlow, Dr Davide Cardella, Dr Alex Lander, Davide Zappalà, Dr Heather Hayes, Dr Helen Bell, Muge Ma, Mochen Dong and Adriana Coricello. Thank you to the PGR department for your invaluable support throughout this process.

Thank you to Dr "Ed" Kalvaitis, Dr Raquel Cruz Samperio, Dr "Pads" Baumann and Dr Emily Schmills for getting me through the rough times. Emily, thank you for keeping me fed, singing and laughing over all these years. I have made a friend for life and couldn't have done this without your support.

Lastly, I would like to thank my friends, family and my partner Nathan. Thank you all for putting up with all the science talk over the years, I promise it doesn't end here. Sorry. Thank you everyone for your encouragement and support, you too Caru and George, our chats helped get me over the finish line!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

aaRS:	Aminoacyl-tRNA synthetase
ATP:	Adenosine triphosphate
BCA:	Bicinchoninic acid
BME:	β-Mercaptoethanol
BRCA1:	Breast cancer type 1 susceptibility protein
CypK:	N-cyclopropene-L-Lysine
DMSO:	Dimethyl sulfoxide
DTT:	Dithiothreitol
DUB:	Deubiquitinase
E. coli:	*Escherichia coli*
EDTA:	Ethylenediaminetetraacetic acid
FT:	Flow through
HEK293:	Human embryonic kidney 293 cell line
HIV:	Human immunodeficiency virus
His Tag:	N-terminal hexa-histidine tag
IEDDA:	Inverse electron-demand Diels-Alder
IPA:	Isopropanol
IPTG:	isopropyl-1-thio-β-D-galactopyranoside
kDa:	Kilodalton
LC:	Liquid chromatography
LUBAC:	Linear ubiquitin chain assembly complex
MS:	Mass spectrometry
MW:	Molecular weight
ncAA:	Non-canonical amino acid
NF-κB:	Nuclear factor kappa B
NSCLC:	Non-small cell lung cancer
o-aaRS:	Orthogonal aminoacyl-tRNA synthetase
OD_{600}:	Optical density at 600 nm
PCR:	Polymerase chain reaction

PMSF:	phenylmethylsulfonyl fluoride
PTM:	Post-translational modification
RBR:	RING-between-RING
RING:	Really Interesting New Gene
RNA:	Ribonucleic acid
SD:	Standard Deviation
SDM:	Site-Directed Mutagenesis
SDS:	Sodium dodecyl sulfate
SDS-PAGE:	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SLE:	Systemic lupus erythematosus
SN:	Supernatant
TAG:	Termination codon (amber codon)
TEMED:	Tetramethylethylenediamine
Ub:	Ubiquitin
UAA:	Unnatural amino acid
UBC:	Ubiquitin Conjugation domain
TEV:	Tobacco Etch Virus

Chapter 1

Introduction

1. INTRODUCTION.

1.1. Ubiquitin and Ubiquitination

Ubiquitination is a reversible post-translational modification (PTM) process that plays crucial roles in most aspects of eukaryotic cellular function. It is one of the most investigated and intriguing PTM processes due to its ability to increase functional diversity by conferring novel properties to proteins. Ubiquitination involves the addition of the small protein ubiquitin (Ub) to targeted sites on a substrate protein through a complex pathway, which includes ubiquitin activating enzymes (E1s), ubiquitin-conjugating enzymes (E2s), and ubiquitin ligases (E3s).[1]

Ubiquitin, a highly conserved 76-amino acid protein, serves as the building block for the ubiquitination process. The covalent attachment of ubiquitin to target proteins can occur as a single ubiquitin moiety or as polyubiquitin chains, where multiple ubiquitin molecules are linked together. This process can confer various functional outcomes, including changes in enzymatic activity, cellular localization, interaction partners, or marking proteins for degradation by the proteasome.[2]

The ubiquitination cascade (**Figure 1.1**) begins with the activation of ubiquitin by E1 ubiquitin activating enzymes, which activate ubiquitin and transfer it to E2 ubiquitin conjugating enzymes. E2 enzymes then interact with E3 ligases to facilitate the transfer of ubiquitin from the E2 enzyme to the substrate protein, often through the formation of an isopeptide bond between the C-terminal glycine residue of ubiquitin and a target lysine residue on the substrate protein.[2]

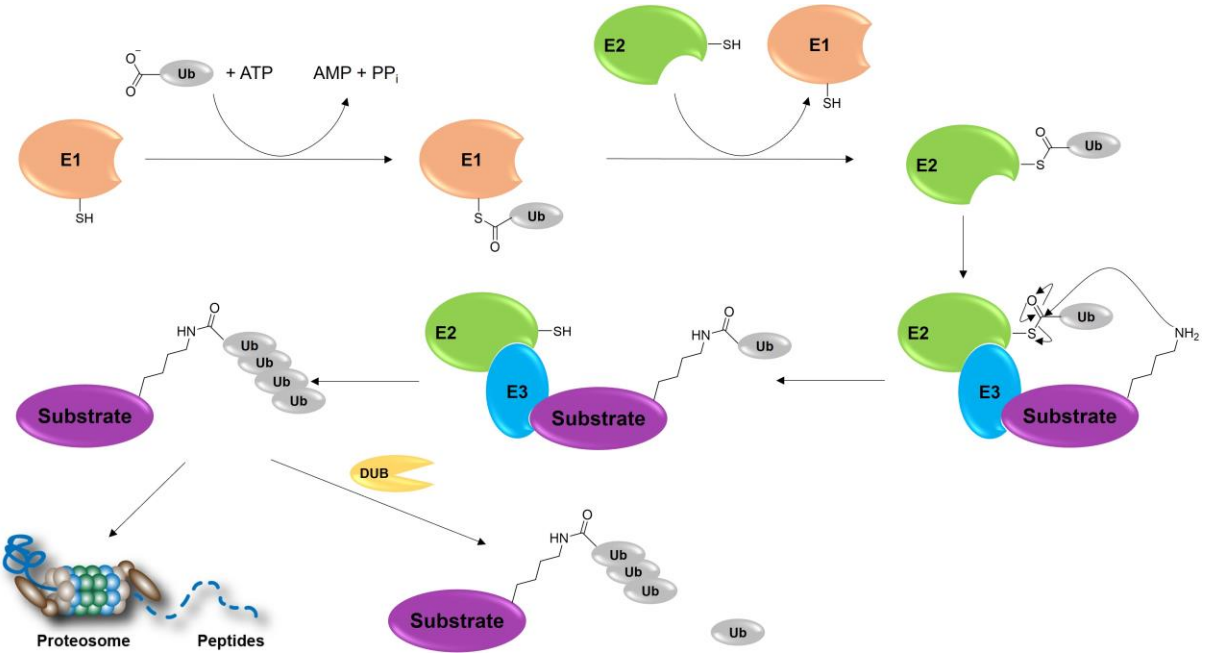


Figure 1.1. Ubiquitination cascade. Ubiquitin activating enzymes (E1s), ubiquitin-conjugating enzymes (E2s), and ubiquitin ligases (E3s), involved in the ubiquitin (Ub) cascade. Ubiquitin can be removed by deubiquitinating enzymes (DUBs).

The process of ubiquitination can have various outcomes for the substrate protein. Through this complex enzymatic machinery, ubiquitination regulates a diverse array of cellular processes, including protein degradation, DNA repair, cell cycle progression, signal transduction, and protein trafficking.[3,4] Understanding the precise mechanisms and functional consequences of ubiquitination is crucial for unravelling the intricate regulatory networks governing cellular physiology.

Once a substrate protein is ubiquitinated, it can undergo different fates depending on the type and length of the ubiquitin chain attached. For instance, ubiquitinated proteins can be targeted for degradation by the proteasome, a large multi-subunit complex responsible for protein breakdown. Proteasomal degradation of ubiquitinated proteins plays a crucial role in regulating protein homeostasis, eliminating damaged or misfolded proteins, and controlling the abundance of specific proteins involved in cellular processes.[5]

It's important to note that ubiquitination is a reversible process. Ubiquitinated substrates can have the ubiquitin molecules removed by deubiquitinating enzymes (DUBs), a family of proteases that specifically cleave the isopeptide bond between ubiquitin and the substrate protein. DUBs play a critical role in maintaining ubiquitin

homeostasis, regulating the duration and intensity of ubiquitin signalling, and ensuring the dynamic control of cellular processes.[4]

Approximately 5% of human genes encode regulators for ubiquitin signalling. This highlights the importance of ubiquitin signalling pathways, as there is a significant evolutionary investment to maintain this activity across a wide array for cellular processes. Knowing this, it is also unsurprising that problems in regulating these pathways can often lead to disease, ranging from involvement in cancers, to neurodegenerative disorders, to autoimmune disease.[6]

Only two E1 enzymes have been identified, the ubiquitin-like modifier-activating enzyme 1 (UBA1) and ubiquitin-like modifier activating enzyme 6 (UBA6). Amongst these two E1 enzymes, UBA1 is responsible for charging around 99% of cellular ubiquitin.[7] As main players in the first step in the ubiquitination pathways, dysregulation in their system can result disrupt cellular homeostasis and lead to a wide variety of disorders, and have been linked with the pathogenesis of human cancer.[8]

E2s play a central role in the ubiquitination process, and, in humans, there are approximately 40 E2 enzymes that play critical roles in deciding the site, extent and topology of ubiquitination. Malfunction of E2s have therefore been associated with different diseases, such as cancer, central nervous system diseases (*e.g.* Parkinson's disease), and autoimmune disorders (*e.g.* rheumatoid arthritis, systemic lupus erythematosus, inflammatory bowel disease, Crohn's disease).[9]

This makes E2 enzymes putative therapeutic targets, however, deciphering the roles of specific E2s in disease development and pathology is challenging. Approaches have been investigated via conventional genetic and/or chemical means. Genetic-based *in vivo* approaches such as gene knockdown or knockout models have a wide coverage and high specificity. However, the poor temporal resolution of genetic inactivation can lead to adaptive compensation that masks the biology. Alternatively, chemical-based methods have also been studied. A common approach in the field is small-molecule inhibitors as they typically act rapidly and are powerful tools for discerning protein function. Nevertheless, the development of selective small-molecule inhibitors against E2s is slow, and minimal are currently available commercially.[10]

E3 ubiquitin ligases are key components of the ubiquitination pathway and play a crucial role in determining substrate specificity and catalysing the transfer of ubiquitin from E2 enzymes to target proteins. They confer substrate selectivity by recognizing specific protein motifs or interacting with other proteins. The human genome encodes hundreds of E3 ligases, which can be categorized into different families based on their structural domains and mechanisms of action.[11]

One well-known family of E3 ligases is the really interesting new gene (RING) finger E3 ligases, which facilitate the direct transfer of ubiquitin from E2 enzymes to substrates.[12] Another prominent family is the homologous to the E6-associated protein C-terminus (HECT) domain E3 ligases, which form an intermediate thioester bond with ubiquitin before transferring it to the substrate.[13] Additional families of E3 ligases include the RBR (RING-between-RING) ligases and U-box ligases, each with distinct mechanisms of action.[14]

E3 ligases have diverse roles in cellular processes and are implicated in various diseases. They participate in the regulation of protein degradation, cellular signalling, DNA repair, cell cycle progression, and immune responses, among others. Dysregulation of E3 ligases can lead to aberrant ubiquitination events, protein accumulation, or degradation of critical signalling components, resulting in the development or progression of diseases.[11]

The relationship between E3 ligases, E2 enzymes, and E1 enzymes is highly interconnected. Different E2 enzymes exhibit specific preferences for certain E3 ligases, and the combination of an E2-E3 pair determines the substrate specificity and ubiquitin linkage type. E3 ligases collaborate with E2 enzymes to ensure efficient ubiquitination and degradation of target proteins.

In terms of E3 ligase inhibitors, there has been considerable interest in developing small-molecule inhibitors targeting specific E3 ligases for therapeutic purposes. Several E3 ligase inhibitors have been identified or designed, primarily focusing on specific E3 ligases involved in cancer-related pathways. These inhibitors aim to disrupt aberrant ubiquitination events or selectively modulate the activity of specific E3 ligases to restore normal cellular processes. An example of a very successful approach in the field is the use of bivalent compounds referred to as PROTACs, or Proteolysis Targeting Chimeras, where one part of the compound binds the protein

target to be degraded and the other side recruits an E3 ligase to mark the target protein for proteasomal degradation. However, it is worth noting that the development of selective and potent E3 ligase inhibitors remains challenging due to the complex mechanisms and structural diversity of E3 ligases.[10,15–17]

Overall, the ubiquitination cascade is tightly regulated yet dynamic process which involves a large number of E3 ligases, E2 enzymes, and E1 enzymes. Dysregulation of these enzymes can disrupt cellular homeostasis and contribute to various disorders, including cancer.[18] The development of selective inhibitors for these enzymes has gained increasing interest as a potential therapeutic approach. However, targeting the enzymes involved in ubiquitination presents significant challenges due to the complexity and diversity of the ubiquitination system. Despite these challenges, ongoing research efforts aim to develop innovative strategies and platforms for specific enzyme inhibition, including the design of small-molecule inhibitors, peptide-based inhibitors, and other targeted approaches. These endeavours hold great promise for uncovering new therapeutic interventions and expanding our understanding of ubiquitin-mediated regulation, ultimately paving the way for improved treatments for a wide range of diseases.

1.2. Expanding on a new strategy for inhibition of cysteine-dependent enzymes

Enzymes play a crucial role in maintaining our health, and dysregulated enzyme activity has been linked to numerous human diseases.[19–21] However, the functions of many enzymes and their potential therapeutic significance remain largely unexplored. Although genomic advancements have revealed statistical associations between enzyme mutations or altered expression and various diseases, they do not definitively establish aberrant enzyme activity as a causal factor.[22] Furthermore, a significant number of human enzymes still lack characterization, leaving their cellular functions and disease relevance unknown.[23] Given the vital importance of enzymes to our well-being, there is an urgent need to unravel these mysteries and assign physiological functions to enzymes in the post-genomics era.

One powerful approach for investigating enzyme activity and assessing its therapeutic modulation potential is the selective and rapid inhibition of target enzymes using small-molecule tools. However, the availability of potent and

selective small molecules is limited to a small fraction of the proteome and often requires expensive development processes involving extensive screening and optimization, without guaranteed success.[24,25] Genetic approaches, such as gene knockdown and knockout, offer specific means to interrogate enzyme function in cells. However, these methods suffer from a prolonged lag time, during which compensatory mechanisms can obscure the biological phenomena being investigated.[26,27] Additionally, enzymes can be large and possess non-catalytic functions that cannot be easily distinguished using genetic approaches alone.

In contrast, chemical or pharmacological inhibition of enzymes using small molecules allows for rapid modulation, providing high temporal resolution for analysing biological processes.[28] Chemogenetic inhibition combines the advantages of both genetic and chemical approaches.[29,30] This approach involves introducing a genetic modification, termed "sensitization," into the enzyme of interest. This modification makes the target enzyme responsive to a designer small molecule derived from an inhibitor that initially lacks selectivity towards the wild-type enzyme or closely related family members. The "bump-and-hole" strategy, which selectively inhibits kinases, is a prominent example of this approach. By introducing a mutation in a conserved bulky amino acid residue in the active site, a "hole" is created that complements a steric "bump" designed into the pan inhibitor, thereby sensitizing the enzyme variant for selective inhibition.[29]

In an effort to expand upon the bump-and-hole strategy, Dr Tsai and Dr Chin developed a complementary approach using bioorthogonal ligand tethering.[31] In this approach, a sensitized target enzyme is generated by incorporating a non-canonical amino acid with a bioorthogonal functionality in proximity to the active site through genetic code expansion. In their work a known pan-kinase inhibitor was repurposed by attaching a complementary bioorthogonal group, such as tetrazine, to the inhibitor (**Figure 1.2**). The strategy relies on the high second-order rate constant exhibited by bioorthogonal inverse electron demand Diels-Alder chemistry, facilitating efficient tethering when the inhibitor conjugate is administered at concentrations that would nominally be sub-inhibitory.[32] This enables selective and rapid inhibition of intracellular kinases, including CRAF, MEK1, MEK2, and LCK, for which no selective small-molecule inhibitors currently exist.

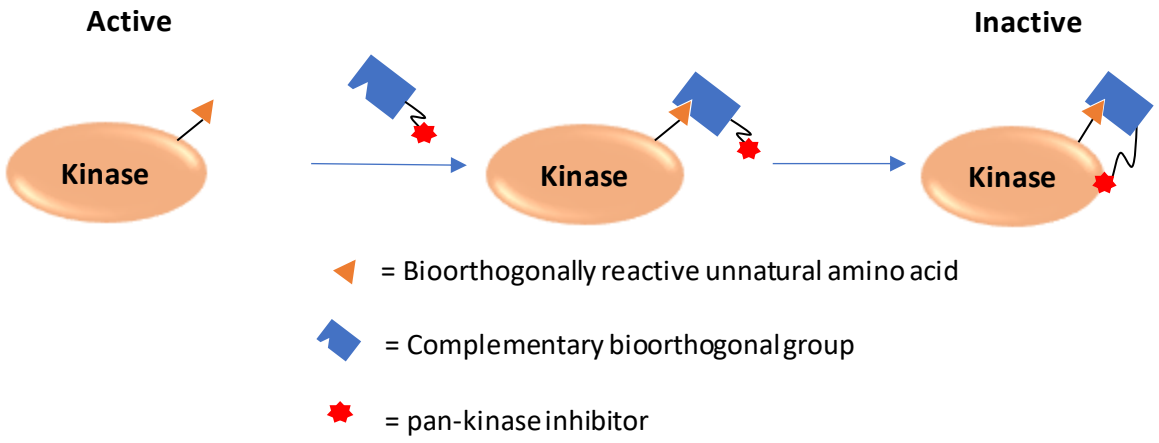


Figure 1.2. Chemical genetic approach for selective inhibition of a protein: Selective modulation of kinases.

MEK1 and MEK2 inhibition was achieved even though they share 82% sequence identity. The similarity between these enzymes makes development of small-molecule inhibitors to selectively target them difficult. Despite this, selective inhibition of both MEK1 and MEK2 was achieved with the bioorthogonal toolkit presented in Figure 1.3. Incorporation of non-canonical amino acids (Figure 1.3A) at specific locations in these enzymes allows for successful inhibition through the tetrazine-containing inhibitor conjugate (Figure 1.3C).

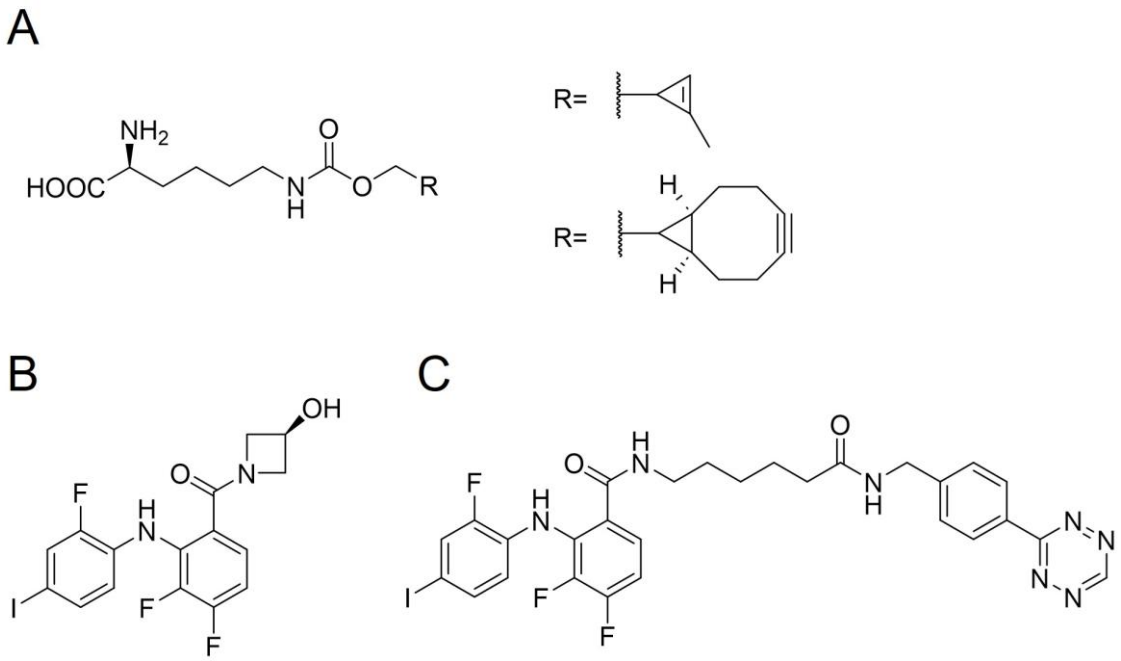


Figure 1.3. Bioorthogonal tools used in Dr. Tsai's kinase research. **A** Non-canonical amino acids with bioorthogonal reactivity. **B** MEK1 and MEK 2 inhibitor. **C** Inhibitor conjugate combining small-molecule inhibitor from **B** and a tetrazine moiety for bioorthogonal reaction with non-canonical amino acids presented in **A**.

The strategy was then demonstrated in a distinct protein kinase family (sharing only 26% sequence identity to MEK1), lymphocyte specific kinase (LCK). For this purpose, a pan-kinase inhibitor, Sunitinib, was used as a template for the creation of another inhibitor conjugate. The Sunitinib was tethered to a tetrazine moiety (Figure 1.4) for bioorthogonal reaction with the non-canonical amino acids presented in Figure 1.3A. The resulting inhibitor conjugate was found to be effective in the same strategy used in MEK1 and MEK2, as well as ultimately proving the flexibility of this strategy with successful use in inhibiting LCK variants.

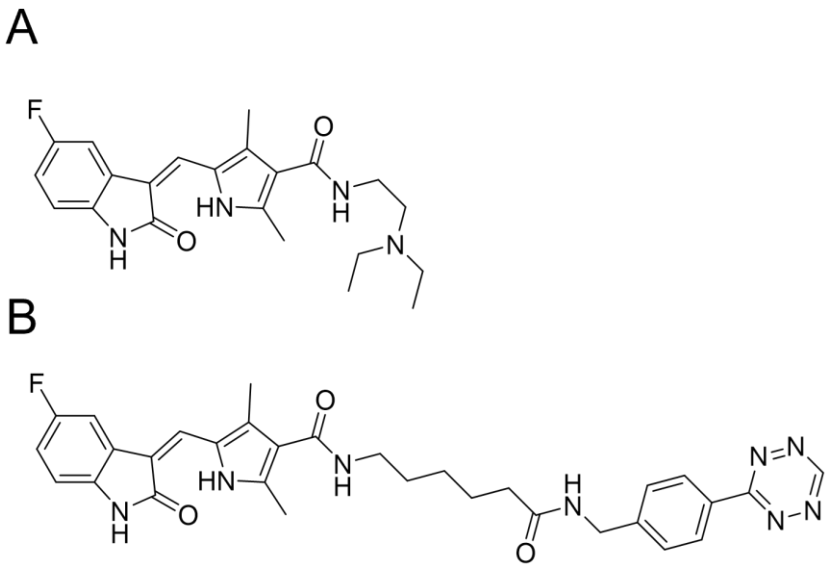


Figure 1.4. A Sunitinib and **B** tetrazine-linked inhibitor conjugate for Suinitinib.

The toolkits described above were even further expanded on through the observation of the effect of linker length and the position of the tethering. As the strategy's success is dependent on fine-tuning of these parameters, one idea that surfaced was the potential to toggle protein function on and off by using photoswitchable linkers between the inhibitor and the tetrazine moieties of the inhibitor conjugates. This approach was successful in demonstrating the first optical toggling of protein function for tethered ligands in live mammalian cells.

However, a limitation of both the bump-and-hole strategy and bioorthogonal ligand tethering approaches is their reliance on repurposing known inhibitors, which restricts their applicability to enzymes for which selective inhibitors are already available.[25] While these methods have shown success in inhibiting specific enzymes, the availability of selective inhibitors is limited, and many enzymes remain poorly characterized or lack specific inhibitors altogether.

The development of selective inhibitors for a particular enzyme often requires extensive screening of compound libraries and subsequent optimization, which is a time-consuming and costly process with no guarantee of success. Furthermore, certain enzymes may have unique structural or functional features that render them challenging targets for the design of selective inhibitors using traditional approaches.

To overcome these limitations, it is crucial to explore alternative strategies that can enable the selective and rapid inhibition of enzymes even in the absence of known inhibitors. To develop a more general approach for the selective and rapid inhibition of any enzyme, this thesis proposes to explore covalent modification of nucleophilic active site residues displayed by a broader enzyme class instead of relying solely on repurposing known inhibitors. This approach would allow for selective and rapid inhibition, even of poorly characterized enzymes. Specifically, the aim is to target the hundreds of enzymes containing a catalytically important cysteine for selective inhibition.

The thiol functional group of catalytically important cysteine residues is crucial for enzyme activity, and its covalent modification can lead to the irreversible inhibition of the enzyme. By developing a strategy that utilizes small-molecule conjugates bearing proximity-dependent thiol-reactive electrophilic warheads, it becomes feasible to selectively and rapidly inhibit enzymes that contain such catalytically important cysteine residues (**Figure 1.5**).

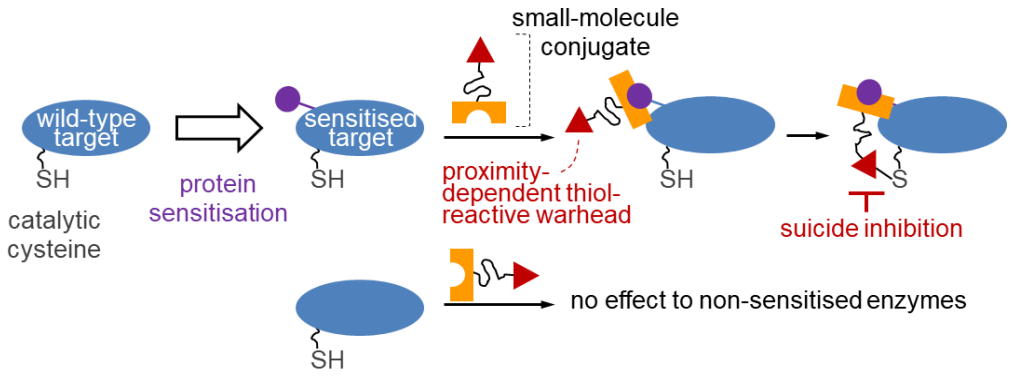


Figure 1.5. Selective inhibition of cysteine-dependent enzymes.

This alternative approach allows us to expand the repertoire of enzymes amenable to selective inhibition, including those that are poorly characterized or lack known inhibitors. Instead of relying on the repurposing of existing inhibitors, this strategy

aims to exploit the inherent reactivity of nucleophilic cysteine residues, opening up possibilities for studying and modulating the functions of a wider range of enzymes.

In summary, by developing a strategy based on covalent modification of nucleophilic active site residues, specifically focusing on enzymes containing catalytically important cysteine residues, we aim to achieve selective and rapid inhibition of poorly characterized enzymes. This approach holds great promise for expanding the repertoire of enzymes amenable to selective inhibition, enabling further exploration of their cellular functions and potential as therapeutic targets.

1.3. Non-canonical amino acid incorporation

In this study, we aim to utilise the incorporation of unnatural amino acids (otherwise known as non-canonical amino acids) into a specific protein of interest. To achieve this, we will explore the process of non-canonical amino acid incorporation and its underlying mechanisms.

To achieve non-canonical amino acid (ncAA) incorporation, we must first take a look at the process of protein translation, the process by which the genetic information stored in DNA is used to synthesize proteins. This process involves the conversion of the information encoded in DNA into messenger RNA (mRNA) by transcription, and the subsequent translation of the mRNA into a specific sequence of amino acids, which are the building blocks of proteins. During translation, the ribosome reads the mRNA sequence and translates it into a corresponding sequence of amino acids, which are brought to the ribosome by transfer RNA (tRNA) molecules. This process occurs in all living organisms and is essential for the growth, development, and function of cells.

Aminoacylation, also known as tRNA charging, is the process of covalently attaching an amino acid to its corresponding tRNA molecule. This reaction is catalyzed by aminoacyl-tRNA synthetases (aaRS), which recognize specific amino acids and their corresponding tRNAs. During aminoacylation, the aaRS first activates the amino acid by forming a high-energy bond between the carboxyl group of the amino acid and adenosine triphosphate (ATP). The activated amino acid is then transferred to the tRNA molecule, forming an aminoacyl-tRNA complex. This charged tRNA can then be used as a substrate for protein synthesis during translation.

There are 20 different aminoacyl-tRNA synthetases, one for each of the 20 standard amino acids. The accuracy and fidelity of the process is essential to ensure that the correct amino acid is added to the growing peptide chain during protein synthesis. This process is a critical step in translating the genetic code from mRNA to protein.

Orthogonal aminoacyl-tRNA synthetase (o-aaRS) and tRNA pairs are genetically encoded aminoacylation systems that are orthogonal, or independent, from the endogenous aminoacylation systems present in cells. These systems are engineered to specifically recognize and charge non-canonical amino acids (ncAAs) onto tRNAs that have been modified to selectively accept these ncAAs. The resulting charged tRNAs can then be used for protein synthesis to incorporate ncAAs into proteins in a site-specific manner.

The term "orthogonal" is used to describe these pairs because they are specifically designed to not cross-react with the endogenous aminoacylation systems in cells, ensuring that only the ncAA-charged tRNA is used for protein synthesis. This allows for precise control over which amino acids are incorporated into proteins and where they are incorporated, enabling the creation of novel proteins with unique properties and functions.

Several different o-aaRS and tRNA pairs have been developed, each with different specificities for different ncAAs. These pairs have been used in a variety of applications, including the creation of biotherapeutics, the study of protein structure and function, and the engineering of proteins with new properties.

One powerful strategy for non-canonical amino acid incorporation is translational control through amber suppression, which will be used in the context of this thesis for protein engineering. This approach involves the use of an amber codon (UAG), which is normally a stop codon in the genetic code, as a sense codon to direct the incorporation of a non-canonical amino acid at a specific site in the protein of interest. Sense codons are also knowns as “coding” or “sense” triplets, a specific sequence of three nucleotides in mRNA that codes for a particular amino acid during protein synthesis. The genetic code is composed of 64 possible codons, of which 61 encode the 20 standard amino acids used in protein synthesis, while the remaining three codons serve as stop signals to terminate protein translation. Sense codons are recognized by tRNA molecules that carry the corresponding amino acids,

allowing for the accurate and specific incorporation of amino acids into the growing polypeptide chain during translation.[33,34]

Amber suppression requires the introduction of an orthogonal tRNA that recognizes the amber codon and is charged with the desired non-canonical amino acid. This orthogonal tRNA is typically derived from a natural tRNA species, engineered to selectively recognize the amber codon through specific modifications in the anticodon region. The orthogonal tRNA is paired with an engineered orthogonal aminoacyl-tRNA synthetase (o-aaRS) that specifically charges the tRNA with the non-canonical amino acid.[35]

To achieve efficient amber suppression, several factors need to be considered. First, the orthogonal tRNA and o-aaRS pair should have minimal cross-reactivity with the endogenous tRNA and aaRS counterparts to ensure specificity. This is crucial to prevent unwanted incorporation of the non-canonical amino acid at other sites in the protein or interference with normal cellular processes.

Second, the efficiency of amber suppression can be influenced by the context surrounding the amber codon. Certain sequence elements, such as the identity of the amino acid preceding the amber codon or the presence of specific RNA secondary structures, can affect the efficiency of tRNA recognition and incorporation of the non-canonical amino acid. Optimizing these sequence elements through rational design or directed evolution approaches can enhance the efficiency and fidelity of amber suppression.[36]

Amber suppression has been successfully employed in various applications. For example, it has been used to introduce non-canonical amino acids with unique chemical properties, such as photo-crosslinking or fluorescent probes, into proteins for structural studies. By incorporating these non-canonical amino acids at specific sites, researchers can obtain precise structural information and gain insights into protein dynamics and interactions.[37]

Additionally, translational control by amber suppression has been utilized to engineer proteins with novel functions or properties. Through the incorporation of non-canonical amino acids, proteins can acquire new enzymatic activities, enhanced stability, altered substrate specificity, or even unnatural post-translational

modifications. These engineered proteins have applications in fields ranging from biocatalysis and drug discovery to materials science and bioengineering.[38,39]

Despite its utility, there are challenges associated with amber suppression. The efficiency of amber suppression can vary depending on the specific non-canonical amino acid and the target protein sequence. Optimization of the orthogonal tRNA and o-aaRS pair, as well as the protein context, may be required to achieve high levels of incorporation. Additionally, the availability and cost of the non-canonical amino acid may limit its widespread use in certain applications.[40]

In summary, translational control through amber suppression offers a powerful strategy for non-canonical amino acid incorporation. By reassigning the amber codon as a sense codon, researchers can precisely introduce non-canonical amino acids into proteins, enabling the study of protein structure, function, and engineering. Continued advancements in orthogonal tRNA and o-aaRS engineering, as well as improvements in the efficiency and diversity of non-canonical amino acids, hold promise for expanding the applications and impact of translational control by amber suppression in various fields of research and biotechnology.

1.4. Aim: back to ubiquitination

Enzymes involved in the ubiquitination pathway play crucial roles in regulating protein homeostasis, cellular signalling, and numerous biological processes. The ubiquitin-proteasome system, consisting of a cascade of enzymatic reactions, orchestrates the attachment of ubiquitin molecules to target proteins, marking them for degradation or modulating their functions. Among the key players in this pathway are the E2 enzymes, which function as ubiquitin-conjugating enzymes.

The E2 enzymes comprise a diverse family of proteins, with multiple isoforms encoded in the human genome. These isoforms share high sequence similarity and possess conserved structural features, reflecting their functional relatedness. Currently, there are approximately 40 known E2 enzymes in humans, with ongoing research continually uncovering new members of this family. However, despite their abundance and importance, the specific functions and individual activities of many E2 enzymes remain poorly understood.[41]

Discerning the distinct activities of each E2 enzyme within this vast family is a significant challenge, as their structural and functional similarities make it difficult to selectively target and modulate their activities. Traditional approaches for studying enzyme function, such as genetic knockdown or knockout, often lack the specificity or temporal resolution necessary to investigate individual E2 enzymes within the complex network of ubiquitination. Furthermore, the development of selective small-molecule inhibitors for each E2 enzyme is a formidable task due to the shared features among these closely related proteins.[10,42]

Herein lies the potential of the inhibition strategy proposed by this thesis. By employing covalent modification of catalytically important cysteine residues, it becomes possible to selectively and rapidly inhibit E2 enzymes, even in the absence of specific inhibitors. The proposed approach offers a unique opportunity to unravel the distinct roles and activities of each E2 enzyme within the ubiquitination pathway. Specifically, the proposed strategy involves the incorporation of a bioorthogonally reactive unnatural amino acid into the E2 enzyme, as illustrated in **Figure 1.6**. This unnatural amino acid carries a complementary bioorthogonal reactive group, along with a proximity-dependent thiol warhead. The thiol warhead undergoes an intramolecular reaction with the E2's catalytic cysteine, forming a covalent bond and rendering the E2 enzyme subsequently inactive. By employing this innovative approach, this thesis aims to create a platform method which will ultimately allow researchers to shed light on the individual functions and activities of various E2 enzymes, facilitating a deeper understanding of their contributions to the ubiquitination pathway, and by extension to cellular processes, signalling pathways and disease mechanisms.

Elucidating the unique functions of each E2 isoform will not only deepen our understanding of the intricacies of the ubiquitination pathway but also shed light on their potential as therapeutic targets. Selective inhibition of specific E2 enzymes may offer new opportunities for modulating protein degradation, altering cellular responses, and designing novel therapeutic strategies for diseases involving dysregulated ubiquitination.

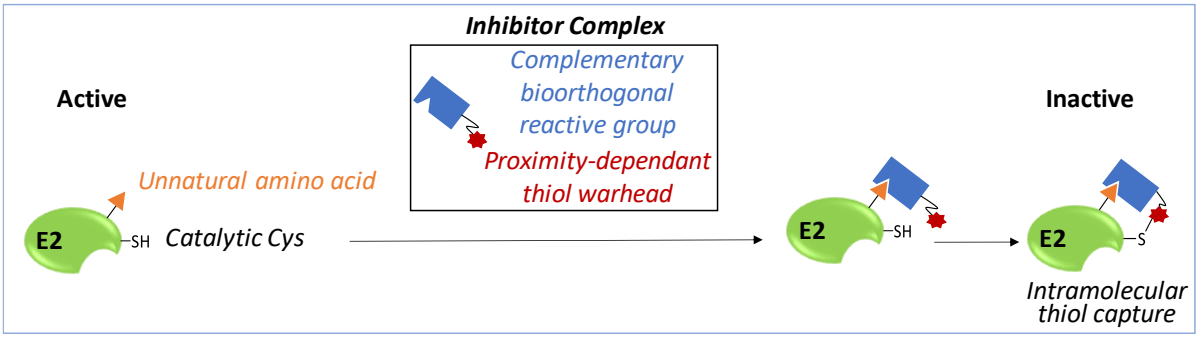


Figure 1.6. Chemical genetic approach for selective cysteine-mediated inhibition of E2 enzymes.

Moreover, the proposed strategy is not limited to the known E2 enzymes but can be extended to the discovery and characterization of new members within this enzyme family. By systematically applying the covalent modification approach to different E2 isoforms, it may become feasible to discern their individual activities and unravel their roles in various cellular contexts.

In summary, the multitude of enzymes involved in the ubiquitination pathway, particularly the closely related E2 enzymes, presents a complex landscape for understanding their distinct functions. The proposed inhibition strategy, based on covalent modification of catalytic cysteine residues, offers a promising avenue for selectively inhibiting and studying individual E2 enzymes.

1.5. Objective breakdown:

The aim of the thesis is to develop a chemogenetic strategy for the rapid and selective inhibition of enzymes in cellular systems using bioorthogonal tethering of inhibitors.

1. Expand the scope of the chemogenetic strategy: The first objective is to broaden the application of the chemogenetic strategy from kinases to cysteine-dependent enzymes. This involves exploring the use of bioorthogonal tethering of electrophilic warheads to target cysteine residues, enabling selective inhibition.
2. Demonstrate selective inhibition in biochemical assays: The thesis aims to provide proof of concept by demonstrating the selective inhibition of specific cysteine-dependent enzymes, such as E2s, in biochemical assays. The

design and synthesis of inhibitor complexes was completed before starting the thesis. This thesis will assess their effectiveness in inhibiting enzyme activity.

3. Investigate the transferability of the strategy: Another objective is to assess the transferability of the chemogenetic strategy to other enzymes within the same family. This exploration will involve identifying suitable target enzymes and optimizing the bioorthogonal tethering approach.
4. Facilitate investigation of cellular functions: The thesis aims to enable the investigation of the cellular functions of the inhibited enzymes. By selectively inhibiting the target enzymes, researchers can study their roles in various cellular processes and pathways, providing insights into their biological functions. For this objective, this thesis will delve into the transfer of this strategy for use in live mammalian cells.
5. Identify optimization opportunities: The thesis aims to identify opportunities for further optimization of the chemogenetic strategy. This includes exploring the use of alternative bioorthogonal amino acids with higher reactivity, optimizing reaction conditions for improved inhibition efficiency, and addressing potential cytotoxicity concerns associated with high inhibitor concentrations.
6. Provide insights for future research: The thesis aims to provide valuable insights and directions for future research in the field of chemogenetic enzyme inhibition. It will discuss the successes, challenges, and limitations encountered during the implementation of the strategy, offering guidance for further refinement and advancement of the approach.

By addressing these objectives, the thesis aims to contribute to the development of a versatile and effective approach for enzyme inhibition in cellular systems, with implications for both research and therapeutic applications.

Chapter 2

ESTABLISHING INHIBITION OF MODEL E2 ENZYME UBE2L3 IN BIOCHEMICAL ASSAYS

2. ESTABLISHING INHIBITION OF MODEL E2 ENZYME UBE2L3 IN BIOCHEMICAL ASSAYS

2.1. Introduction

The human proteome contains approximately 40 E2 enzymes.[43,44] These enzymes are believed to play critical roles in different diseases, including chromosome instability syndromes, cancer predisposition, central nervous system diseases (*e.g.* Parkinson's disease) and immunological disorders (*e.g.* rheumatoid arthritis, systemic lupus erythematosus, inflammatory bowel disease, Crohn's disease).[9] This makes E2 enzymes putative therapeutic targets, however, deciphering the roles of specific E2s in disease development and pathology is challenging by conventional genetic and chemical means. Genetic approaches (*e.g.* knockdown, knockout) have a wide coverage and high specificity, but the poor temporal resolution of genetic inactivation can lead to adaptive compensation that masks the biology. On the other hand, small-molecule inhibitors act rapidly and are powerful tools for discerning protein function, but selective small-molecule inhibitors against most E2s are not yet available.[10,45,46]

Human E2 enzymes share a conserved catalytic domain, but their N- and C-terminal regions can be variable, leading to differences in their overall structure and function. Additionally, different E2 enzymes have unique sequences that can result in differences in their substrate specificity, subcellular localization and regulation. Traditionally, E2s have been classified based on the presence of additional extensions to their UBC domain, the catalytic core. Class I E2s present only the UBC domain, Class II and III E2s present N- or C-terminal extensions respectively, and finally Class IV E2s present both N- and C-terminal extensions to the UBC domain. These extensions provide additional functionality to the E2s, such as stabilising the interaction with E1 enzymes, modulating the activity of paired E3 enzymes or enzyme location within the cell.[41]

Human E2s have also been classed based on phylogenetic analysis in 17 subfamilies.[47] Some authors even suggest classifying them by whether they have an aspartate or serine at a conserved serine/aspartate site, the former being constitutively active and the latter being regulated by phosphorylation.[48] This

classification however does not apply to all E2s equally, even though there seems to be a selective pressure to maintain a serine or aspartate at this position. For example, the first E2 chosen to be a part of this thesis' study, UBE2L3, features a different negatively charged residue in this serine/aspartate site in the form of a glutamate (E118).

The nomenclature and classification of E2 proteins can be somewhat confusing, but new insights on their structural characteristics, functions within the cell and the catalytic systems behind their function have been expanding over the past few decades. As key players in the complex ubiquitination system, more pieces of the puzzle are being revealed. One set of the key pieces is discerning the role of different E2s in the regulation of other protein activity and protein degradation.

Despite this complexity we can home in on their common feature, all E2 enzymes can be said to present a similar core catalytic domain. This domain is composed of around 150 amino acids, and is dubbed the Ubiquitin Conjugation (UBC) domain. This domain is responsible for the covalent attachment of ubiquitin to the E2 enzyme itself, as well as the transfer of ubiquitin to a target substrate. It contains the catalytic cysteine residue that forms a thioester bond with ubiquitin during the conjugation process. Although E2 enzymes are usually described as just the carrier enzymes for Ub from E1 enzymes to E3 enzymes, their single active site helps them engage in the transfer of ubiquitin to a wide variety of different E3 enzymes, and have even been shown to transfer ubiquitin directly to substrate proteins without the aid of E3 enzymes. Their active site allows them to participate in primarily two types of reactions (Figure 2.1). One, this enables them to carry out the transfer of Ub from a thioester to a thiol group in transthiolation reactions. Two, they are able to transfer Ub from a thioester to an amino group.[44,49]

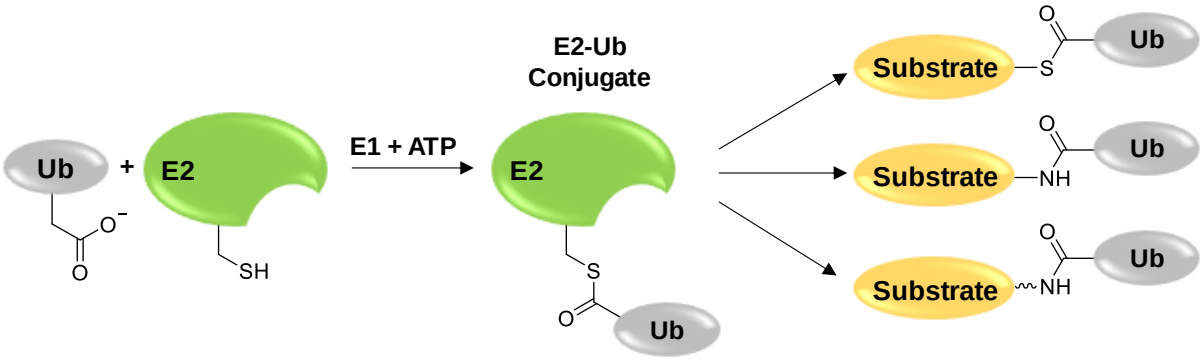


Figure 2.1. In the ubiquitination cascade, ubiquitin's C-terminal Gly is conjugated to E1's active site in an ATP dependent reaction. This activated ubiquitin is then transferred to the catalytic cysteine residue in the E2 enzyme active site, forming a thioester bond. The E2-ubiquitin thioester intermediate has the ability to react with the side chain of a Cys, Lys or N-terminus on the substrate for the transfer.

For proof-of-concept, UBE2L3 (also known as UbcH7) was chosen as the model E2, one of the most abundant E2 enzymes in mammalian cells. As a starting point, UBE2L3 offered readily available structural and biochemical information as research in the field has grown. UBE2L3 has been implicated in the development of various cancers, Parkinson's disease, systemic lupus erythematosus (SLE), rheumatoid arthritis.[9,50,51] For example, the upregulation of UBE2L3 has been detected in non-small cell lung cancer (NSCLC) tissues when compared to non-cancerous tissues. High expression of UBE2L3 was linked to advanced tumour stage and adverse outcomes, and a suppression of UBE2L3 production lowered NSCLC cell growth.[52]

There are still areas of UBE2L3's involvement in various signalling pathways, immune disease and cancer development which require further study. There is also still a need for specific inhibitors of UBE2L3 and a study of their effect and safety in the treatment of the associated diseases.[53] Understanding UBE2L3's cellular function should provide new perspectives and potential strategies for disease diagnostics and targeted therapeutic development. This makes UBE2L3 an excellent target candidate for the development of the inhibition strategy proposed in this work.

In cells, UBE2L3 works together with UBE1 (an E1 enzyme) and HOIP (an E3 ligase) to form linear polyubiquitin chains, and this reaction can be reproduced *in vitro* to monitor the activity of UBE2L3.[54] UBE2L3 forms productive E2-E3 pairs with the E3 complex LUBAC (linear ubiquitin chain assembly complex), of which HOIP is the critical catalytic subunit.

The E3 selected for the polyubiquitination assays in this project is HOIP (HOIL-1-interacting protein), which is the central E3 ligase of the linear ubiquitin chain assembly complex (LUBAC). It is part of the RING-between-RING (RBR) family of RING-type E3 ligases, which forms a covalent intermediate with ubiquitin similar to how HECT-type E3 ligases function, before transferring the ubiquitin to the target substrate, or ubiquitin itself.[54,55] HOIP is able to form free linear polyubiquitin chains without the need for a protein substrate. Using this functionality, an analysis of similar research in the field has shown that the formation of polyubiquitin chains by UBE1, UBE2L3 and HOIP can be tracked via SDS-PAGE analysis.[56] If proven successful, this assay can then further be used to visualise the inhibitory effects being studied on the reaction components.

An example of the visualisation of this polyubiquitin chain formation is shown below (**Figure 2.2**). In this experiment, UBE2L3* was incubated with UBE1 and HOIP in the presence of ATP, leading to the formation of polyubiquitin chains.

UBE2L3* is a variant of the target E2 enzyme where the non-catalytic cysteine residues are mutated to serine, and could also be named UBE2L3(C17S/C137S). This variant serves as a model enzyme to verify the selective inhibition strategy proposed in this work, as it only contains one cysteine residue, the active site cysteine.

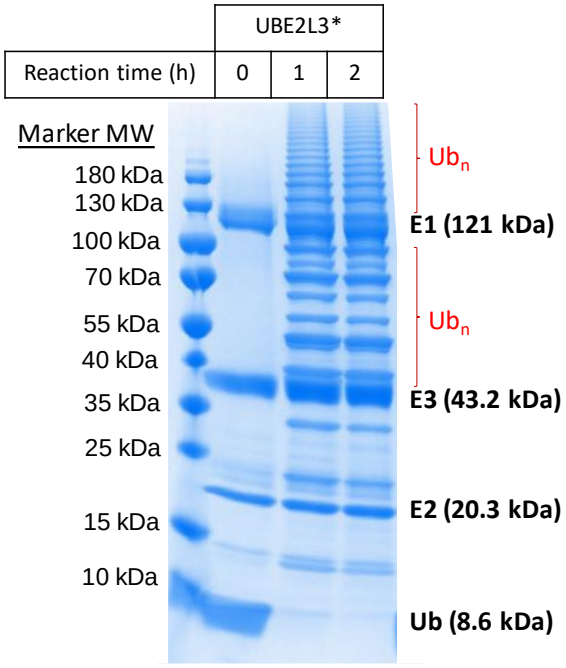


Figure 2.2. Polyubiquitin chain reaction. Polyubiquitin chains are assembled by UBE1, UBE2L3 and HOIP in the presence of ATP.

Based on the crystal structure of UBE2L3, 10 target amino acid residues within the protein sequence were chosen to be replaced with a cyclopropene derivative of lysine (CypK, **1**) via amber stop codon (UAG) suppression,[35,57] generating UBE2L3(XXX-**1**) variants in which XXX denotes the amino acid residue replaced with **1** (**Figure 2.3F**). The selected amino acid residues to be replaced are within 24 Å of UBE2L3's catalytic cysteine residue (C86), and this distance is within the maximum theoretical distance (around 35 Å) needed if **1** and C86 are connected by one of the inhibitor conjugates.

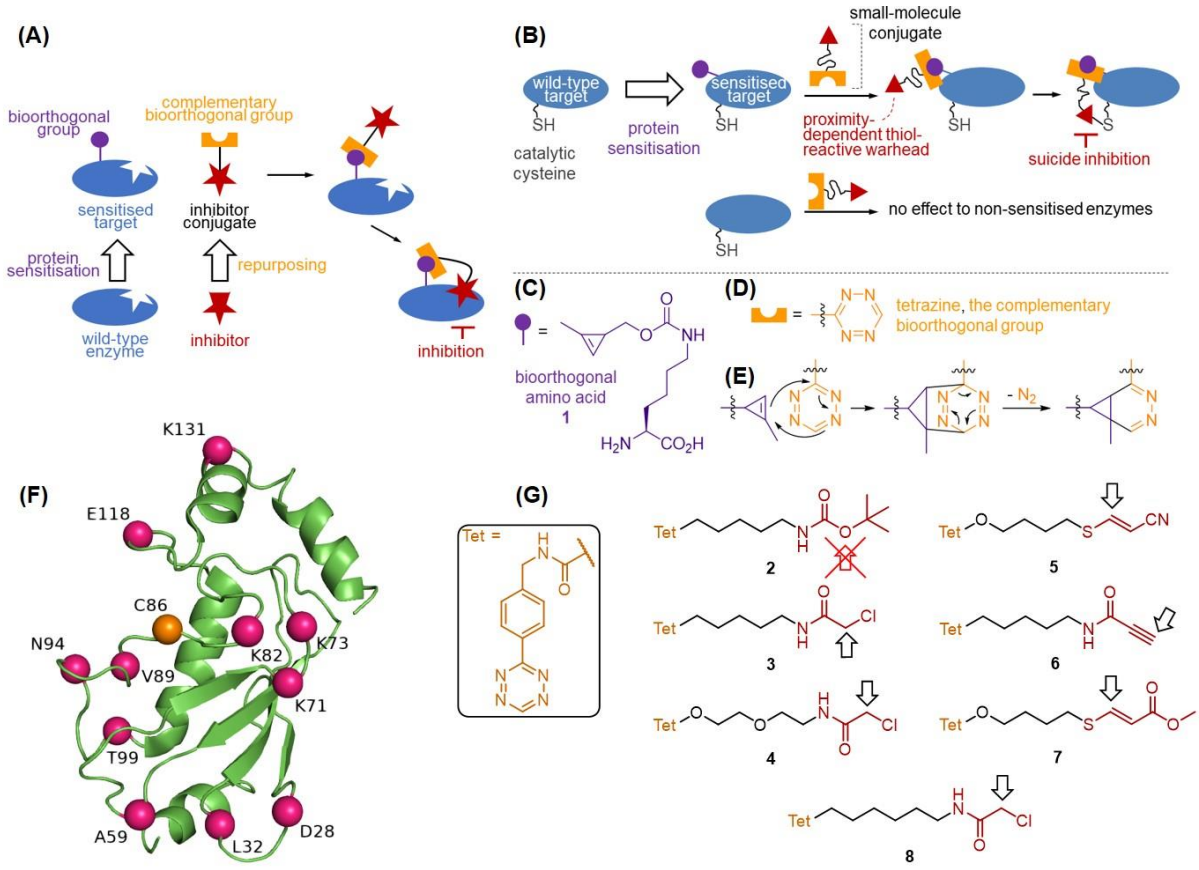


Figure 2.3. Structures of CypK, inhibitor conjugates and UBE2L3.
(A) Chemical genetic inhibition by bioorthogonal ligand tethering. **(B)** Selective inhibition of a cysteine-dependent enzyme. The small molecule conjugate is expected to only inhibit the sensitised target but have no effect to the wild-type and any non-sensitised enzymes. **(C)** Structure of bioorthogonal amino acid **1** intended for use in protein sensitisation. **(D)** Structure of the complementary bioorthogonal group intended for use in the small-molecule conjugate. **(E)** Mechanism of the bioorthogonal tethering. **(F)** Structure of UBE2L3, highlighting the catalytic Cys86 and residues to be substituted by **1**. PDB: 1C4Z. **(G)** Inhibitor conjugates **2**-**8**.

The incorporation of non-canonical amino acids such as **1**, specifically at the amber codon, allows for the generation of UBE2L3 variants with specific modifications. This is because the amber codon, which typically signals the end of protein translation, is instead used as a site for the incorporation of a non-canonical amino acid. In the absence of the appropriate aminoacyl-tRNA synthetase/tRNA pair, translation will terminate prematurely, resulting in a truncated protein. Therefore, it is crucial to use a specific pair of aminoacyl-tRNA synthetase and tRNA that can specifically recognize the non-canonical amino acid and incorporate it at the amber codon site.

In this work, we will utilize a plasmid-based system to introduce the appropriate aminoacyl-tRNA synthetase/tRNA pair and the plasmid DNA encoding UBE2L3

variants with specific modifications. The recombinant production of UBE2L3 variants will be carried out in *E. coli*, which is a commonly used host for protein expression due to its rapid growth rate and well-established genetic tools. Once the plasmids are introduced into *E. coli*, the bacteria will be grown under controlled conditions to produce the desired UBE2L3 variants. The resulting proteins can then be purified and characterized to determine their biological activity and potential applications.

Through the incorporation of CypK, UBE2L3 variants will be sensitised to small-molecule conjugates **2**-**8** (**Figure 2.3G**), which contain both a proximity-dependent thiol-reactive group and a bioorthogonal reactive group complementary to **1**. Upon bioorthogonal tethering (**Figure 2.3E**), the thiol-reactive group on the inhibitor conjugates should form a covalent bond with UBE2L3's catalytic cysteine in a proximity-dependent manner, effectively inhibiting E2 activity. The unnatural amino acid should neither react with any biomolecules under physiological conditions nor affect the activity of the target enzyme.

In recent approaches for kinase inhibition, Dr. Tsai successfully used bioorthogonal tethering by introducing **1** in a kinase and derivatizing an inhibitor with a tetrazine complementary bioorthogonal group (**Figure 2.3D**), enabling rapid reaction with the cyclopropene functionality on **1** ($k \sim 10\ M^{-1}s^{-1}$, **Figure 2.3E**).[31] Cyclopropene lysine was chosen for bioorthogonal tethering due to its higher incorporation efficiency in mammalian cells compared to other tetrazine-reactive bioorthogonal amino acids.[58] Furthermore, selective tethering and inhibition of cyclopropene-containing LCK kinase in mammalian cells were achieved within 20 minutes after incubation with culture media containing a tetrazine-bearing conjugate at 1 mM concentration.[31]

The low reactivity of proximity-dependent thiol-reactive warheads ($k = 10^{-2}$ to $10^{-3}\ M^{-1}s^{-1}$) means that at single-digit micromolar concentrations, the half-life of diffusion-driven cysteine reactivity is on the timescale of months or even years. However, proximity-acceleration by up to five orders of magnitude is possible.[59] Therefore, the conjugate is expected to exhibit high specificity in inhibiting the target enzyme while having negligible off-target activity in cells.

When studying the inhibition by the proposed conjugates, the possibility that inhibition occurs due to steric occlusion of E3 binding by the tethered conjugate cannot be ruled out. As a result, the inhibition by inhibitor conjugate could be due to

one of two outcomes, or a mixture of both: (a) The inhibitor conjugate interferes with E3 binding to E2's conserved binding region, or (b) the covalent bond formed by the inhibitor conjugate with the catalytic cysteine on E2 prevents ubiquitin transfer from E1 to E2, and consequently to E3. To elucidate the mechanism for inhibition (steric occlusion, cysteine modification or both), different inhibitor conjugates will be tested. One set contains different reactive groups with different linker lengths in order to form a covalent bond with the cysteine residue (**3**-**8**) and another inhibitor contains a non-reactive bulky *tert*-butyl group (**2**).

The optimization of this chemical genetics[60] strategy will allow the distinct members of the E2 protein family to be rendered sensitive to a selective and complementary small molecule, which similarly has no effect on other members of the same protein family. This will allow the assessment of the specific biological roles of different E2 enzymes *in vitro*. As a result, this new dual strategy will enhance specific targeting of a chosen variant from a large family of closely related enzymes.

2.2. Cloning, protein expression and purification

2.2.1. Cloning

As part of the work carried out in this study, unless stated otherwise, HOIP and UBE2L3 variants were expressed and purified in the lab.

A combination of pET and pCDF plasmid backbones were used for the expression of UBE2L3 variants, both of them being commonly used plasmids for protein expression in *E. coli* expression systems.

The pET plasmid is widely used for high-level expression of recombinant proteins in *E. coli.* The pET system utilizes the T7 promoter and RNA polymerase for strong and tightly regulated gene expression. The plasmid commonly contains an ampicillin resistance gene for selection in bacteria (although this can be changed for different antibiotic selection) and a His-tag sequence for purification of the recombinant protein. The pET system also offers a variety of vectors with different characteristics such as different selection markers and tags.

The pCDF plasmid is similar to the pET system, but it is designed for co-expression of multiple genes in E. coli. The pCDF system utilizes two different promoters for co-

expression of two genes in a single plasmid. The plasmid also contains an antibiotic resistance gene (spectinomycin in the case of this study).

While pET plasmids are widely used for high-level expression of a single recombinant protein, pCDF plasmids are designed for co-expression of multiple genes in a single plasmid. This is relevant in the case of this study because of the incorporation of the orthogonal aminoacyl-tRNA synthetase and its complementary tRNA in the same plasmid.

Firstly, the gene encoding for HOIP(697-1072) with an N-terminal hexa-histidine tag (His tag) was purchased as a gene fragment from *ThermoFisher GeneArt Strings*. *GeneArt Strings* also offers a codon optimisation tool, which was used to fine-tune the gene for expression in *E. coli*. Codon optimisation was used to avoid a decrease in protein production efficiency due to codon usage bias, HOIP being a human gene expressed in *E. coli*.[61]

To allow for His tag removal after protein purification, a Tobacco Etch Virus (TEV) protease recognition site (ENLYFQG) was included between HOIP and the hexa-histidine tag, permitting TEV cleavage and subsequent removal of the His tag.[62] This fragment was cloned into the *NdeI* and *EcoRI* sites a pET28a vector (kanamycin resistant) to generate the plasmid **pET HOIP**.

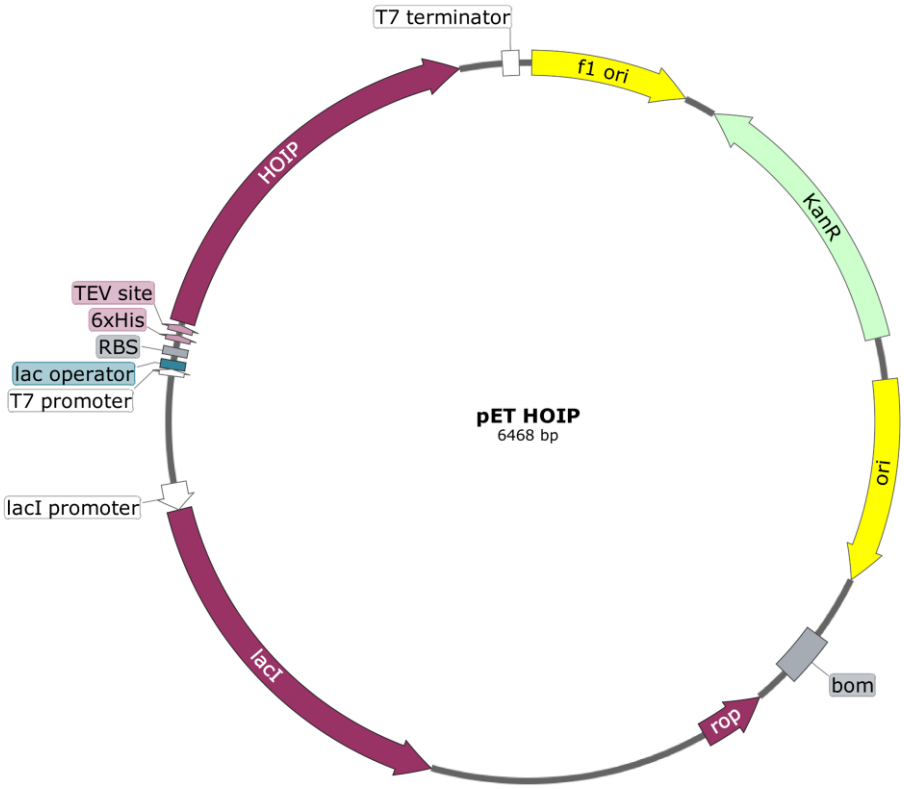


Figure 2.4. Plasmid map for pET HOIP.

Secondly, the first variants of UBE2L3 to be used in this study were provided by Dr Satpal Virdee (University of Dundee) in pET15b vectors. These constructs have the non-catalytic cysteine residues mutated to serine, serving as models to verify catalytic cysteine inhibition, and were named UBE2L3*. The cysteine residues substituted by serine in UBE2L3 were located at positions 17 and 137 in the amino acid sequence (C17S, C137S). These UBE2L3* variants also include the amber stop codon (TAG), which marks the site for the incorporation of unnatural amino acid (CypK, **1**). The site of the TAG codon (amber codon) in the amino acid sequence is denoted by XXX in the plasmid name **pET UBE2L3*(XXXTAG).** In turn, XXX denotes the amino acid residue that will be substituted by **1**. Additional variants, **pET UBE2L3*(59TAG)** and **pET UBE2L3*(73TAG)**, were obtained in-house over the course of this study by mutating the relevant codons to TAG via site-directed mutagenesis (SDM) of **pET UBE2L3***. All UBE2L3* constructs also contain an N-terminal His tag.

In order to express the UBE2L3* variants containing **1**, an orthogonal aminoacyl-tRNA synthetase/tRNA pair was required. The plasmid containing the

synthetase/tRNA pair were readily available in Dr. Tsai's plasmid library as **pCDF MbPylT^{CUA} 3xMbPylRS**.

Lastly, the non-catalytic cysteines C17 and C137 were re-introduced in-house via site-directed mutagenesis, generating a plasmid construct for wild-type UBE2L3, **pET UBE2L3**. Cysteines were also re-introduced in-house into a variant, generating **pET UBE2L3(89TAG)**.

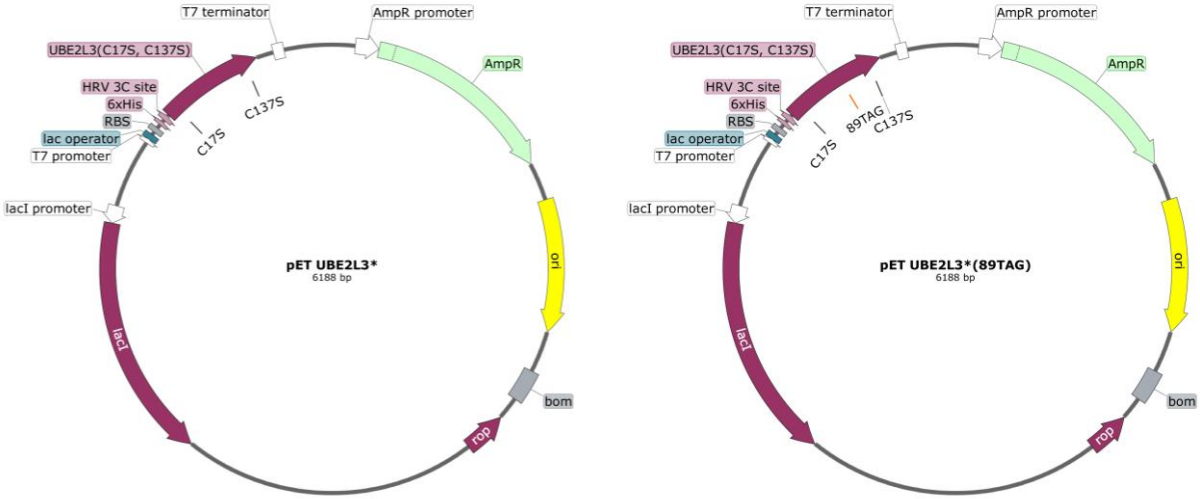


Figure 2.5. Plasmid map for pET UBE2L3* and pET UBE2L3*(89TAG).

2.2.2. Protein expression and purification

Proteins in this section were expressed using the pET vector system. The expression of proteins of interest in this system is regulated by a promoter for T7 RNA Polymerase, which originates from the T7 bacteriophage. Chemically competent *E. coli* cells such as BL21 (DE3), used in this project, have been engineered at a genomic level to produce T7 RNA polymerase, which is controlled by the lacUV5 promoter. The pET vectors include a gene that encodes for the lac repressor protein, lacI. As lacI is expressed, it binds to the lac promoter in the bacterial genomic DNA. This prevents the native RNA polymerase in *E. coli* from binding to the lac promoter and initiating transcription of the T7 RNA polymerase gene. However, this inhibition can be relieved by addition of allolactose or isopropyl-1-thio-β-D-galactopyranoside (IPTG). This releases the lac repressor and allows the transcription and subsequent translation of T7 RNA polymerase. The T7 RNA polymerase is then able to bind to a T7 promoter in the pET vector used, leading to the transcription and translation of the gene of interest. IPTG is favoured over

allolactose, as it is not hydrolysable by the bacterial cell and therefore non-degradable by bacterial metabolic enzymes, allowing for steady control of protein expression.[63,64]

The proteins expressed in-house in order to carry out inhibition assays for UBE2L3 were HOIP, UBE2L3 and its variants.

2.2.2.1. HOIP

Expression of 6xHis-HOIP(697-1072) through **pET HOIP** was carried out in in BL21 (DE3) pLysS cells. More detailed protein expression and purification protocols are detailed in ***Materials and Methods, section 6.6***. Cultures were incubated at 16 °C overnight before harvesting the cells. Cell pellets were lysed and purified via Ni-NTA affinity chromatography. The protein started to elute at imidazole concentrations over 25 mM, which was confirmed by SDS-PAGE (**Figure 2.6A**). The fractions using a concentration of 75-200 mM imidazole were pooled, concentrated and the buffer was exchanged before cleavage of the His tag using TEV protease.

HOIP has several Zn finger motifs, as does the construct used in this study, HOIP(697-1072).[55] With this in mind, the best conditions for removing the His tag were found to be those sensitive to working with zinc fingers, where stronger metal chelators used in standard cleavage reactions such as EDTA and DTT should be avoided. DTT was replaced with a monothiol, β-mercaptoethanol (BME), and EDTA was replaced by a weaker metal chelator, citrate. Cleavage conditions for TEV protease are detailed in ***Materials and Methods, section 6.6.2***. After cleavage of the His tag, the reaction mixture was allowed to drip through NiNTA affinity resin. This allowed for any remaining His tagged protein to bind to the resin. The flow-through (FT) collected then contained untagged HOIP(697-1072). The resin was washed further with cleavage buffer supplemented with increasing amounts of imidazole, samples were prepared and then run on SDS-PAGE (**Figure 2.6B**).

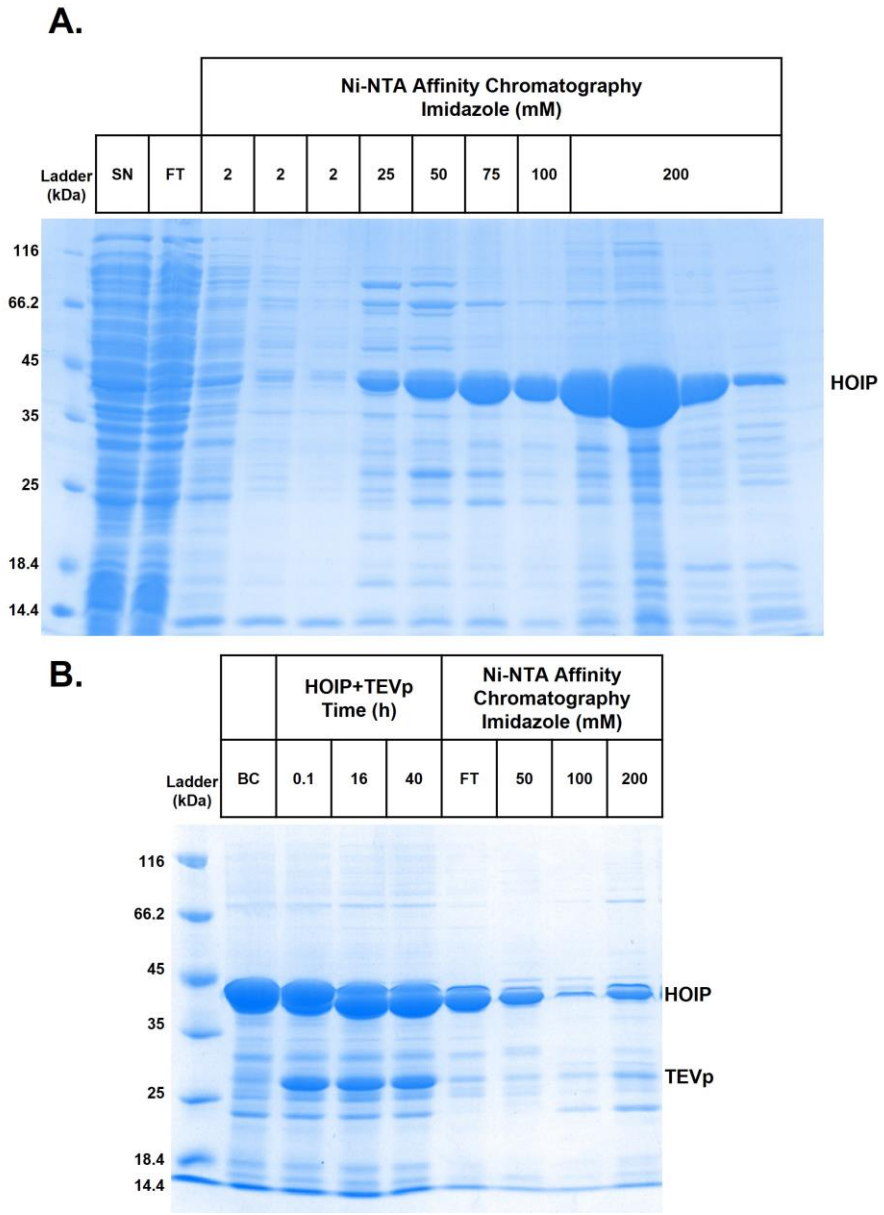


Figure 2.6. SDS-PAGE analysis of 6xHis-HOIP purification and cleavage of the His tag. **(A)** SDS-PAGE analysis of 6xHis-HOIP purification via Ni-NTA affinity chromatography shows that protein appeared to elute from imidazole concentrations as low as 25 mM, carrying with it some higher MW impurities. SN denotes the sonication supernatant and FT denotes Ni-NTA flow-through. **(B)** SDS-PAGE analysis of the cleavage of the His tag on 6xHis-HOIP using TEV protease.

The reaction was trialled at a duration of 40 h, however there was no observable difference in cleavage after the 16 h mark via SDS-PAGE. The activity of the protein contained in the flow-through (FT) and first wash (50 mM imidazole) was determined later via a polyubiquitination assay. In addition, cleavage of the His tag was confirmed via mass spectrometry (**Figure 2.7**).

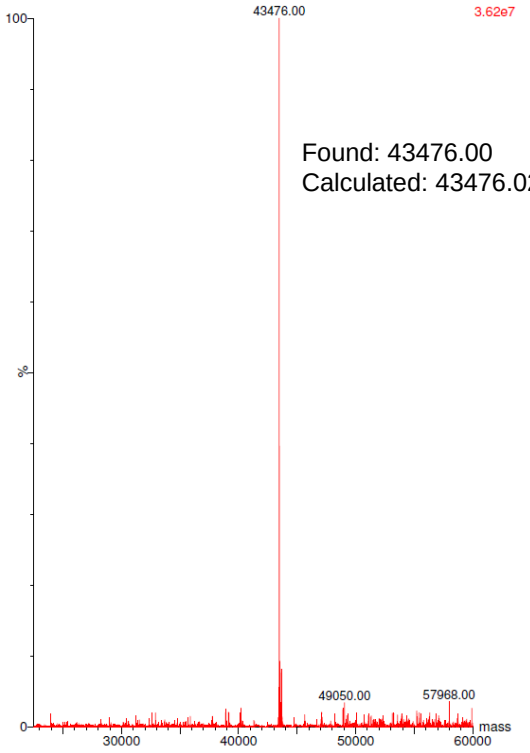


Figure 2.7. Mass spectra of HOIP.
Deconvoluted mass spectra of HOIP indicating species with a molecular mass of 43,476 Da, consistent with the expected mass post-cleavage of the His tag.

In the process of optimising the method for producing HOIP, it was found that the presence of the histidine tag hinders the rate at which ubiquitin is turned over into polyubiquitin chains. In the presence of UBE1, UBE2L3, ubiquitin and ATP, HOIP catalyses to the formation of polyubiquitin chains. A side-by-side comparison of 6xHis-HOIP and untagged HOIP was carried out in the presence of the other components of the model polyubiquitination reaction used in this study for the analysis of UBE2L3* activity (**Figure 2.8**). The analysis showed that while untagged HOIP was able to consume most of the ubiquitin provided in the reaction within 30 minutes, for 6xHis-HOIP the reaction could take up to 2 hours. Therefore, it was decided to ensure cleavage of the histidine tag before using HOIP in the biochemical activity assays for UBE2L3.

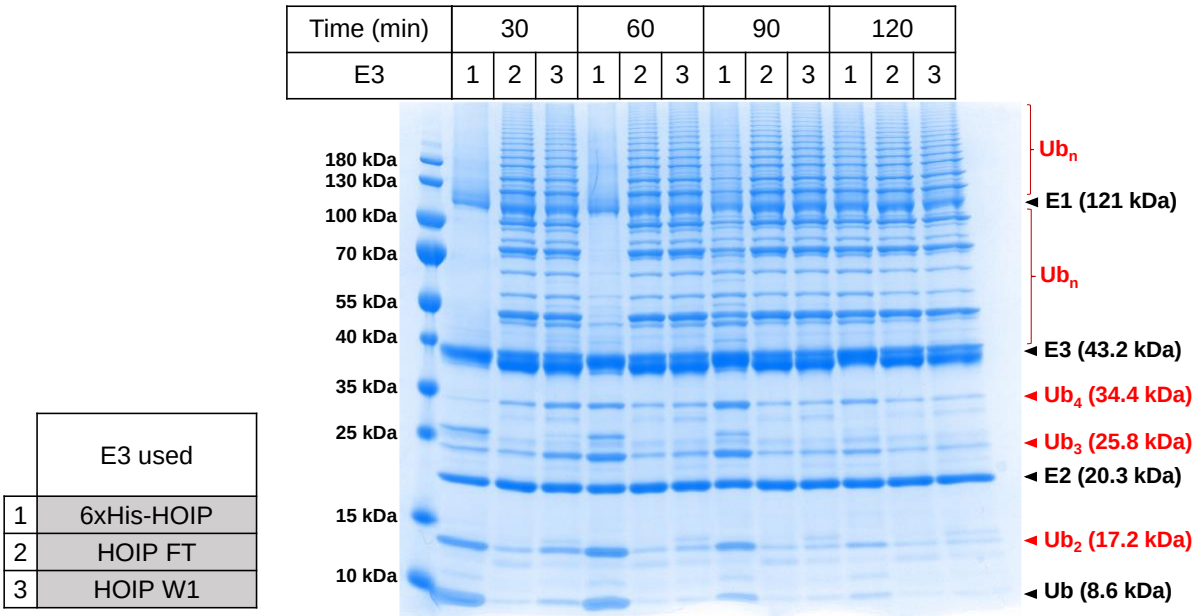


Figure 2.8. SDS-PAGE analysis of HOIP activity via polyubiquitination reaction. Side by side comparison of 6xHis-HOIP (**1**) activity compared to the untagged HOIP fractions represented in **Figure 2.3B** as FT (**2**) and the first wash at 50 mM Imidazole (**3**). Reaction carried out in the presence of ubiquitin, UBE1, UBE2L3 and ATP. Untagged HOIP consumes ubiquitin substrate within the 30-minute mark, whereas His tagged HOIP takes up to 2 hours to reach reaction completion.

2.2.2.2. UBE2L3*

Expression of UBE2L3* through **pET UBE2L3*** was carried out in BL21 (DE3) cells. Protein expression and purification protocols are detailed in ***Materials and Methods, section 6.6***. Cell pellets were lysed and purified via Ni-NTA affinity chromatography (**Figure 2.9**). Purified UBE2L3* was confirmed via mass spectrometry (**Figure 2.10**).

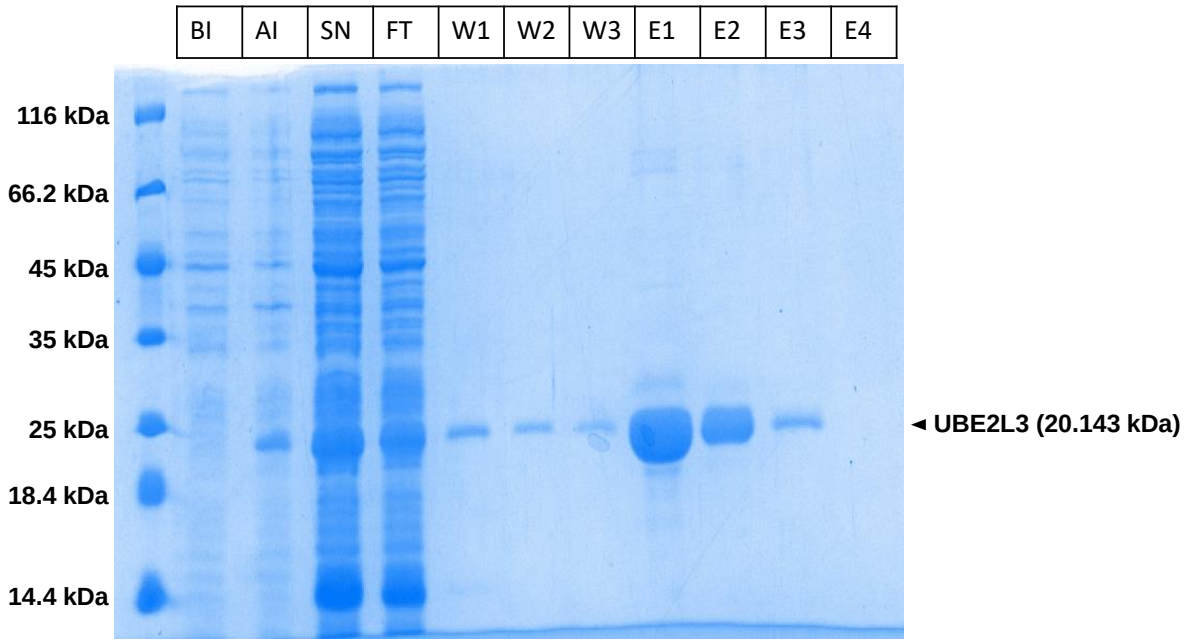


Figure 2.9. SDS-PAGE analysis of UBE2L3* purification.

BI and AI denote before and after induction with IPTG, respectively, showing an appearance of UBE2L3* overexpression after inducing expression with IPTG. After the cells were lysed and centrifuged, the supernatant (SN) is then allowed to pass through Ni-NTA resin, collecting the flow-through (FT) and then washing with low imidazole (25 mM) concentration buffer (W1-3), followed by elution of the His-tagged UBE2L3* in high imidazole (300 mM) buffer (E1-4).

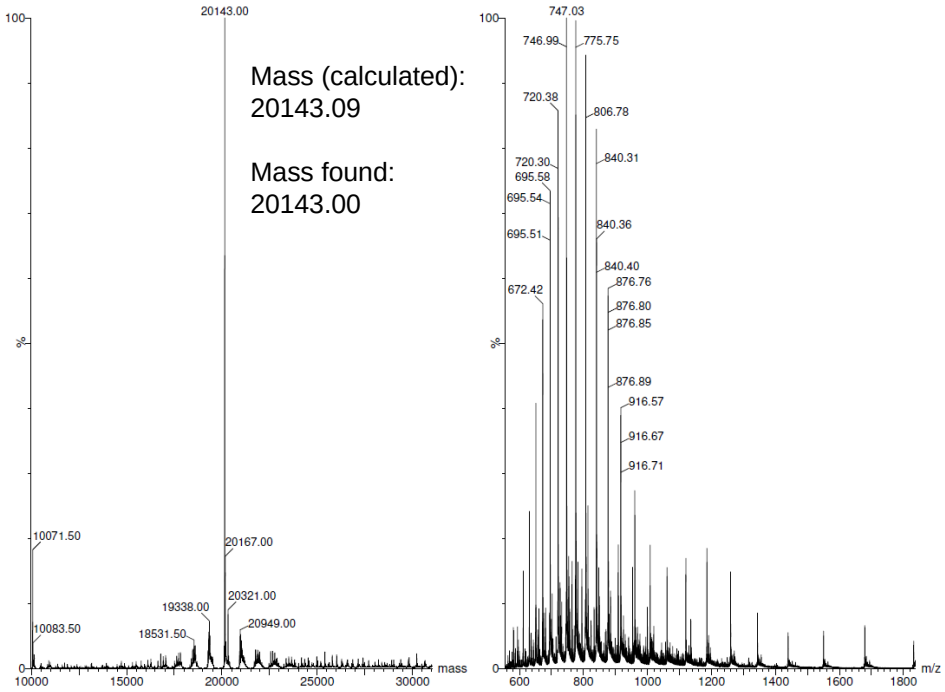


Figure 2.10. Mass spectra of UBE2L3*.

Deconvoluted mass spectrum (left) and m/z signals from intact protein full scan mass spectrum (right) of UBE2L3* indicating species with a molecular mass of 20143.00 Da, consistent with the expected mass.

2.2.2.3. UBE2L3* variants

Initial stocks of UBE2L3* variants were received from Dr Satpal Virdee (University of Dundee). Samples from these stocks were analysed and confirmed by mass spectrometry.

Further expression of UBE2L3* variants through their corresponding plasmids **pET UBE2L3*(XXXTAG)** was carried out in BL21 (DE3) cells. Addition of CypK (**1**, 0.5 mM) was performed prior to induction of the cultures. Protein expression and purification protocols are detailed in ***Materials and Methods, section 6.6***. Cultures were incubated at 37 °C for 4 hours before harvesting the cells. Cell pellets were lysed, purified via Ni-NTA affinity chromatography and fractions were analysed via SDS-PAGE. Variants are named UBE2L3*(XXX-**1**), where the mutation position in the amino acid sequence is denoted by XXX, followed by **1** to mark they contain CypK in that position. Purified UBE2L3*(XXX-**1**) were further confirmed via mass spectrometry.

Interestingly, the UBE2L3* variant with the TAG mutation closest to the C terminal end of the protein, UBE2L3*(135-**1**), showed consistent early termination of protein translation via SDS-PAGE analysis, resulting in a truncation of the protein, both in the initial stocks and in the further attempts to express and purify them. In the figures below, UBE2L3(82-**1**) and UBE2L3(135-**1**) are shown in a side-by-side comparison. Both variants have an equal expected mass, as in both cases a lysine was the residue to be substituted by **1**. Firstly, the SDS-PAGE analysis following the steps to purify these expressed proteins, showed that the protein obtained after Ni-NTA purification for UBE2L3(135-**1**) resulted in a lower molecular weight species than that obtained for UBE2L3(82-**1**). In the figure presented, the SDS-PAGE analysis provides labelled lanes above the gel. The immediate result of cell lysis is centrifuged to clarify the solution from cell debris, the resulting supernatant is labelled SN. The flow-through from the Ni-NTA purification column is labelled FT. The column is then washed, and the obtained fractions are labelled W1-W3. A high imidazole (300 mM) buffer is then used to elute the proteins bound to the Ni-NTA resin, the fractions collected during protein elution are labelled E1-E3 (**Figure 2.11**).

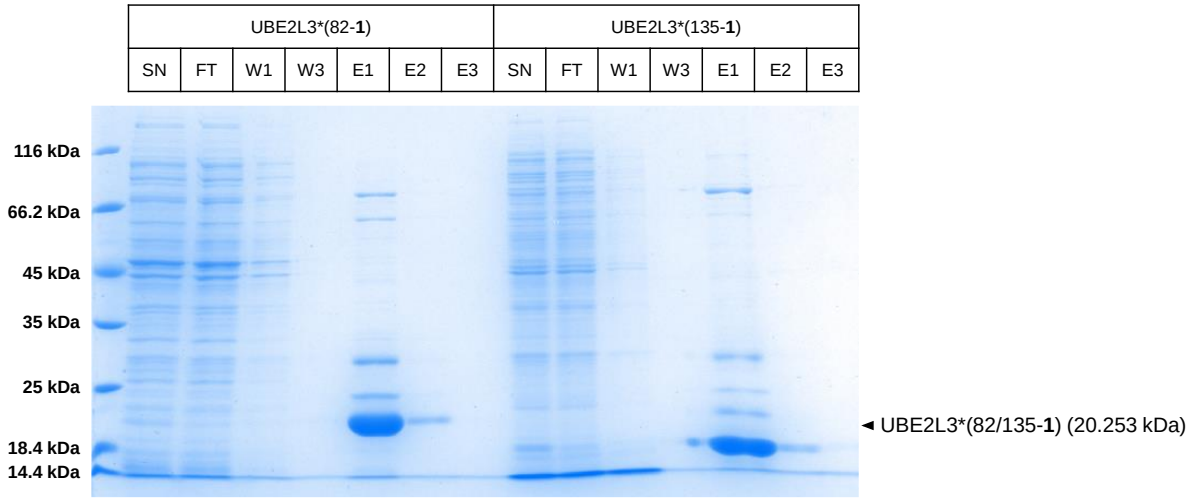


Figure 2.11. SDS-PAGE analysis of UBE2L3* variant purification.
SDS-PAGE analysis of the in-house purification of UBE2L3* variants shows that for UBE2L3*(82-**1**) the size of the protein is consistent with its molecular weight (20.25 kDa), but for the UBE2L3*(135-**1**) variant there are signs that protein translation was terminated early at the TAG codon, resulting in a lower molecular weight than expected (17.76 kDa).

Secondly, both proteins were subject to MS analysis. The LC-MS analysis of UBE2L3*(82-**1**) is shown in the figure below (**Figure 2.12**), and presents an observed mass consistent with its expected mass of 20.25 kDa. LC-MS analysis further confirmed that the UBE2L3*(135-**1**), which has an expected mass of 20.25 kDa, also had a major peak with a mass of 17.76 kDa (**Figure 2.13**). This lower mass protein peak corresponds to the expected mass of the protein resulting from early termination of translation at the TAG codon, the site of the incorporation of unnatural amino acid **1**.

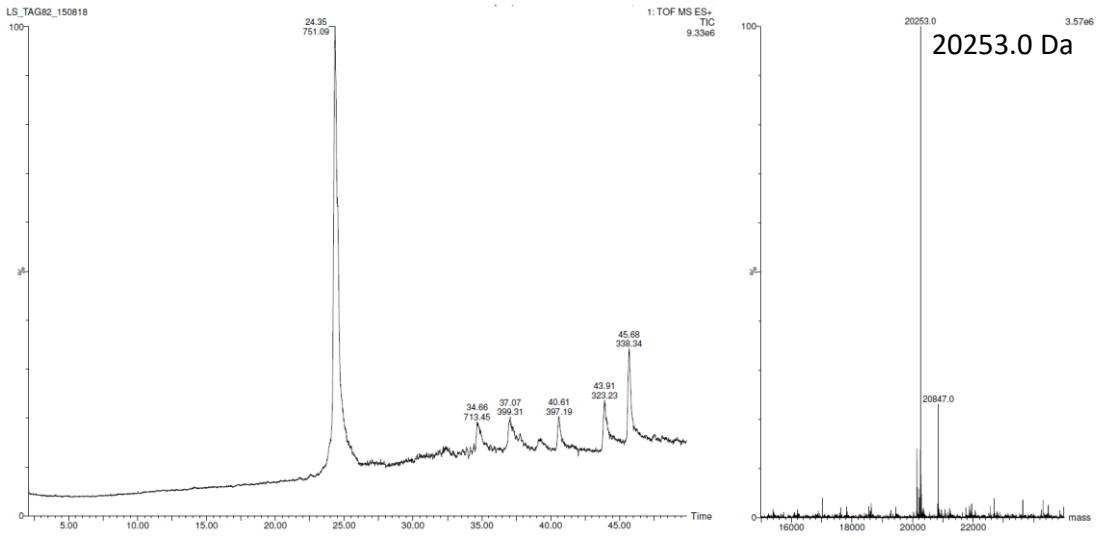


Figure 2.12. LC-MS analysis of UBE2L3*(82-1).
The LC trace shows one major peak, leading to a deconvoluted mass of 20.25 kDa, the expected mass for this species.

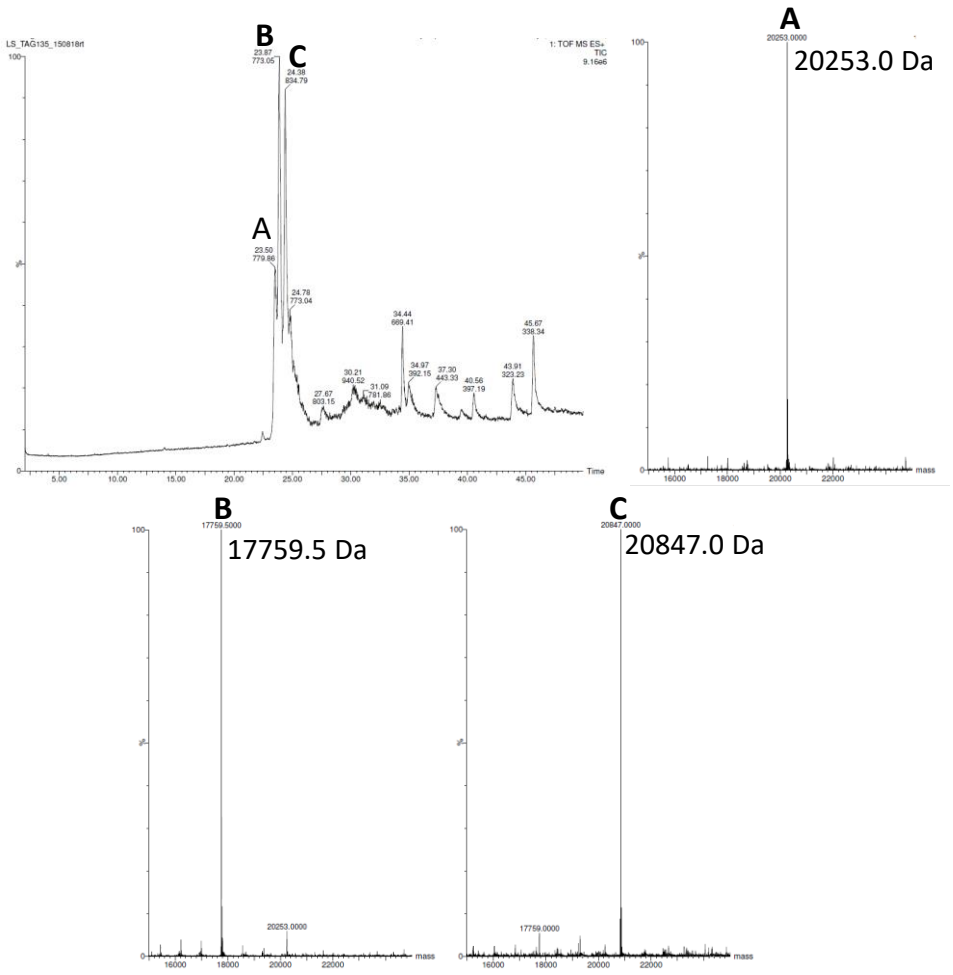


Figure 2.13. LC-MS analysis of UBE2L3*(135-1).
The LC trace shows the existence of multiple species (top left, peaks marked **A**, **B** and **C**), each leading to different masses shown in the individual deconvoluted mass spectra. **A** shows an observed mass equal to the expected mass of the protein, 20.25 kDa. **B** shows a lower mass species which corresponds to the expected mass resulting from early termination of translation instead of incorporation of **1**. Additionally, **C** shows a peak that corresponds to other further modification of the expected protein.

Biochemical activity assays for UBE2L3*

The conditions required to assess the biochemical activity of UBE2L3* have been explored during the work carried out in this thesis. This was performed in order to provide a means by which to visualise the formation of polyubiquitin chains by UBE1, UBE2L3* and HOIP in the presence of ATP. Reactions were carried out in accordance with other research in the field, and subsequently optimising reagent concentrations, reaction times and visualization on SDS-PAGE.[15,54,65,66] Polyubiquitination assay reaction conditions are detailed in ***Materials and Methods, section 6.7***. The objective of these assays, once established, was that they could be

further used to monitor reductions in UBE2L3* variant activity when exposed to the proposed inhibitor complexes **2-8**. Importantly, wild type activity and variant UBE2L3* activity must be comparable in these assays. Consumption of ubiquitin and the formation of polyubiquitin chains by the reaction components can be observed and followed by SDS-PAGE (**Figure 2.14**).

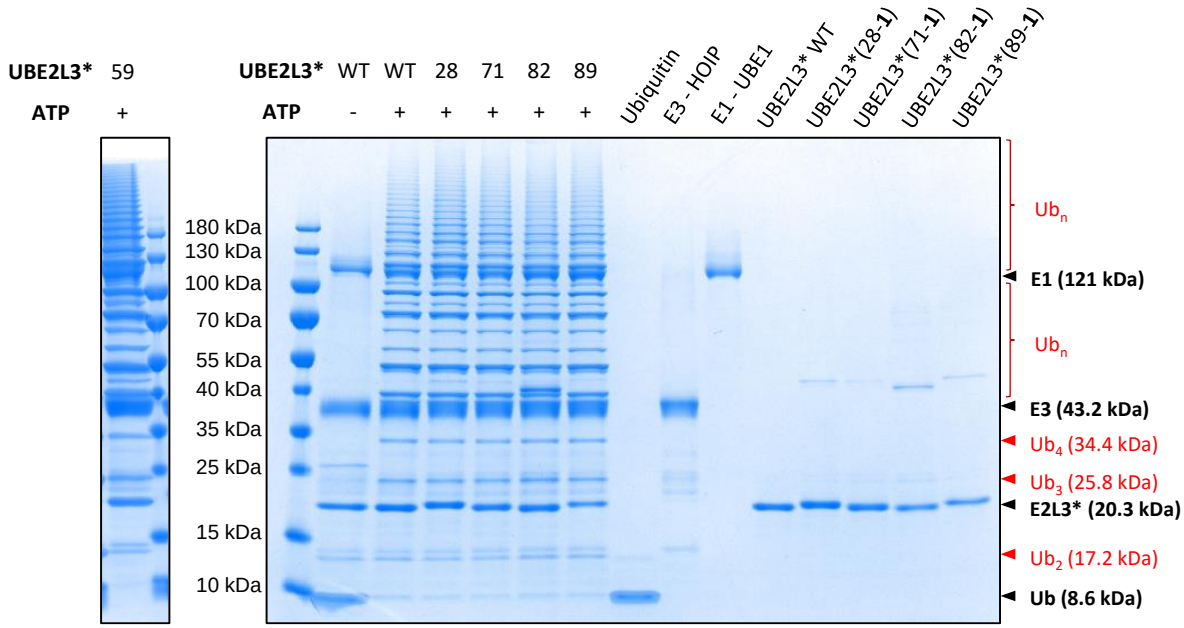


Figure 2.14. SDS-PAGE analysis of UBE2L3* wild type and variant activity.
SDS-PAGE analysis showing the formation of ubiquitin chains by UBE1, UBE2L3* wild type and variants, and HOIP, on the left half of the gel. Note that when there is no ATP added to the reaction, there is no ubiquitin consumption or formation of polyubiquitin chains. Presented on the right half of the gel are the individual protein components of the reactions.

2.3. UBE2L3* variant inhibition assays

Once polyubiquitination reaction conditions were established, purified UBE2L3* mutants containing CypK were incubated with inhibitor conjugates **2-8**, and reaction progression monitored by SDS-PAGE. Inhibitor conjugates (**Figure 2.15**) were previously synthesised within our research group by Dr. Alexander Nödling.

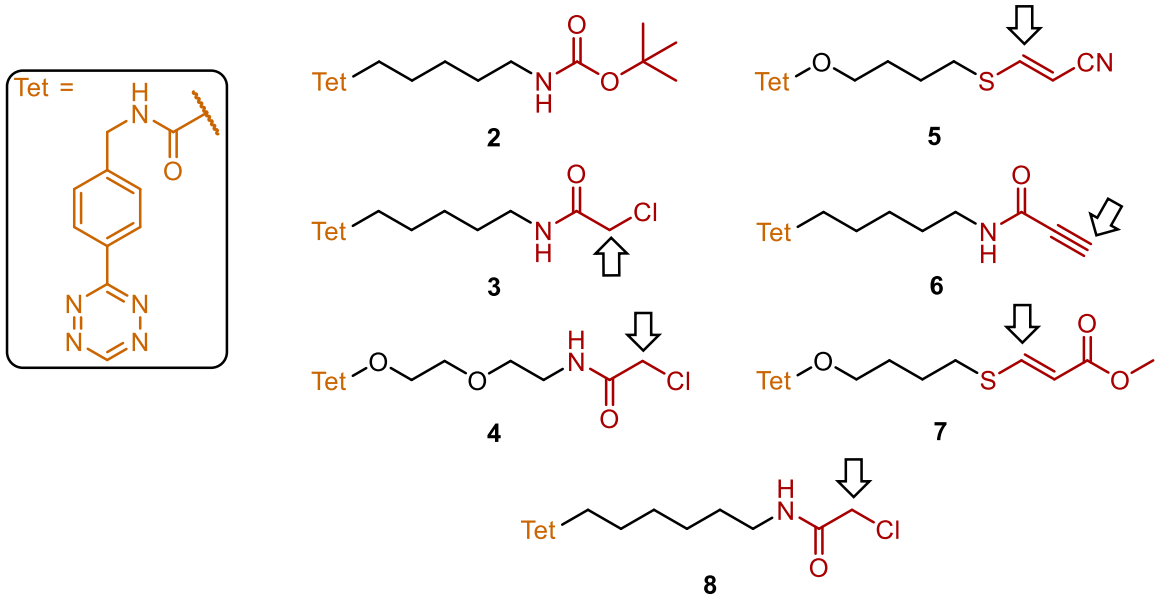


Figure 2.15. Inhibitor conjugates 2-8.
As a reminder, this figure shows all inhibitor conjugates provided by Dr Alexander Nödling have a 3,6-disubstituted-1,2,4,5-tetrazine (Tet) with a proton and phenyl substituent, designed to undergo bioorthogonal tethering reaction with cyclopropene lysine **1**.

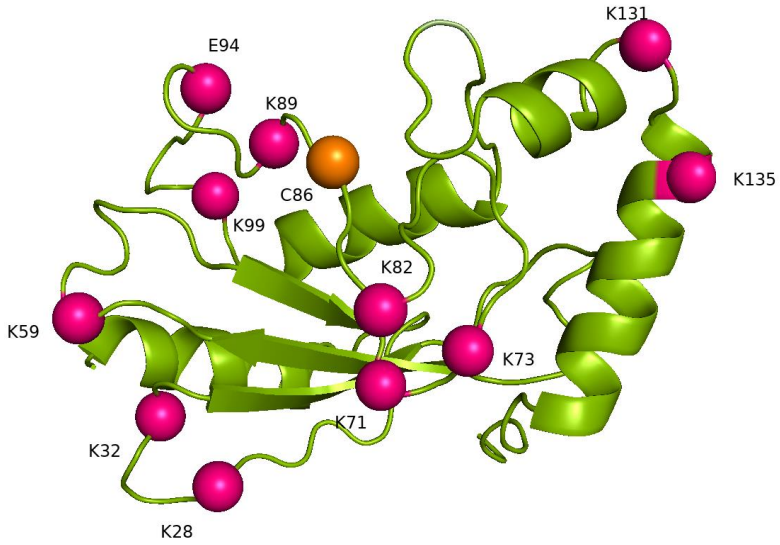


Figure 2.16. Structure of UBE2L3.
Catalytic cysteine residue (C86) is highlighted in orange. The residues that were substituted with **1** are highlighted pink. PDB: 5TTE.

In order to determine whether the inhibitor conjugates (**2**-**7**) were able to reduce the activity of UBE2L3*(XXX-**1**) variants, the aforementioned polyubiquitination reaction was used. Inhibitor conjugate **8** was a later addition and not present in the initial assays. UBE2L3*(XXX-1) mutants that had amino acids in positions 28, 32, 59, 71, 73, 82, 89, 94, 99, 131 or 135 replaced with the bioorthogonally reactive unnatural amino acid **1** (**Figure 2.16**) were screened against the inhibitors. To do this,

UBE2L3*(XXX-**1**) variants were incubated with a tenfold excess of inhibitor conjugates for 16 hours to ensure completion of the bioorthogonal tethering reaction. The inhibitor tethering reaction mixtures were then subjected to the polyubiquitination assay. Conditions for the inhibition assays are detailed in ***Materials and methods, section 6.8***. The initial screening suggested that inhibitor **4** be a potential candidate for inhibiting UBE2L3*(XXX-**1**) variants that have **1** positioned in proximity to the catalytic cysteine, whereas variants that positioned the residue further away, such as UBE2L3*(28-**1**), were unaffected (**Figure 2.17**). Importantly, none of the conjugates showed any inhibitory effect on the wild-type enzyme UBE2L3*, demonstrating the lack of tethering-independent modification of catalytic thiols on UBE1, UBE2L3* or HOIP.

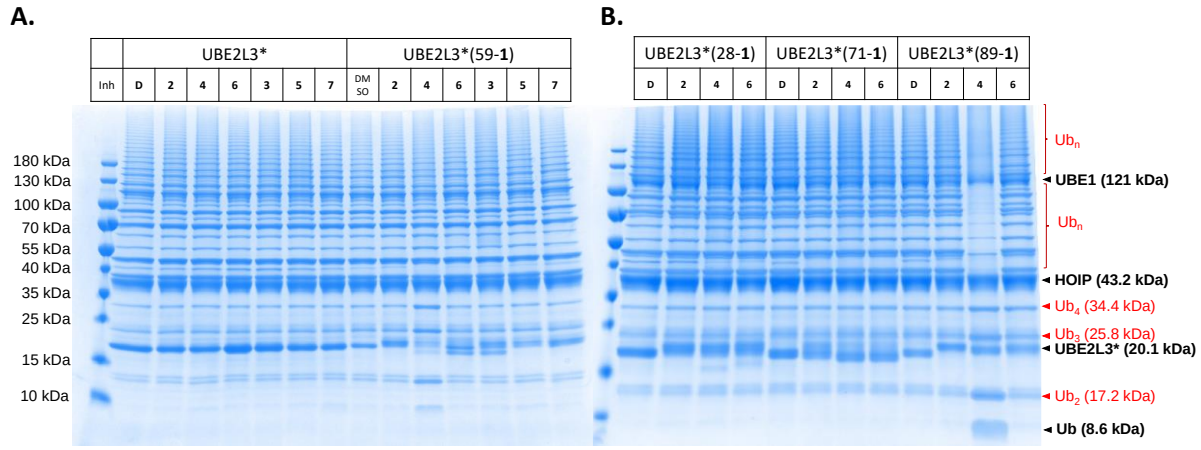


Figure 2.17. SDS-PAGE analysis of UBE2L3* inhibition assays.

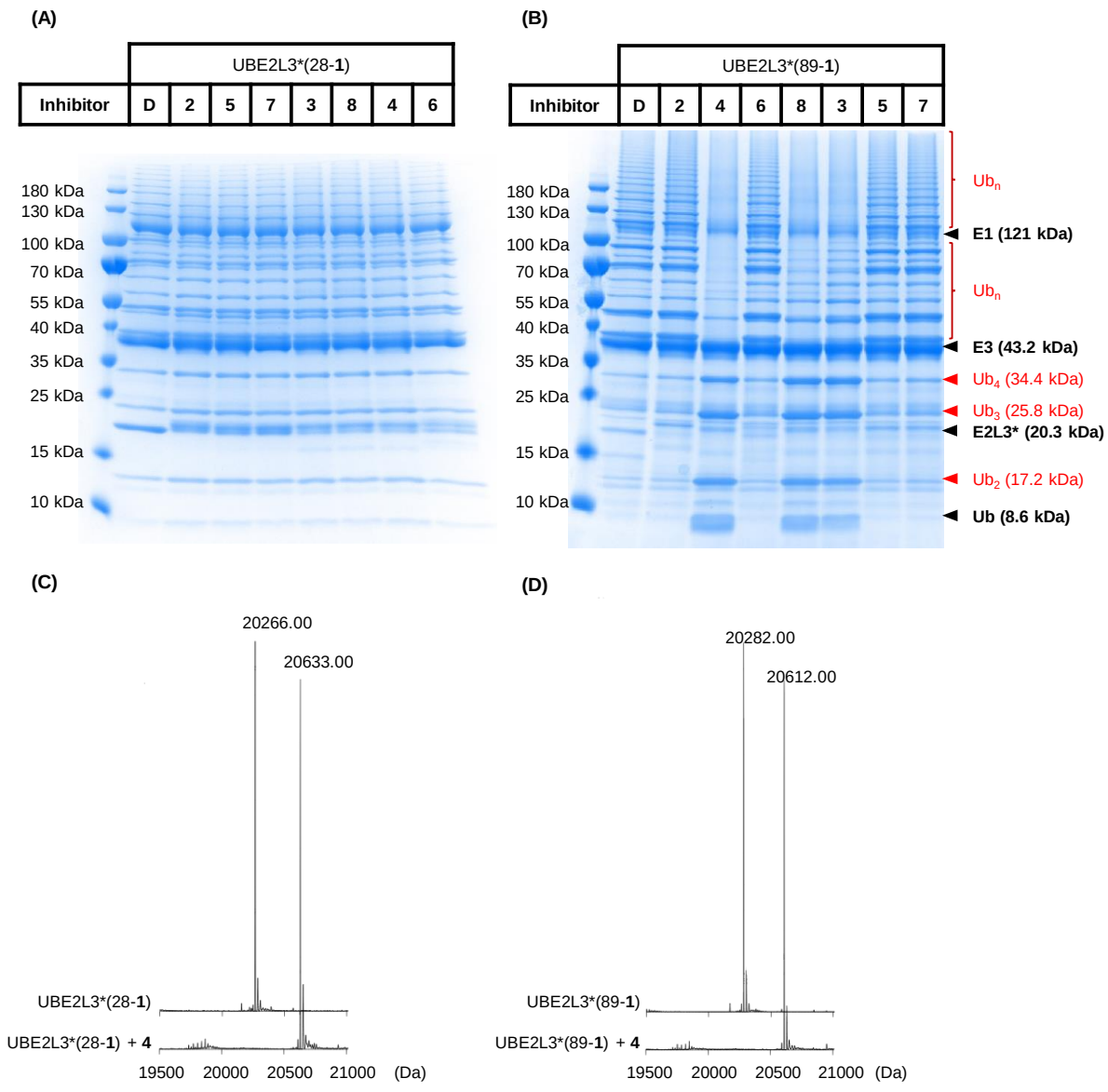
(**A**) Example result from the first set of inhibitor assays, WT UBE2L3* and UBE2L3*(XXX-**1**) variants were screened after incubation with inhibitors **2**-**7**. (**B**) Second set of inhibitor assays showing what effect inhibitor conjugates **2**, **4** and **6** have on UBE2L3* variants where amino acid positions 28, 71 and 89 have incorporated CypK. UBE2L3*(89-**1**) can be inhibited by prior incubation with conjugate **4**.

While all UBE2L3*(XXX-**1**) variants showed activity comparable to the wild-type-like UBE2L3* enzyme in the polyubiquitin assay, activity of UBE2L3*(89-**1**) variant could be consistently inhibited by pre-treatment with conjugate **4** (**Figure 2.18B**). Mass spectrometry confirmed in the case of UBE2L3*(89-**1**) the bioorthogonal tethering of the inhibitor conjugates to the UBE2L3* variant together with covalent modification of the catalytic cysteine residue (**Figure 2.18D**)**,** and similar results were observed for UBE2L3*(99-**1**). These results were then compared to a UBE2L3* variant in which the amino acid substituted by CypK (**1**) is not in close proximity to the catalytic cysteine, UBE2L3*(28-**1**), which was treated with conjugate **4** and analysed via mass

spectrometry. In this case, the biorthogonal tethering of the inhibitor was confirmed, however the loss of chloride that would occur during the covalent modification of the catalytic cysteine was not observed (**Figure 2.18A** and **C**). This is likely due to the longer distance between the catalytic cysteine residue (C89) and the tethering site where **1** substituted the amino acid in position 28.

The inhibitory effect of **4** on UBE2L3*(89-**1**) was further investigated, and the most successful inhibitors were found to be those in which the cysteine warhead consists of a chloroacetamide moiety. In the initial testing, the closely related inhibitor **3**, which has a shorter and aliphatic linker, had also shown signs of an inhibitory effect. And so, the aliphatic counterpart of **4** was synthesised, shown as inhibitor **8** and added for additional screening runs. All three inhibitors with a chloroacetamide moiety had varying degrees of an inhibitory effect when incubated with UBE2L3*(89-**1**). In comparison, inhibitor complex **4** was shown to be the more effective at reducing ubiquitin turnover.

With this information in mind, it was deemed that the inhibition of UBE2L3*****(89-**1**) by inhibitor complex **4** was a result of the intended two-step process. Firstly, the bioorthogonal tethering of the inhibitor complex **4** with unnatural amino acid **1** was successful and observed across different UBE2L3* variants. Secondly, the loss of chloride observed through mass spectrometry supported that the modification of the catalytic cysteine had occurred in a proximity-dependent manner.



(E)

MW (Da)	Before IEDDA	After IEDDA	With cysteine modification	After IEDDA and cysteine modification	Found
UBE2L3*(28-**1**) + **4**	20266.09	20632.89	20624.46	20596.44	20633.00
UBE2L3*(89-**1**) + **4**	20282.05	20648.85	20640.42	20612.40	20612.50

Figure 2.18. SDS-PAGE analysis of UBE2L3* inhibition assays and mass spectra showing CypK tethering to UBE2L3* variants.

Selective inhibition of UBE2L3*(XXX-**1**) by pre-incubation with 50 µM conjugate. DMSO (D) was used as the negative control. **(A)** Conjugates **2**-**8** did not show observable effect to the activity of UBE2L3*(28-**1**). **(B)** Activity of UBE2L3*(89-**1**) can be inhibited to varying degrees by prior incubation with conjugate **3**, **4** or **8**. **(C)** Mass spectrometry analysis showing the product of bioorthogonal reaction, an Inverse electron-demand Diels-Alder (IEDDA) between conjugate **4** and UBE2L3*(28-**1**) but no cysteine modification. **(D)** The same reaction with UBE2L3*(89-**1**) indicated bioorthogonal tethering and covalent modification of the catalytic cysteine residue. **(E)** Expected and observed molecular weight of samples under different conditions.

2.4. Re-introduction of non-catalytic cysteine residues in UBE2L3

With the initial success, it was questioned whether the presence of UBE2L3's native non-catalytic cysteine residues at positions 17 and 137 would interfere with the enzyme inhibition. Thus, the two non-catalytic cysteine residues, C17 and C137, were re-introduced to generate the true wild-type-like variant, referred to herein as UBE2L3(89-**1**). Activity of wild-type UBE2L3 is consistent with the literature[15] and was not affected by conjugate **4**, whereas ubiquitin turnover was successfully inhibited for UBE2L3(89-**1**) pre-incubated with **4** (**Figure 2.19**), indicating that the presence of non-catalytic cysteine residues does not affect the enzyme inhibition. Mass spectrometry further confirmed the bioorthogonal tethering of the conjugate **4**, and covalent modification of a cysteine residue via the substitution of the chloroacetamide's chlorine moiety (**Figure 2.20**).

Another interesting effect of re-introducing the native cysteines to the protein became apparent during the first set of assays carried out with these wild-type UBE2L3 variants. Previous polyubiquitination reactions, under the same conditions, required 2 to 3 hours of incubation to achieve complete or near-complete turnover of the ubiquitin substrate, whereas the wild-type enzymes were considerably more active, turning over the ubiquitin substrate within an hour. This made it necessary to follow the reaction at different time points during the first hour of incubation, as shown in the figure below.

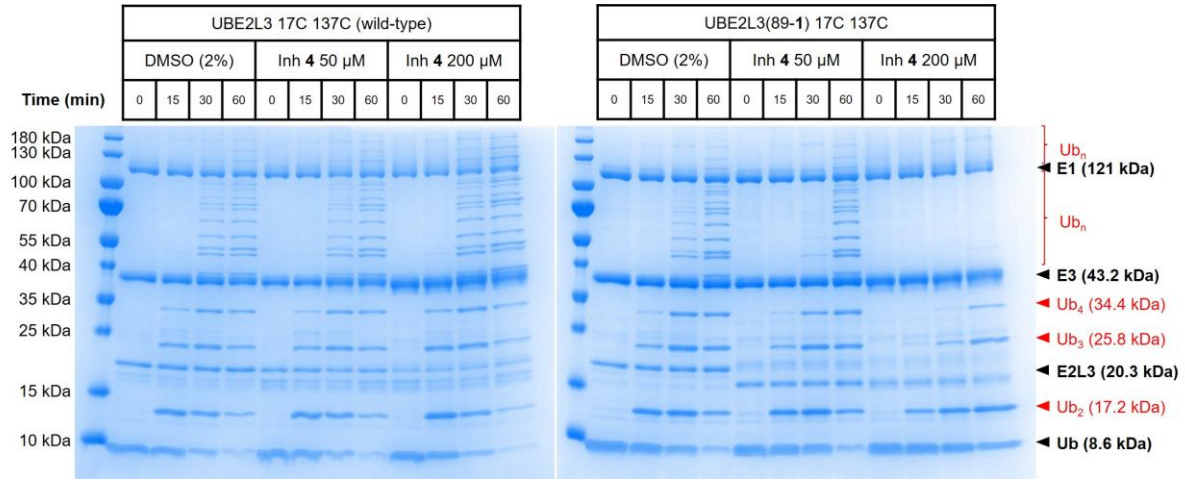


Figure 2.19. SDS-PAGE analysis of UBE2L3 inhibition assays post re-introduction of non-catalytic cysteines.
Polyubiquitination assay after pre-incubation with conjugate **4** or a DMSO control for UBE2L3 wild-type after re-introduction of native cysteines in amino acid positions 17 and 137 the same assay carried out on the UBE2L3(89-**1**) variant.

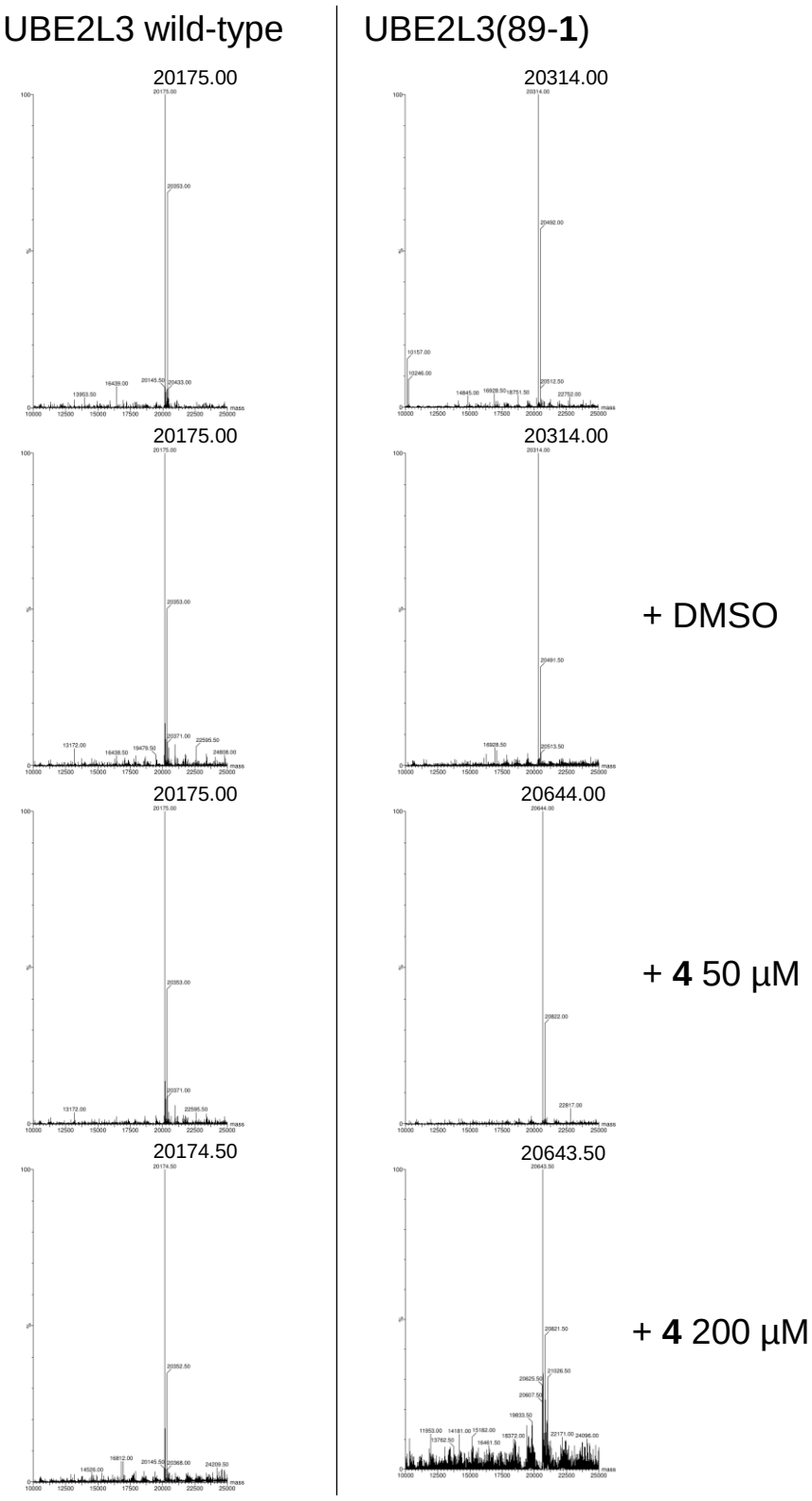


Figure 2.20. Mass Spectrometry data for UBE2L3 and UBE2L3(89-1) after re-introduction of C17 and C137.

Obtained mass for UBE2L3 is 20175 Da and consistent with expected mass. For UBE2L3(89-**1**), the expected mass of 20314 Da is obtained from both the stock and the DMSO control incubation. When incubated with inhibitor **4**, UBE2L3(89-**1**) presents a mass of 20644 Da, in line with expected mass after covalent modification of catalytic cysteine C86.

2.5. Discussion and conclusions

This chapter explored the necessary cloning, protein expression and protein purification steps required for the production of HOIP, UBE2L3 and UBE2L3's variants; all crucial components for the proposed inhibitor screening reactions.

When exploring the production of the E3 HOIP, it was found that the His tag hindered the rate at which HOIP assists in turning over ubiquitin into polyubiquitin chains. This marked an important goal to ensure cleavage of the His tag was confirmed before the use of HOIP in the polyubiquitination assays to determine UBE2L3* activity and potential inhibition of UBE2L3* variants by inhibitors **2**-**8**.

When producing UBE2L3* variants, a curiosity emerged when attempting to produce UBE2L3*(135-**1**), where premature termination of protein elongation at the TAG codon was suspected to be the majority product. The protein's molecular weight is below the expected 20.3 kDa, and closer to the 17.8 kDa mass that would be found due to early termination. LC-MS data showed the major product obtained had the same predicted molecular weight that would be expected from the truncated product. Therefore, it is possible that there is a correlation between early truncation and a closer proximity of the TAG mutation to the C-terminus of the protein. As an additional curiosity, even though the majority product is the truncated protein, polyubiquitination activity was retained (**Figure 2.21**).

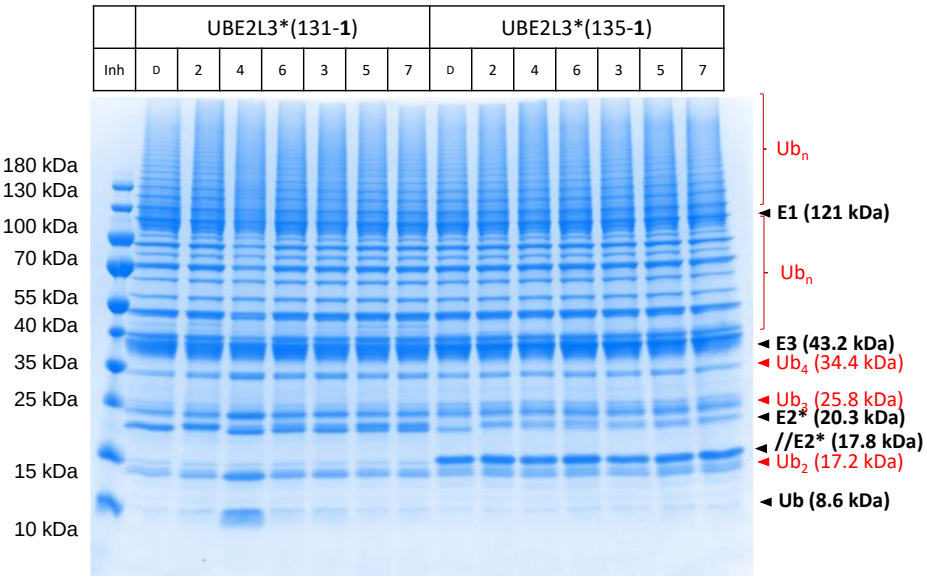


Figure 2.21. SDS-PAGE analysis of inhibitor screen panel for UBE2L3*(131-1) and UBE2L3*(135-1).
This experiment showed that even though UBE2L3*(135-**1**) was observed as a truncated protein (marked as //E2*, 17.8 kDa), it remained active.

The most successful pairing to come out of the UBE2L3* variant inhibition assays was UBE2L3*(89-**1**) when paired with inhibitor complex **4**. Repeated assays were run with this pairing and SDS-PAGE analysis showed consistent inhibition of polyubiquitin chain formation. In addition, mass spectrometry data confirms inhibitor tethering to the protein and supports the possibility of the chloroacetamide warhead in **4** reacting with the catalytic cysteine with the loss of a chloride. In general, more success was found in UBE2L3* variants that placed the unnatural amino acid **1** in close proximity to the catalytic cysteine C86. It is also important to note in the case of inhibitor **4** that when it was tethered to a UBE2L3* variant that places **1** further away from C86, no loss of chloride was observed, as was the case demonstrated with UBE2L3*(28-**1**).

Although bioorthogonal tethering followed by cysteine inhibition was observed through mass spectrometry analysis of UBE2L3*(89-**1**) treated with inhibitor complex **4**, some progression of the polyubiquitination reaction can still be observed through the presence of polyubiquitin chains. This led us to believe that there could still be trace amounts of unmodified enzyme participating in the reaction. Alternatively, it is plausible that the chloroacetyl warhead of inhibitor **4** would react with other nucleophilic amino acid residues non-specifically. This would lead to the UBE2L3* variant continuing to be able to catalyse the reaction, leading to the formation of polyubiquitin chains. This strategy would benefit from further investigation into optimisation of this process to fully inhibit the enzyme without “leaky” polyubiquitination. Nevertheless, the level of polyubiquitination was markedly less than in the control group. In addition, to demonstrate that inhibitor **4** does not react with other available cysteine residues in free solution, further research carried out in Dr. Tsai’s research group successfully demonstrated that inhibitor **4** was inert to reaction with up to a 100-fold molar excess of glutathione, confirming the proximity dependence of the inhibitor warhead.[67]

The re-introduction of the native non-catalytic cysteine residues in C17 and C137 did not appear to affect enzyme inhibition. That being the case, reaction rates for ubiquitin turnover did appear to increase based on the SDS-PAGE analysis, as shorter time-points had to be taken to observe the reaction progression. Observed time for the consumption of the ubiquitin substrate went from 2-3 hours in UBE2L3*

and variants to nearing completion at around the 1 hour mark. The concentration of inhibitor conjugate also had to be increased from 50 µM to 200 µM to observe similar inhibitory effect. This raised the question of whether the cyclopropene lysine **1** was suitable for further cell-based studies, as high concentrations could prove cytotoxic. Initially it was chosen due to its high incorporation efficiency in mammalian cells and Dr. Tsai's previous success in rapidly inhibiting cyclopropene-containing modified kinases.[31,58] The kinase inhibitors used in the aforementioned study followed a similar strategy to the inhibitor complexes used in this study, with the main difference being that here we use cysteine warheads as the inhibitor, where in Dr. Tsai's previous research they were derived from known kinase inhibitors, which became tethered to the target kinase. These kinase inhibitors were likely to have some innate affinity to the target kinase, which they posited would in turn facilitate the bioorthogonal tethering. The conjugates used in this study were not derived from known inhibitors, so it would stand to reason that the tethering kinetics would have a larger impact on the rate of the bioorthogonal reaction. Here, we observe the use of a much higher concentration of the conjugate (50-200 µM vs 1 µM in the previous work), and much longer reaction times (16 hours vs 2 hours). Future work for this strategy should involve the use of a more reactive bioorthogonal amino acid, in order for the complete tethering to be achieved with lower concentration of the inhibitor complexes and in a much shorter timeframe.

The preliminary results shown here clearly demonstrate that it is feasible to selectively inhibit enzyme activity through covalent modification of the catalytic cysteine residue. This chemogenetic approach employed genetic code expansion to incorporate a non-canonical amino acid bearing a bioorthogonal group which enabled the tethering of a proximity-dependent thiol-reactive warhead. This approach should be applicable to other enzymes within the same family, and even other cysteine-dependent enzymes such as ligases, oxidoreductases, phosphatases, and proteases.[68–70]

As a result of the findings in this chapter it was deemed crucial to assess the scope of this strategy. For example, it would be necessary to carry this over to another model E2 protein, ideally carrying over what has been learned in UBE2L3 by placing unnatural amino acid **1** at a similar proximity to the E2's catalytic cysteine, and running the inhibition assay with inhibitor **4**.

Chapter 3

EXPANDING THE SCOPE OF THE SELECTIVE INHIBITION STRATEGY: UBE2D1

3. EXPANDING THE SCOPE OF THE SELECTIVE INHIBITION STRATEGY: UBE2D1

3.1. Introduction

As mentioned in the findings of the previous chapter, the next avenue of research was to assess the scope of the inhibition strategy. For this purpose, another model E2 protein was to be selected, and inhibition of the enzyme was to be tested using our chemical genetic approach.

While researching the literature and working on UBE2L3, it was found that other E2s could catalyse the formation of polyubiquitin chains when paired with the LUBAC component HOIP. Of the human E2s found, UBE2D1, UBE2D2 and UBE2D3 were of particular interest. These E2s hold a similar structure to UBE2L3, while at the same time sharing only around a 33-39% sequence identity with it. On a cellular role level, the key difference between these E2s is that the UBE2D subgroup do not affect NF-κB activation, where UBE2L3 is known to be involved in LUBAC-mediated regulation of the NF-κB response.[9,49,71–73]

The UBE2D family of E2 enzymes are known to interact with the breast cancer type 1 susceptibility protein (BRCA1). BRCA1 is a protein involved in DNA repair pathways and acting as a tumour suppressor linked to breast and ovarian cancer, as well as associated to increased risk of certain lymphomas and leukemias.[74] In this case, although both UBE2Ds and UBE2L3 bind to BRCA1, *in vitro* activity through polyubiquitin chain formation is only observed for UBE2Ds.[75] It would seem that in a cellular environment, each of these E2s fulfil different roles and are involved in different signalling pathways, even if some *in vivo* crosstalk has been observed.

Out of the UBE2D subfamily members it was decided to use UBE2D1, which had similar polyubiquitination reaction results with UBE1 and HOIP in the literature and were documented side by side with UBE2L3. This would provide a great avenue for setting expectations for the results of using UBE2D1 in the polyubiquitination reactions established by the previous chapter. UBE2D1 is ultimately documented to be a promiscuous E2 often used in biochemical and structure studies.[56] This would in turn eliminate the need to procure genes for an alternative E3 ligase, followed by the expression and purification of the new E3 required.

The structures for UBE2L3 and UBE2D1 were aligned for consideration, noting that the corresponding catalytic cysteine in this case is in position C85 of UBE2D1's amino acid sequence (**Figure 3.1**). The comparison study was to be carried out based on the results obtained for the successful UBE2L3 variant UBE2L3(89-**1**), which was inhibited by inhibitor complex **4**. In the case of UBE2D1, this residue to be substituted was its isoleucine at the homologous position 88 (I88). The amber codon UAG would take its place for the incorporation of the same unnatural amino acid, cyclopropene lysine (**1**).

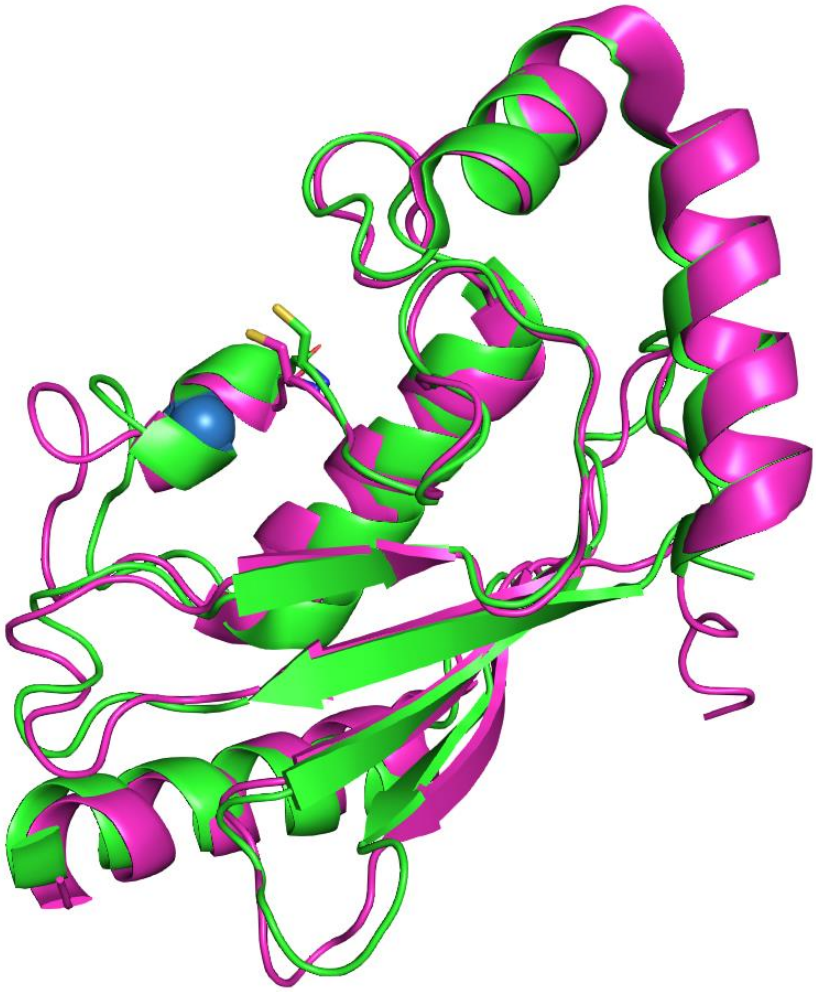


Figure 3.1. Crystal structure of UBE2L3 and UBE2D1.
Structural alignment of UBE2L3 (pink, PDB: 4Q5E) and UBE2D1 (green, PDB: 6D4P). Catalytic cysteines can also be seen around the centre left of the protein, represented as liquorice sticks and coloured yellow. Blue spheres represent the amino acid locations to be substituted in both enzymes, I88 for UBE2D1 and V89 for UBE2L3.

3.2. Cloning

The gene encoding for UBE2D1 with an N-terminal hexa-histidine tag was purchased as a gene fragment from *ThermoFisher GeneArt Strings* (sequence in **Table 6.2**, ***Materials and Methods***), after using their codon optimisation tool in order to optimise the gene for *E. coli* expression. This fragment was cloned into the *XbaI* and *NotI* sites of a pET15b vector, using the vector backbone from **pET UBE2L3***, to generate **pET UBE2D1**.

In order to generate a TAG mutant of UBE2D1 that was homologous to UBE2L3(89-**1**), SDM of **pET UBE2D1** was carried out in order to generate **pET UBE2D1(88TAG)**.

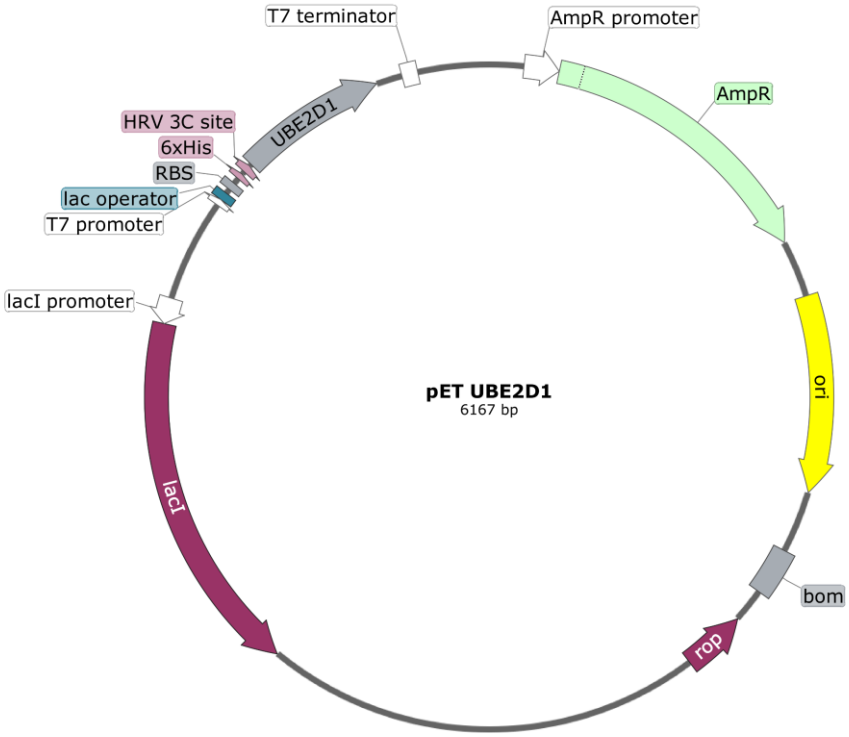


Figure 3.2. Plasmid map for pET UBE2D1.

3.3. Protein expression and purification

Expression of UBE2D1 through **pET UBE2D1** was carried out in BL21 (DE3) cells. Protein expression and purification protocols are detailed in ***Materials and Methods, section 6.6***. Cell pellets were lysed and purified via Ni-NTA affinity chromatography (**Figure 3.3**). Purified UBE2D1 was then confirmed via mass spectrometry (**Figure 3.4**).

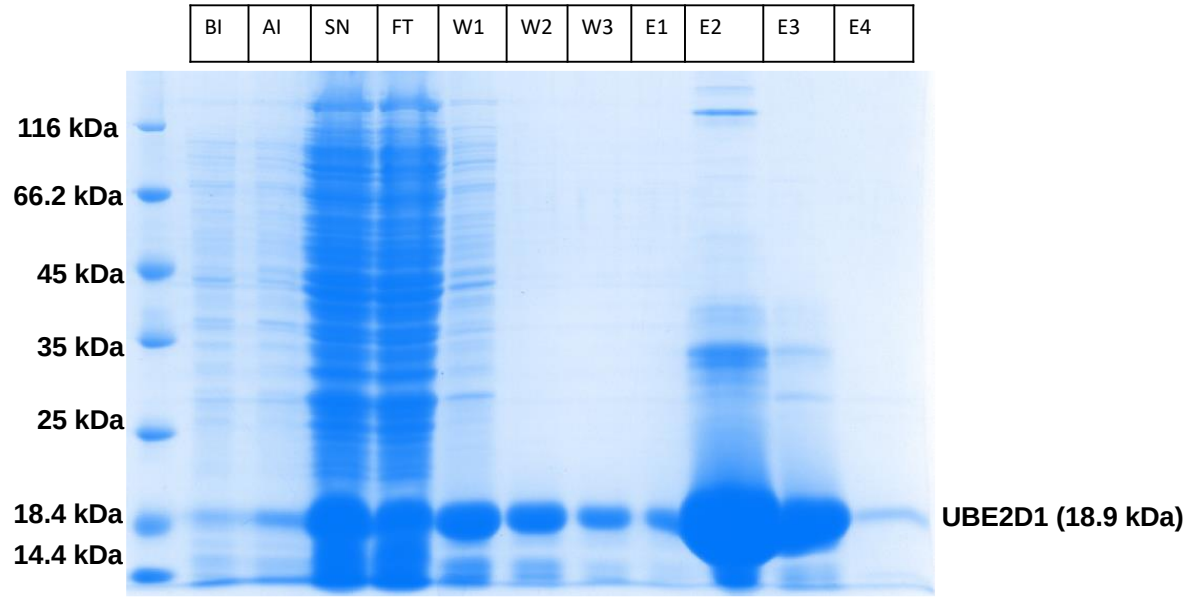


Figure 3.3. SDS-PAGE analysis of UBE2D1 expression and purification.
BI and AI denote before and after induction with IPTG, respectively, showing a faint appearance of UBE2D1 overexpression after inducing expression with IPTG. After the cells were lysed and centrifuged, the supernatant (SN) is then allowed to pass through Ni-NTA resin, collecting the flow-through (FT) and then washing with low imidazole (25 mM) concentration buffer (W1-3), followed by elution of the His-tagged UBE2D1 in high imidazole (300 mM) buffer (E1-4).

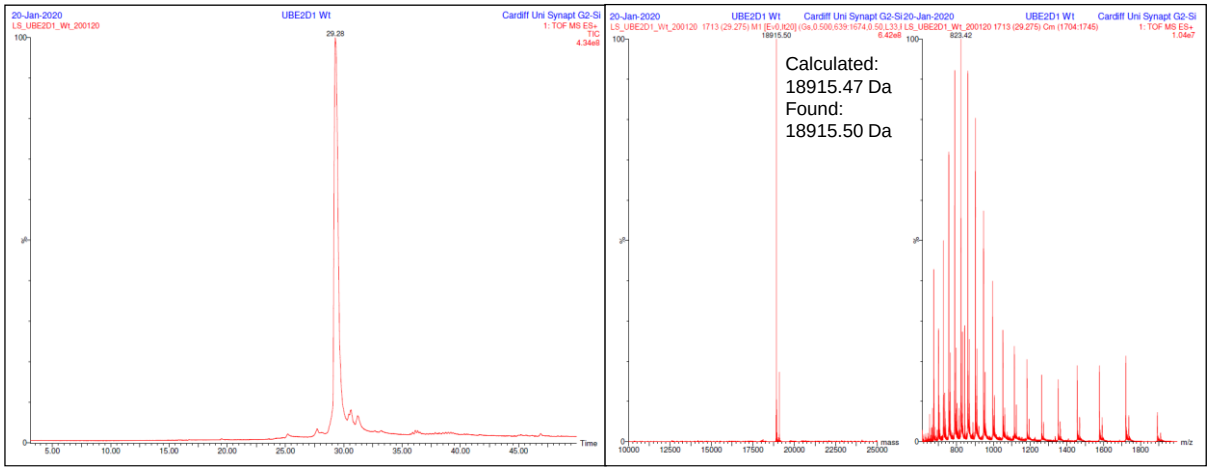


Figure 3.4. LC-MS analysis of UBE2D1.
Purified UBE2D1 was submitted to LC-MS analysis, where the expected mass of 18915.5 Da was found.

Expression and purification of UBE2D1(88-**1**) through **pET UBE2D1(88TAG)** was carried out in BL21 (DE3) cells. Addition of CypK (**1**, 1 mM) was performed prior to induction of the cultures with IPTG. Cultures were incubated at 37 °C for 4 hours before harvesting the cells. Cell pellets were lysed, purified via Ni-NTA affinity chromatography and fractions were analysed via SDS-PAGE (**Figure 3.5**). In this case expression levels and protein yield were noticeably lower, although ultimately

enough to isolate and carry over into polyubiquitination assays. Purified UBE2D1(88-**1**) was further confirmed via mass spectrometry (**Figure 3.6**) and was deemed suitable for use in polyubiquitination reactions.

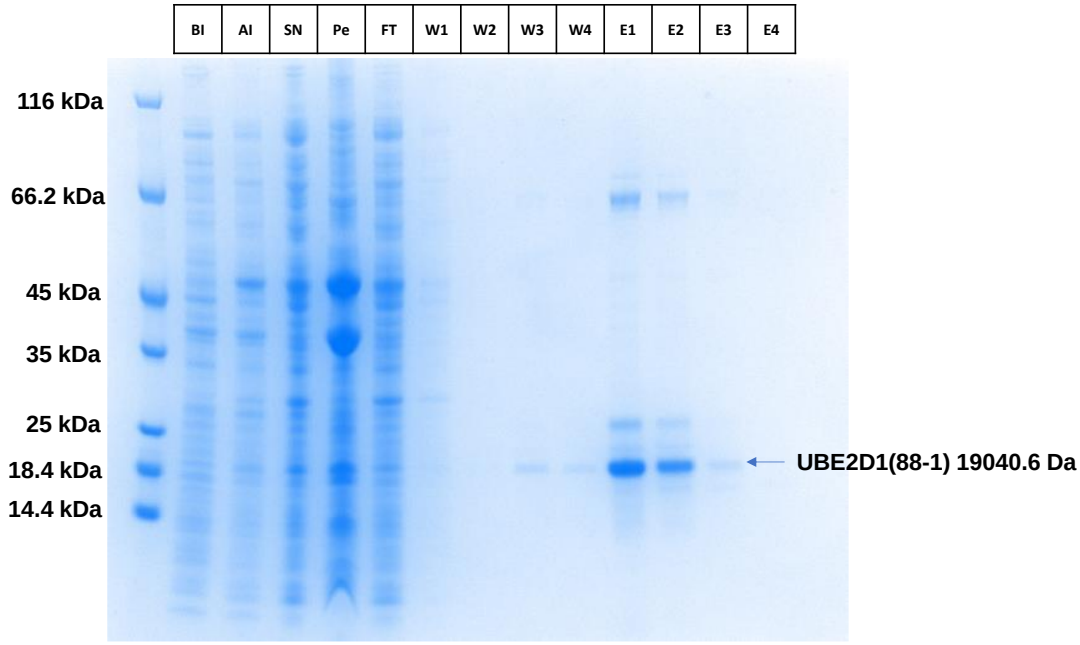


Figure 3.5. SDS-PAGE analysis of UBE2D1(88-1) expression and purification.
BI and AI denote before and after induction with IPTG, respectively, showing a faint appearance of UBE2D1 overexpression after inducing expression with IPTG. After the cells were lysed and centrifuged, the supernatant (SN) is then allowed to pass through Ni-NTA resin, collecting the flow-through (FT) and then washing with increasing concentrations of imidazole buffer (W1-4), followed by elution of the His-tagged UBE2D1(88-**1**) in high imidazole (300 mM) buffer (E1-3).

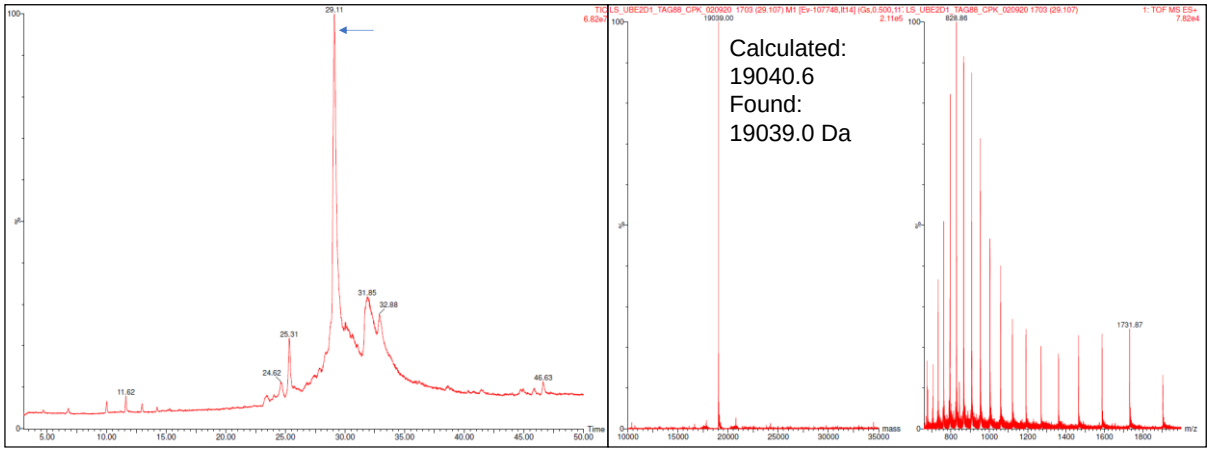


Figure 3.6. LC-MS analysis of UBE2D1(88-1).
Purified UBE2D1(88-**1**) was submitted to LC-MS analysis, where the expected mass was calculated as 19040.6 Da and 19039.0 Da was found.

3.4. Biochemical activity assays and inhibition assays for UBE2D1

The polyubiquitination assay reaction conditions used in this segment were the same as those used in the previous chapter for UBE2L3. The activity of UBE2D1 and its variant UBE2D1(88-**1**) were analysed via SDS-PAGE (**Figure 3.7**). Notably, activity of the UBE2D1(88-**1**) variant seemed more sluggish at the 1 hour mark than its wild-type counterpart. This could be due, in part, to the insertion of **1** into the amino acid sequence. Another important factor to be considered is the purity of the UBE2D1(88-**1**) stock, where background contaminants may have implications on the measured protein concentration of the stock being artificially shifted to higher values. All things considered, the reaction ran to completion at the 2 hour mark and was moved forward for use in the inhibition assay.

In said inhibition assay, both UBE2D1 and UBE2D1(88-**1**) were subjected to pre-incubation with either DMSO as a negative control or inhibitor complex **4**, as completed previously for UBE2L3 variants. The results (**Figure 3.8**) still showed a notable increase in the time it took UBE2D1(88-**1**) to turn over the ubiquitin substrate, 2 hours compared to UBE2D1's 1 hour completion time. That being said, inhibitor **4** had no effect on the wild type UBE2D1, whereas it could effectively inhibit UBE2D1(88-1).

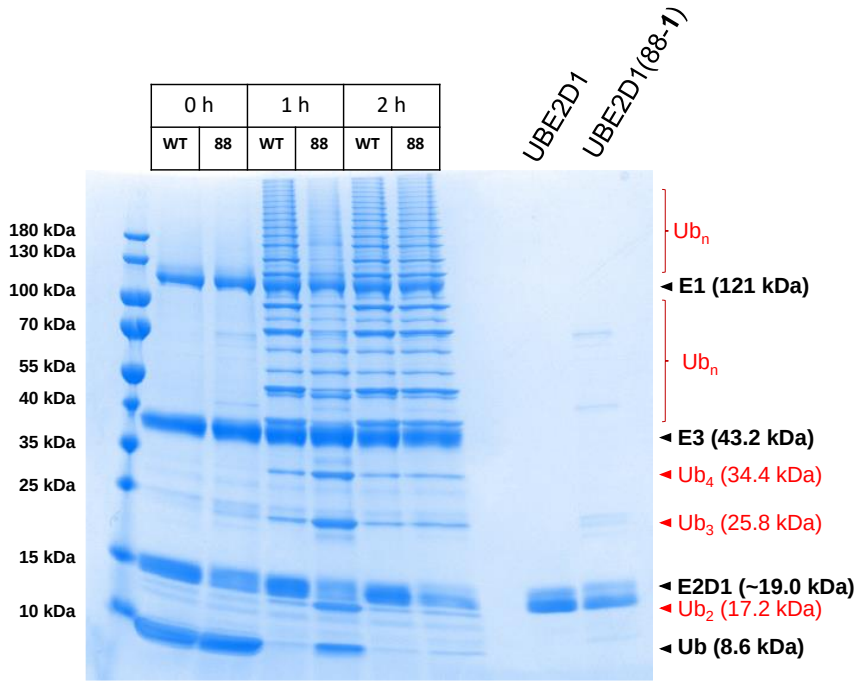


Figure 3.7. SDS-PAGE analysis of UBE2D1 wild type (WT) and variant UBE2D1(88-1) activity.

SDS-PAGE analysis showing the formation of ubiquitin chains by UBE1, UBE2D1 (marked as WT) or UBE2D1(88-**1**) (marked as 88), and HOIP in the presence of ATP over the course of 2 hours. Presented on the rightmost two lanes of the gel are UBE2D1 and UBE2D1(88-**1**) for reference.

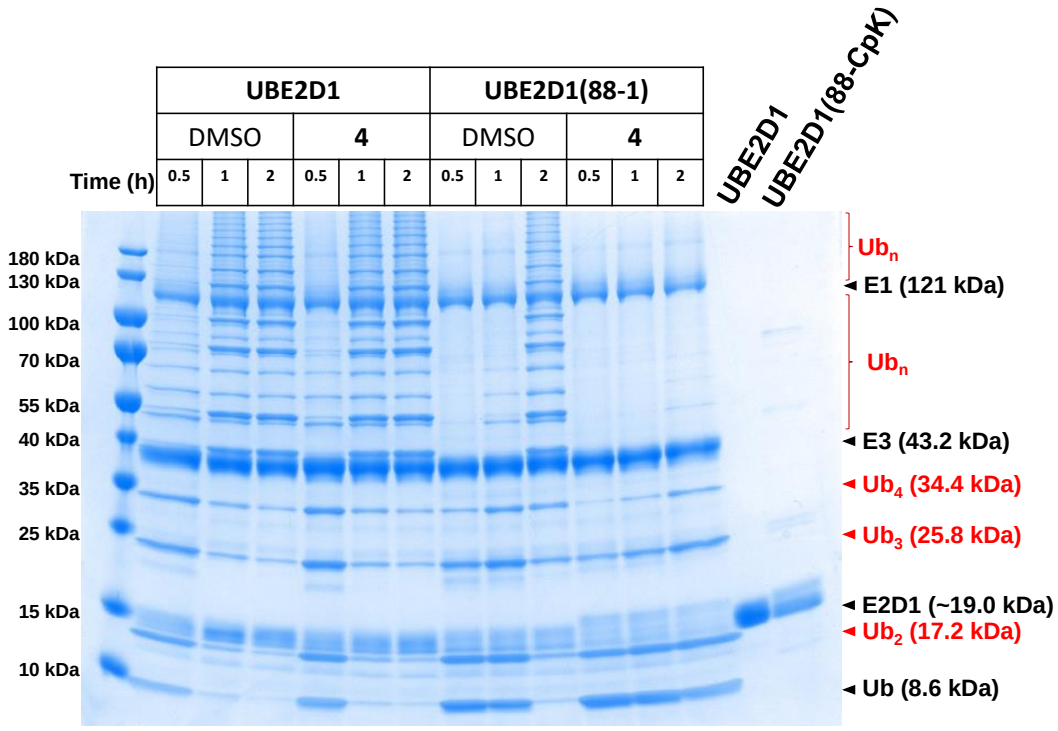


Figure 3.8. SDS-PAGE analysis of UBE2D1 inhibition assays.

Inhibition assay using DMSO as a negative control. Unaltered UBE2D1 is not hindered by the presence of inhibitor conjugate **1**. Activity of variant UBE2D1(88-**1**) is reduced by incubation with **1**.

3.5. Conclusions

In this section the scope and transferability of the selective inhibition of cysteine-dependent enzymes via bioorthogonal tethering was put to the test. Taking successes and lessons learnt from the application of this strategy in UBE2L3, another candidate among the E2 enzyme family was selected. UBE2D1 was the enzyme of choice due to the ease of method transfer regarding the polyubiquitination reactions, as it readily forms linear polyubiquitin chains with UBE1 and HOIP. This cut the time that would have been invested in procuring genes for another E3 ligase and streamlining its expression and purification systems.

Only two E2 variants were required, the wild type UBE2D1 and the variant UBE2D1(88-**1**), the latter being homologous to UBE2L3(89-**1**), which was successfully inhibited in the previous chapter with inhibitor complex **4**.

It was found that while UBE2D1 wild type activity was not affected by pre-incubation with **4**, UBE2D1(88-**1**) activity could be effectively inhibited by it. This demonstrated the transferability of this strategy to other E2 family enzymes. Additionally, these results continue to reinforce the potential of using this strategy for the inhibition of other cysteine-dependent enzymes.

In future work, the expression and purification of UBE2D1(88-**1**) could be further streamlined for higher yields and purity, as its “sluggish” activity compared to wild-type UBE2D1 could be due to impurities in the mixture leading to an erroneous calculation of stock concentration or they could be interfering with reaction progression.

Additionally, during this thesis it was attempted to generate mass spectrometry data after incubation of UBE2D1 and its variant with **4**, similarly to results shown in the previous chapter. The data generated lead to inconclusive results, most likely due to the quality of the protein stock, as the UBE2D1 variant was produced in low yield and likely contained a higher amount of contaminating proteins.

Lastly, the focus of the upcoming section was decided in parallel while working on UBE2L3 and UBE2D1. It was deemed further necessary to investigate inhibition of E2 variants in living mammalian cells by using UBE2L3(89-**1**) as a model together with inhibitor **4**.

Chapter 4

INHIBITION OF UBE2L3 IN LIVING MAMMALIAN CELLS

4. INHIBITION OF UBE2L3 IN LIVING MAMMALIAN CELLS

4.1. Introduction

The end goal for the selective inhibition of cysteine-based enzymes proposed in this work is for it to be transferred to usage in mammalian cell culture systems. This strategy has potential to provide a toolbox for decrypting the cellular functions and disease relevance of human enzymes that rely on catalytically active cysteine residues. The ability to selectively and rapidly inhibit enzyme activity through the proposed combination of genetic and chemical means is attractive, especially in families of enzymes that lack small molecule inhibitors, such as the case for E2 ubiquitin-conjugating enzymes.

The previous sections of this work have described successful modification of E2 enzymes UBE2L3 and UBE2D1 via incorporation of the non-canonical amino acid cyclopropene lysine, **1**. This modification rendered the E2 variants susceptible to inhibition by **4**, which contained a cysteine warhead capable of forming a covalent bond with catalytically active thiols in these enzymes.

Researching ways to measure the activity of UBE2L3 in human cells, the work of Dr Myles Lewis' lab on deciphering the role of UBE2L3 in signal transduction came to the forefront.[76] They demonstrated the involvement of UBE2L3 in the NF-κB signalling pathway. The NF-κB signalling pathway is involved in a variety of cellular processes, including inflammation, immune responses, and cell survival. Dysregulation of this pathway has been linked to various diseases, including cancer, chronic inflammation, autoimmune disorders, and viral infections. In cancer, aberrant NF-κB signalling can contribute to tumour development, progression, and resistance to therapy. Inflammatory diseases such as rheumatoid arthritis and inflammatory bowel disease have also been linked to dysregulated NF-κB activation. Additionally, viral infections such as HIV, hepatitis B and C, and influenza have been shown to manipulate the NF-κB pathway to evade host immune responses.[77–80]

In the context of the work carried out specifically in E2 enzymes, knockdown of UBE2L3 decreases NF-κB activation, whereas overexpression of UBE2L3 leads to upregulation in NF-κB activity. In fact, Dr Lewis established the use of a NF-κB luciferase assay to quantify the activity of UBE2L3 in cells. The cellular assays

developed by the above-described work represented an ideal choice for us to establish selective inhibition of UBE2L3 in cells.

The HEK293-based cell line used for this study is designed to express a luciferase reporter gene under the control of NF-κB response elements (REs). These REs are DNA sequences that are recognized and bound by NF-κB transcription factors, leading to the activation of the reporter gene and production of luciferase.[81] The luciferase gene is placed downstream of NF-κB response elements in the promoter region, so when NF-κB is activated and binds to these response elements, the luciferase gene is expressed and can be measured using a luminometer. The luciferase activity can be measured and used as an indicator of NF-κB pathway activation. HEK293 is a human embryonic kidney cell line commonly used in research. This cell line is relatively easy to culture, has a high transfection efficiency, and are often used as a model system for studying protein-protein interactions and signal transduction pathways.[82]

A visit to Dr. Lewis' facilities was arranged, in Queen Mary University of London's (QMUL) centre for Experimental Medicine and Rheumatology. This presented an excellent opportunity for gaining hands-on experience with human cell culture, while exchanging our work on genetic code expansion and the novel chemical biology technique used throughout this work. During this visit there was a chance to be trained in three key experimental techniques.

First, cell culture techniques including transient transfection of HEK293 cell lines starting with plasmids containing our genes of interest for this study, UBE2L3 and its variants. Second, cell viability assays in the presence of inhibitor complex 4 were performed as a measure of the toxicity of the inhibitors. Third, UBE2L3 inhibition assays were trialled during the visit, as well as functional assays in order to determine the capability of the system.

The cell line used in Dr Lewis' lab was an NF-κB-RE-*luc2P* HEK293 reporter cell line, which enables the measurement of UBE2L3 activity through the nuclear factor kappa B (NF-κB) response signalling pathway. The luciferase gene (*luc2P*) in this cell line acts as a reporter for NF-κB production in the cell, as luciferase gets produced at the same time. Addition of the luciferase substrate, luciferin, generates a

measurable signal in the form of light, which can be captured via an appropriate instrument such as a luminometer in the case of this study.[82]

Upon returning to the laboratories at Cardiff University, additional experimentation was conducted using the NF-κB-RE-*luc2P* HEK293 reporter cell line. The focus of this work was to optimize the plasmid constructs to enhance the production of UBE2L3 *in vivo*. The goal was to increase the yield of UBE2L3 by fine-tuning the construct design. Additionally, the NF-κB activity assay was established to measure the activity of UBE2L3 in relation to the NF-κB signalling pathway.

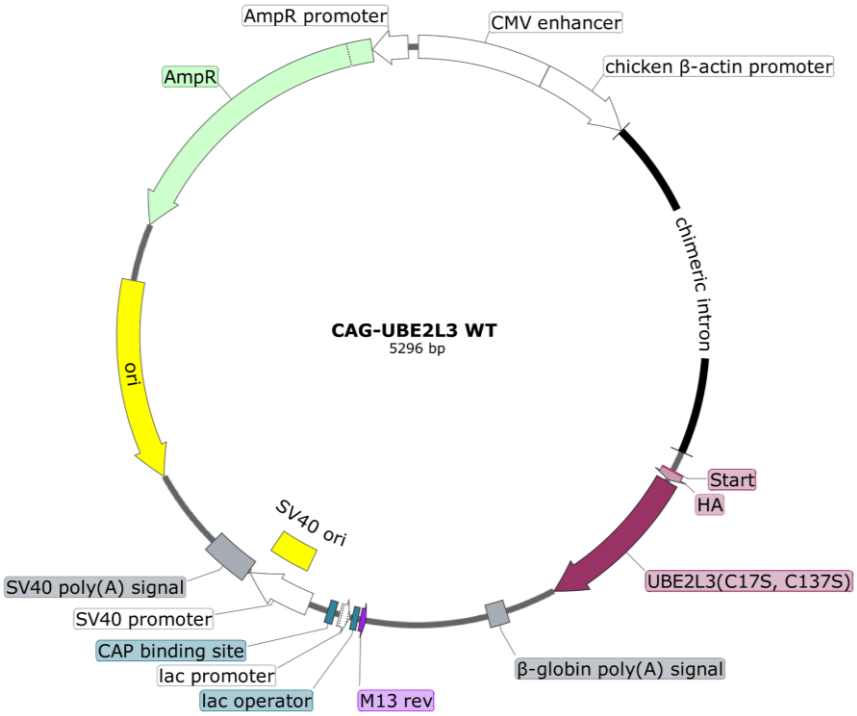
Regrettably, the outcomes of this particular phase of the research were unable to demonstrate the inhibition of UBE2L3 variants in the mammalian cell line as initially anticipated. Nonetheless, representative results obtained during the setup process are presented below, highlighting potential avenues for future investigation. These findings underscore the importance of ongoing research in this area and provide valuable insights for further exploration.

4.2. Cloning

Mammalian expression plasmid constructs for UBE2L3* and variants UBE2L3*(28-**1**), UBE2L3*(89-**1**) and UBE2L3*(99-**1**) were cloned via polymerase chain reaction (PCR) amplification of the genes in their bacterial expression plasmid counterparts, followed by Gibson assembly into a modified piggyBac vector containing the tools necessary for unnatural amino acid incorporation, the aminoacyl-tRNA synthetase and tRNA pair, FLAG-MmPylRS under an EF-1a promoter and 4 copies of $PylT_{CUA(U25C)}$ each under U6 promoters, generating **Plasmid EF1-PylRS CAG-UBE2L3 WT** and variants. Cloning for all mammalian expression plasmids results in UBE2L3* being HA-tagged in order to allow for western blot analysis of expression. To generate plasmids that do not contain MmPylRS and $PylT_{CUA(U25C)}$, the above plasmids were cloned into the *EcoRI* and Eco81I sites of a pCX vector, resulting in **Plasmid CAG-UBE2L3 WT** and variants.

After encountering challenges in visualising the HA-tagged UBE2L3* through western blot analysis, further insights into the mammalian expression of these E2 constructs were sought. This involved the removal of the UBE2L3 Met1 residue via SDM to generate **Plasmid EF1-PylRS CAG-UBE2L3 WT (-Met1)** and its variants,

which retained MmPylRS and $PylT_{CUA(U25C)}$. Similarly, **Plasmids CAG-UBE2L3 WT (-Met1)** and its variants, were generated without the inclusion of the synthetase and tRNA machinery. The objective of these modifications was to investigate whether read-through occurred at the ATG site of the HA tag during translation, resulting in the omission of the HA tag and the initiation of protein translation at UBE2L3*'s ATG site. By eliminating the Met1 residue, insights into the expression dynamics of the UBE2L3* constructs were pursued.



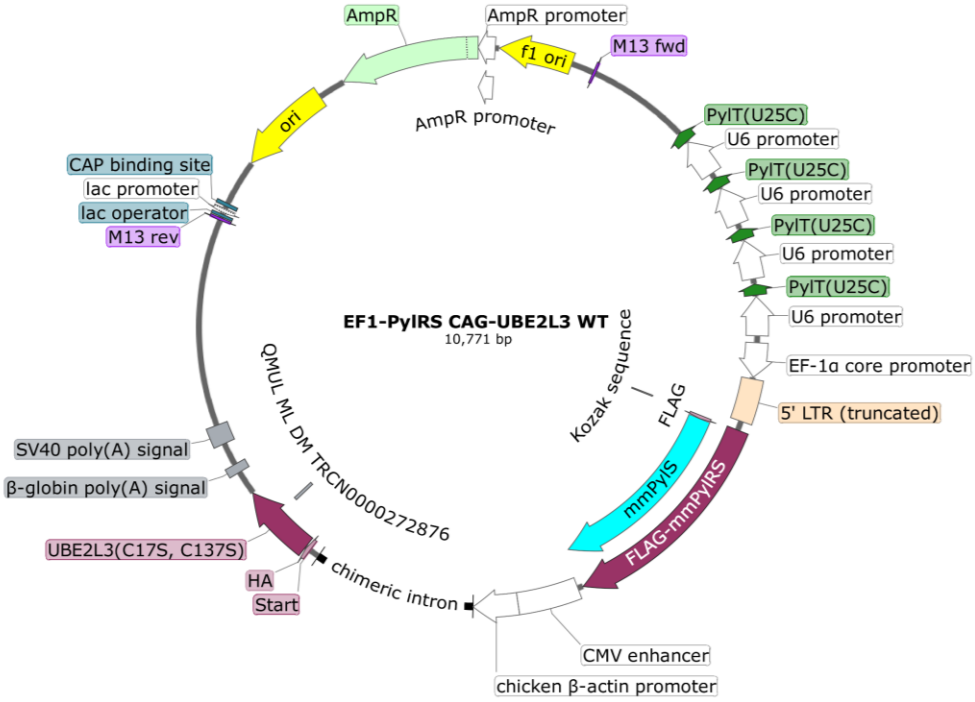


Figure 4.1. Plasmid maps for CAG-UBE2L3 WT and EF1-PylRS CAG-UBE2L3 WT. Notably, the latter highlights the inclusion of FLAG-MmPylRS and four copies of PylTCUA(U25C) as the aminoacyl-tRNA synthetase/tRNA pair for unnatural amino acid incorporation.

4.3. Preliminary work with Dr. Myles Lewis in QMUL

For this subsection of the work carried out in QMUL, unless otherwise stated, the plasmid constructs had been transported from our lab at Cardiff University in order to carry out experiments in Dr. Myles Lewis' laboratories.

The first set of experiments were designed to confirm overexpression levels of the UBE2L3 constructs. To achieve this, the UBE2L3 containing plasmids mentioned below in Table 4.1 were transfected in the NF-κB-RE-*luc2P* HEK293 reporter cell line. An empty vector (EV) was used as a control for baseline levels of the proteins studied, while having the cells undergo the transfection procedure all the same. Transfection was carried out for each plasmid construct. Plasmids that did not contain the synthetase and tRNA required for amber codon (TAG) suppression were co-transfected with plasmid EF1-Flag-MmPylS-IRES-neoR, 4xPylT, abbreviated in this section as SP9. The incorporation of the synthetase/tRNA machinery is critical to allow expression of UBE2L3 variants. Cyclopropene lysine (CypK) **1** was added to cell media and controls with no addition of **1** were also prepared. After transfection, cells were left to grow before harvesting after 48 hours. A BCA (bicinchoninic) protein

assay was used to normalise total protein content of each cell lysate harvested before proceeding to western blot. Western blot analysis was performed using anti-UBE2L3 and anti-HA, as well as anti-actin or anti-GAPDH as a housekeeping gene.

Table 4.1. Plasmids used during the preliminary cell culture work described in this secion

#	Plasmid Name	Notes
LS3	EF1-PylRS CAG-UBE2L3* (C17S C137S)	All LS plasmids contain C17S C137S mutations
LS4	EF1-PylRS CAG-UBE2L3 TAG28	Contains UBE2L3 mutant, MmPylRS and 4x $MmtRNA_{CUA}^{Pyl}(\mathrm{U25C})$. (Synthetase and tRNA pairs required for TAG suppression)
LS5	EF1-PylRS CAG-UBE2L3 TAG89	Contains UBE2L3 mutant, MmPylRS and 4x $MmtRNA_{CUA}^{Pyl}(\mathrm{U25C})$. (Synthetase and tRNA pairs required for TAG suppression)
LS6	EF1-PylRS CAG-UBE2L3 TAG99	Contains UBE2L3 mutant, MmPylRS and 4x $MmtRNA_{CUA}^{Pyl}(\mathrm{U25C})$. (Synthetase and tRNA pairs required for TAG suppression)
LS23	CAG-UBE2L3* WT	All LS plasmids contain C17S C137S mutations
LS24	CAG-UBE2L3 TAG28	Require transfection together with SP9
LS25	CAG-UBE2L3 TAG89	Require transfection together with SP9
LS26	CAG-UBE2L3 TAG99	Require transfection together with SP9
SP9	EF1-Flag-MmPylS-IRES-neoR, 4xPylT	MmPylRS and 4x $MmtRNA_{CUA}^{Pyl}(\mathrm{U25C})$. (Synthetase and tRNA pairs required for TAG suppression)
	HOIL-1L	Plasmid kindly provided from Dr. Lewis stocks
	HOIP	Plasmid kindly provided from Dr. Lewis stocks
	UBE2L3 WT (DM)	Plasmid kindly provided from Dr. Lewis stocks. True WT, *Homo Sapiens.*

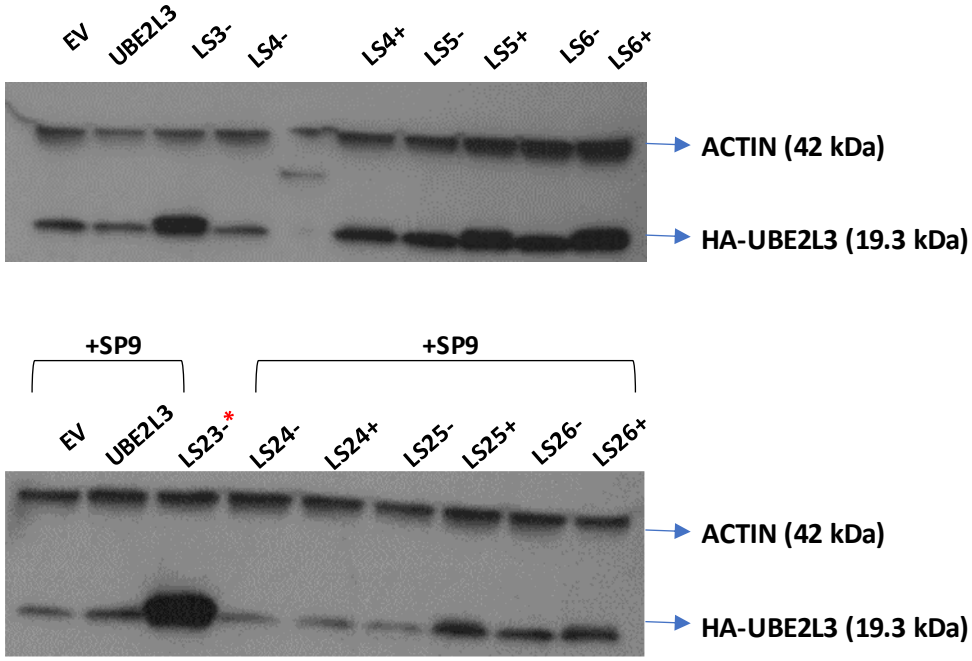


Figure 4.2. Western blot analysis of UBE2L3 expression in NF-κB-RE-luc2P HEK293 reporter cell line. Antibodies used are anti-UBE2L3 and anti-actin (housekeeper gene). Plasmid marked as LS23, which corresponds to CAG-UBE2L3 WT, is not co-transfected with the tRNA/synthetase machinery required for amber codon suppression and incorporation of **1**.

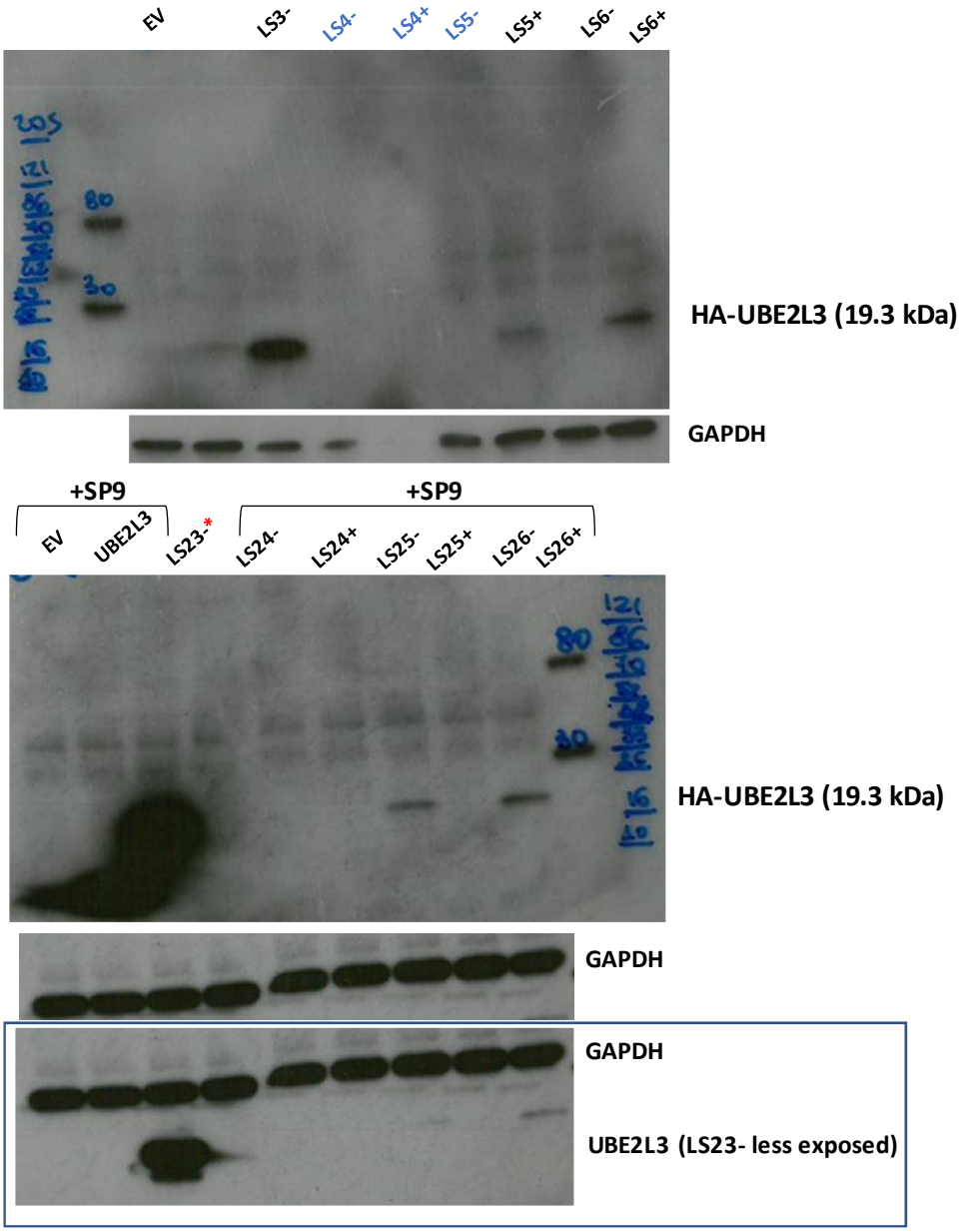


Figure 4.3. Western blot analysis of UBE2L3 expression in NF-κB-RE-luc2P HEK293 reporter cell line. Antibodies used are anti-HA and anti-GAPDH (housekeeper gene). In the top gel, housekeeping gene GAPDH indicates an uneven total protein load, making data for LS4 plasmid unfortunately unreliable in this instance. Marked at the bottom within the blue-lined box is a wider view of the GAPDH blot, showing the above LS23- HA-UBE2L3 band with less exposure after the washes involved in GAPDH staining.

Analysis of the western blots obtained during this first study (**Figure 4.2** and **Figure 4.3**) indicate UBE2L3 overexpression in the reporter cell line is clearly obtained for plasmids LS3 and LS23 (EF1-PylRS CAG-UBE2L3* and CAG-UBE2L3* WT, respectively). Expression levels for CypK-containing variants appear slightly above baseline UBE2L3 expression.

One concern that arose from these assays is that the plasmid stock for UBE2L3 used in Dr. Lewis' laboratory has likely deteriorated or is facing some other issue, as expression levels were comparable to baseline. Protein constructs with a HA-tag required higher exposure times to be revealed on the blot. This led to initial suspicions of the HA tag potentially being skipped during protein translation.

The second series of experiments conducted at QMUL involved cell viability assays to assess the potential toxicity of inhibitor complex **4**. The inhibitor complexes were stored as a DMSO solution, which served as the control. NF-κB-RE-luc2P HEK293 cells were treated with different concentrations of inhibitor **4** (20 μM, 10 μM, and 5 μM) or an equivalent amount of DMSO for a duration of 16 hours.

As shown below in **Figure 4.4**, at a concentration of 20 μM, inhibitor **4** caused a significant decrease in cell viability compared to the controls and baseline, resulting in approximately 77% viability. This observation highlights the importance of considering potential toxicity of inhibitor **4** in future inhibition assays.

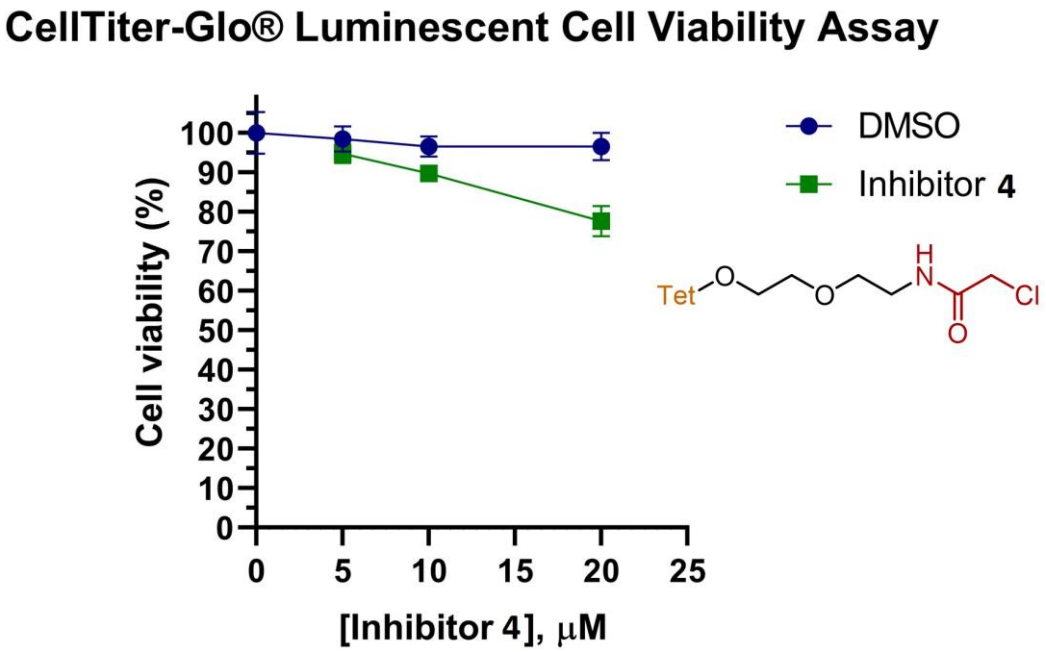


Figure 4.4. Dose response curve obtained during cell viability assays when the NF-κB-RE-luc2P HEK293 reporter cell line was incubated with inhibitor **4** over a period of 16 hours. Data presented as means ± SD, n=3. Graphical representation of data obtained through Graphpad Prism version 8.2.0.

Table 4.2. Cell viability test results. The mean of measured intensities after use of Cell Titer Glo (Promega) is presented in the table below, as well as %viability when compared to the blank sample set. All samples treated with Inhibitor **4** were shown to incurr a significant change in viability when compared to their DMSO control counterparts.

Sample set	Mean Measured Intensity	SD	t	df	Significance (Two-tailed) *p*	Viability (%)
Blank	48946.67	2418.20	NA	NA	NA	100.0
Inh 4, 5 μM	46329.67	1082.97	2.495	9.3	0.032	94.7
0.05% DMSO	48159.33	1432.77				98.4
Inh 4, 10 μM	43901.17	915.08	5.570	9.3	<0.001	89.7
0.10% DMSO	47344.83	1206.54				96.7
Inh 4, 10 μM	37977.00	1707.35	9.812	9.8	<0.001	77.6
0.20% DMSO	47251.67	1564.01				96.5

Results are significant at p<0.05. The t-test was calculated using IBM SPSS Statistics version 29.0.1.1 (244).

In the third and final set of experiments at QMUL, the ONE-Glo™ Luciferase Assay System (Promega) was trialled for monitoring UBE2L3 activity in the reporter cell line. Unfortunately, due to time constraints we were unable to troubleshoot issues that arose during the UBE2L3 inhibition experiments with **4**. One of the takeaways from the troubleshooting experiments lies in the timing at which cells are harvested after transfection. Previous work by Dr. Lewis showed that when expressing UBE2L3, HOIL-1 and HOIP, the NF-κB response activity reaches a peak at the 36-hour mark. As activity is higher at this point it would be better suited to the inhibition assays going forward. It was also revealed that the transfection reagent used in these cultures, Fugene 6, sticks to siliconized pipette tips, therefore we were observing inconsistent transfection rates going down the cell culture plate, as the tip becomes saturated with the transfection reagent whilst pipetting, increasing transfection efficiency. This was measured by transfections with GFP and inspecting the number of cells that fluoresce increasing in the same direction Fugene 6 was added during transfection. This inconsistency led to a high degree of error in initial luciferase reporter assays and first trials of UBE2L3 inhibition assays.

Ultimately, some luciferase assay data was generated at that 36-hour mark, with representative data shown in **Figure 4.5**. After some discussion with Dr. Lewis, the trend seemed to be that the response rates for NF-κB activity were not as high as

they have demonstrated in previous research, and more troubleshooting was to be carried out.

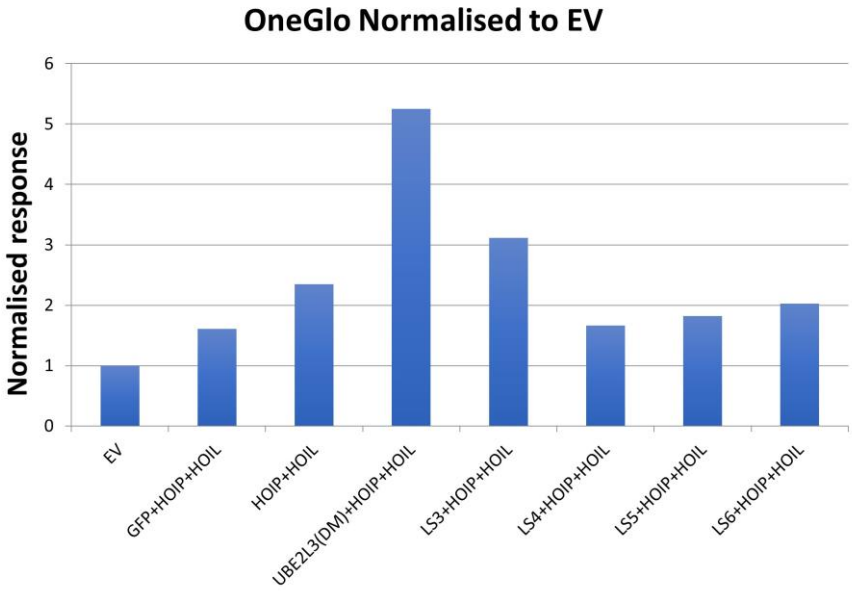


Figure 4.5. Representative results from subjecting NF-κB-RE-luc2P HEK293 cells to the ONE-Glo™ Luciferase Assay System. Cells were harvested 36 hours after transfection. Luciferase response rates were normalised to levels obtained from the empty vector control (EV) and represented in the y axis. The UBE2L3(DM) plasmid used for this assay was obtained from a fresh stock.

Although higher response rates were expected across the board and additional troubleshooting is required, some interesting conclusions could be drawn from the data obtained. The UBE2L3 plasmid construct used by the Lewis group presents a much higher response rate than the UBE2L3 plasmid construct designed during the work carried out in this thesis (LS3 in **Figure 4.5**, corresponding to EF1-PylRS CAG-UBE2L3*), this may be due to the larger size of the plasmid impacting transfection efficiency. Additionally, there appears to be no discernible increase in UBE2L3 activity in cells transfected with plasmids containing UBE2L3* variants UBE2L3*(28TAG), UBE2L3*(89TAG) or UBE2L3*(99TAG) (LS4, LS5 and LS6 respectively in **Figure 4.5**).

The information acquired during this collaboration was enough to conclude that the luciferase reporter assay timing, cell culture and cell transfection methods would require streamlining before its use as a tool to measure the inhibition of the proposed UBE2L3 variants. In addition, the design of the plasmid constructs carrying the genetic information for the enzymes subject to this study may require further

optimisation. In summary, if the expression levels of the proteins of interest could be increased in cells, it may lead to a more streamlined platform for the study of their inhibition using the methods proposed in this study.

4.4. UBE2L3* expression in HEK293 cells

Back at our facilities in Cardiff University, transfection of HEK293 cells with UBE2L3 containing plasmids was tested again (**Table 4.3**). The HEK293 cell line was obtained from Dr. Tsai's stores and the transfection reagent used was Lipofectamine 2000. Several western blot analyses were performed against HA-tagged UBE2L3 using anti-HA antibody. The levels of protein expression seemed to be inconsistent still, and co-transfection with eGFP containing plasmid was trialled side-by-side to visualise transfection more easily under a microscope (**Figure 4.6**). Assays such as this were used to give confidence of the transfection procedure working as intended.

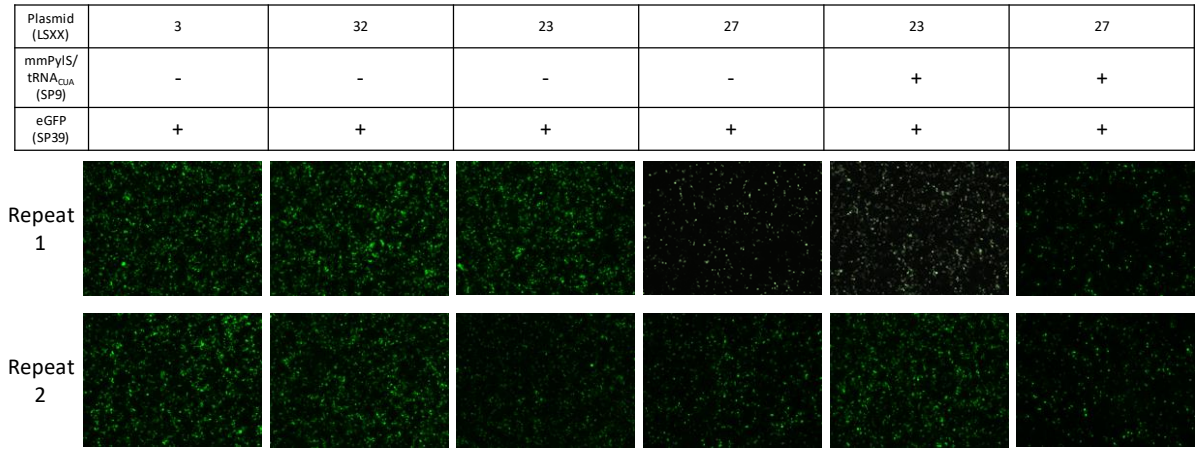


Figure 4.6. Representative example of eGFP expression observed in HEK293 cell cultures when UBE2L3-containing plasmids are co-transfected with e-GFP-containing plasmids.

Table 4.3. Plasmids used during the cell culture work described in this secion

#	Plasmid Name	Notes
LS3	EF1-PylRS CAG-UBE2L3* (C17S C137S)	All LS plasmids contain C17S C137S mutations
LS4	EF1-PylRS CAG-UBE2L3 TAG28	Contains UBE2L3 mutant, MmPylRS and 4x $MmtRNA_{CUA}^{Pyl}(\mathrm{U25C})$. (Synthetase and tRNA pairs required for TAG suppression)
LS5	EF1-PylRS CAG-UBE2L3 TAG89	Contains UBE2L3 mutant, MmPylRS and 4x $MmtRNA_{CUA}^{Pyl}(\mathrm{U25C})$. (Synthetase and tRNA pairs required for TAG suppression)
LS6	EF1-PylRS CAG-UBE2L3 TAG99	Contains UBE2L3 mutant, MmPylRS and 4x $MmtRNA_{CUA}^{Pyl}(\mathrm{U25C})$. (Synthetase and tRNA pairs required for TAG suppression)

#	Plasmid Name	Notes
LS23	CAG-UBE2L3* WT	All LS plasmids contain C17S C137S mutations
LS24	CAG-UBE2L3 TAG28	Require transfection together with SP9
LS25	CAG-UBE2L3 TAG89	Require transfection together with SP9
LS26	CAG-UBE2L3 TAG99	Require transfection together with SP9
LS27	CAG-UBE2L3* WT (-Met)	All LS plasmids contain C17S C137S mutations. This sequence has had the Met1 residue from UBE2L3 removed.
LS32	EF1-PylRS CAG-UBE2L3* (C17S C137S) (-Met)	All LS plasmids contain C17S C137S mutations. This sequence has had the Met1 residue from UBE2L3 removed.
SP9	EF1-Flag-MmPylS-IRES-neoR, 4xPylT	MmPylRS and 4x $MmtRNA_{CUA}^{Pyl}(\mathrm{U25C})$. (Synthetase and tRNA pairs required for TAG suppression)
SP39	pCX eGFP	CAG-eGFP

In order to investigate the inconsistent protein levels observed in western blot analyses using anti-HA, we turned our attention to the availability and integrity of the HA tag itself. The primary concern arose from the protein sequence, which reads as Met-HA-Met-UBE2L3. This raised the possibility of a read-through event during protein translation, where the production of the protein could initiate at UBE2L3's own methionine (Met) residue, bypassing the initial methionine of the HA tag.

To address this potential issue, we designed a series of experiments using specific plasmids. These plasmids were modified to remove the initial methionine (Met) residue from the HA tag region, resulting in plasmids that encoded UBE2L3 without the HA tag and its associated methionine. By comparing the expression and detection of these modified UBE2L3 variants with those containing the intact HA tag, we aimed to determine whether the presence of the HA tag was influencing protein translation and subsequent detection.

Through these investigations, we sought to establish whether the HA tag was being skipped during translation and if this read-through phenomenon could explain the variations in protein levels observed in our western blot analyses. This line of inquiry will provide valuable insights into the role of the HA tag and its impact on UBE2L3 protein expression, allowing us to refine our experimental approaches and interpretations moving forward.

In our pursuit to understand the influence of the HA tag on UBE2L3 protein expression, we conducted a series of experiments involving the deletion of methionine residues from the HA tag region. However, upon analysing the results of western blot analyses, it became evident that the removal of these methionine residues did not lead to any significant increases in protein expression levels. **Figure 4.7** provides representative data from these western blot analyses, demonstrating the lack of substantial changes in protein expression. It does however show lower expression rates when transfection is carried out with the plasmids containing the required tRNA/synthetase machinery.

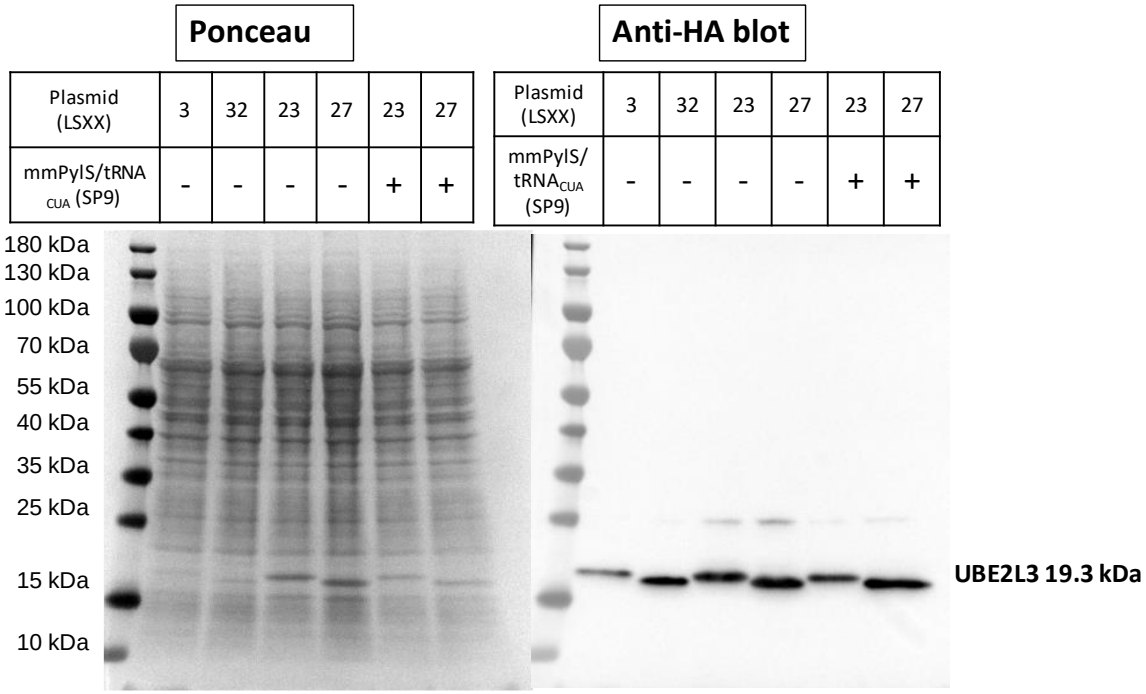


Figure 4.7. Western blot analysis of HA-tagged UBE2L3 production in NF-κB-RE-luc2P HEK293. The total protein is demonstrated on the left via a Ponceau stain, and the anti-HA blot is represented on the right to reveal HA-UBE2L3.

Interestingly, our observations indicated that the plasmids lacking the tRNA/synthetase machinery exhibited more consistent and robust UBE2L3 expression. This observation was consistent across transfections performed with the NF-κB-RE-luc2P HEK293 reporter cell line. When comparing the expression levels of UBE2L3 using plasmids containing both UBE2L3 and the tRNA/synthetase machinery versus plasmids with only UBE2L3 co-transfected with a separate plasmid containing the tRNA/synthetase machinery (**Figure 4.8**), it was apparent that the latter approach yielded higher levels of UBE2L3 expression.

These findings suggest that the presence of the tRNA/synthetase machinery in the same plasmid as UBE2L3 may have adverse effects on UBE2L3 expression or stability. Thus, the separate delivery of UBE2L3 and the tRNA/synthetase machinery via distinct plasmids seems to promote more efficient UBE2L3 expression in our experimental system.

It is worth noting, however, that larger plasmids tend to lead to a decrease in transformation efficiency, which may contribute to the observations described above. Finding the correct balance between plasmid size and number of plasmids used in the same transfection will be critical for future developments of this project.

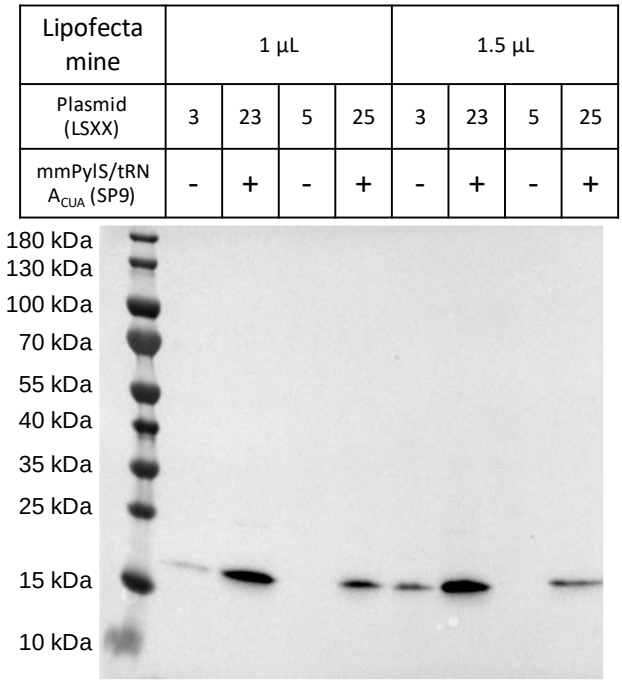


Figure 4.8. Western blot analysis of HA-tagged UBE2L3 production in NF-κB-RE-luc2P HEK293. Plasmids LS3 and LS5 contain the tRNA/synthetase machinery encoded in a single plasmid together with UBE2L3. Plasmids LS23 and LS25 do not contain this machinery and are supplemented by co-transfection with a plasmid which does, SP9.

Of note, the plasmids designed for UBE2L3(89-**1**) expression (LS5 and LS25 in the figures) were showing little or no expression of the target enzyme. This issue was confirmed in several replicate experiments, an example of which can be found in **Figure 4.9**.

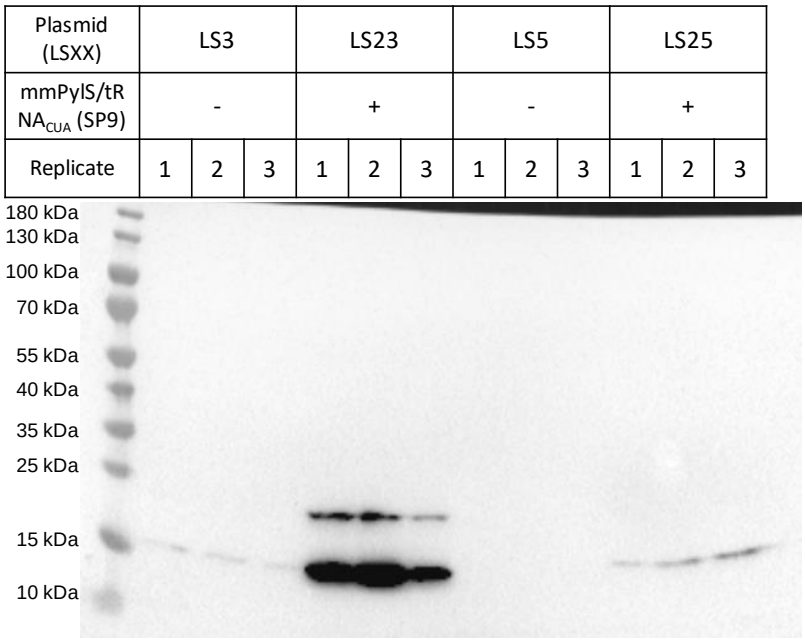


Figure 4.9. Western blot analysis of HA-tagged UBE2L3 production in NF-κB-RE-luc2P HEK293, replicating results shown in **Figure 4.8**.

Further investigation into the root causes was paused at this stage due to time constraints. These results shed light on the factors influencing UBE2L3 expression and highlight the importance of optimizing plasmid design and transfection strategies to achieve desired protein expression levels. The data presented provide valuable insights for the future design and implementation of experiments involving UBE2L3 expression and its functional characterization.

4.5. NF-κB reporter assays

The exploration of the expression system for UBE2L3 variants was accompanied by parallel investigations using the NF-κB-RE-luc2P HEK293 cell line and the ONE-Glo™ Luciferase Assay System.

To ensure the functionality of the NF-κB-RE-luc2P HEK293 cell line, a performance assay was conducted as a validation step. In this assay, the cells were stimulated with TNFα, a potent activator of the NF-κB pathway, followed by the utilization of the ONE-Glo™ Luciferase Assay System. Remarkably, the TNFα stimulation induced a robust increase in luminescence, indicating that the cell line was responding as expected to the NF-κB pathway activation as per vendor guidelines. This confirmed the suitability and responsiveness of the NF-κB-RE-luc2P HEK293 cell line for subsequent experiments.

Building upon this validation, the first and only NFκB luciferase assay within our laboratories at Cardiff University was performed. To assess the impact of UBE2L3 on NF-κB activity, the cells were transfected with UBE2L3 constructs, and the subsequent NF-κB pathway activation was examined. In order to establish appropriate controls, untransfected cells were maintained as a negative control, while cells stimulated with TNFα served as a positive control.

The results of the assay (**Figure 4.10**) revealed that UBE2L3 transfection led to a notable twofold increase in NF-κB activity at 36 hours post-transfection. This finding resonates with UBE2L3 may having a modulatory effect on the NF-κB pathway, potentially influencing downstream signalling events. Intriguingly, transfection with UBE2L3(89-CypK) did not show a significant increase in NF-κB activity, possibly due to lower expression levels of this particular variant. These results are in alignment with the preliminary assays carried out at QMUL, and the prime suspect is still the expression levels for UBE2L3 in these cells. This emphasizes the importance of further investigations and the need to streamline these experiments and their design to obtain more conclusive data.

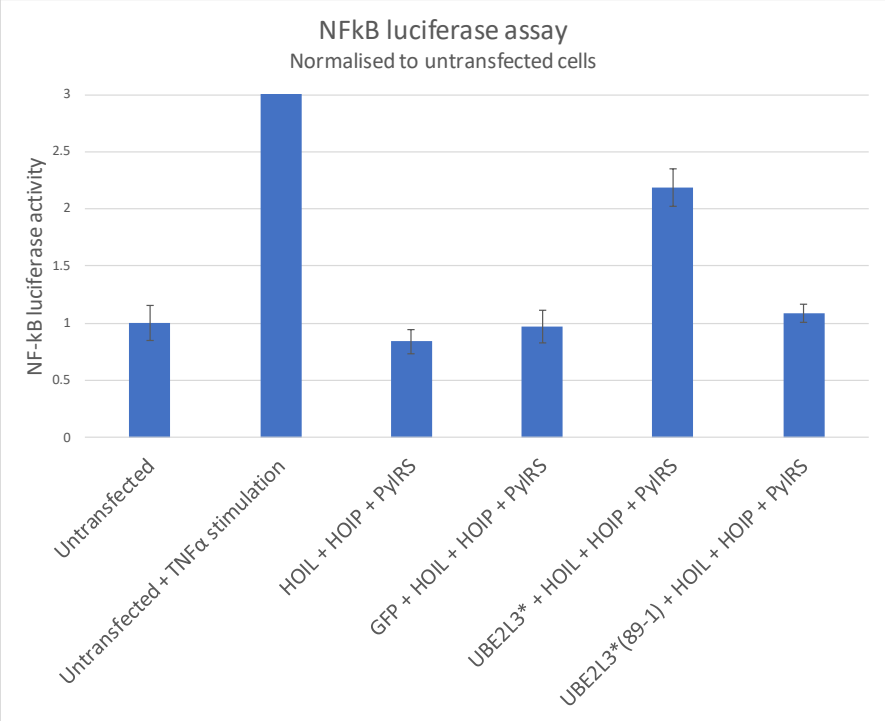


Figure 4.10. Analysis of UBE2L3 influence in NF-κB-RE-luc2P HEK293 cell line, using the ONE-Glo™ Luciferase Assay System. Note that this graph has been "zoomed in" for the purpose of studying the effects of UBE2L3 production, the TNFα stimulation is off the chart, producing over a 77-fold increase in NF-dB production. UBE2L3* was seen to produce a

twofold increase in activity when compared to untransfected cells. Data is presented after normalisation to untransfected cell results. Error bars show ± Standard Error of the Mean.

4.6. Conclusions

In this chapter, we explored a potential platform for monitoring the inhibition of UBE2L3 in living mammalian cells, focusing on its involvement in the NF-κB signalling pathway.

During a visit to Dr. Myles Lewis' lab at Queen Mary University of London, hands-on experience was gained with human cell culture and key experimental techniques were performed. The optimisation of plasmid constructs was tackled to enhance the production of UBE2L3 *in vitro* and the NF-κB activity assay to measure the activity of UBE2L3 in relation to the NF-κB signalling pathway was established.

In collaboration with Dr. Lewis, experiments were conducted using the NF-κB-RE-luc2P HEK293 reporter cell line to confirm the overexpression of UBE2L3 constructs and assess the potential toxicity of inhibitor complex **4**. Western blot analysis demonstrated successful UBE2L3 overexpression, although not to the degree anticipated, and cell viability assays revealed a decrease in cell viability at a concentration of 20 μM of inhibitor **4**, indicating the need to consider toxicity in future inhibition assays.

This work was carried on in Cardiff University, where the NF-κB-RE-luc2P HEK293 reporter cell line was successfully transferred and grown in our facilities. Successful transfections were carried out for production of proteins such as eGFP and UBE2L3*. The cell line was also trialled in its intended use as a reporter of NF-κB upregulation, and the control conditions were successful. It was even able to replicate the upregulation of NF-κB when UBE2L3 is overexpressed in these cells, albeit to a lower degree than reported in past research. These results are still promising for the method's use as a platform for monitoring UBE2L3 activity, or inhibition of its activity, in human cells.

Upon reflection, the work conducted in this chapter came to an end prior to the re-introduction of UBE2L3 native cysteines in amino acid positions 17 and 137. As a result, all cell culture assays were performed using the S17 and S137 mutations instead.

As discussed in Chapter 2, the re-introduction of the native non-catalytic cysteine residues in C17 and C137 in the wild-type UBE2L3 appeared to enhance the enzyme's reaction rates, leading to the formation of polyubiquitin chains in a significantly shorter time frame as observed through SDS-PAGE analysis. If this increased activity observed in the biochemical assays translates to the protein produced in mammalian cells, it is reasonable to assume that it would result in enhanced E2 activity within the cells. It is plausible to theorize that reverting back to its native sequence could have also impacted the enzyme's expression levels or perhaps its folding. Exploring this aspect would be a valuable direction for future investigation, as higher expression levels would likely facilitate the initiation of the inhibition assays.

Unfortunately, the inhibition of UBE2L3 variants in the mammalian cell line was not achieved as initially anticipated. However, the preliminary results obtained during the setup process offer potential avenues for future investigation. These findings emphasize the need for ongoing research in this area and provide valuable insights for further exploration.

Chapter 5

GENERAL CONCLUSIONS

5. CONCLUSIONS

5.1. General Conclusions

The focus of this thesis has been to develop a platform for the selective and rapid inhibition of enzymes containing a catalytic cysteine residue, employing both chemical and genetic approaches. This research aims to address the challenge of targeting proteins that lack small-molecule ligands, as well as entire protein classes that have long been considered “undruggable”.[83,84] By employing innovative strategies, we have sought to overcome these limitations and shed light on the cellular functions and potential disease associations of closely related enzyme family members that may be difficult to selectively inhibit using traditional small molecule inhibitors.

The motivation behind this research stems from the vast number of human proteins that remain unexplored in terms of pharmacological modulation. Despite significant advancements in drug discovery, a substantial portion of the human proteome remains without effective small-molecule ligands. This knowledge gap poses a barrier to understanding the functional roles of these proteins and their implications in various diseases.[85,86] By developing a platform that combines chemical and genetic approaches, we have aimed to bridge this gap and explore the uncharted territory of selectively inhibiting enzymes with catalytic cysteines.

One of the key contributions of this research lies in its relevance to the broader scientific field. By targeting closely related enzyme family members, which may share high structural and functional similarity, we have sought to tackle the challenges associated with selective inhibition. Traditional small molecule inhibitors often struggle to discriminate between closely related enzymes, leading to off-target effects and limited selectivity.[87,88] Our approach, however, offers a promising avenue to overcome these limitations and gain insights into the specific roles of these enzymes in living systems.

Furthermore, our efforts have been driven by the goal of understanding enzyme activity within the context of living systems. By employing both chemical and genetic means to selectively inhibit enzymes, we have aimed to produce a platform that may aid in deciphering their cellular functions and unravelling their potential associations

with various diseases. This integrative approach has the potential to not only enable the investigation of intricate mechanisms underlying enzymatic activity at a cellular level but also provides a foundation for future therapeutic interventions.[89]

In Chapter 2, our research yielded significant achievements in the inhibition of E2 enzymes, specifically UBE2L3. Among the various UBE2L3* variant inhibition assays conducted, the most remarkable outcome was observed with UBE2L3*(89-**1**) in combination with inhibitor complex **4**. Extensive analysis using SDS-PAGE and intact mass spectrometry confirmed the successful tethering of the inhibitor to the protein, providing compelling evidence of its interaction with the catalytic cysteine. These findings underscore the effectiveness of the proposed strategy and its ability to selectively inhibit target enzymes.

Building upon the success achieved in Chapter 2, Chapter 3 further validated the versatility of our approach by targeting another closely related enzyme within the same E2 family, namely UBE2D1. By employing inhibitor complex **4**, we successfully inhibited the activity of UBE2D1(88-**1**). This accomplishment not only expands the scope of our strategy but also emphasizes its potential applicability to a wider range of cysteine-dependent enzymes. The reproducibility of our results across these two different enzymes reinforces the robustness and generalizability of our approach for enzyme inhibition.

These successful achievements in inhibiting UBE2L3 and UBE2D1 through the proposed strategy demonstrate the effectiveness of selectively targeting and inhibiting cysteine-dependent enzymes. By tethering inhibitor complex **4** to the catalytic cysteines of these enzymes, a platform with significant potential for studying enzyme activity, unravelling cellular functions, and investigating disease associations has been established. These findings provide valuable insights into the development of novel methods for deciphering complex biological processes and address the challenges posed by 'undruggable' protein classes and the limited availability of small-molecule ligands. Furthermore, the successful inhibition of UBE2L3 and UBE2D1 serves as a solid foundation for future investigations, fuelling further exploration of this strategy in inhibiting other cysteine-dependent enzymes.

This chemogenetic approach can theoretically be applied to any of the hundreds of cysteine-dependent enzymes, including proteases, phosphatases, oxidoreductases,

isomerases, and ligases. Combining this approach with gene editing could allow for interesting research where genomically encoded proteins are altered to enable rapid and selective modulation of their activity. This integrated strategy holds significant potential for deciphering the biological roles of enzymes for which no selective small-molecule modulators are currently available.

The work carried out for this thesis represents a significant advancement in the development of a platform for selectively and rapidly inhibiting enzymes with catalytic cysteines. By overcoming the limitations associated with “undruggable” protein classes and the lack of small-molecule ligands for many human proteins, valuable contributions have been made to the scientific field. New avenues for studying enzyme activity in living systems have been opened, providing insights that hold great potential for enhancing our understanding of cellular processes and disease mechanisms. These insights lay the groundwork for the discovery of innovative therapeutic strategies and offer promising prospects for future research.

5.2. Future work and closing remarks

Although significant progress has been made in the development of a strategy for selective and rapid enzyme inhibition targeting catalytic cysteines, there are several avenues for future exploration and refinement. The following section outlines key areas that warrant further investigation to enhance the efficacy and applicability of the proposed strategy.

In Chapter 2, the question was raised as to whether cyclopropene lysine (CypK, **1**) would be suitable for the mammalian cell-based studies, mainly because high concentrations could prove cytotoxic. The required concentration of inhibitor conjugates to react with CypK is also an important factor at play, with the biochemical assays showing inhibitor concentrations of up to 200 μM being required to observe inhibitory effects in UBE2L3 variants. Considering that inhibitor **4** was found to have a degree of cytotoxicity at 20 μM concentration in HEK293 cells during initial studies in Chapter 4, it is reasonable to suggest that an alternative unnatural amino acid and inhibitor pair should be considered.

Further work within Dr. Tsai’s research group justified the choice of the cyclopropene and tetrazine reactive pair due to its reaction rate and highlighting that incubation time with the inhibitors could potentially be reduced from 16 hours explored during

this thesis down to 4 hours.[67] A greater reduction to reaction times could be achieved by using an amino acid such as exo-bicyclo[6.1.0]nonyne lysine. This has the potential to not only reduce reaction times, but to also reduce the amount of inhibitor required to successfully inhibit the studied enzyme variants.

Another finding from further work carried out within Dr. Tsai's research group revealed that the more promising inhibitor conjugate candidates, including **4**, were not cell permeable.[67] Although the cell permeability of these inhibitors poses an issue that will be interesting to work on in future experiments, these inhibitors could still be considered for selective inhibition of extracellular cysteine-dependent enzymes such as cysteine cathepsins.[68] Regardless, the inhibitor conjugates would benefit from further development, perhaps increasing their hydrophobicity to help in rendering them cell permeable.

To summarize the findings of Chapter 4, although the NF-κB reporter assay promises to be a great platform for measuring UBE2L3-induced activity in human cells. However, a major challenge encountered is the need for consistent overexpression of the UBE2L3 variants and the tRNA/synthetase machinery required for producing these variants by incorporating the unnatural amino acid CypK. The production of enzymes containing CypK was found to be significantly lower than native protein containing canonical amino acids. Improving the turnover of CypK-containing protein will be crucial for demonstrating their inhibition in cells.

While the re-introduction of native cysteines in UBE2L3 was not performed in this cell culture chapter, the biochemical assays in Chapter 2 indicated that reverting to the native sequence could enhance enzyme activity. Exploring the impact of the native sequence on expression levels and folding represents a valuable direction for future research. Higher expression levels could facilitate the initiation of inhibition assays.

The selective inhibition of cysteine-based enzymes, as proposed in this work, has the potential to be transferred for usage in mammalian cell culture systems. This strategy can provide a toolbox for decrypting the cellular functions and disease relevance of human enzymes that rely on catalytically active cysteine residues. The ability to selectively and rapidly inhibit enzyme activity through the proposed combination of genetic and chemical means is attractive, especially in families of

enzymes that lack small molecule inhibitors, such as E2 ubiquitin-conjugating enzymes.

In conclusion, although the inhibition of UBE2L3 in live mammalian cells was not achieved in this study, the groundwork has been laid for further investigations. The combination of genetic and chemical means for selective and rapid enzyme inhibition holds promise for decrypting the functions and disease relevance of cysteine-based enzymes. Future studies can build upon these findings to refine the methodology and explore alternative approaches for inhibiting UBE2L3 and other enzymes in living mammalian cells.

Chapter 6

MATERIALS AND METHODS

6. MATERIALS AND METHODS

All chemicals and growth media components, unless otherwise specified, were purchased from Fisher Scientific, Fluorochem, Sigma-Aldrich, Melford, New England Biolabs or VWR and were used without further purification. All solutions, including media components, were prepared following manufacturer's guidelines.

UBE1 was kindly provided by Satpal Virdee, Axel Knebel and team members of the MRC PPU Protein Production and Assay development facility.

Ultrapure water was obtained from an Elga® PURELAB Chorus 2 system.

6.1. Buffers, solutions and media

6.1.1. pH measurements

Measurement of pH was carried out at room temperature using a Jenway 3510 pH meter together with a general-purpose SJ 113 pH electrode. The electrode was stored in 3 M KCl and calibrated daily using a 3-point calibration with standards at pH 4.0 (phthalate buffer, Fisher Scientific #J/2820/15), 7.0 (phosphate buffer, Fisher Scientific #J/2855/15) and 10.0 (borate buffer, Fisher Scientific #J/2885/15). Adjustments to pH were carried out using 1 M NaOH or 1 M HCl unless specified otherwise.

6.1.2. Lysogeny broth (LB) liquid media

LB media composition: 10 g/L tryptone, 10 g/L NaCl and 5 g/L yeast extract.

To prepare 1 L of LB liquid media, 10.0 g tryptone, 10.0 g NaCl and 5.0 g yeast extract were dissolved in 800 mL of ultrapure water. The solution was then made up to 1000 mL with ultrapure water and autoclaved for sterilisation at 121 °C for 20 min at a pressure of 2.0 bar. LB liquid media could then be stored at room temperature, sealed to avoid contamination.

6.1.3. Lysogeny broth (LB) agar

LB agar composition: 10 g/L tryptone, 10 g/L NaCl, 5 g/L yeast extract and 1.5% (w/v) agar.

To prepare 500 mL of LB agar, 5.0 g tryptone, 5.0 g NaCl, 2.5 g yeast extract and 7.5 g agar were dissolved in 400 mL of ultrapure water. The solution was then made up to 500 mL with ultrapure water and autoclaved for sterilisation at 121 °C for 20 min at a pressure of 2.0 bar. LB agar could then be stored at room temperature, sealed to avoid contamination.

LB agar plates were made by first microwaving LB agar to liquefy, allowing to cool to touch and mixing with antibiotic stocks as required before pouring into sterile petri dishes. The agar was allowed to cool and solidify at room temperature, and could then either be used or stored at 4 °C.

6.1.1. Antibiotic stock solutions

Antibiotic stock solutions used in this study were created as a 1000x stock composed of either 100 mg/mL Ampicillin in ultrapure water, 50 mg/mL Kanamycin in ultrapure water or 25 mg/mL Chloramphenicol in ethanol. These stock solutions were then stored at -20 °C.

6.1.2. Competent cell preparation buffers

B1 composition: 30 mM KOAc, 50 mM $MnCl_2$, 100 mM KCl, 10 mM $CaCl_2$ and 15% (v/v) glycerol in ultrapure water.

B2 composition: 100 mM MOPS, 75 mM $CaCl_2$ and 10 mM KCl in ultrapure water.

The first buffer, B1 buffer, was prepared as a 200 mL stock by dissolution of 0.589 g KOAc, 1.259 g $MnCl_2$, 1.49 g KCl and 0.222 g $CaCl_2$ in an initial 100 mL of ultrapure water. The pH was adjusted to 5.8 with acetic acid before 30 mL of glycerol was added to the mixture.

The second buffer, B2 buffer, was prepared as a 200 mL stock by dissolution of 0.418 MOPS, 1.66 g $CaCl_2$ and 0.149 g KCl in ultrapure water. The pH was adjusted to 7.0 before addition of 30 mL glycerol to the mixture.

Both solutions were made up to their final volume of 200 mL with ultrapure water, and lastly filtered sterilised with a 0.22 μm syringe filter under sterile conditions into sterile 50 mL falcon tubes. These solutions could then be stored at 4 °C until used, or for longer term at -20 °C.

6.1.3. TAE buffer

A 50x concentrate of TAE buffer for agarose gel electrophoresis was prepared by dissolving 242.0 g Tris base and 18.6 g Na_2EDTA in 800 mL of ultrapure water, followed by addition of 57.1 mL of glacial acetic acid. The solution was made up to 1 L with ultrapure water.

A working stock of TAE buffer was made by diluting 20 mL of the 50x stock with 980 mL of ultrapure water.

6.1.4. SDS resolving buffer

A stock of 4x SDS-PAGE resolving buffer was prepared as a solution containing 1.5 M Tris-HCl pH 8.8 with 0.4% (w/v) SDS.

To prepare 150 mL of 4x stock, 27.23 g Tris base was dissolved in 100 mL of ultrapure water. Then, pH was adjusted to 8.8 using 1 M HCl, followed by addition of 1.5 g SDS. The volume was brought up to 150 mL with ultrapure water and solution could be stored at room temperature.

6.1.5. SDS stacking buffer

A stock of 4x SDS-PAGE stacking buffer was prepared as a solution containing 0.5 M Tris-HCl pH 6.8 with 0.4% (w/v) SDS.

To prepare 100 mL of 4x stock, 6.0 g Tris base was dissolved in 80 mL of ultrapure water. Then, pH was adjusted to 6.8 with 1 M HCl, followed by addition of 1.0 g SDS. The volume was brought up to 100 mL with ultrapure water and solution could be stored at room temperature.

6.1.6. SDS-PAGE running buffer (TGS)

A stock of 10x SDS-PAGE running buffer was prepared as a solution containing 250 mM Tris-HCl pH 8.3, 1.92 M glycine and 1% (w/v) SDS. To prepare a 1 L stock, 30.3 g Tris base, 144.0 g glycine and 10.0 g SDS were dissolved in 1 L of ultrapure water.

6.1.7. SDS-PAGE loading dye

A stock of 4x SDS-PAGE loading dye was prepared as a solution containing 0.2 M Tris-HCl pH 6.8, 8% (w/v) SDS, 0.4 % (w/v) bromophenol blue, 40% (v/v) glycerol and 0.4 M DTT.

6.2. Bacterial culture

6.2.1. Bacterial strains

All bacterial strains used in this thesis originated from *E. coli* and are chemically competent cells. For protein expression, both BL21(DE3) and BL21(DE3)pLysS competent cell strains were utilised. For cloning, Stbl3 and DH5α were used.

6.2.2. Preparation of chemically competent cells

Chemically competent *E. coli* strains for protein expression, BL21(DE3) and BL21(DE3)pLysS were readily available as stocks produced by other members of the research group. As for cloning strains, Stbl3 strain was obtained from ThermoFisher (One Shot™ Stbl3™) and DH5α strains were obtained both from stocks available to the research group and from New England Biolabs as NEB® 5-alpha.

Bacterial cell stocks were plated on LB agar and incubated overnight at 37 °C. Single bacterial colonies were selected to inoculate 5 mL of LB medium in a 50 mL centrifuge tube (falcon tubes) in order to generate a bacterial starter culture. This starter culture was incubated overnight at 37 °C with continuous shaking at 180 RPM.

To prepare the main culture, 200 mL of sterile LB media in a conical flask was inoculated with 1 mL of the starter culture and incubated at 37 °C with continuous shaking until the culture reached an optical density (OD) within the range of 0.6 - 0.7. The flask was then removed from the incubator and cooled on ice for 30 minutes to suspend the cell growth phase.

The culture was divided into 50 mL falcon tubes and centrifuged at 1000 g for 30 minutes at a constant temperature of 4 ºC using a Hettich Rotina 420R centrifuge. The supernatant was carefully removed under sterile conditions, and the resulting

cell pellets were re-suspended in 30 mL of buffer B1. The suspension was incubated on ice for 1 hour.

Following the incubation, the cultures were centrifuged again at 1000 g for 30 minutes at a constant temperature of 4 ºC. The supernatant was discarded, and the cell pellets were re-suspended in 5 mL of buffer B2. Under sterile conditions, the 5 mL suspension was divided into sterile 1.5 mL Eppendorf tubes as 50 µL aliquots. These aliquots were snap frozen in liquid nitrogen and stored at -80 ºC until further use.

New stocks of chemically competent cells were screened for antibiotic resistance by spreading on LB agar plates containing antibiotic controls used in this work (ampicillin, kanamycin, or chloramphenicol). In order to assess a sufficient degree of transformation efficiency for each new batch, chemical transformation with pUC19 was carried out as a control.

6.2.3. Chemical transformation of competent cells

To a 50 µL aliquot of chemically competent cells, thawed on ice, 10-30 ng of plasmid DNA was added and gently mixed with the aliquot, which was then allowed to rest on ice for 30 minutes. In order to induce plasmid uptake, the competent cells were then subjected to a heat shock for 45 seconds in a water bath heated to 42 °C, followed by chilling on ice for 2 minutes. Following this, 1 mL of LB liquid media was added to the cell suspension and gently mixed, followed by a recovery step performed by incubating the cells at 37 °C for 1 hour with constant shaking at 800 g on a ThermoMixer® C (Eppendorf) with 1.5 mL SmartBlock^{TM} attachment. Post-recovery, cell suspensions were either centrifuged in a microcentrifuge and resuspended in LB media before plating, or 50-100 µL of the suspension was directly plated on LB agar containing antibiotics for plasmid selection. Plates were incubated at 37 °C overnight.

6.3. Gene sequences and plasmids

Some of the plasmids used in this PhD research were drawn from the plasmid library of Dr. Yu-Hsuan Tsai (pCDF MbPylT^{CUA} 3xMbPylRS, EF1-Flag-MmPylS-IRES-neoR 4xMmPylT$^{CUA(U25C)}$, pCX eGFP), others were obtained from Dr. Satpal Virdee (UBE2L3* wild type and TAG mutants), Prof. Rudolf Allemann (TEV protease) and

Dr. Myles Lewis (as mentioned in Chapter 4.3). UBE2D1 and HOIP(697-1072) were purchased as gene fragments from *ThermoFisher GeneArt Strings*. A complete list of all plasmids can be found in **Table 6.1**.

All plasmid sequences were confirmed by DNA Sanger sequencing performed by Eurofins Genomics.

Table 6.1. List of plasmids used in this study; denoting the plasmid name, genes carried, vector backbone and antibiotic resistance for bacterial propagation.

Plasmid	Gene	Vector	Antibiotic resistance
pET HOIP	His_{6x}-TEV-HOIP(697-1072)	pET28a	Kanamycin
pET UBE2L3*	His_{6x}-UBE2L3*	pET15b	Ampicillin
pET UBE2L3*(28TAG)	His_{6x}-UBE2L3*(28TAG)	pET15b	Ampicillin
pET UBE2L3*(32TAG)	His_{6x}-UBE2L3*(32TAG)	pET15b	Ampicillin
pET UBE2L3*(59TAG)	His_{6x}-UBE2L3*(59TAG)	pET15b	Ampicillin
pET UBE2L3*(71TAG)	His_{6x}-UBE2L3*(71TAG)	pET15b	Ampicillin
pET UBE2L3*(73TAG)	His_{6x}-UBE2L3*(73TAG)	pET15b	Ampicillin
pET UBE2L3*(82TAG)	His_{6x}-UBE2L3*(82TAG)	pET15b	Ampicillin
pET UBE2L3*(89TAG)	His_{6x}-UBE2L3*(89TAG)	pET15b	Ampicillin
pET UBE2L3*(94TAG)	His_{6x}-UBE2L3*(94TAG)	pET15b	Ampicillin

pET UBE2L3*(99TAG)	His_{6x}-UBE2L3*(99TAG)	pET15b	Ampicillin
pET UBE2L3*(131TAG)	His_{6x}-UBE2L3*(131TAG)	pET15b	Ampicillin
pET UBE2L3*(135TAG)	His_{6x}-UBE2L3*(135TAG)	pET15b	Ampicillin
pET UBE2L3	His_{6x}-UBE2L3	pET15b	Ampicillin
pET UBE2L3(89TAG)	His_{6x}-UBE2L3(89TAG)	pET15b	Ampicillin
pET UBE2D1	His_{6x}-UBE2D1	pET15b	Ampicillin
pET UBE2D1(88TAG)	His_{6x}-UBE2D1(88TAG)	pET15b	Ampicillin
EF1-PylRS CAG-UBE2L3 WT	MmPylT$^{CUA(U25C)}$ FLAG-MmPylRS HA-UBE2L3*	piggyBac	Ampicillin
EF1-PylRS CAG-UBE2L3 TAG28	MmPylT$^{CUA(U25C)}$ FLAG-MmPylRS HA-UBE2L3*(28TAG)	piggyBac	Ampicillin
EF1-PylRS CAG-UBE2L3 TAG89	MmPylT$^{CUA(U25C)}$ FLAG-MmPylRS HA-UBE2L3*(89TAG)	piggyBac	Ampicillin
EF1-PylRS CAG-UBE2L3 TAG99	MmPylT$^{CUA(U25C)}$ FLAG-MmPylRS HA-UBE2L3*(99TAG)	piggyBac	Ampicillin
CAG-UBE2L3 WT	HA-UBE2L3*	pCX	Ampicillin
CAG-UBE2L3 TAG28	HA-UBE2L3*(28TAG)	pCX	Ampicillin
CAG-UBE2L3 TAG89	HA-UBE2L3*(89TAG)	pCX	Ampicillin

CAG-UBE2L3 TAG99	HA-UBE2L3*(99TAG)	pCX	Ampicillin
EF1-PylRS CAG-UBE2L3 WT (-Met1)	MmPylT$^{CUA(U25C)}$ FLAG-MmPylRS HA-(-Met1)UBE2L3*	piggyBac	Ampicillin
EF1-PylRS CAG-UBE2L3 TAG28 (-Met1)	MmPylT$^{CUA(U25C)}$ FLAG-MmPylRS HA-(-Met1)UBE2L3*(28TAG)	piggyBac	Ampicillin
EF1-PylRS CAG-UBE2L3 TAG89 (-Met1)	MmPylT$^{CUA(U25C)}$ FLAG-MmPylRS HA-(-Met1)UBE2L3*(89TAG)	piggyBac	Ampicillin
EF1-PylRS CAG-UBE2L3 TAG99 (-Met1)	MmPylT$^{CUA(U25C)}$ FLAG-MmPylRS HA-(-Met1)UBE2L3*(99TAG)	piggyBac	Ampicillin
CAG-UBE2L3 WT (-Met1)	HA-(-Met1)UBE2L3*	pCX	Ampicillin
CAG-UBE2L3 TAG28 (-Met1)	HA-(-Met1)UBE2L3*(28TAG)	pCX	Ampicillin
CAG-UBE2L3 TAG89 (-Met1)	HA-(-Met1)UBE2L3*(89TAG)	pCX	Ampicillin
CAG-UBE2L3 TAG99 (-Met1)	HA-(-Met1)UBE2L3*(99TAG)	pCX	Ampicillin
EF1-Flag-MmPylS-IRES-neoR, 4xPylT	MmPylT$^{CUA(U25C)}$, FLAG-MmPylRS	NA	Ampicillin
pCDF MbPylT^{CUA} 3xMbPylRS	MbPylT$^{CUA(U25C)}$, MmPylRS	pCDF	Spectinomycin

The sequences of the genes of interest in this project are found below in **Table 6.2.**

Table 6.2. Gene sequences of proteins employed in the project.

His_{6x}-TEV-HOIP(697-1072)	**Purchased from ThermoFisher GeneArt Strings** Cloning sites: *NdeI-EcoRI*
	CATATGGGTAGCAGCCATCACCATCATCATCATAGCAGCG GTGAAAACCTGTATTTTCAAGGTGGTAGCAGTCAAGAATG TGCAGTTTGTGGTTGGGCACTGCCGCATAATCGTATGCA GGCACTGACCAGCTGTGAATGTACCATTTGTCCGGATTGT TTTCGTCAGCATTTTACCATTGCGCTGAAAGAAAAACACA TCACCGATATGGTTTGTCCGGCATGTGGTCGTCCGGATC TGACCGATGATACCCAGCTGCTGAGCTATTTTAGCACCCT GGATATTCAGCTGCGTGAAAGCCTGGAACCGGATGCCTA TGCACTGTTTCATAAAAAACTGACCGAAGGTGTTCTGATG CGTGATCCGAAATTTCTGTGGTGTGCACAGTGTAGCTTTG GCTTTATCTATGAACGTGAACAGCTGGAAGCAACCTGTCC GCAGTGTCATCAGACCTTTTGTGTTCGTTGTAAACGTCAG TGGGAAGAACAGCATCGTGGTCGTAGCTGTGAAGATTTT CAGAATTGGAAACGTATGAACGATCCGGAATATCAGGCA CAAGGTCTGGCAATGTATCTGCAAGAAAATGGTATTGATT GCCCGAAATGCAAATTTAGCTATGCCCTGGCACGTGGTG GTTGTATGCATTTTCATTGTACCCAGTGTCGTCATCAGTTT TGTAGCGGTTGTTATAATGCCTTCTACGCCAAAAACAAAT GTCCGGAACCGAATTGTCGTGTGAAAAAAGCCTGCATG GTCATCATCCGCGTGATTGTCTGTTTTATCTGCGTGATTG GACCGCACTGCGTCTGCAGAAACTGCTGCAGGATAATAA TGTGATGTTTAATACCGAACCGCCTGCCGGTGCGCGTGC AGTTCCTGGTGGTGGCTGTCGTGTTATTGAACAGAAAGAA GTTCCGAACGGCCTGCGTGATGAAGCATGTGGTAAAGAA ACTCCGGCAGGTTATGCAGGTCTGTGTCAGGCACATTATA AAGAATATCTGGTGAGCCTGATTAATGCCCATAGCCTGGA TCCGGCAACACTGTATGAAGTTGAAGAACTGGAAACCGC AACCGAACGTTATCTGCATGTTCGTCCGCAGCCGCTGGC AGGCGAAGATCCGCCTGCATATCAAGCACGTCTGCTGCA

	GAAATTGACCGAAGAAGTGCCGCTGGGTCAGAGCATTCC GCGTCGTCGTAAATAAGAATTC
His_{6x}-UBE2L3*	**Obtained from Dr. Satpal Virdee** Cloning sites: *BamHI-NotI*
	ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGC CTGGAAGTTCTGTTCCAGGGGCCCGGATCCATGGCGGCC AGCAGGAGGCTGATGAAGGAGCTTGAAGAAATCCGCAAA TCTGGGATGAAAAACTTCCGTAACATCCAGGTTGATGAAG CTAATTTATTGACTTGGCAAGGGCTTATTGTTCCTGACAA CCCTCCATATGATAAGGGAGCCTTCAGAATCGAAATCAAC TTTCCAGCAGAGTACCCATTCAAACCACCGAAGATCACAT TTAAAACAAAGATCTATCACCCAAACATCGACGAAAAGGG GCAGGTCTGTCTGCCAGTAATTAGTGCCGAAAACTGGAA GCCAGCAACCAAAACCGACCAAGTAATCCAGTCCCTCATA GCACTGGTGAATGACCCCCAGCCTGAGCACCCGCTTCGG GCTGACCTAGCTGAAGAATACTCTAAGGACCGTAAAAAAT TCTCTAAGAATGCTGAAGAGTTTACAAAGAAATATGGGGA AAAGCGACCTGTGGACTAAGCGGCCGCAGATCCAGATCC
His_{6x}-UBE2L3	**Obtained via SDM of His_{6x}-UBE2L3*** Cloning sites: *BamHI-NotI*
	ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGC CTGGAAGTTCTGTTCCAGGGGCCCGGATCCATGGCGGCC AGCAGGAGGCTGATGAAGGAGCTTGAAGAAATCCGCAAA TGTGGGATGAAAAACTTCCGTAACATCCAGGTTGATGAAG CTAATTTATTGACTTGGCAAGGGCTTATTGTTCCTGACAA CCCTCCATATGATAAGGGAGCCTTCAGAATCGAAATCAAC TTTCCAGCAGAGTACCCATTCAAACCACCGAAGATCACAT TTAAAACAAAGATCTATCACCCAAACATCGACGAAAAGGG GCAGGTCTGTCTGCCATAGATTAGTGCCGAAAACTGGAA GCCAGCAACCAAAACCGACCAAGTAATCCAGTCCCTCATA GCACTGGTGAATGACCCCCAGCCTGAGCACCCGCTTCGG

	GCTGACCTAGCTGAAGAATACTCTAAGGACCGTAAAAAAT TCTGTAAGAATGCTGAAGAGTTTACAAAGAAATATGGGGA AAAGCGACCTGTGGACTAAGCGGCCGCAGATCCAGATCC
His6x-UBE2D1	**Purchased from ThermoFisher GeneArt Strings** Cloning sites: *XbaI-NotI*
	TCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACC ATGGGTAGCAGCCATCACCATCATCATCATAGCAGCGGT CTGGAAGTTCTGTTTCAAGGTCCGGGTAGCATGGCGCTG AAGAGGATTCAGAAAGAATTGAGTGATCTACAGCGCGATC CACCTGCTCACTGTTCAGCTGGACCTGTGGGAGATGACT TGTTCCACTGGCAAGCCACTATTATGGGGCCTCCTGATAG CGCATATCAAGGTGGAGTCTTCTTTCTCACTGTACATTTTC CGACAGATTATCCTTTTAAACCACCAAAGATTGCTTTCACA ACAAAAATTTACCATCCAAACATAAACAGTAATGGAAGTAT TTGTCTCGATATTCTGAGGTCACAATGGTCACCAGCTCTG ACTGTATCAAAAGTTTTATTGTCCATATGTTCTCTACTTTG TGATCCTAATCCAGATGACCCCTTAGTACCAGATATCGCA CAAATCTATAAATCAGACAAAGAAAAATACAACAGACATG CAAGAGAATGGACTCAGAAATATGCAATGTAAGCGGCCG C
FLAG-MmPylRS	**Obtained from Dr. Yu-Hsuan Tsai**
	TCTAGAGCTAGCGTTTAAACTTAAGCTTGCCACCATGGAC TACAAGGACGACGACGACAAGATGGACAAGAAGCCCCTG AACACCCTGATCAGCGCCACAGGACTGTGGATGTCCAGA ACCGGCACCATCCACAAGATCAAGCACCACGAGGTGTCC CGGTCCAAAATCTACATCGAGATGGCCTGCGGCGATCAC CTGGTCGTCAACAACAGCAGAAGCAGCCGGACAGCCAGA GCCCTGCGGCACCACAAGTACAGAAAGACCTGCAAGCGG TGCAGAGTGTCCGACGAGGACCTGAACAAGTTCCTGACC AAGGCCAACGAGGACCAGACCAGCGTGAAAGTGAAGGT GGTGTCCGCCCCCACCCGGACCAAGAAAGCCATGCCCAA GAGCGTGGCCAGAGCCCCCAAGCCCCTGGAAAACACCG

	AAGCCGCTCAGGCCCAGCCCAGCGGCAGCAAGTTCAGC CCCGCCATCCCCGTGTCTACCCAGGAAAGCGTCAGCGTC CCCGCCAGCGTGTCCACCAGCATCTCTAGCATCTCAACC GGCGCCACAGCTTCTGCCCTGGTCAAGGGCAACACCAAC CCCATCACCAGCATGTCTGCCCCTGTGCAGGCCTCTGCC CCAGCCCTGACCAAGTCCCAGACCGACCGGCTGGAAGT GCTCCTGAACCCCAAGGACGAGATCAGCCTGAACAGCGG CAAGCCCTTCCGGGAGCTGGAAAGCGAGCTGCTGAGCC GGCGGAAGAAGGACCTCCAGCAAATCTACGCCGAGGAAC GGGAGAACTACCTGGGCAAGCTGGAAAGAGAGATCACCC GGTTCTTCGTGGACCGGGGCTTCCTGGAAATCAAGAGCC CCATCCTGATCCCCCTGGAGTACATCGAGCGGATGGGCA TCGACAACGACACCGAGCTGAGCAAGCAGATTTTCCGGG TGGACAAGAACTTCTGCCTGCGGCCCATGCTGGCCCCCA ACCTGTACAACTACCTGCGGAAACTGGATCGCGCTCTGC CCGACCCCATCAAGATTTTCGAGATCGGCCCCTGCTACC GGAAAGAGAGCGACGGCAAAGAGCACCTGGAAGAGTTTA CAATGCTGAACTTTTGCCAGATGGGCAGCGGCTGCACCA GAGAGAACCTGGAATCCATCATCACCGACTTTCTGAACCA CCTGGGGATCGACTTCAAGATCGTGGGCGACAGCTGCAT GGTGTACGGCGACACCCTGGACGTGATGCACGGCGACC TGGAACTGTCTAGCGCCGTCGTGGGACCCATCCCTCTGG ACCGGGAGTGGGGCATCGATAAGCCCTGGATCGGAGCC GGCTTCGGCCTGGAACGGCTGCTGAAAGTCAAGCACGAC TTTAAGAACATCAAGCGGGCTGCCAGAAGCGAGAGCTAC TACAACGGCATCAGCACCAACCTGTGATGAGGATCCGCG GCCGC
U6 **MmPylT$^{CUA(U25C)}$**	**Obtained from Dr. Yu-Hsuan Tsai** GACAAGTGCGGTTTTTCCTAGTTGGGCAGGAAGAGGGCC TATTTCCCATGATTCCTTCATATTTGCATATACGATACAAG GCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACAC AAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAAT TTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATG

	GACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTT CTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGGA AACCTGATCATGTAGATCGAACGGACTCTAAATCCGTTCA GCCGGGTTAGATTCCCGGGGTTTCCGGACAAGTGCGGTT TTTCCTAGTT
MbPylRS	**Obtained from Dr. Yu-Hsuan Tsai**
	ATGATGGATAAAAAACCGCTGGATGTGCTGATTAGCGCG ACCGGCCTGTGGATGAGCCGTACCGGCACCCTGCATAAA ATCAAACATCATGAAGTGAGCCGCAGCAAAATCTATATTG AAATGGCGTGCGGCGATCATCTGGTGGTGAACAACAGCC GTAGCTGCCGTACCGCGCGTGCGTTTCGTCATCATAAATA CCGCAAAACCTGCAAACGTTGCCGTGTGAGCGATGAAGA TATCAACAACTTTCTGACCCGTAGCACCGAAAGCAAAAAC AGCGTGAAAGTGCGTGTGGTGAGCGCGCCGAAAGTGAAA AAAGCGATGCCGAAAAGCGTGAGCCGTGCGCCGAAACC GCTGGAAAATAGCGTGAGCGCGAAAGCGAGCACCAACAC CAGCCGTAGCGTTCCGAGCCCGGCGAAAAGCACCCCGA ACAGCAGCGTTCCGGCGTCTGCGCCGGCACCGAGCCTG ACCCGCAGCCAGCTGGATCGTGTGGAAGCGCTGCTGTCT CCGGAAGATAAAATTAGCCTGAACATGGCGAAACCGTTTC GTGAACTGGAACCGGAACTGGTGACCCGTCGTAAAAACG ATTTTCAGCGCCTGTATACCAACGATCGTGAAGATTATCT GGGCAAACTGGAACGTGATATCACCAAATTTTTTGTGGAT CGCGGCTTTCTGGAAATTAAAAGCCCGATTCTGATTCCGG CGGAATATGTGGAACGTATGGGCATTAACAACGACACCG AACTGAGCAAACAAATTTTCCGCGTGGATAAAAACCTGTG CCTGCGTCCGATGCTGGCCCCGACCCTGTATAACTATCT GCGTAAACTGGATCGTATTCTGCCGGGTCCGATCAAAATT TTTGAAGTGGGCCCGTGCTATCGCAAAGAAAGCGATGGC AAAGAACACCTGGAAGAATTCACCATGGTTAACTTTTGCC AAATGGGCAGCGGCTGCACCCGTGAAAACCTGGAAGCG CTGATCAAAGAATTCCTGGATTATCTGGAAATCGACTTCG AAATTGTGGGCGATAGCTGCATGGTGTATGGCGATACCC

	TGGATATTATGCATGGCGATCTGGAACTGAGCAGCGCGG TGGTGGGTCCGGTTAGCCTGGATCGTGAATGGGGCATTG ATAAACCGTGGATTGGCGCGGGTTTTGGCCTGGAACGTC TGCTGAAAGTGATGCATGGCTTCAAAAACATTAAACGTGC GAGCCGTAGCGAAAGCTACTATAACGGCATTAGCACGAA CCTGTAA
MbPylT^{CUA}	**Obtained from Dr. Yu-Hsuan Tsai**
	TTCTCAACATAAAAAACTTTGTGTAATACTTGTAACGCTAG ATCTGGGAACCTGATCATGTAGATCGAATGGACTCTAAAT CCGTTCAGCCGGGTTAGATTCCCGGGGTTTCCGCCA

6.4. General cloning methods

6.4.1. Plasmid DNA purification

To prepare plasmid stocks, chemically competent *E. coli* Stbl3 or DH5α were chemically transformed (Section 6.2.3) and spread on LB agar plates containing the relevant antibiotic for selection. Culture plates were incubated overnight at 37 °C. A single colony was isolated from the plate and used to seed 5 mL of LB medium, which was further incubated overnight at 37 °C with continuous shaking.

The resulting culture was then centrifuged at 1000 g for 20 minutes at 4 °C to pellet the bacterial cells using a Hettich Rotina 420R centrifuge. Plasmids were extracted using QIAprep Spin Miniprep Kit (Qiagen #27106). Following the manufacturer's protocol, the supernatant was carefully removed, and the pellet was resuspended in 250 μL of buffer P1. Subsequently, 250 μL of buffer P2 was added and the mixture was shaken to obtain a blue and viscous solution. To neutralize P2 buffer, 350 μL of buffer N3 was added and samples were shaken gently until blue colour disappeared and a white precipitate was formed.

The contents were transferred to a 1.5 mL Eppendorf tube and centrifuged at 20,000 g for 20 minutes in a microcentrifuge to pellet the debris. The supernatant was carefully aspirated and added to a Qiagen miniprep spin column. The resulting supernatant was passed through the column via centrifugation at 20,000 g for 1

minute in a microcentrifuge, and the flow-through in the collection tube was discarded. The column was washed with 500 μL of PB buffer, centrifuged again for 1 minute at 20,000 g, and the flow-through was discarded. It was then washed with 750 μL of PE buffer by centrifuging for 1 minute at 20,000 g, and the flow-through was discarded. The collection column was removed and placed in a 1.5 mL Eppendorf tube, and 50 μL of EB buffer or ultrapure water was added. The column was incubated with EB buffer (or ultrapure water) for 2 minutes at room temperature before centrifuging for 1 minute at 20,000 g. The column was then discarded, and the concentration of plasmid DNA was measured using a NanoDrop One (ThermoFisher, #ND-ONE-W) before storage at -20 °C.

Aliquots (5-10 μL) of plasmid stocks were also stored in a plasmid archive at -80 °C.

6.4.2. General PCR and SDM methods and programmes

Some genes and plasmids were amplified via polymerase chain reaction (PCR) during this project. Site-directed mutagenesis (SDM) was performed using the same cycle criteria used for PCR reactions.

There were two different thermocyclers available in the laboratory and they were used interchangeably: 3PrimeG PCR Thermocycler (Techne) or TC-512 Gradient thermal cycler (Techne). The thermophilic polymerase used was PrimeSTAR DNA polymerase (Takara, #R010A), which was purchased individually and as a 2x Master Mix (PrimeSTAR MAX DNA Polymerase). The reaction for PrimeSTAR contained:

Table 6.3. PCR mixture for PrimeSTAR DNA polymerase.

Reaction components	V (μL)	Master Mix V (μL)
DNA template (2 ng/μL)	**0.25**	**0.25**
Primer FWD[a] (10 μM)	**1**	**1**
Primer REV[b] (10 μM)	**1**	**1**
5x Buffer (incl. Mg^{2+})	**10**	**-**
dNTP[c]	**4**	**-**
PrimeSTAR polymerase	**0.5**	**-**

Reaction components	V (µL)	Master Mix V (µL)
PrimeSTAR Max Master Mix[d]	**-**	**25**
diH_2O	**Up to 50***	**Up to 50***

[a] FWD = forward (5' to 3'). [b] REV = (3' to 5').[c] dNTP = deoxyribose nucleotide triphosphate. [d] The Master Mix contains dNTPs and the polymerase in the appropriate buffer with $MgSO_4$. * Final reaction volume of 50 µL. Volume of diH_2O depends on the volume of reactants.

The thermocycler programme employed for PrimeSTAR DNA polymerase was:

Table 6.4. PCR conditions for PrimeSTAR DNA polymerase.

Step	Temperature (°C)	Time (s)	Number of cycles
Denaturing	98	10	30
Annealing	55	5	
Extending	72	1 min/kbp or 5 s/kbp[a]	

[a] Extension time depends on the size of the desired PCR product, around 1 minute per 1 kbp (kilo base pair) for PrimeSTAR DNA Polymerase or around 5 seconds per 1 kbp for PrimeSTAR MAX DNA Polymerase.

The result of PCR amplification of full plasmids was always subsequently digested with *DpnI* (Thermo Fisher #10819410) for 1 hour at 37 °C to eliminate the parental plasmid from the mixture. The PCR product was then purified via PCR clean-up kit (Qiagen) and 5 µL of the reaction mixture was used to transform Stbl3 or DH5α competent cells, where endogenous homologous recombination of overlapping sequences would produce the required plasmid product. For subsequent Gibson Assembly applications, PCR product DNA was purified via gel extraction from a 1% (w/v) agarose gel.

6.4.3. Agarose gel electrophoresis

To prepare agarose gels (1% w/v) 1 g of agarose powder was suspended in 100 mL of 1x TAE buffer and heated to boiling with the aid of a microwave. Once safe to handle, 5 µL of SYBR Safe nucleic acid stain (Invitrogen, #S33102) was added to

the solution, which was swirled to mix thoroughly. The solution was then poured into a gel cast with a comb to mould the wells, and allowed to rest until solid for at least 30 minutes.

Gels were submerged in 1x TAE buffer in a gel tank. DNA samples were generally mixed with FastDigest Green loading dye buffer (10x, ThermoScientific #B72) and loaded into the wells. GeneRuler 1kb Plus DNA Ladder (ThermoScientific #SM1332) was generally used as the control ladder, adding 5 μL to a well.

Electrophoresis was performed by applying a constant voltage of 120V over 35 minutes or until clear DNA fragment or plasmid separation was visualised. The gel imaging was carried out using a Chemi-Doc XRS gel imager (BioRad) with an XcitaBlue™ conversion screen.

6.4.4. Agarose gel extraction

If DNA fragments were to be separated via agarose gel electrophoresis before further use, gel extraction was performed by cutting bands of interest from the gel using a scalpel knife. The cut out gel was then transferred to 1.5 mL microcentrifuge tubes and weighed.

Gel slices were dissolved in QG buffer (Qiagen or in-house 20 mM Tris-HCl, 5.5 M guanidine thiocyanate, pH 6.6) by adding 300 μL of buffer per 100 μg of gel and incubating at 50 °C for 10 minutes. The resulting solution was purified via Qiagen miniprep spin column by transferring to the column, centrifuging for 1 minute at 20,000 g using a microcentrifuge, discarding flow through and then washing column with 750 μL of Qiagen PE buffer via centrifugation again for 1 minute. The DNA adhered to the column was then eluted into a 1.5 mL microcentrifuge tube by adding 50 μL of Qiagen EB buffer or ultrapure water and centrifuging for another 1 minute. DNA concentrations were measured by NanoDrop.

6.4.1. Restriction enzymes

For restriction enzyme digestions, all enzymes used were purchased from the ThermoFisher FastDigest catalogue. All digests were performed as per manufacturer recommendations.

6.4.1. DNA fragment T4 ligation

Where cloning was performed via restriction digest sites, fragments were assembled by T4 ligation (ThermoFisher #EL0011) following manufacturer recommendations. Reaction products were used directly to transform either Stbl3 or DH5α chemically competent cells.

6.4.2. Gibson assembly (GA)

In cases where the assembly of multiple DNA fragments was required, NEBuilder (New England Biolabs, #E2621S) was used as per manufacturer recommendations and the reaction products were used directly to transform either Stbl3 or DH5α chemically competent cells.

6.5. Detailed cloning procedures

Plasmid constructs were generated utilizing the previously described standard procedures outlined in the methods section. Specific details regarding the primer sequences and any variations in experimental conditions are provided below.

6.5.1. Primer sequences

The primers listed in this section were used for cloning of gene constructs described throughout sections 6.5.2-6.5.9, where primers are referenced by the names given in Table 6.5.

Table 6.5. List of primers for cloning.

Name	**Sequence (5' to 3')**	**Aim**
LSP1	GTGCTGTCTCATCATTTTGGCAAAGAATTCGCCACCATGtacccatatgacgttccagattacgctACGCGTATGGCGGCCAGCAGGAG	EF1-PylRS CAG-UBE2L3 WT and variants
LSP2	CAGCCTGCACCTGAGGAGTGGCGGCCGCTTAGTCCACAG	EF1-PylRS CAG-UBE2L3 WT and variants
LSP3	CAGCCTGCACCTGAGGAGTGGCGGCCGCTTAGTCCACAG	CAG-UBE2L3 (-Met)

Name	Sequence (5’ to 3’)	Aim
LSP4	TGACGTTCCAGATTACGCTACGCGTGCGGCCAGCAGGAGGCTGATG	CAG-UBE2L3 (-Met)
LSP5	ATGGAAGTATTTGTCTCGATTAGCTGAGGTCACAATGGTCACCAG	UBE2D1 mutations to TAG89 analogue
LSP6	ATCGAGACAAATACTTCCATTACTGTTTATG	UBE2D1 mutations to TAG89 analogue
LSP7	AGAATCGAAATCAACTTTCCATAGGAGTACCCATTCAAACCACCGAAG	pET UBE2L3*(59TAG)
LSP8	TGGAAAGTTGATTTCGATTCTGAAG	pET UBE2L3*(59TAG)
LSP9	ATTTAAAACATAGATCTATCACCCAAACATCGACGAAAAG	pET UBE2L3*(73TAG)
LSP10	TGTTTTAAATGTGATCTTCGGTGGTTTGAATG	pET UBE2L3*(73TAG)
LSP11	AGCTTGAAGAAATCCGCAAATGTGGGATGAAAAACTTCCGTAACATCC	UBE2L3 S17C mutation
LSP12	TTTGCGGATTTCTTCAAGCTCC	UBE2L3 S17C mutation
LSP13	CTAAGGACCGTAAAAAATTCTGTAAGAATGCTGAAGAGTTTACAAAGAAATATGG	UBE2L3 S137C mutation
LSP14	AGAATTTTTTACGGTCCTTAGAGTATTCTTC	UBE2L3 S137C mutation

6.5.2. pET HOIP

The gene fragment encoding HOIP(697-1072) with an N-terminal hexa-histidine tag (His tag) was acquired from ThermoFisher (GeneArt Strings) as a commercial

product. This gene fragment was put through ThermoFisher's Gene Optimization process in order to carry out codon optimisation for expression in *E. coli*. To enable cleavage of the His tag, a tobacco etch virus (TEV) protease recognition site (ENLYFQG) was incorporated between the tag and the HOIP sequence. Subsequently, the gene fragment was cloned into the pET28a vector at the *NdeI* and *EcoRI* restriction sites.

6.5.3. pET UBE2L3*(XXXTAG) variants

Where new pET UBE2L3*(XXXTAG) variants were required, such as pET UBE2L3*(59TAG) and pET UBE2L3*(73TAG), mutation of the relevant codons to TAG via SDM was carried out using pET UBE2L3* as a template.

For pET UBE2L3*(59TAG), plasmid pET UBE2L3* was used as a template and amplified through PCR reaction with primers LSP7 and LSP8 to introduce the TAG mutation in amino acid position 59 of UBE2L3* via SDM. PCR products were digested with *DpnI* and transformed into DH5α chemically competent cells.

For pET UBE2L3*(73TAG), plasmid pET UBE2L3* was used as a template and amplified through PCR reaction with primers LSP9 and LSP10 to introduce the TAG mutation in amino acid position 73 of UBE2L3* via SDM. PCR products were digested with *DpnI* and transformed into DH5α chemically competent cells.

6.5.4. pET UBE2L3 and pET UBE2L3(89TAG). Reintroduction of C17 and C137.

Mutations S17C and S137C were carried out via SDM in a sequential manner, using pET UBE2L3* and pET UBE2L3*(89TAG) as templates. The first round of PCR was carried out with primers LSP11 and LSP12 in order to reintroduce S17C. PCR products were treated with *DpnI* to remove template DNA, followed by wash on a miniprep column. DH5α cells were transformed with the PCR product and plasmids generated were confirmed via sequencing. The second round of PCR used these previous products as templates and was carried out with primers LSP13 and LSP14 in order to reintroduce S137C. PCR products were treated with *DpnI* to remove template DNA, followed by wash on a miniprep column. DH5α cells were transformed with the PCR product and plasmids generated were confirmed via sequencing.

6.5.5. pET UBE2D1

The gene encoding for UBE2D1 was purchased as a gene fragment from ThermoFisher GeneArt Strings as a commercial product. This gene fragment was put through ThermoFisher's Gene Optimization process to carry out codon optimisation for expression in *E. coli*. This fragment was subjected to restriction enzyme digest at the *XbaI* and *NotI* sites.

The same vector backbone as pET UBE2L3* (a pET15b vector) was subjected to restriction enzyme digest at the *XbaI* and *NotI* sites.

The UBE2D1 gene insert fragment and the pET15b vector were isolated via agarose gel electrophoresis and extraction before being subjected to T4 ligation. Plasmids were propagated with DH5α chemically competent cells.

6.5.6. pET UBE2D1(88TAG)

Plasmid pET UBE2D1 was used as a template and amplified through PCR reaction with primers LSP5 and LSP6 to introduce the TAG mutation in amino acid position 88 of UBE2D1 via SDM. PCR products were digested with *dpnI* and transformed into Stbl3 chemically competent cells.

6.5.7. EF1-PylRS CAG-UBE2L3 WT and UBE2L3 variants (28TAG, 89TAG and 99TAG)

This plasmid was developed using an existing template in our laboratory as the backbone, originally containing 4xMmPylT$^{CUA(U25C)}$, EF1-FLAG-MmPylRS and CAG-eGFP.[90] This plasmid was used to create a 4xMmPylT$^{CUA(U25C)}$, EF1-FLAG-MmPylRS vector by removal of the eGFP gene through restriction enzyme digest at *EcoRI* sites.

For the purposes of this project, pET UBE2L3* or variants (28TAG, 89TAG and 99TAG) were used for amplification of the UBE2L3 gene region via PCR using primers LSP1 and LSP2. This gene was then subcloned, for each variant, into the 4xMmPylT$^{CUA(U25C)}$, EF1-FLAG-MmPylRS derived vector by Gibson Assembly, generating plasmids EF1-PylRS CAG-UBE2L3 WT and variants containing mutations at UBE2L3 sites 28TAG, 89TAG and 99TAG. Plasmids were propagated with Stbl3 chemically competent cells.

6.5.8. CAG-UBE2L3 WT and variants

This plasmid was developed using an existing template in our laboratory as the backbone, originally containing CAG-eGFP (pCX-eGFP).[90]

pCX-eGFP was subjected to restriction enzyme digest at Eco81I and *EcoRI* sites to generate the vector backbone. The gene insert was generated via restriction enzyme digest of either the EF1-PylRS CAG-UBE2L3 WT plasmid or its variants (28TAG, 89TAG and 99TAG) at the *EcoRI* and Eco81I sites. Inserts and vector were isolated via agarose gel electrophoresis and extraction before being subjected to T4 ligation. Plasmids were propagated with Stbl3 chemically competent cells.

6.5.9. EF1-PylRS CAG-UBE2L3 WT (-Met1), CAG-UBE2L3 WT (-Met) and variants

The gene encoding for UBE2L3* in CAG-UBE2L3 WT (or variants TAG28, TAG89 and TAG99) was amplified via PCR using primers LSP3 and LSP4, removing UBE2L3* Met1 residue from the gene. PCR products were run on agarose gel and bands for gene insert extracted. The vector backbone from CAG-UBE2L3 WT was generated via restriction enzyme digest with *NotI* and *MluI*. The PCR products and vector backbone were assembled by Gibson Assembly and resulting plasmids were propagated with DH5α chemically competent cells.

Similarly, to yield EF1-PylRS CAG-UBE2L3 WT (-Met1), the gene encoding for UBE2L3* (or variants TAG28, TAG89 and TAG99) was amplified via PCR using primers LSP3 and LSP4, removing UBE2L3* Met1 residue from the gene. PCR products were run on agarose gel and bands for gene insert extracted. The vector backbone from EF1-PylRS CAG-UBE2L3 WT was generated via restriction enzyme digest at the *MluI* and *NotI* sites. The PCR products and vector backbone were assembled by Gibson Assembly and resulting plasmids were propagated with DH5α chemically competent cells.

6.6. Protein expression and purification

6.6.1. SDS-PAGE analysis

For SDS-PAGE analysis of polyubiquitination reactions, Novex™ Tris-Glycine Mini Gels, WedgeWell™ Format, with a polyacrylamide percentage gradient of 4-20% were used (Invitrogen #XP04205BOX). For these analyses, the protein stain of choice was also purchased as InstantBlue® Coomassie Protein stain (abcam #ab119211).

For all other applications, SDS-PAGE gels were hand cast using Mini-PROTEAN® Tetra Cell Casting Module (Bio-Rad, #1658021), with a standard concentration of 12% acrylamide. The resolving gel was prepared by mixing 3.40 mL ultrapure water, 2.50 mL 4x SDS resolving buffer stock, 4.00 mL of 30% acrylamide/bis-acrylamide (37.5:1, Fisher Scientific #15474099), 100 μL of a 10% (w/v) ammonium persulfate solution and 10 μL of TEMED. This mixture was quickly transferred to the casting module and covered with a layer of isopropanol (IPA). Once polymerisation was complete, IPA was removed. The stacking gel was prepared at 4% acrylamide concentration by mixing 2.90 mL ultrapure water, 1.25 mL 4x SDS stacking buffer stock, 0.83 mL of 30% acrylamide/bis-acrylamide, 50 μL of a 10% (w/v) ammonium persulfate solution and 5 μL of TEMED. This mixture was added to the casting module above the set resolving gel. A comb was immediately placed, before polymerisation completed, in order to create 10-15 sample wells. These set gels could be stored at 4 °C until use.

Samples for SDS-PAGE were prepared by mixing with either NuPAGE LDS Sample Buffer (4x concentrate, Invitrogen #NP0007) or 4x SDS loading dye prepared in-house. The mixed samples were then heated at 95 °C for 5 minutes, followed by cooling on ice before loading onto gels.

The SDS-PAGE gels were run in a gel running tank with 1x SDS-PAGE running buffer. Gels were loaded with 5 μL of molecular weight marker in one well and samples were loaded in up to 5-15 μL volumes per well. Electrophoresis was carried out at a constant current of 35 mA and a voltage cap of 250 V over 45-60 min, or until the bromophenol blue in sample buffers reached the bottom of the gel.

Gels were then removed from their casts and stained with in-house Coomassie Blue or InstantBlue® Coomassie Protein stain to visualise protein bands, with image data being captured on a Chemi-Doc XRS gel imager.

6.6.2. General protein expression procedure

Unless otherwise specified, proteins were expressed using the following general procedure. Plasmids encoding for the protein of interest were transformed into chemically competent *E. coli* bacterial cells BL21(DE3) or BL21(DE3)pLysS and grown on an LB agar plate with appropriate antibiotic selection. A single colony from the plate was then used to inoculate 5-10 mL of LB medium, also supplemented with the appropriate antibiotic for selection, and incubated at 37 °C overnight (16 hours). The resulting culture was used the next day to inoculate a larger volume of LB medium (100-1000 mL) containing the appropriate antibiotic for selection. The volume of starter culture used to inoculate a larger volume of LB for culture could be 2-5% (v/v).

Inoculated cultures were then incubating at 37 °C until reaching an optical density at 600 nm (OD_{600}) value in the range of 0.5-0.7. Once this point was reached, protein expression was induced by the addition of IPTG (0.5 mM) and cultures continued incubation at either 37 °C for 4 hours or 20 °C overnight (16 hours) before harvest. Cells were harvested by centrifugation at 5000 g and 4 °C for 20 minutes using a Hettich Rotina 420R centrifuge, and the supernatant was discarded. The resulting bacterial cell pellet could then be stored at -20 °C for future use, or used immediately in purification procedures. Samples were typically taken before induction and just before harvesting for use in SDS-PAGE analysis.

6.6.1. General protein purification procedure via Ni-NTA affinity chromatography

Unless stated otherwise, cell pellets were suspended in a solution containing 20 mM Tris-HCl pH 7.4, 25 mM imidazole, and 150 mM NaCl. To aid in the lysis process, the suspension was supplemented with lysozyme (1 mg/mL), 1 mM $MgCl_2$, phenylmethylsulfonyl fluoride (0.1 mM), and DNase (5 μg/mL). Subsequently, sonication was employed to lyse the cells while chilled with an ice bath and applying 5 second ON busts with 15 second recovery, using a microtip at 39% amplification

over a period of 10 minutes. Lysates were centrifuged at 5000 g and 4 °C for 20 minutes using a Hettich Rotina 420R centrifuge and supernatant was collected and filtered through a 0.45 μm pore syringe filter.

Ni-NTA resin (Bio-Rad) was equilibrated with lysis buffer (without lysozyme, $MgCl_2$, PMSF or DNase) before pouring into a gravity flow column (Bio-Rad). Cleared supernatant was then applied to the Ni-NTA resin and allowed to flow through. The resin was then typically washed with a wash buffer containing 20 mM Tris-HCl, pH 8.0, 150 mM NaCl and 25 mM imidazole. Once the flow through ran clear and no protein was detected via NanoDrop measurement of protein content at A_{280} (1 Abs = 1 mg/mL), the protein was eluted from the resin using elution buffer (typically wash buffer with 300 mM imidazole, unless otherwise stated, taking care to maintain a pH of 8.0).

Fractions were collected throughout the purification process at a typical volume of 1-1.5 mL and samples of them were analysed via SDS-PAGE.

Protein concentrations were measured using NanoDrop UV-Vis at A_{280}. Extinction coefficients estimates were calculated using Expasy-ProtParam (https://web.expasy.org/protparam).

6.6.2. HOIP expression and purification

His-tagged HOIP(697-1072) was produced in *E. coli* BL21(DE3)pLysS cells by inducing the expression with IPTG at a concentration of 0.4 mM. Additionally, the cultures were supplemented with $ZnCl_2$ at a concentration of 0.1 mM once the OD_{600} reached 0.6. The cultures were then incubated overnight (16 hours) at 16 °C before harvesting the cells.

To begin the purification process, the cell pellets were resuspended in a buffer solution containing 50 mM Tris-HCl (pH 8.0), 2 mM imidazole, 1 μM $ZnCl_2$, 150 mM NaCl, and 5 mM 2-mercaptoethanol. The resuspended cells were supplemented with lysozyme (1 mg/mL), phenylmethylsulfonyl fluoride (PMSF, 0.1 mM), and DNase (5 μg/mL), and then lysed using sonication. The initial purification step was carried out using Ni-NTA affinity chromatography (Ni-NTA resin and gravity flow column, Bio-Rad).

This Ni-NTA affinity chromatography was carried out by utilising a stepwise gradient of imidazole eluent at 2, 25, 50, 75 and 100 mM imidazole before eluting at 200 mM imidazole (taking care to maintain a pH of 8.0). The eluted protein was combined, and buffer was exchanged for a buffer appropriate for the TEV protease cleavage of the His tag. The chosen buffer consisted of 20 mM Tris-HCl pH 7.4, 10 μM $ZnCl_2$, 200 mM NaCl, 5 mM citrate and 5 mM 2-mercaptoethanol.

TEV protease cleavage of the His tag was carried out at a 1:5 protease to HOIP ratio and allowed to proceed in a 15 mL centrifuge tube overnight (16 hours) while on a roller at 4 °C. The reaction mixture was then allowed to drip through Ni-NTA resin in order to remove TEV protease and remaining His-tagged HOIP, and the flow through was collected, concentrated and buffer exchanged for polyubiquitination assay buffer before storage at -80 °C.

6.6.3. UBE2L3* expression and purification

Expression of UBE2L3 through **pET UBE2L3*** in BL21 (DE3) cells was induced with IPTG (0.5 mM) when cultures reached an optical density at 600 nm (OD_{600}) of 0.6. Cultures were incubated at 37 °C for 3 hours before harvesting the cells. Cell pellets were lysed and purified via Ni-NTA affinity chromatography.

6.6.1. UBE2L3 expression and purification

Expression of UBE2L3 through **pET UBE2L3** in BL21 (DE3) cells was induced with IPTG (0.5 mM) when cultures reached an optical density at 600 nm (OD_{600}) of 0.6. Cultures were incubated at 20 °C overnight (16 hours) before harvesting the cells. Cell pellets were lysed and purified via Ni-NTA affinity chromatography.

6.6.2. UBE2D1

Expression of UBE2D1 through **pET UBE2D1** in BL21 (DE3) cells was induced with IPTG (0.5 mM) when cultures reached an optical density at 600 nm (OD_{600}) of 0.7. Cultures were incubated at 20 °C overnight (16 hours) before harvesting the cells. Cell pellets were lysed and purified via Ni-NTA affinity chromatography.

6.6.3. UBE2L3* variants and UBE2D1(88-1). Incorporation of unnatural amino acid 1.

Expression of UBE2L3* variants and UBED1(88-**1**) through their corresponding plasmids **pET UBE2L3*(XXXTAG)** in BL21 (DE3) cells was induced with IPTG (0.5 mM) when cultures reached an optical density at 600 nm (OD_{600}) of 0.6. Addition of CypK (0.5 mM) was performed prior to induction, when cultures reached an OD_{600} of 0.4. Cultures were incubated at 37 °C for 4 hours before harvesting the cells. Cell pellets were lysed and purified via Ni-NTA affinity chromatography.

6.7. Polyubiquitination assays

Polyubiquitination reactions were conducted using the following components: 1 μM UBE1, 5 μM UBE2L3 or UBE2D1, 5 μM HOIP(697-1072), 40 μM ubiquitin, and 10 mM ATP. These components were combined in a buffered solution (dubbed polyubiquitination buffer) consisting of 50 mM HEPES-NaOH, 150 mM NaCl, and 20 mM $MgCl_2$, with a pH of 7.5. The reactions were then incubated at 37°C for a duration of 3 hours, unless stated otherwise. After the incubation period, 12 μL aliquots were taken, and 4 μL loading buffer (ThermoFisher, #NP0007) was added. The mixtures were subsequently heated at 95°C for 5 minutes to denature the proteins. Finally, the samples were loaded onto Novex™ 4–20% Tris-Glycine Mini Protein Gels, specifically designed in a 15-well WedgeWell™ format, for subsequent analysis using SDS-PAGE.

6.8. E2 Inhibition assays

The E2 enzymes were subjected to incubation with a concentration of 50 μM (unless specified otherwise) of inhibitor complexes **2**-**8** at a temperature of 25°C for a duration of 16 hours. This incubation step took place in a buffered solution consisting of 50 mM HEPES-NaOH, 150 mM NaCl, and 20 mM $MgCl_2$, with a pH of 7.5 (polyubiquitination buffer). Following the incubation, the subsequent polyubiquitination assay was conducted as previously described, incorporating the addition of E1 (UBE1), E3 (HOIP), ubiquitin, and ATP. After the incubation at 37°C, samples were collected and subjected to analysis using SDS-PAGE to assess the results of the polyubiquitination reaction.

6.9. Liquid chromatography-mass spectrometry (intact mass)

LC-MS analysis was conducted by the School of Chemistry Analytical Services at Cardiff University, using a Waters Synapt G2-Si quadrupole time-of-flight mass spectrometer coupled with a Waters Acquity H-Class UPLC system. The column employed for the analysis was an Acquity UPLC protein BEH C4 (300 Å, 1.7 μm × 2.1 mm × 100 mm) operated in the reverse phase mode and maintained at a column temperature of 60°C. A gradient elution method was utilized, starting with 95% solvent A and transitioning to 35% solvent A over a period of 50 minutes. Solvent A consisted of 0.1% (v/v) formic acid in ultrapure water, while solvent B comprised 0.1% (v/v) formic acid in acetonitrile. The LC-MS data acquisition was performed in the positive electrospray ionization mode, and subsequent data analysis was carried out using Waters MassLynx software version 4.1. Deconvolution of the protein charged states was achieved using the maximum entropy 1 (MaxEnt 1) processing software.

6.10. HEK293 cell culture

HEK293 cells (ECACC General Collection #85120602) were made readily available and subjected to regular testing for mycoplasma contamination. The NF-κB-RE-*luc2P* HEK293 reporter cell line was generously provided by Dr. Myles Lewis. The cells were cultured in T75 flasks at 37 °C in a 5% CO_2 atmosphere. The growth medium used was DMEM (Fisher Scientific #11574516) supplemented with 10% FBS (v/v, Fisher Scientific #11573397). The cells were maintained as a sub-confluent monolayer and were split when they reached 80-85% confluency.

To initiate the cell splitting process, the cells were first washed with PBS. Subsequently, trypsinization was performed by adding 1 mL of trypsin (0.25%, Fisher Scientific #11560626) to the cells. After incubating for a short duration, 200 µL of the trypsin-cell suspension was transferred to a new T75 flask containing 12 mL of fresh DMEM supplemented with 10% FBS (v/v). The cells were then allowed to proliferate and establish a new sub-confluent monolayer. This routine procedure ensured the maintenance and propagation of the HEK293 cell lines.

6.10.1. HEK293 transfection

The procedure for HEK293 and NF-κB-RE-*luc2P* HEK293 cell line transient transfection carried out in the facilities at Cardiff University were as follows.

Cells were seeded at an approximate density of $1x10^6$ cells per well in a 24-well plate (Corning #10380932). The cells were incubated at 37 °C in a 5% CO_2 atmosphere for 24 hours or until they reached 90% confluency.

For transfection, each well was prepared as follows: 1.5 μL of Lipofectamine 2000 (Life Technologies #10696343) was mixed with 50 μL of OPTIMEM (Fisher Scientific #11058021) and allowed to incubate at room temperature for at least 5 minutes and no longer than 25 minutes. Meanwhile, 500 ng of plasmid was diluted in 50 μL of OPTIMEM. The plasmid solution was then combined with the initial lipofectamine and OPTIMEM mixture, resulting in a final volume of 100 μL. This transfection solution was incubated at room temperature for 20 minutes.

The growth medium in the wells was replaced with fresh DMEM supplemented with 10% (v/v) FBS. Additionally, if transfecting with plasmids that contain UBE2L3 variants, a 100 mM stock solution of CypK (**1**) was added to achieve a final working concentration of 1 mM in the media. The plasmid containing lipofectamine solution (transfection solution) was added drop-wise to each well, and the plate was further incubated at 37 °C in a 5% CO_2 atmosphere for 36-48 hours.

Cellular imaging was performed using a Zeiss AxioCam MRm microscope camera and the ZEN imaging software (version 2.3). Images were captured at a magnification of 40x, covering representative regions of the entire well. GFP fluorescence was detected when applicable using the Zeiss FSet 38 green fluorescence filter (excitation 470/40 and emission 525/50), with a constant exposure time of 550 ms per image.

6.10.2. NF-κB-RE-*luc2P* HEK293 reporter assay

The NF-κB-RE-luc2P HEK293 reporter assay was performed with the ONE-Glo™ Luciferase Assay System (Promega # E6110) as per manufacturer specifications. TNFα stimulation was used as a positive control by adding TNFα (Merck #T0157) to a final concentration in culture wells of 20 ng/mL and allowed to incubate for 5 hours

prior to the ONE-Glo Luciferase assay. Luminescence was measured using the PHERAstar LUM plus module, reading at 0.5 seconds/well.

6.11. Western blotting

Once ready for harvest cell culture media was carefully aspirated from the wells, and the cells were washed twice with 1x PBS. To extract cellular proteins, a 50 μL solution of RIPA Buffer (Sigma #R0278) supplemented with 1% (v/v) protease inhibitor cocktail (Sigma-Aldrich, #P8340) was added to the centre of each well. The plate was then placed on ice and incubated for 10 minutes. Subsequently, the cells were gently scraped from the well surface, and the entire contents were transferred to a 1.5 mL microcentrifuge tube.

The cell lysates were pelleted by centrifugation at 20,000 g for 10 minutes at 4 °C in a microcentrifuge, and 45 μL of the resulting supernatant was combined with 15 μL of NuPAGE™ LDS Sample Buffer (Invitrogen #NP0007). The samples were denatured by heating at 95 °C for 5 minutes and loaded onto a Novex™ 4–20% Tris-Glycine Mini Protein Gel for SDS-PAGE. PageRulerTM Prestained Protein Ladder (Thermo Scientific #26616) was used as the molecular weight ladder for western blot. After electrophoresis was performed, the gel was imaged using a Chemi-Doc XRS gel imager.

Proteins from the gel were transferred onto a nitrocellulose membrane (BioRad #1704158) using a Trans-Blot Turbo Transfer System (Bio-Rad #1704150) running at mixed molecular weight setting. To verify the successful transfer of proteins, the membrane was stained with Ponceau S (0.1% w/v Ponceaus S, 5% v/v acetic acid in ultrapure water) and imaged again. After confirmation, the membrane was blocked with PBST (0.05% v/v Tween 20 in PBS) containing 5% (w/v) milk powder at 18 °C for 1 hour on a rocker with gentle agitation. Subsequently, the membrane was incubated overnight at 4 °C with a primary antibody such as anti-HA (Sigma-Aldrich #05-904) at a dilution of 1:1000 (v/v) in a 5% (w/v) milk PBST solution.

Following the primary antibody incubation, the membrane was washed three times with PBST (10 mL per wash, 5 minutes per wash). Subsequent washing steps followed the same procedure. The membrane was then incubated with a secondary antibody complementary to whichever primary antibody is used. In the case of anti-HA, the anti-mouse secondary antibody (Thermo Fisher #32430) was used at a

dilution of 1:1000 (v/v) in a 5% (w/v) milk PBST solution for 1 hour at 18 °C, followed by another round of washing. To visualize the protein bands, the membrane was treated with Clarity Max^{TM} Western ECL Substrate (Bio-Rad #1705062). The chemiluminescent signal was detected using the Bio-Rad Chemi-Doc XRS system and data was processed using the ImageLab software (Bio-Rad, version 6.0).

7. REFERENCES

1. Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu Rev Biochem* **67,** 425–479 (1998).
2. Pickart, C. M. Mechanisms underlying ubiquitination. *Annu Rev Biochem* **70,** 503–533 (2001).
3. Al-Hakim, A., Escribano-Diaz, C., Landry, M.-C., O'Donnell, L., Panier, S., Szilard, R. K. & Durocher, D. The ubiquitous role of ubiquitin in the DNA damage response. *DNA Repair (Amst)* **9,** 1229–1240 (2010).
4. Komander, D. & Rape, M. The ubiquitin code. *Annu Rev Biochem* **81,** 203–229 (2012).
5. Finley, D. Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu Rev Biochem* **78,** 477–513 (2009).
6. Franklin, T. G. & Pruneda, J. N. A high-throughput assay for monitoring ubiquitination in real time. *Front Chem* **7,** 816 (2019).
7. Hyer, M. L., Milhollen, M. A., Ciavarri, J., Fleming, P., Traore, T., Sappal, D., Huck, J., Shi, J., Gavin, J., Brownell, J., Yang, Y., Stringer, B., Griffin, R., Bruzzese, F., Soucy, T., Duffy, J., Rabino, C., Riceberg, J., Hoar, K., Lublinsky, A., Menon, S., Sintchak, M., Bump, N., Pulukuri, S. M., Langston, S., Tirrell, S., Kuranda, M., Veiby, P., Newcomb, J., Li, P., Wu, J. T., Powe, J., Dick, L. R., Greenspan, P., Galvin, K., Manfredi, M., Claiborne, C., Amidon, B. S. & Bence, N. F. A small-molecule inhibitor of the ubiquitin activating enzyme for cancer treatment. *Nat Med* **24,** 186–193 (2018).
8. Hwang, J.-T., Lee, A. & Kho, C. ubiquitin and ubiquitin-like proteins in cancer, neurodegenerative disorders, and heart diseases. *Int J Mol Sci* **23,** 5053 (2022).
9. Alpi, A. F., Chaugule, V. & Walden, H. Mechanism and disease association of E2-conjugating enzymes: Lessons from UBE2T and UBE2L3. *Biochemical Journal* **473,** 3401–3419 (2016).
10. Cohen, P. & Tcherpakov, M. Will the ubiquitin system furnish as many drug targets as protein kinases? *Cell* **143,** 686–693 (2010).
11. Yang, Q., Zhao, J., Chen, D. & Wang, Y. E3 ubiquitin ligases: Styles, structures and functions. *Molecular Biomedicine* **2,** 23 (2021).
12. Dove, K. K. & Klevit, R. E. RING-Between-RING E3 ligases: Emerging themes amid the variations. *J Mol Biol* **429**, 3363-3375 (2017).
13. Sluimer, J. & Distel, B. Regulating the human HECT E3 ligases. *Cellular and Molecular Life Sciences* **75,** 3121-3141 (2018).
14. Reiter, K. H. & Klevit, R. E. Characterization of RING-Between-Ring E3 ubiquitin transfer mechanisms. *Methods in Molecular Biology* **1844,** 3-17 (2018).
15. Johansson, H., Isabella Tsai, Y.-C., Fantom, K., Chung, C.-W., Kümper, S., Martino, L., Thomas, D. A., Eberl, H. C., Muelbaier, M., House, D. & Rittinger, K. Fragment-based covalent ligand screening enables rapid discovery of inhibitors for the RBR E3 ubiquitin ligase HOIP. *J Am Chem Soc* **141,** 2703–2712 (2019).
16. Fajner, V., Maspero, E. & Polo, S. Targeting HECT-type E3 ligases – insights from catalysis, regulation and inhibitors. *FEBS Lett* **591,** 2636–2647 (2017).

17. Jevtić, P., Haakonsen, D. L. & Rapé, M. An E3 ligase guide to the galaxy of small-molecule-induced protein degradation. *Cell Chem Biol* **28,** 1000–1013 (2021).
18. Popovic, D., Vucic, D. & Dikic, I. Ubiquitination in disease pathogenesis and treatment. *Nat Med* **20,** 1242–1253 (2014).
19. Dhanwani, R., Takahashi, M. & Sharma, S. Cytosolic sensing of immuno-stimulatory DNA, the enemy within. *Curr Opin Immunol* **50,** 82–87 (2018).
20. Dantuma, N. P. & Bott, L. C. The ubiquitin-proteasome system in neurodegenerative diseases: Precipitating factor, yet part of the solution. *Front Mol Neurosci* **7,** 70 (2014).
21. Sreedhar, A. & Zhao, Y. Dysregulated metabolic enzymes and metabolic reprogramming in cancer cells. *Biomed Rep* **8,** 3–10 (2018).
22. Makin, S. The amyloid hypothesis on trial. *Nature* **559,** S4–S7 (2018).
23. Simon, G. M. & Cravatt, B. F. Activity-based proteomics of enzyme superfamilies: Serine hydrolases as a case study. *J Biol Chem* **285,** 11051–11055 (2010).
24. Oprea, T. I., Bologa, C. G., Brunak, S., Campbell, A., Gan, G. N., Gaulton, A., Gomez, S. M., Guha, R., Hersey, A., Holmes, J., Jadhav, A., Jensen, L. J., Johnson, G. L., Karlson, A., Leach, A. R., Ma'ayan, A., Malovannaya, A., Mani, S., Mathias, S. L., McManus, M. T., Meehan, T. F., von Mering, C., Muthas, D., Nguyen, D.-T., Overington, J. P., Papadatos, G., Qin, J., Reich, C., Roth, B. L., Schürer, S. C., Simeonov, A., Sklar, L. A., Southall, N., Tomita, S., Tudose, I., Ursu, O., Vidović, D., Waller, A., Westergaard, D., Yang, J. J. & Zahoránszky-Köhalmi, G. Unexplored therapeutic opportunities in the human genome. *Nat Rev Drug Discov* **17,** 317–332 (2018).
25. Schreiber, S. L., Kotz, J. D., Li, M., Aubé, J., Austin, C. P., Reed, J. C., Rosen, H., White, E. L., Sklar, L. A., Lindsley, C. W., Alexander, B. R., Bittker, J. A., Clemons, P. A., de Souza, A., Foley, M. A., Palmer, M., Shamji, A. F., Wawer, M. J., McManus, O., Wu, M., Zou, B., Yu, H., Golden, J. E., Schoenen, F. J., Simeonov, A., Jadhav, A., Jackson, M. R., Pinkerton, A. B., Chung, T. D. Y., Griffin, P. R., Cravatt, B. F., Hodder, P. S., Roush, W. R., Roberts, E., Chung, D.-H., Jonsson, C. B., Noah, J. W., Severson, W. E., Ananthan, S., Edwards, B., Oprea, T. I., Conn, P. J., Hopkins, C. R., Wood, M. R., Stauffer, S. R., Emmitte, K. A., Brady, L. S., Driscoll, J., Li, I. Y., Loomis, C. R., Margolis, R. N., Michelotti, E., Perry, M. E., Pillai, A. & Yao, Y. Advancing biological understanding and therapeutics discovery with small-molecule probes. *Cell* **161,** 1252–1265 (2015).
26. Shogren-Knaak, M. A., Alaimo, P. J. & Shokat, K. M. Recent advances in chemical approaches to the study of biological systems. *Annu Rev Cell Dev Biol* **17,** 405–433 (2001).
27. El-Brolosy, M. A. & Stainier, D. Y. R. Genetic compensation: A phenomenon in search of mechanisms. *PLoS Genet* **13,** e1006780 (2017).
28. Arrowsmith, C. H., Audia, J. E., Austin, C., Baell, J., Bennett, J., Blagg, J., Bountra, C., Brennan, P. E., Brown, P. J., Bunnage, M. E., Buser-Doepner, C., Campbell, R. M., Carter, A. J., Cohen, P., Copeland, R. A., Cravatt, B., Dahlin, J. L., Dhanak, D., Edwards, A. M., Frederiksen, M.,

Frye, S. V, Gray, N., Grimshaw, C. E., Hepworth, D., Howe, T., Huber, K. V. M., Jin, J., Knapp, S., Kotz, J. D., Kruger, R. G., Lowe, D., Mader, M. M., Marsden, B., Mueller-Fahrnow, A., Müller, S., O'Hagan, R. C., Overington, J. P., Owen, D. R., Rosenberg, S. H., Ross, R., Roth, B., Schapira, M., Schreiber, S. L., Shoichet, B., Sundström, M., Superti-Furga, G., Taunton, J., Toledo-Sherman, L., Walpole, C., Walters, M. A., Willson, T. M., Workman, P., Young, R. N. & Zuercher, W. J. The promise and peril of chemical probes. *Nat Chem Biol* **11,** 536–541 (2015).

29. Islam, K. The bump-and-hole tactic: Expanding the scope of chemical genetics. *Cell Chem Biol* **25,** 1171–1184 (2018).
30. Tsai, Y.-H., Doura, T. & Kiyonaka, S. Tethering-based chemogenetic approaches for the modulation of protein function in live cells. *Chem Soc Rev* **50,** 7909–7923 (2021).
31. Tsai, Y.-H., Essig, S., James, J. R., Lang, K. & Chin, J. W. Selective, rapid and optically switchable regulation of protein function in live mammalian cells. *Nat Chem* **7,** 554–561 (2015).
32. Morgan, C. W., Dale, I. L., Thomas, A. P., Hunt, J. & Chin, J. W. Selective CRAF inhibition elicits transactivation. *J Am Chem Soc* **143,** 4600–4606 (2021).
33. Yanofsky, C. Establishing the triplet nature of the genetic code. *Cell* **128,** 815–818 (2007).
34. Gardin, J., Yeasmin, R., Yurovsky, A., Cai, Y., Skiena, S. & Futcher, B. Measurement of average decoding rates of the 61 sense codons *in vivo*. *Elife* **3,** e03735 (2014).
35. Nödling, A. R., Spear, L. A., Williams, T. L., Luk, L. Y. P. & Tsai, Y.-H. Using genetically incorporated unnatural amino acids to control protein functions in mammalian cells. *Essays Biochem* **63,** 237–266 (2019).
36. Young, T. S., Ahmad, I., Yin, J. A. & Schultz, P. G. An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J Mol Biol* **395,** 361–374 (2010).
37. Nödling, A. R., Spear, L. A., Williams, T. L., Luk, L. Y. P. & Tsai, Y.-H. Using genetically incorporated unnatural amino acids to control protein functions in mammalian cells. *Essays Biochem* **63,** 237–266 (2019).
38. Wang, L., Xie, J. & Schultz, P. G. expanding the genetic code. *Annu Rev Biophys Biomol Struct* **35,** 225–249 (2006).
39. Liu, C. C. & Schultz, P. G. Adding new chemistries to the genetic code. *Annu Rev Biochem* **79,** 413–444 (2010).
40. Hickey, J. L., Sindhikara, D., Zultanski, S. L. & Schultz, D. M. Beyond 20 in the 21st century: Prospects and challenges of non-canonical amino acids in peptide drug discovery. *ACS Med Chem Lett* **14,** 557–565 (2023).
41. Wijk, S. J. L. & Timmers, H. T. M. The family of ubiquitin-conjugating enzymes (E2s): Deciding between life and death of proteins. *The FASEB Journal* **24,** 981–993 (2010).
42. Harper, J. W. & King, R. W. Stuck in the middle: Drugging the ubiquitin system at the E2 step. *Cell* **145,** 1007–1009 (2011).
43. Wertz, I. E. & Wang, X. From discovery to bedside: Targeting the ubiquitin system. *Cell Chem Biol* **26,** 156–177 (2019).

44. Stewart, M. D., Ritterhoff, T., Klevit, R. E. & Brzovic, P. S. E2 enzymes: More than just middle men. *Cell Res* **26,** 423-440 (2016).
45. Huang, X. & Dixit, V. M. Drugging the undruggables: Exploring the ubiquitin system for drug development. *Cell Res* **26,** 484–498 (2016).
46. Chang, S.-C., Zhang, B.-X. & Ding, J. L. E2-E3 ubiquitin enzyme pairing - partnership in provoking or mitigating cancers. *Biochim. Biophys. Acta, Rev. Cancer* **1877,** 188679 (2022).
47. Michelle, C., Vourc'h, P., Mignon, L. & Andres, C. R. What Was the Set of Ubiquitin and Ubiquitin-Like Conjugating Enzymes in the Eukaryote Common Ancestor? *J Mol Evol* **68,** 616–628 (2009).
48. Valimberti, I., Tiberti, M., Lambrughi, M., Sarcevic, B. & Papaleo, E. E2 superfamily of ubiquitin-conjugating enzymes: Constitutively active or activated through phosphorylation in the catalytic cleft. *Sci Rep* **5,** 14849 (2015).
49. Sheng, Y., Hong, J. H., Doherty, R., Srikumar, T., Shloush, J., Avvakumov, G. V, Walker, J. R., Xue, S., Neculai, D., Wan, J. W., Kim, S. K., Arrowsmith, C. H., Raught, B. & Dhe-Paganon, S. A human ubiquitin conjugating enzyme (E2)-HECT E3 ligase structure-function screen. *Mol. Cell. Proteomics* **11,** 329–341 (2012).
50. Kim, T., Bae, S.-C. & Kang, C. Synergistic activation of NF-κB by TNFAIP3 (A20) reduction and UBE2L3 (UBCH7) augment that synergistically elevate lupus risk. *Arthritis Res Ther* **22,** 93 (2020).
51. Agik, S., Franek, B. S., Kumar, A. A., Kumabe, M., Utset, T. O., Mikolaitis, R. A., Jolly, M. & Niewold, T. B. The autoimmune disease risk allele of UBE2L3 in african american patients with systemic lupus erythematosus: A recessive effect upon subphenotypes. *J Rheumatol* **39,** 73–78 (2012).
52. Ma, X., Zhao, J., Yang, F., Liu, H. & Qi, W. Ubiquitin conjugating enzyme E2 L3 promoted tumor growth of NSCLC through accelerating p27kip1 ubiquitination and degradation. *Oncotarget* **8,** 84193–84203 (2017).
53. Zhang, X., Huo, C., Liu, Y., Su, R., Zhao, Y. & Li, Y. Mechanism and disease association with a ubiquitin conjugating E2 Enzyme: UBE2L3. *Front Immunol* **13,** 793610 (2022).
54. Smit, J. J., Monteferrario, D., Noordermeer, S. M., van Dijk, W. J., van der Reijden, B. A. & Sixma, T. K. The E3 ligase HOIP specifies linear ubiquitin chain assembly through its RING-IBR-RING domain and the unique LDD extension. *EMBO J* **31,** 3833 (2012).
55. Lechtenberg, B. C., Rajput, A., Sanishvili, R., Dobaczewska, M. K., Ware, C. F., Mace, P. D. & Riedl, S. J. Structure of a HOIP/E2~ubiquitin complex reveals RBR E3 ligase mechanism and regulation. *Nature* **529,** 546 (2016).
56. Martino, L., Brown, N. R., Masino, L., Esposito, D. & Rittinger, K. Determinants of E2-ubiquitin conjugate recognition by RBR E3 ligases. *Sci Rep* **8,** 68 (2018).
57. de la Torre, D. & Chin, J. W. Reprogramming the genetic code. *Nat Rev Genet* **22,** 169–184 (2021).
58. Schmied, W. H., Elsässer, S. J., Uttamapinant, C. & Chin, J. W. Efficient multisite unnatural amino acid incorporation in mammalian cells via optimized pyrrolysyl tRNA synthetase/tRNA expression and engineered eRF1. *J Am Chem Soc* **136,** 15577–15583 (2014).

59. Stanley, M., Han, C., Knebel, A., Murphy, P., Shpiro, N. & Virdee, S. Orthogonal thiol functionalization at a single atomic center for profiling transthiolation activity of E1 activating enzymes. *ACS Chem Biol* **10,** 1542–1554 (2015).
60. Knight, Z. A. & Shokat, K. M. Chemical Genetics: Where genetics and pharmacology meet. *Cell* **128,** 425–430 (2007).
61. Burgess-Brown, N. A., Sharma, S., Sobott, F., Loenarz, C., Oppermann, U. & Gileadi, O. Codon optimization can improve expression of human genes in *Escherichia coli*: A multi-gene study. *Protein Expr Purif* **59,** 94–102 (2008).
62. Raran-Kurussi, S., Cherry, S., Zhang, D. & Waugh, D. S. in (ed. Burgess-Brown, N. A.) 221–230 (Springer New York, 2017). doi:10.1007/978-1-4939-6887-9_14
63. Mierendorf, R. C., Morris, B. B., Hammer, B. & Novy, R. E. in *Nucleic Acid Protocols Handbook, The* (ed. Rapley, R.) 947–977 (Humana Press, 2000). doi:10.1385/1-59259-038-1:947
64. Wurm, D. J., Veiter, L., Ulonska, S., Eggenreich, B., Herwig, C. & Spadiut, O. The *E. coli* pET expression system revisited—mechanistic correlation between glucose and lactose uptake. *Appl Microbiol Biotechnol* **100,** 8721–8729 (2016).
65. Pao, K.-C., Wood, N. T., Knebel, A., Rafie, K., Stanley, M., Mabbitt, P. D., Sundaramoorthy, R., Hofmann, K., van Aalten, D. M. F. & Virdee, S. Activity-based E3 ligase profiling uncovers an E3 ligase with esterification activity. *Nature* **556,** 381–385 (2018).
66. Yuan, L., Lv, Z., Atkison, J. H. & Olsen, S. K. Structural insights into the mechanism and E2 specificity of the RBR E3 ubiquitin ligase HHARI. *Nat Commun* **8,** 211 (2017).
67. Spear, L. A., Huang, Y., Chen, J., Nödling, A. R., Virdee, S. & Tsai, Y.-H. Selective inhibition of cysteine-dependent enzymes by bioorthogonal tethering. *J Mol Biol* **434,** 167524 (2022).
68. Vidak, E., Javoršek, U., Vizovišek, M. & Turk, B. Cysteine cathepsins and their extracellular roles: Shaping the microenvironment. *Cells* **8,** 264 (2019).
69. Morreale, F. E. & Walden, H. Types of ubiquitin ligases. *Cell* **165,** 248-248.e1 (2016)
70. Fomenko, D. E., Marino, S. M. & Gladyshev, V. N. Functional diversity of cysteine residues in proteins and unique features of catalytic redox-active cysteines in thiol oxidoreductases. *Mol Cells* **26,** 228–235 (2008).
71. Kirisako, T., Kamei, K., Murata, S., Kato, M., Fukumoto, H., Kanie, M., Sano, S., Tokunaga, F., Tanaka, K. & Iwai, K. A ubiquitin ligase complex assembles linear polyubiquitin chains. *EMBO J* **25,** 4877–4887 (2006).
72. Lorick, K. L., Jensen, J. P. & Weissman, A. M. in *Ubiquitin and Protein Degradation, Part A* **398,** 54–68 (Academic Press, 2005)..
73. Brzovic, P. S., Lissounov, A., Christensen, D. E., Hoyt, D. W. & Klevit, R. E. A UbcH5/Ubiquitin Noncovalent Complex Is Required for Processive BRCA1-Directed Ubiquitination. *Mol Cell* **21,** 873–880 (2006).
74. Friedenson, B. The BRCA1/2 pathway prevents hematologic cancers in addition to breast and ovarian cancers. *BMC Cancer* **7,** 152 (2007).

75. Brzovic, P. S. & Klevit, R. E. Ubiquitin transfer from the E2 perspective: Why is UbcH5 so promiscuous? *Cell Cycle* **5,** 2867–2873 (2006).
76. Lewis, M. J., Vyse, S., Shields, A. M., Boeltz, S., Gordon, P. A., Spector, T. D., Lehner, P. J., Walczak, H. & Vyse, T. J. UBE2L3 polymorphism amplifies NF-κB activation and promotes plasma cell development, linking linear ubiquitination to multiple autoimmune diseases. *Am. J. Hum. Genet.* **96,** 221–234 (2015).
77. Karin, M. & Ben-Neriah, Y. Phosphorylation meets ubiquitination: The control of NF-κB activity. *Annu Rev Immunol* **18,** 621–663 (2000).
78. Tak, P. P. & Firestein, G. S. NF-κB: a key role in inflammatory diseases. *Journal of Clinical Investigation* **107,** 7–11 (2001).
79. Perkins, N. D. The diverse and complex roles of NF-κB subunits in cancer. *Nat Rev Cancer* **12,** 121–132 (2012).
80. Hayden, M. S. & Ghosh, S. NF-κB, the first quarter-century: remarkable progress and outstanding questions. *Genes Dev* **26,** 203–234 (2012).
81. Chen, L.-F. & Greene, W. C. Shaping the nuclear action of NF-κB. *Nat Rev Mol Cell Biol* **5,** 392–401 (2004).
82. Miller, S. C., Huang, R., Sakamuru, S., Shukla, S. J., Attene-Ramos, M. S., Shinn, P., Van Leer, D., Leister, W., Austin, C. P. & Xia, M. Identification of known drugs that act as inhibitors of NF-κB signaling and their mechanism of action. *Biochem Pharmacol* **79,** 1272–1280 (2010).
83. Backus, K. M., Correia, B. E., Lum, K. M., Forli, S., Horning, B. D., González-Páez, G. E., Chatterjee, S., Lanning, B. R., Teijaro, J. R., Olson, A. J., Wolan, D. W. & Cravatt, B. F. Proteome-wide covalent ligand discovery in native biological systems. *Nature* **534,** 570–574 (2016).
84. Crews, C. M. Targeting the undruggable proteome: The small molecules of my dreams. *Chem Biol* **17,** 551–555 (2010).
85. Santos, R., Ursu, O., Gaulton, A., Bento, A. P., Donadi, R. S., Bologa, C. G., Karlsson, A., Al-Lazikani, B., Hersey, A., Oprea, T. I. & Overington, J. P. A comprehensive map of molecular drug targets. *Nat Rev Drug Discov* **16,** 19–34 (2017).
86. Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. How many drug targets are there? *Nat Rev Drug Discov* **5,** 993–996 (2006).
87. Lipinski, C. & Hopkins, A. Navigating chemical space for biology and medicine. *Nature* **432,** 855–861 (2004).
88. Knight, Z. A. & Shokat, K. M. Features of selective kinase inhibitors. *Chem Biol* **12,** 621–637 (2005).
89. Schenone, M., Dančík, V., Wagner, B. K. & Clemons, P. A. Target identification and mechanism of action in chemical biology and drug discovery. *Nat Chem Biol* **9,** 232–240 (2013).
90. Suzuki, T., Asami, M., Patel, S. G., Luk, L. Y. P., Tsai, Y.-H. & Perry, A. C. F. Switchable genome editing via genetic code expansion. *Sci Rep* **8,** 10051 (2018).