DOCTOR OF PHILOSOPHY

Design and Synthesis of an E3 ligase Activity-Based Probe and its Application for the Discovery of a New Class of E3 Ligase

Pao, Kuan-Chuan

Award date:
2018

University of Dundee

Design and Synthesis of an E3 ligase Activity-Based Probe and its Application for the Discovery of a New Class of E3 Ligase

Kuan-Chuan Pao

A thesis submitted for the degree of Doctor of Philosophy

MRC Protein Phosphorylation and Ubiquitylation Unit

University of Dundee

June 2018

I. Content

II. List of Figures

III. List of Schemes and Tables

IV. Acknowledgements

This thesis is not only the story about my four-years research in MRC-PPU but also a journey for becoming a better scientist. The journey was only possible to begin with the kind help from many others. First and foremost, I would like to thank my supervisor, Dr. Satpal Virdee, who gave me the chance to carry out this amazing trip. He not only provided his full support for my PhD career but also showed me the right path to become a good scientist. I highly appreciate his kindness and willing to teach me all he knows and that he never tired of my questions. I would also like to thank the postdoc, Dr. Mattew Stanley, who helped me through all the tough tasks since my first day in the lab. I could not finish my thesis without his great help and support. I would also like to thank Paul Murphy, one of my best friends at Dundee, who helped me deal with the cultural shocks and introduced me to different communities. It was quite a good time for the past four years.

MRC-PPU is a big family where you can always find the help from everyone. I would like to thank all the PIs in the unit for their precious input into my projects, especially to Dr. Miratul Muqit who helped us to build a model to test our probes. I would also like to thank my thesis committee, Dr. Ron Hay and Dr. Sarah Ten Have, who keep me on the right track for my PhD; I would also like to thank all the fantastic facilities in the unit which provided great help for my projects, especially to Dr. Axel Knebel for the proteins preparation and Dr. Nicola Wood for cloning work; I would also like to thank Mass facility, Tissue-Culture facility, media kitchen and the lab management team.

People, are always the most precious and important factor in a great research unit. I obtained a huge amount of help from numerous people in the PPU and school. I would like to thank Yu-Chiang, Cong, Christoph, Stefan, Sid, Sam, Matthew,

Catherine, Nicola, Sambit, Roy, Emma and Julien who taught me how to do all the assays I was not familiar with in the beginning. These skillful postdocs truly helped me a lot in my PhD project. They are not only my colleagues in the unit but also my good friends who showed me how to do good person and people. When you talk about the people, you cannot miss your friends. I do not think I can obtain my PhD without all the help from my friends. I would like to thank Yosua who shared time with me during basketball and on a trip to the U.S; I would like to thank Kristin, Polly, Maithili, Karim, Luke. F, Luke. H, Katie, Sophie, Maria, Lambert, Andrew, Atul, Pragya, Kevin to share the great memory with me at Dundee.

The last but not the least, I would like to thank all my family who give me 100% support to pursue my dream. I know you are all with me in my mind even though I was alone abroad. I would also like to thank my grandpa, who always told me I would become a doctor one day even though he cannot witness this himself. Finally, I would like to thank my fiancé, Jin-Feng, who is going to be my wife this coming September. This PhD journey would not even happen without your great support and company. You showed me how good I can be in the science world and never lost faith in me. This degree is the first baby we have and I will do everything to protect it.

The path to take a PhD is full of difficulties and challenges, being humble toward the science and maintain the hunger toward the world are the attitude for success. I hope I can apply all the things I learnt from Dundee and make this world a better one.

23th, May, 2018 (Dundee)

V. Declarations

I hereby declare that the following thesis is based on the results of investigations conducted by myself, and that this thesis is of my own composition. Work other than my own is clearly indicated in the text by reference to the researchers or their publications. This thesis has not in whole or in part been previously presented for a higher degree.

Kuan-Chuan Pao

I certify that Kuan-Chuan Pao has spent the equivalent of at least nine terms in research work in the School of Life Sciences, University of Dundee and that he has fulfilled the conditions of the Ordinance General No. 14 of the University of Dundee and is qualified to submit the accompanying thesis in application for the degree of Doctor of Philosophy.

Dr. Satpal Virdee

VI. List of Publications

1. **Pao, K-C.**, Wood, N., Knebel, A., Rafie, K., Stanley, M., Mabbit, P., Sundaramoorthy, R., Hofmann, K., van Aalten, D., Virdee, S., Activity-based E3 ligase profiling uncovers an E3 ligase with esterification activity, *Nature*. (2018), 556, 381-385.

2. **Pao, K-C**., Stanley, M., Han, C., Lai, Y-C., Murphy, P, Balk, K., Wood, N., Corti, O., Corvo, J-C., Muqit, M., Virdee, S., Probes of ubiquitin E3 ligases enable systematic dissection of parkin activation, *Nat. Chem. Biol.* (2016), 12, 324-331.

3. Han, C*., **Pao, K-C***., Kazlauskaite, A., Muqit, M. M. and Virdee, S., A Versatile Strategy for the Semisynthetic Production of Ser65 Phosphorylated Ubiquitin and Its Biochemical and Structural Characterisation., *ChemBioChem*, (2015), 16, 1574-1579.

*Joint first authors

VII. Abbreviations

AVS	Activated vinylsulfide
AR	Ankyrin repeat-containing domain
GPCR	G-protein-coupled receptor
LHR	Linker helix region
PD	Parkinson’s disease
PTM	Post translation modification
RLD	RCC1-like GEF domain
TAT	Trans-activator of transcription
THPTA	Tris(3-hydroxypropyltriazolylmethyl)amine
uPA	Urokinase plaminogen activator
BME	2-mercaptoethanol
AMPK	5' AMP-activated protein kinase
TAMRA	5-carboxytetramethylrhodamine
ABPs	Activity-based probes
ABPP	Activity-based protein profiling
AOMK	Acyloxymethyl ketone
APC/C	Anaphase-promoting complex/ cyclosome
ANKIB1	Ankyrin repeat- and IBR domain-containing 1
ARIH1	Ariadne RBR E3 Ubiquitin Protein Ligase 1
APF-1	ATP-dependent proteolysis factor 1
ATG8	Autophagy-related gene 8
BH3	Bcl-2 homology 3
BSA	Bovine serum albumin
CCCP	Carbonyl cyanide *m*-cholorophenylhydrazone
CPPs	Cell-penetrating peptides
RLD	Chromosome condensation 1 (RCC)-like domain
CuAAC	Copper-catalyzed Azide-Alkyne Cycloaddition
CRLs	Cullin RING Ligases
DUBs	Deubiquitinase
DMSO	Dimethyl sulfoxide
DDR	DNA damage response
DLK	Dual leucine zipper bearing kinase
DMEM	Dulbecco's Modified Eagle Medium
ER	Endoplasmic reticulum
ERAD	Endoplasmic-reticulum associated degradation
EGFR	Epidermal growth factor receptor

ENaC	Epithelial sodium channel
FBS	Fetal bovine serum
FPs	Fluorophosphonates
GEF	Guanine nucleotide-exchange factor
HOIL	Haem-oxidized IRP2 ubiquitin ligase 1
HC	Heavy chain
HPLC	High-performance liquid chromatography
FAT10	HLA-Fadjacent transcript 10
HOIP	HOIL-1-interacting protein
HECT	Homologous to the E6AP Carboxyl Terminus
HEK293	Human embryonic kidney 293
HA-tag	Human influenza hemaglutinin tag
IBR	In-Between-RING
IGF-1	Insulin-like growth factor 1
ISG15	Interferon-stimulated gene 15
IPTG	Isopropyl thio-β-D-galactoside
IKK	IκB kinase, inhibitor of kappa B kinase
LDD	Linear ubiquitin chain determining domain
LC-MS	Liquid chromatography-mass spectrometry
LDS	Lithium dodecyl sulfate
LPA	Lysophosphatidic acid
MHC	Major histocompatibility complex
MBP-PINK1	Maltose-binding protein (MBP)-tagged PINK1
MAP3K12	MAPK kinase kinase 12
MS	Mass spectrometry
MARCH	Membrane-associated RING-CH
MEM	Minimum essential medium
MAGL	Monoacylglycerol lipase
MAGEs	Monoalkylglycerol ethers
MEF	Mouse embryonic fibroblast
MudPIT	Multidimensional Protein Identification Technology
MYCBP2	MYC binding protein 2
NDFIP	Nedd4 family-interacting protein
NEDD8	Neural precursor cell-expressed, developmentally downregulated protein 8
NMNAT2	Nicotinamide Mononucleotide Adenyl Transferase 2
NEAA	Non-essential amino acids
NS-1 lg	NS-1 immunoglobulin

NMR	Nuclear magnetic resonance
OA	oligomycin and antimycin A
OTULIN	OTU domain DUB with linear linkage specificity
OTUB2	OTU Domain-Containing Ubiquitin Aldehyde-Binding Protein 2
PHR	PAM-Highwire-Rpm-1
*Ph*PINK1	*Pediculus humanus corporis* PINK1
PMSF	Phenylmethane sulfonyl fluoride
PLC	Phospholipase C
p-Parkin	Phosphorylated-Parkin
p-Ub	Phosphorylated-ubiquitin
PARP1	Poly(ADP-ribose) polymerase 1
His-tag	Polyhistidine-tag
PTMs	Post translational modifications
PARL	Presenilins-associated rhomboid-like protein
ProcK	Propargyloxycarbonyl-L-lysine
PKC	Protein kinase C
PINK1	PTEN-induced kinase-1
PIMs	PUB interaction motifs
pUBH	p-Ub-binding helix
RBD	RAE1 binding domain
RING	Really Interesting New Gene
RCC1	Regulator of Chromosome Condensation 1 repeat domain
REP	Repressor element of parkin
RCR	RING-Cys-Relay
RBR	RING-Between-RING
RNAi	RNA interference
SHARPIN	SHANK-associated RH domain interactor
SEC-MALS	Size exclusion chromatography with multi-angle light scattering
SUMO	Small Ub-like modifier
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
MESNA	Sodium 2-mercaptoethanesulfonate
SENP1	SUMO1/Sentrin Specific Peptidase 1
TCRα	T-cell antigen receptor α-chain
LUBAC	Linear ubiquitination assembly complex
TDAE	Tosyl substituted-Doubly Activated Ene
TRPV1	Transient receptor potential vanilloid receptor 1
*Tc*PINK1	Tribolium castaneum PINK1
TEA	Triethylamine

TCEP	Tris(2-carboxyethyl)phosphine
THPTA	Tris(3-hydroxypropyltriazolylmethyl)amine
TSC	Tuberous Sclerosis Complex
TRIAD1	Two RING fingers and a DRIL
UFM1	Ub-fold modifier1
Ub	Ubiquitin
UCHL3	Ubiquitin Carboxyl-Terminal Esterase L3
UFD	Ubiquitin fusion degradation
UIM	Ubiquitin interacting motif
UBA	Ubiquitin-associated domain
Ub-Lys	Ubiquitinated lysine
Ub-Ser	Ubiquitinated serine
UBDs	Ubiquitin-binding domains
UBM	Ubiquitin-binding motif
UBLs	Ubiquitin-like proteins
Ub-SR	Ubiquitin-thioester
VME	Vinyl methyl ester
VS	Vinyl methyl sulfone
WD40	WD40/YVTN repeat-like-containing domain
ZNF	Zinc finger

VIII. Amino Acid Code

Amino acid	Three letter symbol	Single letter symbol
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V
Any amino acid	Xaa	X

Chapter 1. Ubiquitylation: Function and Regulation

1.1 *Ubiquitylation*

1.1.1 The history of the ubiquitin discovery

Ubiquitylation is an important post-translational modification that connects ubiquitin (mono- or poly-) to various substrate proteins through a covalent bond[1]. Ubiquitin was initially identified in 1975 as a heat-stable polypeptide ~8,500Da in size[2]. It was found to be able to degrade denatured globin in reticulocyte lysates in an ATP-dependent manner[3]. Ubiquitin (Ub) was so named for its ubiquitous expression within the cells of higher order eukaryotes. It is found with high sequence conservation in animals, plants and yeast. The polypeptide has later been discovered as an important component for ATP-dependent proteolysis and was named ATP-dependent proteolysis factor 1 (APF-1)[4]. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and size exclusion analysis revealed that APF-1 was covalently conjugated to numerous cellular proteins and demonstrated tolerance to heat and high-pH conditions. This was the first time that ubiquitin was considered as an essential constituent of an apparent proteolysis pathway. Hershko *et al.*[5] further demonstrated APF-1 was heavily conjugated to authentic proteolysis substrates and for first time elucidated the important components of the ubiquitin system, for example, ligases (E3 ligases) and amidases (deubiquitinase, DUBs). These breakthrough discoveries have since paved the way for the identification and further deconvolution of the role and function of many important protein players within this ATP-dependent proteolytic pathway. This has allowed scientists to further elucidate and understand important aspects of biology and physiology through the role of this important post-translational modification.

Since the 1980s, we have learned more about ubiquitin and the proteolytic system in which it plays a crucial role. The structure of ubiquitin was solved by nuclear magnetic resonance (NMR) and X-ray crystallography by several different groups[6-9] (**Fig 1.1**). Ubiquitin is composed of 76 amino acids and displays a compact β-grasp fold with a 3.5 turn α-helix, a short 3_{10}-helix and mixed β-sheet components. The conserved hydrophobic core stabilizes the globular fold and allows it to survive in non-physiological conditions, for example, extremes of pH and temperature. Conserved solvent exposed patches and a highly-flexible C-terminal tail enable ubiquitin to covalently react and non-covalently interact with a huge body of proteins to progress various cell signaling pathways[8]. Moreover, the canonical Ile44 hydrophobic patch on the Ub has been demonstrated as an important region for interacting with E2s and multiple ubiquitin binding domains (UBDs)[10].

1.1.2 Divergent functions of Ubiquitin-like proteins (UBLs)

The distinguished β-grasp structure of ubiquitin and its conserved amino acids sequence allowed the identification of further proteins which shared similar biochemical and structural properties. These Ub-like proteins (UBLs)[11], though similar from a structural point of view (sharing the ubiquitin fold), however, can be functionally categorized into two groups. The first group includes SUMO (small Ub-like modifier), ATG8 (autophagy-related gene 8), NEDD8 (neural precursor cell-expressed, developmentally downregulated protein 8), ISG15 (interferon-stimulated gene 15), FAT10 (HLA-adjacent transcript 10), and UFM1 (Ub-fold modifier1) (**Figure 1.1**). These UBLs can also be conjugated to specific substrate proteins via a ATP-dependent enzymatic cascade similar to ubiquitin. Despite the regulation of protein turnover, UBLs also control various cellular function; e.g:

SUMO-substrates form complexes with chromatin and regulate DNA repair[12]; NEDD8 regulates cullin RING (Really Interesting New Gene) E3 ligase activity by neddylation of a conserved lysine residue[13]; ATG8 is conjugated to phosphatidylethanolamine[14] (PE) and determines the size of autophagosomes[15] which is the precursor of autophagy[16].

The second group consist of proteins containing a UBL domain that are frequently found as part of multi-protein complex structures which prevent the protein itself to be conjugated to substrates. Moreover, the hydrophobic property from UBL domain allows these UBLs to mediate various protein-protein interactions. For example, the UBL domain in DNA repair protein HHR23A, which displays high affinity towards the second ubiquitin-interacting domain of S5a in the proteasome, enable HHR32A to play an important role involved in nucleotide excision repair via UBL-binding[17].

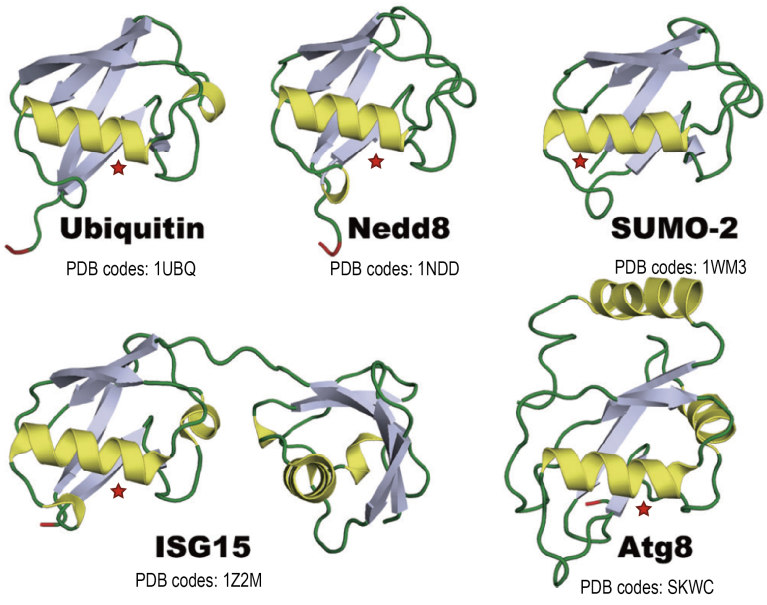


Figure 1.1 structural comparison of ubiquitin and ubiquitin-like proteins.
Ub and UBL in cartoon representation. α-helices are in yellow, β-strands are in grey and loops are in green. UBL all share the similar α-helices with Ub which make them ubiquitin-like protein. (Figure was reproduced from Ronau et al., 2016[18])

1.1.3 The enigma of ubiquitin system

Typically, ubiquitylation is the biochemical event describing the conjugation of the C-terminal Gly76 of ubiquitin to the ε-NH_2 group of a Lys residue in a substrate protein via an isopeptide bond. In addition, the seven lysine residues in ubiquitin itself can also serve as substrates, allowing the formation of different types of ubiquitin oligomers and polymers, referred to as polyubiquitin chains. Alternatively, the N-terminal methionine of ubiquitin can also form a linear peptide linkage with ubiquitin. Combinations of all eight different possible ubiquitin chain types[19] (Met1, K6, K11, K27, K29, K33, K48, K63) act as a complicated code system to regulate cell function and signaling cascades.

The complexity of the Ub code system exceeds that of the eight different chain types[20]. The simplest type is mono-ubiquitylation in which a single ubiquitin is conjugated to a single lysine ε-NH_2 upon a substrate. Mono-ubiqutination has been demonstrated to regulate multiple cellular processes such as DNA repair, gene expression [21,22]. The earliest mono-ubiquitylation event was found on histone H2A[23]. The most immediate effect of mono-ubiquitylationin on substrates is to alter protein-protein interaction which can either activate a downstream cascade[24], or prevent the interactions between substrates and their respective binding partner. This principal can serve as a regulation mechanism[25]. On the other hand, mono-ubiquitylation of multiple lysines within a substrate protein is able to trigger endocytosis[26]. However, it is the role played by oligomeric and polymeric ubiquitin chains that make this system of biological signaling both elegant and complicated.

The eight different ubiquitin chain types can be further categorized into two main groups, homotypic chains or heterotypic chains. In homotypic ubiquitin chains, the ubiquitin proteins are conjugated via the same lysine position to the Gly76 of a ubiquitin C-terminus. In heterotypic ubiquitin chains, the ubiquitin

proteins can be conjugated via various lysine positions to form a so-called branched Ub-chain. The classical and most abundant chain type is the K48-chain[27,28] which has been traditionally associated with targeting proteins to the 26S proteasome and their subsequent processing by proteolysis[29,30]. The K48 chain was first found to regulate the degradation of cell-cycle proteins but direct evidence that demonstrated the recognition of the proteasome toward K48-linked proteins was provided in 2000[31]. Consistent with the vital role in proteolysis, the K48 chain is generated by several different E3s which are known to be involved in protein degradation, for example, the conserved acidic loop on the E2 cdc34 and the cullin RING ubiquitin ligase SCF^{Cdc4} tightly regulate K48 chain formation upon the Sic1 during the cell cycle[32]; Lu *et al.* further demonstrated that cyclin-B1, the substrate of another cullin E3, the anaphase-promoting complex (APC/C) E3 ligase, is rapidly degraded by the proteasome when two lysines are both modified with a di-K48 chain (*two* ubiquitin proteins are conjugated together via the lysine 48 position on second Ub) rather than with a single tetra-K48 chain on one lysine[33]. E2 and E3 enzymes transfer ubiquitin to substrates and will be discussed later.

Instead of playing a similar role to K48 chains in regulating proteolysis, K63 chains interact with multiple proteins and have been shown to regulate various signaling pathways[20]. Li *et al.* first reported that the RING E3 ligase TRAF6 can activate IKK (IκB kinase, inhibitor of kappa B kinase) by assembling a K63 polyubiquitin chain[34]. Other groups also demonstrated that K63 chains are involved in DNA damage repair[35,36], oxidative stress responses[37], anti-viral immune responses[38,39] and can even mediate mRNA splicing and hence regulate protein translation[40,41]. Both K48 and K63 ubiquitin chains are the most well studied and are abundant chain types in mammalian cells. The remaining five chain types are classed as atypical poly-Ub chains.

A series of proteomic analyses revealed enormous roles for atypical ubiquitin chains and additional codes have been added[27,29,42]. The Met-1 chain (M1) is a unique chain type because of its special linear structure[19,43] and has been demonstrated thus far to be made by a single specific E3 ligase complex[44] (The linear ubiquitylation assembly complex**,** LUBAC[45,46]) Furthermore, it was discovered that the coordination of LUBAC and M1 chain formation are associated with NF-κB activation[45,47-49]. An interesting feature of the M1-chain is the difficulty to detect M1-linked conjugates without activation of the inflammatory signaling cascade[47]. In addition to the growing body of work in the area on M1 chains in the NF-κB signaling pathway, the regulation in pathogen infection[50,51] and cell death[52] have also been shown to be mediated by LUBAC-associated M1 chains.

The roles of K11 chains in cells are poorly understood. One exception is that K11-ubiquitylation has been shown to be increased during mitosis and G1-phase. Depletion of APC/C E3 largely prevents K11 chain formation indicating that it is a major source of K11-Ub chains[53]. Although the K11 chains have been linked to a proteolytic role, the homotypic K11 chain modification is not sufficient to trigger proteasome-dependent protein degradation of APC/C susbtrates. Mixed K48/K11 heterotypic ubiquitin chains have been demonstrated to accelerate proteolysis compared to K11 chains alone[54]. K11 chains have also been found to mediate the HIF-1α-dependent pathway,[55] and immune response after viral infection[56].

In contrast to K48, K63, K11 and M1 chains, the remaining four chain types (K6, K27, K29 and K33) do not seem to be made by a highly Ub-chain specific enzymatic system which further complicates their studies. However, certain regulatory roles associated with these chains have been presented. Quantitative proteomics has indicated that K6- and K33-Ub chains are specifically increased after DNA damage[57]. Moreover, the K6 chains have also been verified in their

ability to mediate proteasome-dependent protein degradation[58,59] and more recently have been shown to play a role in the response to mitochondrial damage or mitochondria-specific autophagy[60,61] (mitophagy). Michel *et al.* developed a K6-linked affimer with high selectivity and revealed that upon mitochondrial depolarization, the outer mitochondrial membrane protein mitofusin-2, is modified with K6 chains in a HUWE1 E3-dependent manner[62]. Little is known about K27-linked chains so far and the E3 ligases responsible have only been revealed recently[63,64]. Several reports indicate that K27 chains can regulate neuronal protection by branching with K29 chains[65], activate DNA damage response (DDR)[64] and control autoimmunity through regulation of the NF-κB activity in dendritic cells[66]. K29 chains were initially shown to regulate the ubiquitin fusion degradation (UFD) pathway in yeast by branching with K48 chains[67,68]. Two different groups established methods to prepare bulk amounts of K29 chains and solved the crystal structure[69,70]. Furthermore, K29 chains are reported to mediate lysosomal degradation in mammalian cells[71,72]. Al-Hakim *et al.* also demonstrated that K29 chains accompanied with K33 chains can modulate 5' AMP-activated protein kinase (AMPK) activity[73]. Additionally, several groups have demonstrated that K33 chains can mediate cell trafficking[74], T-cell regulation[75,76] and TBK1 degradation[77].

The discovery and study of the various roles played by all eight ubiquitin chain types in the regulation of various cellular functions still requires further maturation and development in order to fully understand how the ubiquitin code is 'read' by the aforementioned ubiquitin-binding domains (UBDs) and how the code is 'written' in an orchestrated, context dependent manner by the various E2/E3 pairs. Moreover, the majority of these studies are based on either protein overexpression systems or *in-vitro* biochemical strategies, thus the biological and

physiological relevance of these studies must be interpreted with caution. Recently, another interesting and important research area in the ubiquitin field was uncovered, the direct post translation modification (PTM) of ubiquitin[78,79], this increases the complexity of the ubiquitin code to a higher level. The most striking and exciting example is phosphorylation on Ser65 in ubiquitin[80-82] (Ser65-p-Ub). This modification serves not only to change the well-known structure of ubiquitin but has also been demonstrated to affect the preference of DUBs or UBDs toward these modified-ubiquitin moieties[83-85].

1.2 *Regulation of the ubiquitin conjugation cascade*

There are three major steps involved in the ubiquitylation of target proteins[86,87] (**Figure 1.2**). First, ubiquitin is activated by an ATP-dependent ubiquitin-activating enzyme (E1), resulting in the reaction of an activated Ub-resulting in the formation of a thioester bond between the C-terminus of ubiquitin and the active site cysteine in E1s[88] (E1~Ub, ~ refers to labile thioester bond). E1~Ub will then recruit an ubiquitin conjugating enzyme (E2) and promote the transfer of ubiquitin to the E2 (E2~Ub) forming a thioester-linked intermediate[89]. The E2~Ub is then recruited by E3 ubiquitin ligases that mediate substrate specificity[90]. E3s belonging to the RING (Really Interesting New Gene) family can transfer ubiquitin from E2~Ub to substrates directly[91]; however, E3 ligases belonging to the HECT (Homologous to the E6AP Carboxyl Terminus)[92] and RBR (RING-Between-RING) families[87] form a covalent thioester intermediate (E3~Ub) through E2-E3 transthiolation and mediate transfer of ubiquitin to the lysine residue on substrates. The enzyme trio (E1, E2 and E3) can be considered as a ubiquitylation '*writers*'. The ubiquitin signal can be decoded by downstream effector proteins. These proteins contain the ubiquitin binding domains (UBDs)[93,94] considered as the '*readers*'; the Ub/UBLs

specific proteases which remove the Ub or Ub chains from substrate proteins are the '*erasers*', such as deubiquitinases (DUBs)[95,96]. The detailed mechanism and function of each enzyme will be discussed below (vide infra).

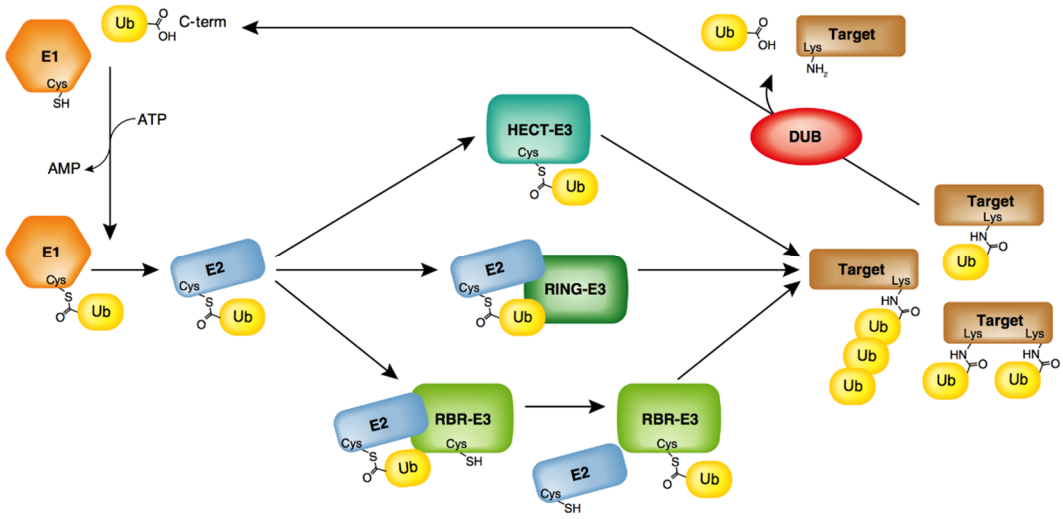


Figure 1.2 The proposed mechanism of ubiquitin transfer by E1, E2 and E3 enzymes. The ubiquitin transfer cascade in cartoon. Three distinct E3 ligase families perform differently in transferring Ub to substrates. (Figure was reproduced as a whole from Smit *et al.*, 2014[87])

Using the same analogy, if Ub or Ub chains are the 'words' in the ubiquitin code, the next question is how do cells 'write' the sentence to allow proteins to communicate with each other. The workflow of the aforementioned enzyme trio, 'E1-E2-E3', perform the key reactions to accomplish ubiquitin modification. It has been demonstrated by multiple groups that the disruption of this cascade or mutation in one of the enzymes will lead to numerous diseases[97-100]. Understanding the detailed molecular mechanism of how these enzymes regulate ubiquitin transfer will undoubtedly expand the number of therapeutic targets for drug discovery against disease states where ubiquitylation or 'E1-E2-E3' is of vital importance. A key question that has puzzled the field for a long time is how the

activity of these key enzymes is regulated? The following is a brief review of this important enzyme cascade in the ubiquitin system with subsequent focus on the final component, the 'E3 ligase'.

1.2.1 Initiation by E1 activating enzymes

The ubiquitin proteolysis system was discovered to be an ATP-dependent process[4,5]. ATP is an important component for triggering ubiquitin activating enzymes, E1s, namely UBA1 and UBA6 in vertebrates[101,102]. UBA1 is specific for ubiquitin transfer and UBA6 can transfer both Ub and FAT10[103] to their cognate E2s. Other UBL E1s also share mechanistic but not sequence similarities in activating UBL cascade. All the E1s are in the range of approximately 120 kDa and feature three major important domains to transfer ubiquitin with high fidelity: pseudodimeric adenylation domain, cysteine domain with catalytic cysteine residue, and ubiquitin-fold domain (UFD) that interacts with downstream E2s[104]. E1s associate with Mg^{2+}•ATP and Ub/UBL through their adenylation domain and adenylate the C-terminal glycine of Ub/UBL to form Ub-AMP and pyrophosphate. Previous structural studies have shown that the E1 is in an open conformation and that the catalytic cysteine residue is far from the adenylation site at this step[105-107]. After the Ub-AMP has formed, pyrophosphate will be released from the adenylation site. There is a significant conformational change (130° rotation) in this domain to allow the Cys domain to bring the E1s catalytic cysteine in close proximity of the C-terminal Ub/UBL[108]. This conformational transition is extremely important for ubiquitin transfer since it brings the E1s catalytic cysteine residue 35Å closer to the Ub/UBL C-terminus. After the rotation, the catalytic cysteine attacks the Ub-AMP and forms a transient thioester bond between the Ub Gly76 and the Cys sulphydryl of E1. Once the thioester bond has formed, the E1s Cys

domain rotates back to an open configuration and frees up the adenylation domain for the next round of adenylation[109]. Early studies into E1 activity revealed that a doubly loaded E1 was better in recruiting E2s, although the mechanism was unknown[110]. The transthiolation mechanism of Ub transfer between E1/E2 has only revealed over the last few years[88,104]. The mystery before these studies was how the E1s catalytic cysteine positioned itself in close enough proximity with the E2s catalytic cysteine. According to other UBL E1 structures, the doubly loaded NEDD8 E1/E2 structure and doubly loaded SUMO E1/E2 structure all showed that the two catalytic cysteines are 15-20 Å apart from each other[109,111] thus incompatible with transthiolation. Olsen *et al.* compared the structure of 'Ub E1/Ub/ATP,Mg' alone with 'Ub E1/Ub/ATP,Mg associated with Ub E2 Ubc4' and revealed that a certain rotation of UFD domain on E1 is needed for E1-E2 interaction[104]. In addition, the catalytic cysteine on E1 is hindered by a loop located in the Cys domain but this Cys cap is disordered upon contact with an E2. The E1-E2 complex crystal structure revealed multiple contacts between the E2 and the E1s UFD and Cys domains, with nearly 40% of E2 surface area buried in these domains[104]. However, due to the transient nature of the thioester bond, the Ub~E1/E2 structure remains elusive which further emphasizes the need for a tool to determine the transthiolation activity of E1 enzymes.

Most knowledge related to E1 activity regulation has relied heavily on molecular biology studies, especially from a structural point of view. We still know little about how E1 activity is regulated *in vivo*. Several groups demonstrated that mutation or depletion of E1 in cells would affect the cell cycle and proliferation[112-115]. UBA1 exists in most of mammalian cell lines and appears to be constitutivly active [102]. Analysis of the level of Ub~E2s indicated that the E1s are activated and sufficient to trigger the ubiquitylation cascade. Therefore, E2-E3 and E3-substrate

interactions are vital steps in controlling ubiquitin modification. This thesis focuses on developing the tools to study E2-E3 interaction and monitoring E3 ligase activity.

1.2.2 E2s are cascade intermediates

1.2.2.1 General features of E2s

There are nearly 40 Ub/UBL-conjugating E2 enzymes (E2s) in the human genome and most of their structures have been solved already. Transthiolation and aminolysis are the two major reactions which E2s perform. As the description in the title of Stweart *et al.* review[116], E2s are not only middle men passing the message from E1s to E3s. They also play the key roles in transferring ubiquitin to substrate proteins, building the correct ubiquitin chains and regulating the function of other proteins beyond the canonical ubiquitin mechanism[117,118]. There is a conserved core structure among all E2s which contains roughly 150 amino acids, termed UBC domain. This domain adopts an α/β-fold composed of four α-helices and a four stranded antiparallel β-sheet. Although UBC domain features a high similarity 'core' structure in E2s, certain E2s also have N- or C- terminal extensions. The function of these is poorly understood but in some cases they facilitate E2s interacting with the ~700 different E3s[116]. As stated previously, E1s appear to be constitutively active in proliferating cells and produce a steady turnover of Ub loaded E2s. Downstream E2~Ub activity is then tuned based on the E3 and/or substrate that is encountered in a given pathway. The interaction within E3s will be covered in later sections.

One of the E2s primary roles is to transfer thioester-linked Ub to the lysine residues on substrates in a RING E3 mediated manner. However, E3-independent discharge onto high concentration of free lysine is observed[89]. Therefore, monitoring the discharge of ubiquitin from E2s to free lysine can indicate the

intrinsic activity of E2s without being confused by other components used in different activity assays[89]. Although the cysteine-dependent ubiquitylation mediated by E2, Pex4p[119,120], was not discovered by the free amino acids assay. Wenzel *et al.* obtained the exciting data that UBCH7, which supports activity of the presumed RING E3 HHARI cannot discharge onto lysine. This raised the possibility that HHARI belonged to a mechanistically novel clan of E3 that underwent transthiolation with E2~Ub[121]. The loaded UBCH7 (UBE2L3)~Ub performs exclusive reactivity towards cysteine rather than lysine which redefines the function of this E2. UBCH7 is now considered a functional E2 for HECT and RBR E3s which can transfer Ub from itself to the catalytic cysteine on E3s. UBCH7 distinguishes itself from other E2s that genuinely target lysine as their only substrate. The E2 UBE2J2 has also been reported to undergo non-lysine ubiquitylation with RING E3 ligase mK3, instead modifying serine/threonine residues[122].

1.2.2.2 The distinct 'closed' and 'open' E2~Ub conformations

Early crystal studies showed that a canonical binding surface on UBC domains direct E2s toward many different E3s. Multiple residues on helix 1 and loop 4 and 7 mediate the E2s binding with HECT and RING E3s[123,124]. If the E3 binding region is similar on E2 enzymes, how is ubiquitin selectively transferred to lysine or cysteine? The highly flexible C-terminal region (72-76, RLRGG) on Ub offers the ability to swing the Ub molecule while conjugated with E2s. Several groups performed various experiments to demonstrate that Ub~E2 moves from an 'open state' to 'closed state' upon interaction with RING E3 ligases[125-127]. The Ub~E2 conjugates react significantly faster toward lysine residues when accompanied by RING E3 ligase[121]. The E2~Ub:RING E3 crystal structures further revealed that the Ub is folded back toward the E2 to expose its C-terminal tail, making it favorable

for substrate lysine attack[125]. Moreover, the D87 and D117 residues on UBCH5 E2s, which have been indicated to be essential for the deprotonation of the poorly nucleophilic ε-NH2 group on the substrate lysine residue are missing in the HECT/RBR specific E2, UBCH7. Unlike the RING E3 ligases, the HECT or RBR E3s do not promote an E2~Ub closed conformation. The crystal structures of E2~Ub with Nedd4-2 (HECT E3) and HOIP (RBR E3) proved this theory[128,129]. Moreover, the binding with HECT or RBR prevents the E2~Ub closed conformation, which also blocks its intrinsic activity towards free lysines[118].

1.2.2.3 The 'backside effect' sterically activates E2s

The E2-E3 interaction is not the only intermolecular interaction to regulate E2s activity. Recently, the backside attachments of ubiquitin to E2s has been shown to increase E2~Ub activity for transferring ubiquitin to acceptors[130]. This area is located opposite from the catalytic site and can be contacted by free Ub, the Ub on E2~Ub, mono or poly-Ub chains on E3s or substrates. There were some early studies on E2/E3 allosteric backside binding effects before the crystal structures were reported. These reports demonstrated low affinity of Ub toward E2~Ub but still showed the increase in activity of ubiquitin transfer[131,132]. Biochemical and structural data have revealed that Ser22 on UBCH5 and the Ile44 hydrophobic patch on Ub are the key factors in this Ub/E2~Ub interaction. However, the opposite result was also found by another E2, UBE2E3, for which the backside effect governs the ability to make polyUb chains[133]. Autoubiquitylation of lysine residues on E2 itself has also been showed to have a regulatory effect on E2 activity. Instead of enhancing the activity of E2s, these modifications inhibit the ubiquitin transfer from E2 *in vitro*[134,135]. On the whole, considering that there are more than 600 RING E3 ligases needed to work with 40 different E2s to either make free

ubiquitin chains or conjugate mono- or poly- ubiquitin to substrates, the regulation of E2s activity may need to be controlled. However, the mechanism for E2 regulation is poorly characterized.

1.2.3 E3 confers substrate specificity

The aim for this thesis is to study the mechanism of E3 ligases whose activity is clearly regulated in many cases. E3 transthiolation activity further plays a key role in ubiquitin transfer and regulation can occur at this step. However, a good mechanism-based tool to study the tranthiolation activity between E2s and E3s is missing. In order to develop a good tool to help us to dissect E3 activity, understanding how E3 ligases generally work is the first step. E3s can be categorized into three subfamilies based on their distinct structure and mechanism of how they catalyze ubiquitin transfer to substrates. In general, RING E3s interact with E2~Ub and regulate the ubiquitin transfer **indirectly**. On the other hand, transthiolation between E2~Ub and HECT and RBR E3s allow these E3s to transfer ubiquitin to substrates **directly** based on their activity.

1.3 *The E3 ligases are more than a messenger but a creator*

1.3.1 RING E3 ligases

1.3.1.1 The RING domain binds E2s

RING E3 ligases were first discovered with the RING motif from Ring1 protein (really interesting new gene 1)[136]. The conserved sequence in canonical RING E3s is Cys-X_2-Cys-$X_{(9\text{-}39)}$-$Cys_{(1\text{-}3)}$-His-$X_{(2\text{-}3)}$-Cys-X_2-Cys-$X_{(4\text{-}48)}$-Cys-X_2-Cys (X can be any amino acid) (**Figure 1.3**) and adopts a cross-braced arrangement with Zn ions[91]. Structural analysis reveals that this motif resides in the core of RING E3s and the

association with Zn^{2+} yields a globular and rigid platform for recruiting E2s[137]. Considering its essential role in RING E3 composition, mutations in these Zn^{2+}-binding domains usually abolishes RING E3 activity. Another exception is U-Box proteins which adopt a RING-like structure for E2 interaction and transfer ubiquitin with the same mechanism as RING E3s but do not coordinate with Zn^{2+}-ions[138,139]. Interestingly, RING E3s were not discovered as ubiquitin E3 ligases at first but were identified through involvment in regulating DNA damage. Many discoveries of RING E3s were related to DNA research[91]. The E3 activity was later demonstrated with histone ubiquitylation promoted by Rad18.

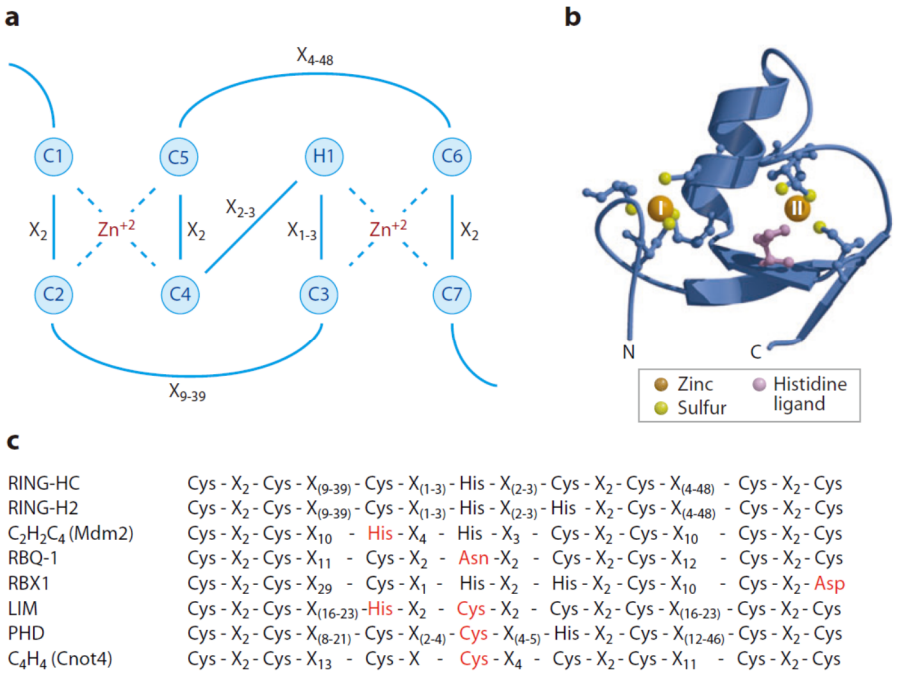


Figure 1.3 RING Finger domain adopts a highly conserved cross-braced arrangement.
(a) The sequence of RING-HC domain with Zn^{2+}. The first cysteine coordinated with Zn^{2+} is ordered C1 and so on. The H1 refers to the histidine binding with Zn^{2+}. (b) RING domain from c-Cbl in ribbon diagram with the N as N-terminus and C as C-terminus. (c) The alignment of different RING-like sequences with canonical RING domain. LIM and PHD share high similarity in sequence with RING E3 but do not retain E3 activity. (Figure was reproduced from Deshaies *et al.*, 2009[91])

1.3.1.2 RING E3s work collectively

According to bioinformatic analyses there are more than 600 RING E3s in the human genome and nearly half of them are not yet well understood[140]. The interesting and also difficult part to studying RING E3s is their diversity of action (**Figure 1.4**). The single functional polypeptide RING E3s (RING monomer) can be found in CBL[123], CNOT4[141] and RNF38[130]. However, the most-well known feature of RING E3s is their ability to form homodimers or heterodimers. For example, cIAP[142], RNF4[143], TRAF6[144], BIRC7[126], CHIP[145] and Prp19[146,147] are all be found as homodimeric RING E3s in cells. These homodimeric RING E3s have intrinsic activity to react with Ub~E2 but the mechanisms of dimer formation are quite different. The hydrophobic patches and N-terminal hairpin on CHIP RING E3 mediate the U-box/U-box intermolecular reaction and substrate binding domain[148] while others simply dimerize via RING domain interaction. In contrast, some E3s such as BRCA1-BARD1, BMI1-RING1B and MDM2-MDMX are known to exist in heterodimeric form. BARD1 from this category does not have intrinsic activity to interact with E2s but is able to enhance the stability of a dimer complex or substrate recognition *in vivo*[149,150]. The domains on RING E3s responsible for dimerization are also varied. In some RING E3s, the sequence outside the RING domain plays an important role for dimerization, as seen in BRCA1-BARD1 and RING1-Bmi1 complexes. The others can form a dimer via their RING domains which are generally located in the extreme C-terminus of these proteins, for example, MDM2-MDMX and RNF4 homodimer[143,151]. The complexity of RING E3s is more than just the dimer. The Cullin RING Ligase (CRLs) subfamily including the anaphase promoting complex/cyclosome (APC/C) are reported as huge complexes with E3 activity. CRLs are built up of multiple proteins including RING E3 (in most cases RBX1 or RBX2), a cullin protein (CUL1, CUL2, CUL3, CUL4A/4B, CUL5 or CUL7) and a substrate

receptor[152,153]. Due to the variety of the substrate receptors and the combination of different cullin proteins, the substrate number for CRLs is enormous. Due to the high variety of different RING E3s forms, the detection of RING E3s activity has proven to be challenging.

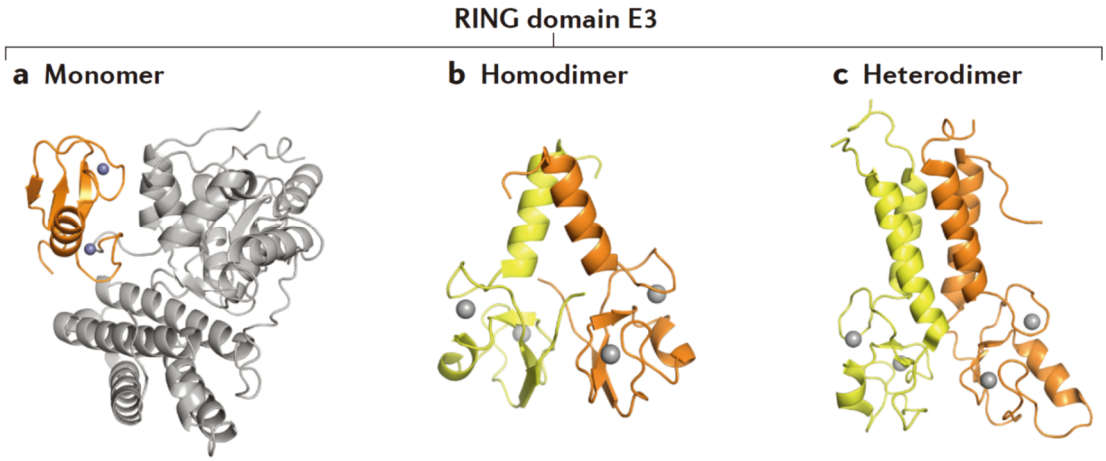


Figure 1.4 RING E3 ligases work in various form from monomer to dimer.
The crystal structure of different RING E3 in either **a**: monomer, CBL RING, PDB:1FBV, **b**: Homodimer, cIAP2 RING E3, PDB:3EB5, **c**: Heterodimer, BRCA1(orange)/BARD1(yellow) RING E3, PDB:1JM7. The grey sphere represents a Zn ion that coordinates with RING domain. The homodimer adopts a different structural architecture compared to the hererodimer. (Figure was reproduced from Buetow *et al.*, 2016[137])

1.3.1.3 Regulation of RING E3s

The most important function of RING E3s is to mediate ubiquitin transfer from E2~Ub to substrates. The mechanism of how RING E3s distinguish the loaded E2~Ub from unloaded E2 and how the enhanced reactivity of E2~Ub towards substrate lysine is promoted via RING E3s were only be reported recently. RING E3s used to be considered as only the cofactors for helping E2 to transfer ubiquitin to the correct substrate residues. The early reported crystal data showed little differences between E2s and E2:RING complexes[123,154,155]. Moreover, the catalytic cysteine on E2 is also far from the E2 binding site on RING E3s. Considering the

highly dynamic nature of E2~Ub[156] and the enhancement of ubiquitin transfer from E2~Ub while RING E3s are presented , it was suggested that there is an unknown allosteric effect between RING E3s and E2s. The main shortcoming in earlier structures was the lack of ubiquitin on E2s which is important for E2:E3 interaction[157,158].

The structure of the thioester-linked ubiquitin on E2 in complex with the dimeric RNF4 E3 demonstrated the dynamic nature of the so called 'open' conformation of this conjugate[132]. The significant advance in our understanding of RING E3 mechanism is that three landmark papers reported structures of E2-Ub:RING E3s in 2012 and provided the molecular detail of their action[125,126,159]. In contrast to the previous E2~Ub structures which were in open conformations, the 'closed' conformation was discovered by all three groups (**Figure 1.5**). The Ile44 hydrophobic patch on the donor ubiquitin (the ubiquitin on E2 referred to as the donor, the ubiquitin as a substrate containing the target lysine is referred to as the acceptor) interacts with α helixes on the E2 to stabilize the ubiquitin in a folded back position and allows the C-terminal tail to be exposed to the substrate in the E2 active site. RING E3s prime the ubiquitin transfer from E2s by keeping the flexible ubiquitin in an active mode which allows the lysine on substrates to attack the thioester bond more easily. Consistent with the mechanism proposed, several E2:Ub interface mutations showed a decrease in ubiquitin transfer and disruption of the closed E2~Ub conformation[159]. Moreover, by forming a hydrogen bond with Gln92 on E2 and Arg72 on ubiquitin, the Arg residue in loop 2 which is on the RING domain plays a vital role in priming the ubiquitin transfer[157,160,161]. These findings greatly expand our knowledge about the mechanism of ubiquitin being transferred by RING E3s. However, it has been reported that several E2s which work with RING E3s are also capabled to transfer ubiquitin in an E3-independent manner[157,160,161].

Further study about the detail mechanism of how these two enzymes mediate ubiquitin transfer is still needed.

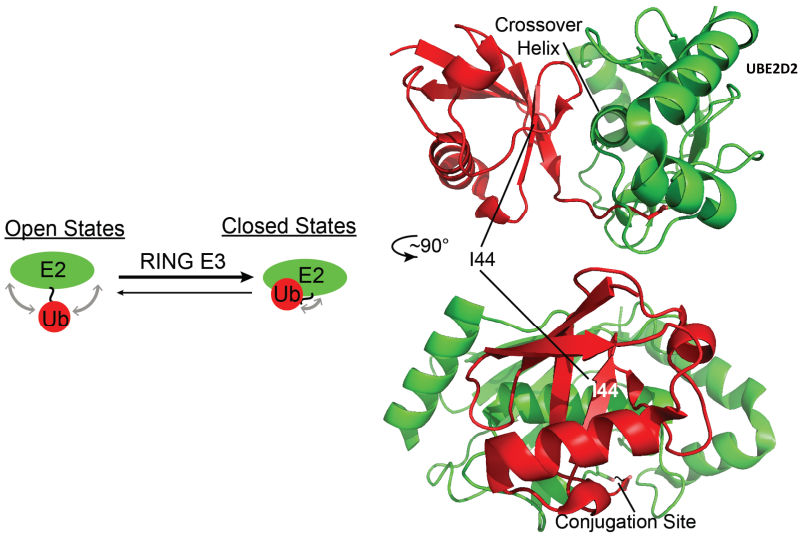


Figure 1.5 E2~Ub adopts a close conformation whilst binding with RING E3.
The highly flexible tail on Ub makes E2~Ub highly dynamic. However, once it has bound with RING E3 ligase, it adopts a 'closed' state for incoming lysine attack from substrate proteins. The UBE2D2~Ub oxyester conjugate forms a closed conformation upon binding with a RING E3 BIRC7 (not shown) (PDBID:4AUQ). (Figure was reproduced from Stewart *et al.*, 2016[116])

The concept that RING E3s prime the E2~Ub to transfer ubiquitin toward substrate is well accepted, but how these RING E3s are regulated or activated is still not fully understood. One of the mechanisms to regulate RING E3s is through dimer or complex formation, as discussed above. The post translational modification on RING E3s has also currently been reported. The Cbl RING E3 family (c-Cbl, Cbl-B and Cbl-C) is reported to be regulated by phosphorylation. Phosphorylation of a highly conserved tyrosine residue at position 371 (Tyr371) promotes a conformational change of the linker helix region (LHR) and allows the RING-bound E2 to present the catalytic site towards the substrate[162]. This

phosphorylation event is also involved in stabilizing the closed E2~Ub conformation for priming ubiquitin transfer[163]. Despite autoubiquitylation on RING E3s enhancing their ligase activity[131,164], the post-translational modification on the subunit of CRL complex can also regulate ligase activity. All the cullin proteins have been reported to be modified by NEDD8, one of the UBL proteins, to regulate CRL activity[165,166]. The neddyation of cullin proteins induces a conformational change to potentially free up the RBX1 RING domain and allow it to bring E2~Ub in proximity to substrates[154,167]. E2~Ub levels are in sufficient supply within the cells which makes regulation of RING E3 activity extremely important considering the huge number of RING E3s (nearly 600). A good tool to measure the activity of RING E3s will be a big step forward to accelerate research on this topic.

1.3.2 HECT E3 ligases

1.3.2.1 HECT E3s

HECT E3 ligases were the first type of E3 ligase to be discovered[168] and in humans there are ~28 enzymes ranging between 80 kDa to more than 500 kDa in size[169] (**Figure 1.6**). The mechanism for ubiquitin transfer via HECT E3s is distinct to that of RING E3s. The ubiquitin transfer from E2 to catalytic cysteine on HECT is via a transthiolation reaction to a specific cysteine residing within the HECT E3 and then subsequent Ub transfer to a lysine side chain on a substrate via an aminolysis reaction. These two distinct reactions are responsible for one single ubiquitin transfer and as such, makes the study and deconvolution of HECT E3 activity and regulation challenging.

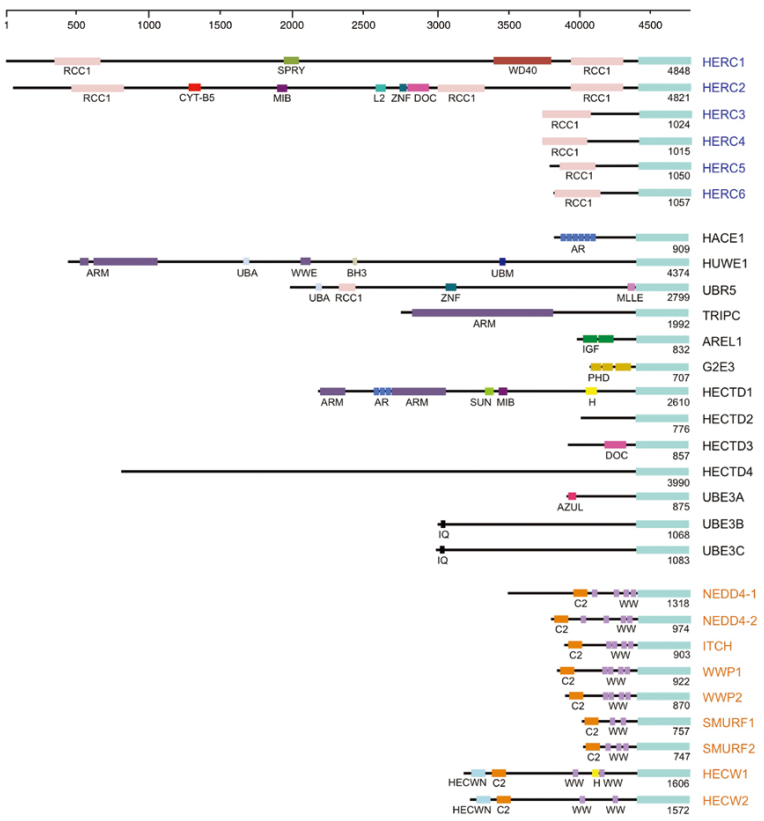


Figure 1.6 Overview of three subtypes of HECT E3s.

The N- and C- terminal extensions from the HECT domain not only regulate HECT E3 activity but also largely expand the substrate numbers for HECT E3s. **Abbreviation**: RCC1, Regulator of Chromosome Condensation 1 repeat domain; SPRY, B30.2/SPRY (SPIA and RYanodine Receptor) domain; WD40, WD40/YVTN repeat-like-containing domain; CYT-B5, Cytochrome B5-like heme/steroid binding domain; MIB, MIB-HERC2 domain; L2, ribosomal protein L2 domain; ZNF, zinc finger; DOC, APC10/DOC domain; AR, ankyrin repeat-containing domain; ARM, armadillo-type fold domain (overlaps in TRIPC with a predicted WWE domain); UBA, ubiquitin-associated domain; WWE, WWE domain; BH3, BCL-2 homology region 3 domain; UBM, ubiquitin-binding motif; MLLE, Mademoiselle/PABC domain; IGF, immunoglobulin-like fold; PHD, PHD-type zinc finger; SUN, SAD1/UNC domain (overlaps in HECTD1 with predicted galactose-binding-like domain); AZUL, AZUL domain/ N-terminal zinc-binding domain; IQ, IQ domain/ EF-hand binding site; C2, C2 domain; WW, WW domain (Figure was reproduced as a whole from Lorenz *et al.*, 2018[170])

There are three subtypes of HECT E3s based on their structural diversity within the substrate recognition domain. The Nedd4/Nedd4-like HECT E3 subfamily contain a membrane-targeting C2 domain tethered to several WW

domains N-terminal to the HECT domain[171,172]. In contrast, six E3s (HERC1 to 6) within HERC subfamily contain RCC1-like domains which are N-terminal to the HECT domain. The remaining HECT E3s are also predicted to bear varied protein architectures before the HECT domain and are likely to be important for mediating protein-protein interactions (**Fig 1.6**).

Understanding how HECT E3s work on a molecular level is currently a difficult task to achieve as there are no available full length structures of any HECT E3s (either by crystallography or NMR analysis). Also, many regions within HECT E3s are predicted to be disordered in nature, thereby hampering efforts towards complete structural characterization. However, by combining the growing body of biochemical data relating to HECT E3 activity and by solving and analyzing the structures of HECT- domain only or appropriately truncated HECT E3s, we will be able to better understand HECT E3 ligase activity, regulation and mechanism.

1.3.2.2 Nedd4 HECT E3s

The Nedd4 family is the most well studied category which contains nine members and all possess C2 and WW domains. The C2 domain was discovered in several different proteins including protein kinase C (PKC)[173], cytosolic phospholipase A_2[174] and synaptotagmin[175] and was first identified as a Ca^{2+}-dependent phospholipid binding domain in PKC. It contains approximately 130 amino acids and folds into two four-stranded β-sheets. Multiple conserved aspartate amino acids reside in the loop regions between these β-strands, bind Ca^{2+} and mediate the C2 interaction[176,177]. Plant *et al.* reported that the C2 domain in Nedd4 is responsible for localizing the Nedd4 from the cytosol to specific cellular fractions such as membranes in response to Ca^{2+} treatment[178]. Experiments demonstrated that the C2 domain is able to bind not only Ca^{2+} but also phospholipids and even other

proteins. The other domain present in Nedd4 is the WW domains (2-4 repeats N-terminal to HECT domain), which has been reported as the substrate recognition site[179,180]. The WW domain is so-named as there are two conserved tryptophan residues within its structure that are commonly spaced 20-22 residues apart. The domain adopts a compact module with three-stranded anti-parallel β-sheets that forms a hydrophobic ligand binding groove. It is found to bind exclusively to proline-rich motifs such as PPxY[181] (x can be any amino acid), PPLP[182], and phosphoserine/threonine (pSer/pThr) that are followed by a proline residue[183]. WW domains fall into four classes based on the recognized motif they interact with. It has been found that WW domain in Nedd4 E3s particularly favors PY motifs (PPxY). However, Lu *et al.*[183] also demonstrated the possibility of WW domain to recognize pSer/pThr which broadens the substrate members of the Nedd4 HECT family. Importantly, some WW domains reside within the same E3 and may function with entirely different substrates. As such, these proteins architectures would enable these enzymes to act as an interaction hub, enabling interaction with various adaptor proteins and substrates for specific signaling pathways[184].

1.3.2.3 HERC-family and miscellaneous HECT E3s

HERC E3s are a family of large proteins which range from 120 kDa to 500 kDa. The characteristic domain of HERC E3s is the chromosome condensation 1 (RCC)-like domain (RLD). The RLD is composed of a seven-bladed β-propeller fold which itself is a 51-68 amino acid repeat unit. The RLD acts as a guanine nucleotide-exchange factor (GEF) that interacts with GTPases or interacts with chromatin via H2A and H2B. However, only in the case of the large HERC E3s (HERC1 and HERC2), has the above function been demonstrated. The small HERC E3s do not retain GEF activity. Convincing experimental data concerning the function and structure of HERC E3s

is currently absent from literature due to their often large size and to the presence of multiple domains of hitherto unknown function. The miscellaneous HECT E3s, have divergent N-terminal domains and display preferences toward interacting partners not found in previously described HECT E3 families[185]. For example, the UBR5(EDD) contains a UBR 1-like zinc finger and PABC domain which allows HECT E3 mediated regulation of the activity of the PABP transcription factor[186]; HUWE1, has a WWE and a BH3 domain N-terminal to the HECT domain.

1.3.2.4 General mechanism of Ub transfer by HECT E3s

Ubiquitin transfer by HECT E3s involves two sequential steps. Firstly transthiolation occurs to transfer thioester-linked ubiquitin from E2s to HECT E3s and then secondly, aminolysis, enabling the Ub-modification of substrate proteins. Initial structural studies of E2:HECT protein complexes indicated the distance between the E2 catalytic cysteine and the catalytic cysteine of HECT E3 are too far from each other to allow cysteine-cysteine juxtaposition and subsequent transthiolation to occur. Significant conformational changes must be undertaken to allow a catalytically competent protein complex to be formed for transthiolation and subsequent ubiquitin transfer. It has recently been demonstrated that the structural movement of HECT E3s is actually regulated by two conserved lobes and a linker region within the HECT domain. The HECT domain is a bi-lobed structure that contains N-lobe which is essential for E2 binding and a C-lobe harboring the catalytic cysteine[187]. The highly flexible hinge region which connects these two lobes allows the C-lobe to adopt different conformations when interacting with substrates. The other conserved region among all HECT domains is the '-4Phe' region which is located towards the C-terminus and is proposed to be an important region for HECT-mediated isopeptide bond formation[188]. This region is found to

only be structurally ordered when co-crystalized with ubiquitin and a substrate peptide as they enable stabilization of the N-lobe[189]. Mutation of the C-lobe Phe806 on Rsp5 dramatically impairs the polyubiquitylation process (**Figure 1.7**).

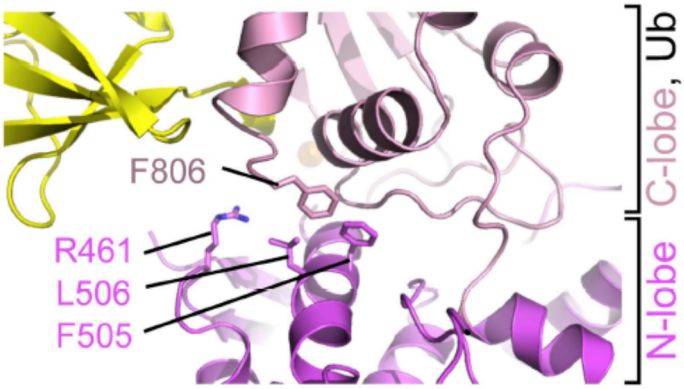


Figure 1.7 The close-up view of the crystal structure for '-4Phe' position between N- and C-lobe.

The structure of the interface between $Rsp5^{WW3\text{-}HECT}$, Ub and $Sna3^{C}$ indicated how F806 interacts with residues on N-lobe (R461, L506 and F505). (Figure was reproduced from Kamadurai *et al.*, 2013[189])

The first step to transfer ubiquitin by HECT E3 is the interaction with E2~Ub. The crystal structures for E2:HECT[187] and E2~Ub:HECT[128] revealed the interface between them. The major differences between these two structures are the rotation of C-lobe in HECT domain and the non-covalent interaction between Ub and C-lobe when E2~Ub is bound to HECT E3. Huang *et al.* solved the structure of UBCH7:E6AP and showed that the catalytic cysteine 820 is located in the junction connecting the two lobes. Both lobes provide important amino acid contacts within this broad cleft, with N-lobe residues being broadly polar and negatively charged whereas C-lobe residues are widely conserved and are hydrophobic in nature. Comparing the two structures, the HECT-binding domain from the E2 is conserved among them but the E2 recognition region upon HECT binding domain

is varied. In addition, Phe (62 on UBH5B, 63 on UBCH7) on loop4 is important for E2 interaction with the hydrophobic patch on some HECT domains (**Figure 1.8a**). The mutations made at these sites often disrupt the transthiolation between E2s and HECT E3s but these sites appear more important and consistent with HECT E3s rather than RING E3s[190]. In addition, the wide range of this hydrophobic region among HECT E3s may explain the reason why the same E2 is capable of interacting with numerous HECT E3s. The above mentioned structures also indicate the selectivity of HECT E3s toward E2s. Nedd4L Trp742 which is conserved in the Nedd4 family has been found to form a hydrogen bonding interaction with UBCH5C Ser94. This is the same region within HECT E3s that enables interaction with the UBCH4/5 and UBE2E families[191,192]. Other HECT E3s such as HECTD1, HERC3, HERC6 and HECTD6 also have the corresponding residue acting as Trp742 to interact with the residue homologous to Ser94 in UBCH5C. As mentioned above, the catalytic cysteine on the E2 is found to be >40 Å away from the catalytic cysteine in the HECT domain in the H7:E6AP structure, making transthiolation impossible. However, in UBCH5B~Ub:Nedd4L structure, the two cysteines are within <8 Å and implies that a certain conformational change occurred, enabling cysteine-cysteine juxtaposition. A complete and accurate mechanistic explanation of how E2~Ub triggers the rotation of the HECT E3 C-lobe and the predicted E2 binding interactions are all likely to play an important role and contribute to catalysis.

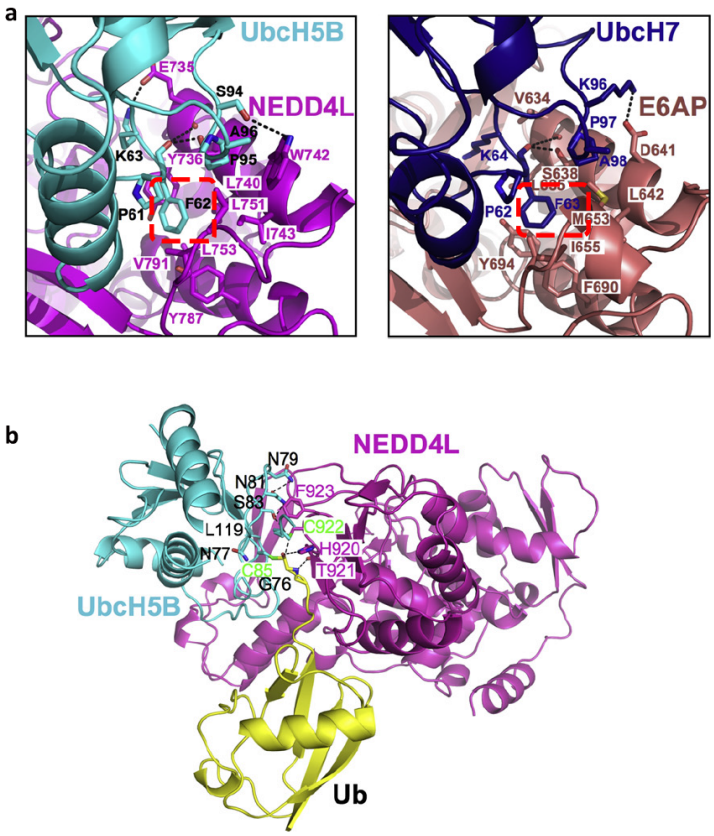


Figure 1.8 The close-up view of the crystal structure of E2:HECT complex.
(a) Interactions between UBCH5B (cyan, black labels) and HECT NEDD4L N lobe (magenta) (left) and interactions between UBCH7 (navy) and HECT E6AP (salmon), in same orientation (Right). The red frame highlights the F62/F63 positon on E2s which is essential for E3 binding. (b) The model of UBCH5B~Ub with NEDD4L HECT domain. 4° rotation on the hinge linker is predicted to adopt catalytic Cys on E2 in proximity with catalytic Cys on HECT domain. (Figure was reproduced from Kamadurai *et al.*, 2009[128])

In the predicted modeling complex reported by Kamadurai *et al.*, the E2 can be positioned more closely to HECT domain in order to approach catalytic cysteine for competent transthiolation by rotating the hinge linker by approximately 4°[128]. Furthermore, the rotation in this model had minimal effect upon the interface of E2~Ub and the HECT domain. The conserved region was discovered by the predicted model and agreed with the results of the mutational analysis of important HECT E3 residues. The Leu119 on UBCH5B was found to interact with His920, Phe923 and catalytic Cys922 (**Figure 1.8b**). Although the exact role of the

residues found by modeling require further investigation, the interactions between donor ubiquitin on E2 and the HECT E3 C-lobe in the crystal structure indicate that the mechanism to prime ubiquitin transfer is largely mediated by the donor ubiquitin. Unlike the contacts made with RING E3 ligases, E2~Ub adopts as opened conformation when contacting with HECT E3s in the pre-transthiolation state. The Ub tail is then stabilized by the interaction with the C-lobe. According to the crystallographic data, the lle36, Leu71 and Leu73 on Ub coordinate with Leu916 and Met943 on the HECT to enable the interface between them. Hydrogen bonds are also discovered between Gly35 on Ub with Lys90 on HECT and Gln40 on Ub with Leu916 on HECT. The hydrophobic patch on Ub including Ile36 and Leu71 both play important role on mediating the interaction. Mutaion of both Ile36 and Leu71 perturb ubiquitin transfer to Nedd4L, ITCH and Rsp5[128]. Clearly the presence of the ubiquitin component in E2~Ub plays a role in inducing conformational changes in the E3 that are required for catalysis.

HECT E3s have been experimentally determined to enable both mono-ubiquitylation and poly-ubiquitylation of protein substrates. After ubiquitin has been transferred to the HECT domain, the interaction between C-lobe and Ub is still preserved but Ub is orientated in an extended conformation and extends past the β-sheet of the HECT C-lobe. The C-lobe then rotates ~130° and turns it away from E2 binding site for thioester-linked donor ubiquitin to be attacked by substrate lysine[189] (**Figure1.9**). The well-known "-4Phe, F806" resides in this region at the C-lobe and collaborates with Phe505 and Leu506 to orientate the two lobes into appropriate position for substrate ubiquitination. The mutation on Phe806 severely impairs the ligation but can be rescued by mutating the L506W or L506F upon the N-lobe[189] (**Figure 1.7**). However, the conserved region on the N-lobe is only preserved in Nedd4 family HECT E3s which raises the possibility for other

mechanisms of ubiquitin transfer being possible with respect to miscellaneous HECT E3s[169]. The HECT:substrate structure also reveals that the loop in N-lobe containing Asp495 is important for ubiquitin ligation. [193]. The ubiquitin transfer is assisted by the noncovalent Ub binding sites on the N-lobe which have been demonstrated to be essential for chain elongation but not for transthiolation with E2~Ub[194]. This region is distal to the E3~Ub which makes unsuitable contacts with ubiquitin in order to accept ubiquitin from the donor. Experimental data suggests that this site can coordinate with ubiquitylated substrate and promote the chain elongation[195].

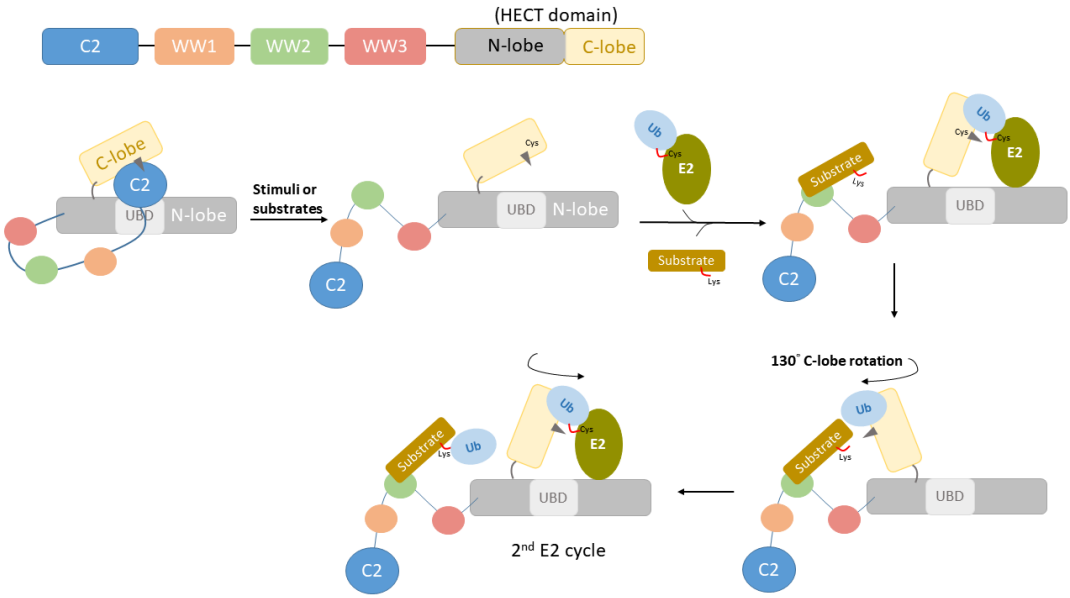


Figure 1.9 HECT E3s in motion.

The C-lobe on the HECT domain is blocked by C2 domain at rest in an auto-inhibited form. By induction from stimuli or binding with substrates, the HECT domain is freed from inhibition and is able to interact with E2~Ub. C-lobe will rotate 130° upon binding with E2~Ub and present the donor ubiquitin toward substrates. (Figure was reproduced from Fajner *et al.*, 2017[185])

The intrinsic activity of HECT E3s to make poly-ubiquitin chains is independent of E2s although the universal mechanism for the chain type preference of HECT

E3s is still poorly understand. By far, only the C-terminal region from Nedd4 Rsp5 has been shown to be critical for K63 chain formation. The replacement the last 60 residues on Rap5 to the region corresponded to E6AP can manipulate the chain formation from K63 to K48[196].

1.3.2.5 The regulation of HECT E3s activity

The activity of HECT E3s must be regulated since the loaded E2~Ub can adopt a catalytically competent orientation for ubiquitin transfer once it associates with a E3. As a result, in order to maintain an appropriate level of E3s or substrates within the cell, some HECT E3s have been found to be in an auto-inhibited state whilst at rest[197-199]. Again, the Nedd4 HECT E3 family is the most well-studied example for auto-inhibitory mechanisms of HECT E3s. In addition, the N-terminal C2 domain on Smurf2 has been demonstrated to interact with its HECT domain to prevent E2~Ub binding[200]. Through in-solution NMR studies and biochemical experiments, the region on the C2 domain which interacts with phosphoinositides overlaps with the region found to interact with the HECT domain. NMR binding assays, accompanied with isoleucine mutations upon the C2 domain, revealed that the C2 domain also masks the catalytic cysteine on the HECT domain important for ubiquitin thioester formation. The proposed mechanism has been further developed and supported by the observation of increased Ub thioester formation using ΔC2-truncated Smurf2 in biochemical experiments. Moreover, deletion of the C2 domain or by interrupting the interaction of C2-HECT both enhance the autoubiquitylation of Smurf2[200]. The same laboratory further demonstrated the disruption of these regions protects Smurf2 from degradation in cells and enhances its activity towards its substrate, RhoA. Smad7, which is known to enhance Smurf2 activity[201,202], also shows competition with C2 domain binding.

The PY motif on SMAD7 will interact with the WW domain on Smurf2 and release the C2 domain from the HECT domain for E2~Ub binding[203] (**Figure 1.9**). The dual role for the C2 domain expands the possibility that other domains within HECT E3s might play different roles for activity regulation and expand upon the hypothesis that substrate binding to HECT E3s may induce E3s activation.

WW domains have also been demonstrated to regulate the activity of HECT E3s. Gallagher *et al.* first indicated that JNK1 binds the HECT E3 ITCH and phosphorylates Ser199, Ser232 and Thr222 which frees the HECT domain from WW and PRR domain intramolecular inhibition[199]. Small PY motif-containing membrane proteins, NDFIP1 and NDFIP2 (Nedd4 family-interacting proteins), were later shown to activate ITCH activity by unmasking the catalytic cysteine rather than through E2 binding. NDFIP is able to interact with WW domains and prime for ubiquitin thioester formation[198]. In addition to the WW domain itself, the linker between them also participates in the autoinhibition of HECT E3s. The recently solved structure indicated that the 26-residue α-helical linker (named 2,3-linker) between WW2 and WW3 domain on WWP2 E3 is able to inhibit WWP2 activity by interacting with both the N- and C- lobe[204]. The C-terminal part of 2,3-linker interacts with the hinge connecting the two lobes and significant limits the flexibility of the C-lobe important for ubiquitin transfer. Furthermore, phosphorylation of Tyr369 and Tyr392 were also found to enhance the autoubiquitylation of WWP2. WWP1 and ITCH share these two residues which implies the same inhibitory mechanism exists within this HECT E3 family[205]. However, the Nedd4 E3 also contains a structurally similar linker after the WW1 domain but it does not retain the same auto-inhibitory effect. In conclusion, although the field has learned many details from HECT E3 structure studies, this auto-inhibitory mechanism might only apply to certain examples.

Since a number of HECT E3s have been shown to have evolved an auto-inhibitory mechanism to control enzymatic activity, the question remains, 'how are they activated?' Firstly, conformational changes are induced by E2~Ub binding. As discussed in an earlier section, the HECT domain itself can adopt conformations that are able to accept E2~Ub binding and that are compatible with transthiolation with E2~Ub via rotation of the hinge region. However, in general, there is no universal mechanism to explain the regulation of HECT E3 activity, especially one that encompasses the varied roles played by the extended domains N-terminal to the HECT domain. That is, inhibition of activity by intramolecular interaction or increasing activity via presenting substrates to the HECT domain. In the case of Nedd4, there are several mechanisms to explain the activation of this enzyme. It has been found that increased Ca^{2+} levels can activate Nedd4 by relieving of the C2 domain inhibition effect[206,207]. Wang *et al.* demonstrated Ca^{2+} can compete with the binding of the C2 domain with the HECT domain in Nedd4-1/2 ligases by increasing Ca^{2+} levels by ionomycin treatment (which activates endogenous Nedd4 *in vivo*)[206]. They further showed that a ligand for epidermal growth factor receptor (EGFR), carbachol, can also activate endogenous Nedd4 in cells. Escobedo *et al.* further discovered that Ca^{2+} interacts with three loops (β1-β2, β5-β6, and β7-β8) on the C2 domain by structural analysis[207]. Nedd4L is mainly located in the cytoplasm but will translocate to the membrane after Ca^{2+} stimulation. Phospholipase C (PLC) cleaves PIP_2 and generates IP_3 which can later diffuse into the endoplasmic reticulum (ER) and binds with IP_3 receptor which is a Ca^{2+} gated channel. Interestingly, the IP_3 can also interact with C2 and PH domains which direct HECT E3s to stay in the cytoplasm. The Nedd4-type HECT E3s can also be activated by ubiquitin non-covalently bound on an area called the 'exosite'[208-210]. The conserved hydrophobic patch (IIe44 patch) and C-terminal region (Leu71 and

Lue73) of ubiquitin can interact with this region (micromolar-K_D). The inhibitor which specifically targets this area abolishes the activity of Nedd4-1[211]. Zhang *et al.* recently developed a new tool, ubiquitin variants, to further study this molecular mechanism in detail[212].

E6AP has low E3 activity *in vitro* but it has also been found to be activated by another HECT E3 ligase, HERC2, via the RCC1-like domain 2 which is independent of its ligase activity[213]. Furthermore, the E6AP HECT domain trimeric complex that was identified through solving the E6AP crystal structure also demonstrated the potential for E6AP to regulate its own activity[187], although firm data to support this conclusion are not available to date. The mutation of the essential residues for trimer formation lost enzymatic ligase function[214,215]. The interfaces between trimer complex revealed that F727, R626, D543 and Y533 are all important for activity[216]. Other mechanisms of HECT E3 are not yet well characterized but the HECT E3 HUWE1 can undergo dimerization to inhibit enzymatic activity and might be activated by donor ubiquitin association. Interestingly, the tumor suppressor, p14ARF, showed the ability to downregulate HUWE1 activity in cells by shifting HUWE1 population into dimeric or oligomeric states[217].

1.3.2.6 The physiological role of HECT E3s

The limited knowledge regarding regulation of HECT E3 activity does not prevent understanding of the physiological roles played by this family of enzymes. The Nedd4-like family was the first identified family of HECT E3s. In Nedd4-1 deficient mice, growth retardation and abnormal development in nervous and cardiovascular systems are observed[218]. The tremendous growth retardation observed is believed to be due to the decrease of insulin-like growth factor 1 (IGF-1) levels and the increased level of adaptor protein Grab10. Nedd4-1 apparently

positively regulates the IGF-1 level and insulin signaling pathway to mediate animal development. *Nedd4-1*$^{-/-}$ mice also demonstrate misregulation of neuron formation and branching of retinal ganglion axons[219,220]. Nedd4 KO leads to increased level of thrombospondin-1 (Tsp-1) which is able to regulate angiogenesis[221]. The increased level of epithelial sodium channel (ENaC) in Nedd4-2 KO mice causes severe lung collapse and death after birth[222]. The PY motifs in ENaC are the substrate for the WW-domain in Nedd4-2 and the interruption between these regions also abolish the endocytosis and degradation of ENaC which leads to abnormal sodium level in cells.

WWP1 is another HECT E3 that also contains four WW domains and the gene knockdown has been demonstrated to arrest tumor cell growth[223]. In addition, the knockout of WWP1 also reverses the TNF-induced osteoblast differentiation arrest of mesenchymal stem cells[224]. On the other hand, WWP2 can regulate the important transcription factor for cell pluripotency, OCT4, by ubiquitylation. This has led to studies being undertaken to ascertain whether WWP2 can be manipulated for cell growth[225]. Smurf1 and Smurf2 have also been studied and have been shown to regulate the TGF-β signaling pathway by directly ubiquitylating Smad proteins. However, knock out mice later proved the redundancy of Smurf1/2 in this pathway and showed no phenotype in early ages[226].

Other HECT family HERC E3s have also demonstrated various roles in regulating multiple cellular function. HERC2 has recently been shown to be involved in DNA metabolism by collaborating with RNF8 and RNF168[227]. The RING E3 BRCA1 has also been shown to be a substrate for HERC2 which allows HERC2 to indirectly regulate the cell cycle[228]. Other miscellaneous HECT3s have also been shown to play important roles in the cell. For example, the expression levels of UBE3A contribute to the neurodevelopmental disorder Angelman syndrome and

autism[229]. HUWE1 has also been identified in multiple signaling pathways and is responsible for cell proliferation and apoptosis. It is overexpressed in various cancer types and is considered a good target for drug development especially through its regulation of p53 and Miz1 stability[230,231]. Though being the first class of E3s to be discovered in 1995, a huge amount of data has demonstrated that HECT E3s play important roles and functions in cell biology[92]. However, such data has been heavily based on structural biology or cell biology approaches. The goal of this thesis therefore is to apply knowledge garnered in both biological disciplines to help better understand the roles and functions played by HECT E3s.

1.3.3 RING-Between-RING (RBR) E3 ligase

1.3.3.1 General features of RBR E3s

In contrast to the previous E3 families described, RBR E3s utilize a distinct RING/HECT hybrid mechanism to transfer ubiquitin from E2s to substrates. RBR E3s contain a RING1 domain, an IBR(In-Between-RING) domain and a RING2 domain. General RING1 recruits E2~Ub and positions the thioester-linked Ub in close proximity to the catalytic cysteine on RING2. The ubiquitin is then transferred to the RING2 domain and subsequently to protein substrates. Unlike RING E3s, which indirectly regulate ubiquitin transfer via mediation of E2 interactions, RBR E3s are directly involved in substrate recognition and ubiquitin transfer events[87,232]. RBR E3s were discovered from RING E3 sequence alignments where three Zn^{2+}-binding domains showed a distinct alignment which was in contrast to canonical RING E3s[68,233]. Nearly a decade later Wenzel *et al.* determined that RBR E3s belong to another E3 subfamily apart from RING E3s[121] since the RING1 domain shares a similar structure and function with canonical RINGs. The discovery of bone fide catalytic cysteine on different RBR E3s and the trapping of Ub-E3 intermediates

further emphasized the inherent structural and mechanistic differences between RING E3s and RBR E3s. Interestingly, to date, no isolated IBR or RING2 domain has been identified in a protein in nature which again highlights the unique, collective biological features of RBR E3s.

1.3.3.2 General structure of RBR E3s

The RBR region is the core structure of RBR E3s and is the main region responsible for ligase activity. It is normally located at the C-terminus of the protein and is often found with a variety of N-terminal extensions or domains which are used to interact with adaptor or substrate proteins[232]. (**Figure 1.10**) The ANKIB1 (ankyrin repeat- and IBR domain-containing 1) and Dorfin RBR E3s are the only two exceptions; where the RBR domain sits in the center or N-terminal part of protein. Like the HECT E3 family, the variety of N-terminal extension architectures allow RBR E3s to interact with different substrates, be regulated by other adaptor proteins, or post translational modifications. For example, the Ser65 phosphorylation of the N-terminal UBL (ubiquitin-like) domain on Parkin activate its ligase activity[234]; the UBL domain of HOIL-1 (haem-oxidized IRP2 ubiquitin ligase 1) on other hand recruits HOIP (HOIL-1-interacting protein) via interaction with a UBA (ubiquitin-associated) domain and together with SHARPIN (SHANK-associated RH domain interactor) forms the LUBAC complex for linear ubiquitin chain formation. The acidic residues on the N-terminal region of HHARI (ARIH1, Ariadne RBR E3 1), TRIAD1 (two RING fingers and a DRIL) and TRIAD3 enable RBR E3 activation through association with other E3s, namely CRLs[235]. Instead of activating RBR E3 ligase activity, the domains outside of the core RBR structure adopt an inhibitory role. The RING0 domain on Parkin has been shown to block the accessibility of catalytic cysteine on RING2[236]. The C-terminus of the Ariadne

domain on HHARI, TRIAD1 and ANKIB1 also play a role in intramolecular auto-inhibitory machinery[237]. Further research is needed upon structural and mechanistic aspects of RBR E3 function since the majority of structure-function studies carried out only focus on a sub-set of the RBR E3 family (that is Parkin, HOIP and HHARI). Little is known about other RBR E3 family members, for example, Dorfin, RNF144 A/B, ANKIB1 and PARC[232,238].

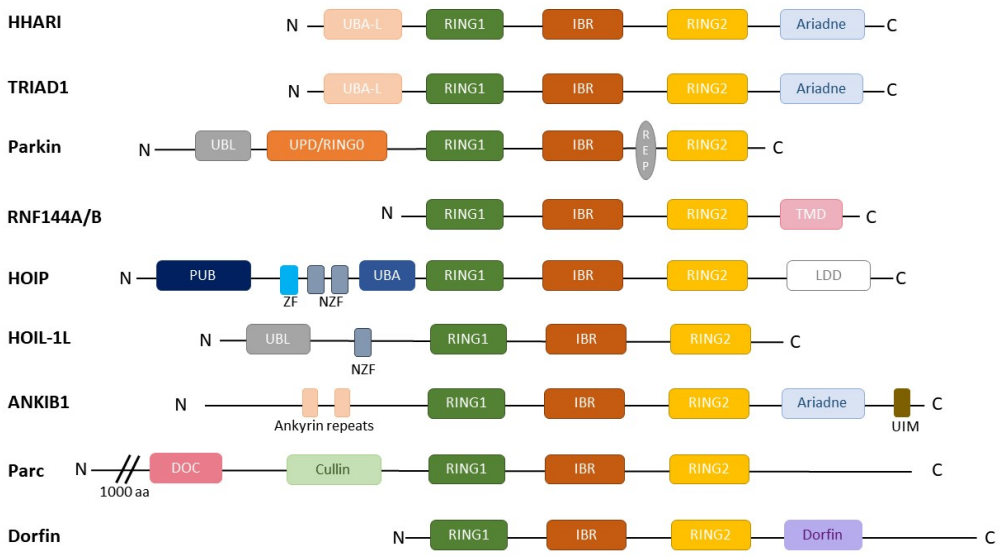


Figure 1.10 Domain architecture of the RBR E3 ligases. (Figure was reproduced from Dove *et al.*, 2017[238])

Most of the solved RBR E3 structures contain a reasonably conserved core RBR domain which allows some insight to be gained from their analysis. The three domains exist in order RING1-IBR-RING2 but might structurally display in different orientations depending in which form they exist (active- or inactive-conformation). The conserved feature of these three domains is they each are all cysteine-rich and coordinate two Zn^{2+} ions. RING1 adopts a typical C3HC4 cross-brace fold which is similar to canonical RINGs and enable them to functionally interact with E2s. Although the name, RING2, was ascribed due to the similarity of sequence to the

RING domain, later structural analysis revealed that neither the IBR or RING2 adopts the same structure as RING1 or other canonical RING domain. The IBR domains and regions N- or C-terminal to them show extensive interaction with the RING2 domain in solved crystallographic structures. It has been suggested that the IBR domain might interact with ubiquitin during transfer or mediate the inhibitory domain (as the linker region is dynamic in nature), but the exact role for IBR is still not clear[232]. The RING2 domain contains the active site of RBR E3s, the catalytic cysteine, and it has been experimentally determined that this cysteine is not involved in Zn^{2+} coordination. The ubiquitin is transferred from E2s to the RING2 and forms a thioester bond between C-terminal tail and catalytic cysteine. The detailed mechanism of subsequent ubiquitin transfer from RING2 to substrate is not fully understood. The RBRE3 HOIP which contains a linear ubiquitin chain determining domain (LDD) is the only RBR E3 where insights into substrate recognition have been found[239].

1.3.3.3 The ubiquitin transfer mechanism of RBR E3s

As mentioned above, RBR E3s accept ubiquitin from E2~Ub and the Ub~RBR E3s intermediate then goes on to further react with substrates. Previous studies have demonstrated that E2s which are able to function with RBR E3s can also function with RING E3s. It has raised the question, 'how do RBR E3s determine whether to undergo transthiolation or an aminolysis reactions?'. We will undoubtedly learn more from structural studies of E2~Ub:RBR complexes, as currently there are only two complexes with solved structures[129,240]. However, the distinct selectivity toward HECT/RBR E3s from the E2, UBCH7, may shed some light on this question. In fact, UBCH7 plays an important role in the discovery of RBR E3s not only because it was the very first E2 enzyme to be mapped to E3s binding partners but because

UBCH7~Ub shows high preference in discharging Ub toward cysteine rather than lysine[121,123]. UBCH7 has been shown to function with HECT E3s but Wenzel *et al.* later applied an E3-independent assay to demonstrate that UBCH7~Ub exclusively discharges Ub onto cysteine[121]. Together with the data showing that UBCH7 is functional with Parkin and HHARI further identified that RBR E3s demonstrate a RING/HECT hybrid mechanism. Similar results were also provided for other RBR members soon after their initial publication[241,242].

E2~Ub exhibits a highly dynamic structure and reactivity profile toward substrates and is able to interact with both RING and RBR to elicit substrate ubiquitylation by different mechanism. As we know E2~Ub displays to a 'closed' conformation, once bound with canonical RING E3s. Where the ubiquitin folds back in a closed conformation toward the E2 and displays the ubiquitin thioester bond in a favorable position to be react with an incoming lysine[125,126]. However, the RBR E3s do not support this conformation to accept E2~Ub[243]. RING1 domain does not contain the linchpin Arg residue which is an important and a conserved region in canonical RING domain, used to anchor the closed conformation[159]. Although not explaining how RING1 domain coordinate the incoming E2~Ub to favor transthiolation, the recently solved E2~Ub:RBR structures showed the E2~Ub adopts an extended and 'open' conformation which reduce the likelihood of proceeding by aminolysis[129,244]. The determinant of the E2~Ub conformation that enable transthiolation over direct aminolysis have not yet been confirmed since the unbound UBCH7~Ub complex exists in a mainly closed conformation in solution but capable of transthiolation with free cysteine amino acid[121,243]. The UBCH7~Ub has been demonstrated to exist in an open conformation when bound with RING1 in the UBCH7~Ub/HHARI complex. In fact, it is not the E2 itself that determines the next move but the extension of the 2nd Zn^{2+}-loop on HHARI that

may contribute to disfavoring of the closed E2~Ub conformation. The deletion of two residues in this loop abolish the ability of HHARI to stabilize the E2~Ub open conformation. The steric effect from the extended loop is conserved among some RBR E3s but not all since the extra residues do not exist in Parkin and other RBR E3s. In addition, the affinity toward E2s is extremely low for an inactived Parkin compared to a fully activated Parkin which implies that Parkin might not use the exact same mechanism as HHARI. However, we cannot rule out the possibility that other RBR E3s may use a similar mechanism to mediate E2~Ub binding with RING1 since only two structures of complex have been solved so far.

In contrast to UBCH7~Ub/HHARI structure, UBCH5B~Ub/HOIP structure provides some complementary information about RBR E3 interactions with E2s. Both E2~Ubs adopt an open conformation in these structures but the Ub moiety on H5B contains more non-covalent contacts with HOIP. These interactions have been proposed to maintain the open conformation of E2~Ub since H5B~Ub is found to retain an open conformation when uncomplexed. Once the E2~Ub binds with RING1, the Ub is able to transfer to the catalytic cysteine on the RING2 domain. Little is known about how ubiquitin is transferred to RING2 and what kind of interactions occur between Ub and the RBR E3 since most of the RBR E3s are in an auto-inhibited state in the context of full-length protein. Furthermore, several crystal and NMR structures showed that the distance between catalytic cysteine on E2s and RBR E3s is significantly greater than would be chemically allowed for a successful transthiolation reaction to occur (>55 Å in Parkin case and >65 Å in HHARI case) (**Figure 1.11**). Large and specific conformational changes have to occur during the ubiquitin transfer to allow transthiolation. The highly dynamic IBR domain and its linkers support this hypothesis since different orientations of the IBR domain and linker have been discovered in different structures for the same

protein, but the various extra domains out of the core RBR domain might also cause the IBR domain to behave differently, from a structural point of view. The RING2 domain displays similar features among RBR E3s such as the catalytic triad (consisting of the catalytic cysteine and neighbouring His and Glu) which are found in Parkin[236,245], HHARI[237] and HOIP[239]. Also, these catalytic regions are all buried by other domains (RING0 for Parkin, Ariaden for HHARI) by hydrophobic interactions. A good tool to stabilize the highly flexible active RBR E3s is urgently needed for both structure and activity regulation studies of RBR E3s.

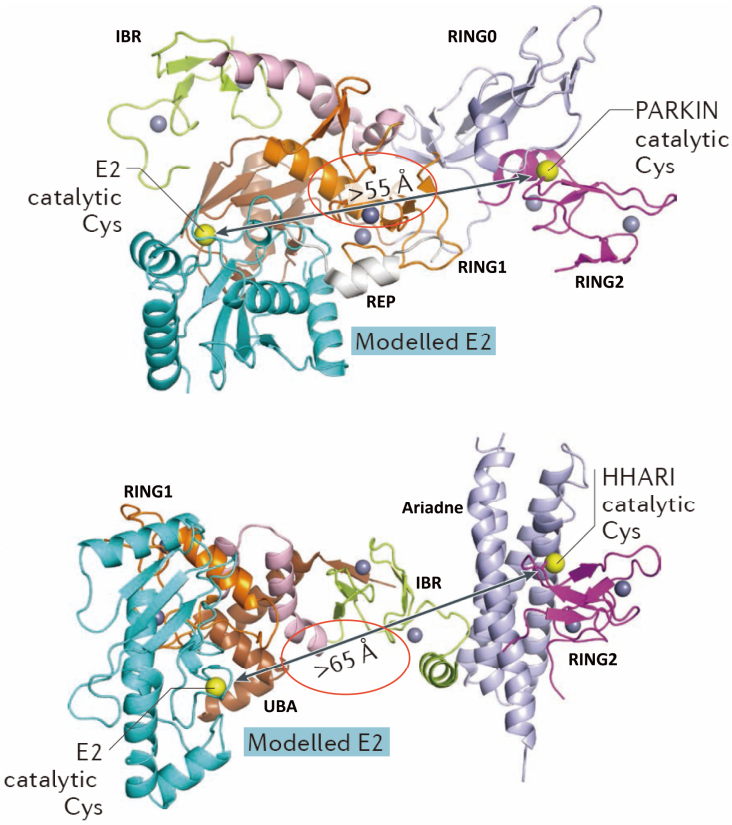


Figure 1.11 Structures of PARKIN and HHARI with a modelled ubiquitin-conjugating enzyme.

The model was built by docking RING1 domain from Parkin (PDB:5C23) or HHARI (PDB:4KBL) with the structure of RNF38 RING domain bound to E2 (PDB: 4V3L). The distance between two catalytic cysteines on E2 and E3 are presented and highlighted. In order for successful transthilation, a conformational change spanning nearly 60 Å is implied. Moreover, the clash on UBL domain on Parkin was observed when E2 was closed. (Figure was reproduced from Buetow *et al.*, 2016[137])

1.3.3.4 Activity regulation of RBR E3s

It has been proven by biochemical assays and currently solved structures that the RBR E3s are in the auto-inhibited conformation at the rest stage[234,241,242]. Apparently, the extended domains N- or C-terminal to RBR core involved in either blocking the E2s binding or regulating the accessibility of catalytic cysteine on RING2. How to prime the activity of RBR E3s has been heavily studied recently although still focused only on few of them. Also, most of the mechanisms are based on structural data interpretation and domain deletion/truncation in *in vitro* assays. The up-stream stimulation for ligase activity is still poorly understood.

1.3.3.4.1 Unique domain architecture inhibits Parkin activity but the dual phosphorylation activates it

Parkin has been found in loss-of-function mutation for its E3 activity in various types of Parkinson's disease (PD). To understand the biology behind of it, especially for the regulation of E3 ligase activity, many groups have contributed. Like other RBR E3s, the full-length Parkin structures solved to date has been indicated that it is in an auto-inhibited form with five domains in a tight embrace[236,245]. The Parkin protein comprises an N-terminal the UBL domain followed by a flexible short linker. The RING0 (or UBD) domain, which is unique to Parkin, and IBR domain reside in the middle of Parkin architecture. Before the RING2 domain, a disordered segment which contains a α-helix connects IBR with the RING2 domain. Since 2013, crystal structures have been solved and indicated that Parkin is inhibited and stabilized via several hydrophobic interactions[236,245-247].

First, the hydrophobic region around Ile44 residue on the UBL domain, which share similarity with ubiquitin, interacts with the RING1 domain and block the E2 binding with Parkin since the RING1 has been proven to be the site that recruits

E2~Ub (**Figure 1.12**). The mutation or truncation of UBL can relieve the structure and increase Parkin activity *in vitro*[248]. UBL has also been shown able to interact with SH3 and UIM (ubiquitin interacting motif) domains which indicate that the competition for the UBL binding may also retain Parkin activity. The following RING0 domain coordinates with Zn^{2+} in hairpin structure and this unique segment occludes the catalytic cysteine in RING2 domain for ubiquitin transferring. The IBR and RING2 domains in Parkin are quite conserved within RBR E3s but the linker between them, now referred to as REP domain (repressor element of Parkin), has been shown to bind with the RING1 domain and to block the E2s interaction. To overcome this effect, one single amino acid residue mutation in the REP region, W403A, has been shown to increase the Parkin activity both *in vitro* and in cells[245]. The mutations associated with PD have also been found to either disrupt the zinc coordination for the structure maintaining or to block the E2 binding. The G430D and C431F mutations on the RING2 domain even directly abolish Parkin activity[246,249].

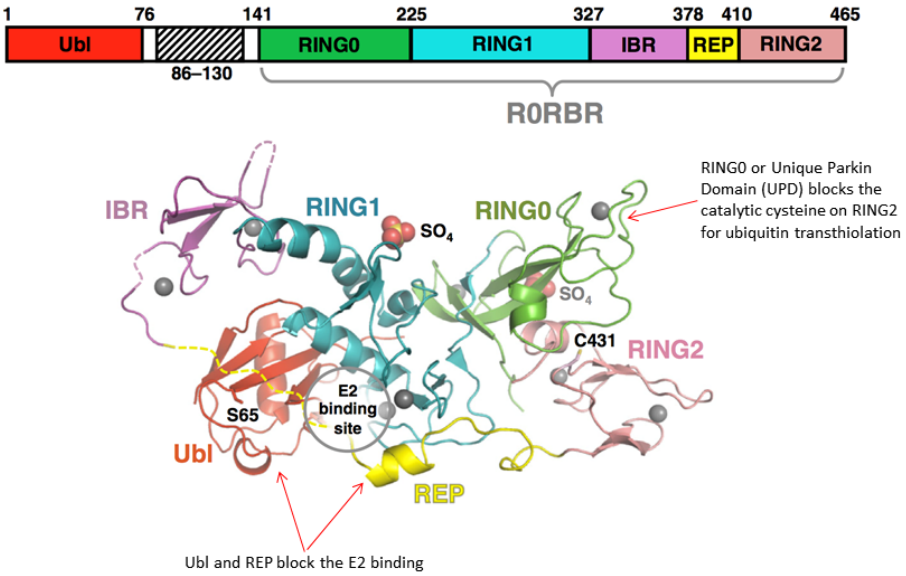


Figure 1.12 The unique domains in Parkin regulate auto-inhibition.
Parkin is a 465 amino acids E3 ligase which adopts several different domains in structure. The Ubl domain outside the RBR core together with the REP domain which resides between the IBR and RING2 block the interaction between E2 and Parkin. The RING0 further occludes the catalytic cysteine for ubiquitin access. *this structure is missed a.a 86-130 for crystalization purposes. (Figure was reproduced from Sauvé *et al.*, 2015[250])

How to efficiently activate Parkin has been a question in the field for a long time. Early *Drosophila* studies indicated that PINK1 (PTEN-induced kinase-1) is associated with Parkin activity by serving as the up-stream protein in mitochondrial quality control signaling pathway[251,252]. Interruptions in PINK1-Parkin pathway lead to the dysfunction of mitochondria maintenance. The mutations on both PINK1 and Parkin proteins have also been discovered in PD patients[253]. PINK1 is constitutively being cleaved at mitochondria by mitochondrial intramembrane protease PARL (Presenilins-associated rhomboid-like protein). The remaining PINK1 N-terminal domain will subsequently translocate to the cytosol and be degraded by the N-end rule proteolysis[254]. However, once the mitochondria sense the damage signal such as depolarization by harsh CCCP reagent (carbonyl cyanide m-cholorophenylhydrazone) treatment, PINK1 is stabilized and

accumulates on the outer mitochondrial membrane. Parkin is then be recruited to the mitochondria and phosphorylated by PINK1[255-257]. PINK1 directly phosphorylates the Ser65 on the UBL domain both *in vitro* and in cells and liberates the auto-inhibition mechanism caused by the UBL domain[234,249,255]. It frees the RING1 domain for the E2~Ub engagement. Deletion of the UBL domain also showed the enhancement of Parkin activity *in vitro*[249]. However, the phospho-mimetic S65E Parkin cannot bypass the PINK1 requirement for Parkin translocation to mitochondria together with the data that S65A Parkin still has the trace amount of activity in cells[80] indicated that PINK1 might work with other proteins to activate Parkin.

The breakthrough discovery was provided by three different groups that ubiquitin can be phosphorylated by PINK1 and the phosphorylated-ubiquitin (p-Ub) can further prime the Parkin activity in a feed-forward pathway[80-82]. PINK1 phosphorylates Ub at the same position as the UBL domain which makes isolated phospho-UBL also capable to activate Parkin *in vitro*. The introduction of Ser65 mutation on UBL and ubiquitin both impaired Parkin activity[80]. Furthermore, Parkin has high affinity toward p-Ub by 400nM and resulted in 20-fold enhancement for its phosphorylation by PINK1 when p-Ub was supplied[82,258]. The PD-associated mutation site on Parkin, Lys161 and Lys211, together with the Arg163 residue form a pocket for p-Ub binding and was further be confirmed by crystal structure and *in vitro* assays[84,258]. The p-Ub interacts with the RING0 domain and its phospho-group and Ile44 hydrophobic patch further react with the RING1 and IBR domain. The disruptions on the p-Ub-binding helix (pUBH) showed the decrease of Parkin activity[84]. Recently, several different Parkin structures either without the UBL domain (presumed as active) or conjugated with p-Ub provide the molecular mechanism for priming Parkin activity[247,259]. In brief, the p-Ub binding

within the Parkin can release the UBL domain from core structure which allow it to be phosphorylated by PINK1. Once the UBL domain is phosphorylated, it relieves the auto-inhibited machinery by displacing RING0 and REP domains allowing RING1 to interact with incoming E2~Ub. Although the structure of the E2~Ub:Parkin complex has not yet been solved, it has been hypothesized that Parkin becomes flexible in structure once activated and performs a conformational change to bring the RING2 domain in close proximity with the E2 catalytic site (**Figure 1.13**).

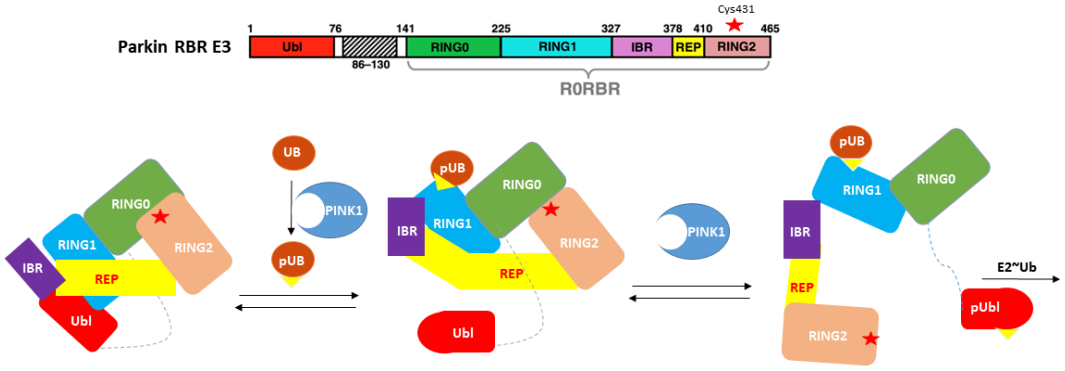


Figure 1.13 Dual-phosphorylation on Parkin UBL domain and Ub primes Parkin activity. Parkin E3 exists in an auto-inhibited conformation at resting stage. In the current model, once Ub has been phosphorylated by PINK1, pUb is able to interact with Parkin RING1 domain and triggers a conformational change of Parkin which frees the Ubl domain from Parkin core structure. The Ubl domain is then further be phosphorylated by PINK1 and induces another conformation change to expose the catalytic cysteine for incoming E2~Ub. The catalytic cysteine 431 is then no longer blocked and is ready for E2-E3 transthiolation. Both phosphorylation events on Parkin and Ub prime the activity of Parkin. Red star denotes catalytic cysteine 431 and yellow triangle denotes phospho-Ser65. (Figure was reproduced from Trempe *et al.*, 2013[245])

It is not yet clear the order of these phosphorylation events in cells since the binding with p-Ub enhances the phosphorylation of Parkin; the p-Parkin also has higher affinity toward p-Ub. In addition, who recruits the first ubiquitin to be

phosphorylated by PINK1 and make this p-Ub to be able to recruit Parkin is still unknown. However, it is possible that the two phosphorylation events proceed simultaneously and with the help from other proteins. Although it is clear that the PINK1 activates Parkin E3 activity *in vitro*, most of the cellular experiments are accomplished by use of overexpression system. Furthermore, the lack of phenotype in PINK1/Parkin mutant mice also brings the attention in field. It remains possible that other regulation mechanisms exist in the more physiological relevant conditions. More mechanistic studies in how phosphorylations contribute to the Parkin activity will be discussed in the latter chapter.

1.3.3.4.2 Activation of the only linear ubiquitin chain maker---HOIP and its binding partners from the LUBAC complex

LUBAC complex composed by two RBR E3s, HOIP (also termed RNF31) and HOIL-1L, and one adaptor protein SHARPIN. It has been the only protein which is able to make linear ubiquitin chain (M1-chain) in vertebrate to date[44]. The HOIP E3 is the main E3 ligase for catalyzing M1-chains assembly and is the only component from LUBAC to show robust E3 activity *in vitro*[239]. Similar to other RBR E3s, HOIP adopts an auto-inhibited state in the full-length protein. The HOIP RBR core alone can make a free chain *in vitro* but lost the activity while one N-terminus extended domain, UBA domain, is linked[239]. The isolated N-terminus domains can also inhibit HOIP activity in trans[87]. However, the detailed auto-inhibitory mechanism is still not yet fully understood since the lack of a full-length HOIP structure. Intriguing, HOIP is always coordinated with HOIL-1L and SHARPIN in cells as a LUBAC complex since no isolated HOIP protein has been detected after purification[44]. This implied that the intramolecular interactions between LUBAC components may regulate the E3 activity of HOIP.

The UBL domain on HOIL-1L has been found to activate HOIP E3 activity by interacting with the UBA domain and exposes the RBR core on HOIP for incoming E2~Ub[260]. Several binding complexes for the UBA-UBL crystal have been solved by Yagi *et al.*[260]. These data combined with the deletion mapping assays further indicated that the helices α6, α7, α8 and α9 on HOIP UBA domain are involved in recognizing the UBL surface on HOIL-1L which prevents the UBA inhibitory effect on HOIP. Although SHARPIN has the similar UBL domain for binding with HOIP E3, the exact molecular mechanism studies remain to be solved. Liu *et al.* recently solved the crystal structures for $UBA_{HOIP}/UBL_{SHARPIN}$ and $UBA_{HOIP}/UBL_{HOIL\text{-}1L}$ and demonstrated that HOIL-1L and SHARPIN individually interacts with a different region of the UBA domain on HOIP[261]. Furthermore, HOIL-1L and SHARPIN are individually capable to activate HOIP but are also able to synergistically activate HOIP via enhancing the E2 loading[261]. In contrast to other RBR E3s which need to be activated from the autoinhibited state, the LUBAC complex is constitutively activated by HOIP/HOIL-1L or HOIP/SHARPIN interactions which mentioned above. However, the question remains, how is the activity of LUBAC regulated?

LUBAC is the only E3 complex capable of making M1-chains, there are also only two DUBs that have been found to manage to cleave the M1-chains, OTULIN (OTU domain DUB with linear linkage specificity) and CYLD. In addition, it has been found that the N-terminal domains of HOIP E3 are able to react with these two DUBs to regulate the activity of the LUBAC[262,263]. There is also a PUB domain (Peptide:N-glycanase/UBA or UBX-containing proteins) in HOIP which is responsible for mediating interactions between the PUB interaction motifs (PIMs) of OTULIN DUB and spermatogenesis-associated protein 2 (SPATA2). SPATA2 is able to recruit CYLD to the LUBAC complex for chain cleavage once it binds with HOIP[264]. Based on the structures, these two proteins bind the same region in HOIP and both

regulate chain formation *in vitro*. The conserved tyrosine residue on the PIMs has even been found to be phosphorylated and may regulate the binding with the DUBs[262,264]. A proposed model of OTULIN/LUBAC regulation is once the signaling pathway is activated by stimuli, the phosphorylation on Tyr56 will disassociate the OTULIN and LUBAC interaction and lead to M1-chains accumulation. However, the kinase and phosphatase responsible for this event are awaited to be discovered.

Although the mechanism of LUBAC activity regulation is not fully understood, how is HOIP able to selectively make the M1-chains and how does it interact with E2~Ub have been demonstrate by recently solved crystal structures[129,239]. The E2~Ub:$HOIP_{RBR\text{-}LDD}$ is the first and to date the only solved crystal structure for the E2-RBR E3 complex while RBR is in active form[129]. It revealed that the E2~Ub is truly in the open conformation when engaged with RBR E3s. It also uncovered many interactions between the RBR domain and E2~Ub, especially for the ubiquitin moiety since the other E2~UB:RBR show little contact between Ub and the core RBR. Notably, the HOIP adopts a conformation that brings E2~Ub in proximity of the catalytic cysteine in the RING2 domain which is important for transthiolation. In addition, it demonstrated that the non-covalent contacts between ubiquitin and the RING1-IBR linker and IBR domain and also showed less E3 activity when mutated in these regions. However, this model was built on two separate asymmetric molecules that includes one RING1-IBR and RING2-LDD element respectively. In addition, the auto-inhibited HOIP structure is still unavailable which prevents us from comparing the differences between them. As a result, we still cannot formally draw an accurate model to describe how the RING2 domain adopts close proximity toward E2~Ub from inactivated form to activated form.

The selectivity for linear chain formation via HOIP E3 comes from the linear ubiquitin chain-determining domain (LDD)[241,242]. The $HOIP_{RBR\text{-}LDD}$ domain complexed with a donor ubiquitin and an acceptor ubiquitin in structure indicated the mechanism of the M1-chains formation[239]. The RING2-LDD surface orientates the acceptor ubiquitin to the position where only the N-terminal amino group of it is directed toward the HOIP catalytic cysteine Cys885, but not the other lysine residues in ubiquitin. The ZF1 domain is also involved in presenting the acceptor ubiquitin to the correct position. Moreover, the conserved histidine His887 acts as a general base to deprotonate the N-terminal amino group on the acceptor ubiquitin for a nucleophilic attack of the donor ubiquitin thioester bond[242]. The donor ubiquitin hydrophobic patch and its extended C-terminal tail interact with the β-hairpin on LDD and RING2 domain to stabilize the Ub on the RING2 and to guide the Gly76 in proximity of the catalytic cysteine. In conclusion, the distinct structure and architecture of HOIP E3 contributes mainly to the specificity of linear ubiquitin chain formation, but how to coordinate the activity between LUBAC and DUBs for regulating the M1-chains formation is stillnor fully understood.

1.3.3.4.3 RBR E3s are RING-linked--- Activity regulation of HHARI and TRIAD1 with CRLs

HHARI, TRIAD1, ANKIB1 and PARC/CUL9 all feature an Ariadne domain C-terminal to their RBR core structure. It has been proven that the Ariadne domain occludes the catalytic cysteine on the RING2 domain of HHARI[235,237,240]. The mutations or deletion of the Ariadne domain both activate HHARI *in vitro*[237]. Intriguingly, the binding with E2~Ub does not trigger a conformational change of HHARI moiety for bringing the RING2 domain in proximity of E2, the Ariadne domain still blocks the RING2 domain preventing E2-E3 transthiolation. A recent study showed the

W386A mutation on the RING2 domain or the Y531A on the Ariadne domain can both disrupt the inhibitory machinery for HHARI. The Y531A mutation can rescue the activity of the catalytic cysteine access and enhance the autoubiquitylation of HHARI, but it is still unable to transfer ubiquitin to the substrate proteins[265]. How does the RING2 domain free itself from the Ariadne domain binding is still not clear.

The other distinct character for these RBR E3s is the partnership with cullin RING E3 ligases (CRLs). It has been shown that the neddylated-CRLs can activate HHARI and TRIAD1 in a neddylation-dependent manner[235,265]. TRIAD1 can be activated by CUL5/RBX-2 machinery while HHARI can work with CUL1, CUL2, CUL4 and CUL4A. Notably, RBX1 can work with all CULs but RBX-2 is only able to react with CUL5. There is still no hint about which CULs teams up with PARC since there is already a cullin domain in PARC itself. As CULs play the vital roles in HHARI/TRIAD1 activity regulation, it has also been demonstrated that the Neddylation of the CULs affects the interaction between CRLs and RBR E3s. The treatment of NEDD8 E1 enzyme inhibitor, MLN4924, abolishes the TRAID1/CUL5 and HHARI/CUL1 interactions in cells but does not affect UBCH7 binding with TRIAD1[235,265]. The recent structural and biochemical studies both indicated that the predicted NEDD8 interacting sites on the UBA domain are important for the ubiquitin transfer from UBCH7 to HHARI[240]. As a result, the interactions between these two E3s are speculated to be essential to adopt the Ariadne-RBR E3s into a favorable conformation compatible with ubiquitin transfer[265].

The latest discovery of HHARI E3 is the partnership with the CRLs to prime the substrate poly-ubiquitination[265]. Scott *et al.* reported that when HHARI was activated by the Neddylated-CUL1 or CUL3, HHARI then primes the first mono-ubiquitylation on the substrates which CRLs recruit. CRLs then use their intrinsic RING E3 activity to extend the ubiquitin chains[265]. This discovery greatly broadens

our imagination about the regulation of E3 ligases activity since we generally interpretate the data by the very last step, the ubiquitylation. The outcome of the ubiquitylation may not simply be regulated by one E3 enzyme or one machinery but might involve synergetic cooperation of multiple E3s.

1.3.3.4.4 Physiological role of RBR E3s and their partner proteins

The clarification of RBR E3s as a new subfamily of E3 ligases has only recently been defined[121]. The above discussed RBR E3s are the only members which have been sufficiently studied. The other RBR E3s such as Dorfin, TRIAD3, RNF144A/B are not easy to assay for their E3 activity *in vitro* which makes them difficult to study. Furthermore, although all of them contain a RBR core structure, there is no universal activity regulation mechanism for them due to the divergent domains extensions. However, we still can gain some insight for these RBR E3s from their partner proteins and the cellular function which they have been involved.

Parkin has been demonstrated to mediate the substrates for regulating autophagy and neurodegeneration[266,267]. It has also been demonstrated to interact with G-protein-coupled receptors (GPCR) for the dopaminergic signalling and dopamine receptor[268,269]. Although the main study of Parkin E3 is focus on neurodegeneration disease and mitophagy, more studies have also indicated an involvement of Parkin in the cell proliferation and cancer regulation[270,271]. HHARI shares some substrates with Parkin which are PD-related but has also been shown to target mRNA cap-binding protein and regulate cell proliferation[272,273]. Less-studied RNF144A has been shown to regulate DNA damage-induced cell apoptosis by regulating its dimerization[274,275]. The recent studies also reported that RNF144A can regulate cancer cells proliferation and resistance towards a PARP1 (Poly(ADP-ribose) polymerase 1) inhibitor[276,277]. The RNF144B also showed the ability to

promote IL-1β mRNA expression by LPS (lipopolysaccharide) inducement[278]. The giant RBR E3 PARC/CUL9 contains the cullin sequence and an RBR core in the same polypeptide. The detailed relationship between these two domains is unknown but PARC has been proven to regulate the cell survival by mediating the level of cytochrome C[279]. Parkin is not the only RBR E3 to associated with neurodegeneration diseases, RNF19A/Dorfin has also been linked to synphilin-1 regulation and co-localizes with Lewy bodies [280,281].

Clearly, RBR E3 ligases play the coordinated roles in cell biology via their complicated regulation mechanism. To date, we have only got a glimpse of their mechanism. How much impact the RBR E3s have under physiological model needs more careful and detailed studies. The development of an accurate and facile method to detect their activation state is important to achieve.

Chapter 2. Activity-Based Probes

2.1 Unravelling protein identity and function---The case for Activity Based Probes

Enzymes of diverse function and biological activity are tightly orchestrated and regulated within cells in order to enable and maintain the complexity of life. It has been demonstrated that enzyme misregulation can lead to the onset of various diseases[97,282,283]. Therefore, understanding how enzyme function is regulated not only allows us to better understand fundamental biology, but also demands the development of new tools with which to address and combat human diseases. Directly monitoring the cellular activity of an enzyme or enzyme family is challenging. This is particularly important for enzymes expressed at low levels within the cell and which can also be subjected to regulation of their activity by post translational mechanism[284]. For example, activation via post translational modification or through interaction with specific cofactors[285,286]. As previously discussed, several HECT and RBR E3 ligases become activated when constituent intramolecular inhibitory domains become phosphorylated. For example, phosphorylation of the Ubl domain in Parkin relieves its autoinhibited structure, allowing Parkin to interact with E2~Ub. In addition, p-Ub binding to Parkin largely enhances the activation process of Parkin and primes the E3 for activity by altering the structure of Parkin E3[84,247]. Tightly regulated protein activation machineries have also been found and described for various cellular protease systems[287]. It is therefore not surprising that most proteases have been found in the auto-inhibited forms due to their intrinsic activity is to degrade their target proteins which must be tightly regulated spatiotemporally. As a consequence, methods used to measure mRNA levels or protein abundances do not provide quantitative

information relating to intrinsic protein activity or function only abundance information that does not always correlate with activity.

It can be argued that the detection of a protein's substrate(s) is an alternative way to study. However, the fast turnover rate of substrates, the lack of knowledge about bone fide substrates for a given protein, and the fact substrates can be regulated by different enzymes, all limit the utility of this approach[288]. In addition, substrate degradation normally involves numerous co-factors in a complex cellular environment which prevent clear and unambiguous data interpretation.

The ubiquitin cascade system is a very good example of this ambiguity since substrate ubiquitylation is co-regulated by three different enzymes (E1s, E2s and E3s)[86]. Misregulation of the ubiquitin system often cannot be unambiguously attributed to the activity of a single enzyme within the cascade. This is particularly pertinent when the activity readout (Ub chain formation or proteosomal degradation) is somewhat removed from the source of the misregulation. Moreover, substrate ubiquitylation, often used as a method to detect and infer misregulation within the ubiquitin cascade might only be observed (or enhanced to allow substrate detection) when accompanied by proteosomal inhibition. In many cases, small-molecule proteosomal inhibititor has been demonstrated to alter or modify the activity of other enzymes, making it difficult to attribute direct effects upon enzyme activity[289]. On the other hand, there are the examples of substrates being regulated by several different upstream enzymes. Thus, a probe molecule which could target the catalytic site of specific protein family, in order to monitor protein activity, would be a valuable tool to better understand the complex interactions and relationships involved in the regulation of enzyme activity and function.

2.2 What are Activity-Based Probes (ABPs) and how to make them?

2.2.1 Where did this concept originate?

Activity-based probes (ABPs) are typically small-molecule chemical entities that aim to target an enzyme of interest in an activity-dependent manner. They are designed to mechanistically interact with a specific target enzymes or protein enzyme and frequently covalently label the protein active site as a means to monitor protein activity and function. The covalent labelling of a protein by ABPs is normally achieved by having mutually reactive chemical handles present on the ABPs and on the protein of interest, most commonly, an electrophilic moiety on the ABPs and a native nucleophilic residue within the active site of the enzyme[290]. On the other hand, photocrosslinkers can be incorporated into the probe if the catalytic nucleophile in the target enzyme is missing [291]. The photocrosslinker produces a reactive radical intermediate which is capable of forming covalent bonds with nearby atoms when irradiated with UV light. As a consequence, the binding affinity of these probes toward the target proteins determines the selectivity and labelling efficiency of the photocrosslinker-conjugated probes.

The idea of ABPs was first described in 1961 by Barnard *et al.* in which they demonstrated the use of a radioactive chemical probe for cytochemistry study, which was a forerunner to modern ABPs[292]. Although the term called 'ABPs' is relatively new within scientific literature[293,294], researchers in the 1970s applied the concept of ABPs in their work. However, limitations of the technology and a lack of biological understanding often prevented early research at this chemical biological interface from adequately addressing complex biological questions[295-297]. In addition, early ABP-type studies were often based on small-molecule inhibitor development, for example, towards hydrolytic enzymes such as serine hydrolase, chymotrypsin and trypsin rather than use as discrete tools to directly address

biological questions related to those enzymes such as activity regulation[248,298,299]. Up until the year 2000, a wide variety of selective inhibitors targeting serine, cysteine and threonine proteases had been demonstrated[300] which not only provided insight and greatly aided drug discovery programs, but also provided the foundational knowledge to design and synthesise activity-based probes.

In general, there are three essential components to an activity-based probe (**Figure 2.1**); The *reactive group*, which is normally referred to as a 'warhead', is responsible for covalent, irreversible attachment of the ABP to the catalytic site in the enzyme of interest; A *recognition element* is typically required to deliver the ABP to the catalytic site of an active enzyme; and the reporter tag, which allows the formed ABP-protein complexes to be retrieved, enriched or visualized[288,290,293,301-303]. The pioneering research performed by the laboratories of *Dr. B. Cravatt* and *Dr. M. Bogyo* have provided many of the conceptual and technical insights that have underpinned the ABP field. To date, multiple types of ABPs have been developed for targeting numerous of enzyme classes[288,290,302,304-307].

2.2.2 The Warhead (Reactive functional group)

The functionality of the warhead is an important aspect in ABP design. The choice of warhead is often protein-target dependent but generally a directly-reactive electrophile is utilised when probes are designed for labelling serine, cysteine or threonine amino acid side-chains within an active site. The choices of electrophile can also be varied based on the target amino acids. In contrast, for enzymes which do not contain catalytic nucleophilic residues, for example, aspartyl proteases and metalloproteases, photocrosslinkers have been utilised for labelling in a mechanism-dependent manner[304,308,309]. However, the majority of the ABPs to

date make use of an electrophilic warhead[310]. The key element for 'warhead' design is that it should be sufficiently reactive for targeting the specific active site amino acids of interest but unreactive enough to prevent off-target labelling. As a result, one of the main challenges for the field is to find this balance between selectivity and reactivity.

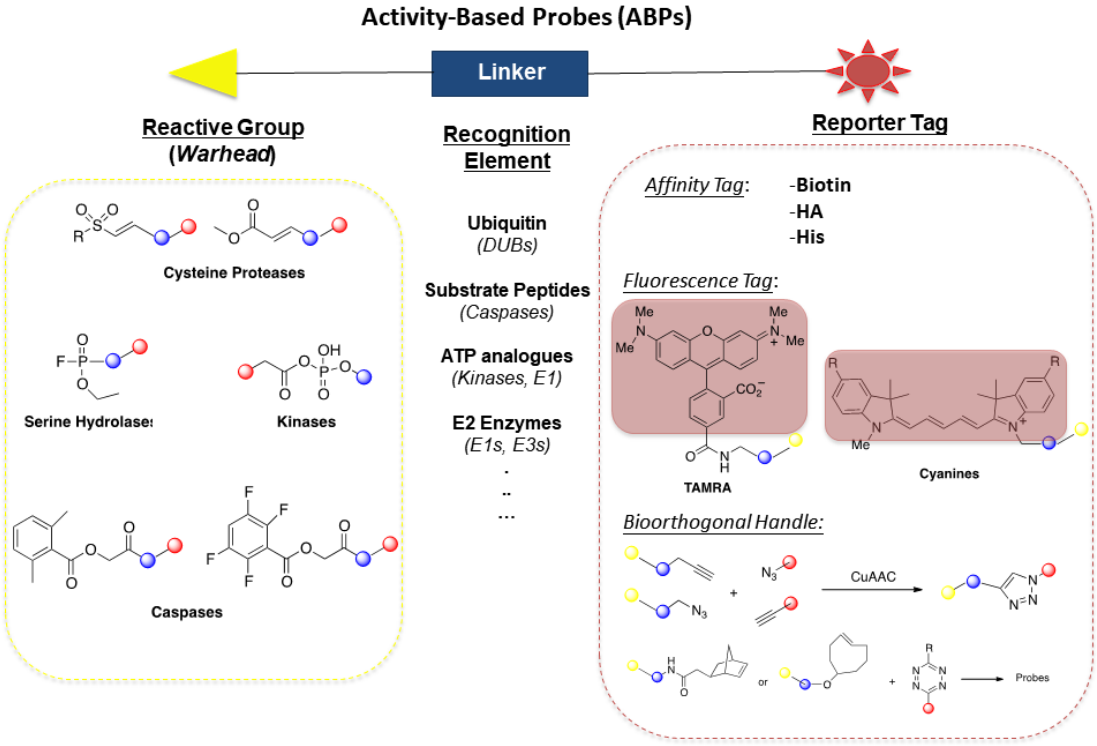


Figure 2.1 Activity-based probe construction.
Activity-based probes (ABPs) are composed of three essential elements: Reactive group, Recognition element and Reporter tag. A wide variety of warheads can be incorporated into ABPs for labelling different types of enzymes. Based on the purpose of the assays, affinity tags or fluorescence tags can be introduced upon ABPs for monitoring enzyme activity via proteomic methods or in-gel scan. In order to enhance ABPs specificity, the substrates peptides or full-length proteins can be used as the backbone for the ABPs. The introduction of bioorthogonal handles upon the ABPs allows post probe-labelling tag-conjugation. (Figure was reproduced from Niphakis *et al.*, 2014[302])

The labelling efficiency of ABPs is typically based on two factors: the affinity of the recognition element and the reactivity of the warhead with respect to the

protein target. The warhead is required to be in close proximity of the target site on proteins for the covalent reaction to occur between the catalytic nucleophile and the electrophilic 'warhead'. The affinity of the probes toward the target proteins (K_D) drives the recruitment of the ABP and the target protein (reversible interaction). If an ABP cannot contact with its target protein at the correct location and with suitable geometry to allow chemical crosslinking then the degree to which the 'warhead' can react is diminished significantly or completely. For example, the E2 binding site in the autoinhibited SMURF2 HECT E3 is blocked by its C2 domain which prevents the interaction with the E2~Ub [185]. Therefore, if a chemically engineered E2 was chosen as a recognition element, the warhead in the ABP would not be delivered to the active site of the autoinhibited SMURF2. When ABPs sit in the correct region on target protein, the irreversible covalent attachment is then driven by the rate of the warhead reaction and the catalytic nucleophile. High reactivity of the electrophile enhances the labelling rate of the ABP but needs to be balanced with the affinity of the recognition element to ensure optimal selectivity. For example, a weakly reactive warhead might not work with a low affinity recognition element but might be sufficient if the affinity was relatively high. Ideally, K_{on} and K_{off} rates should also be considered[311].

Sometime, promiscuous Cys-reactive warheads with no recognition element have utility, for example, such ABPs can serve as positive controls in inhibitor screening experiments of cysteine-containing proteins and can be used in a competitive role in profiling the off-target sites modified by covalent inhibitors[312]. Generic cysteine-reactive ABPs have found utility in determining the reactivity profile of cysteine and specific cysteine residues upon proteins due to their high intrinsic reactivity toward thiol nucleophiles. Catalytic cysteine residues often display higher reactivity compared to non-catalytic cysteines with a non-catalytic

function[312]. This often comes as a consequence of being situated in proximity to amino acids chain residues (for example, histidine), which serve as a general base, rendering the cysteine thiol group more nucleophilic in character.

PTMs or conformational changes of proteins which expose the catalytic cysteine for labelling result in highly reactive warheads undergoing labelling in a dose-dependent manner, allowing researchers to distinguish the roles of different cysteines within the same protein. On the other hand, the ABP AWP28, which contains a low reactivity acyloxymethyl ketone electrophile (**Figure 2.2**) as the warhead and the low affinity recognition element as the back bone, is a highly selective towards caspase-1[313].

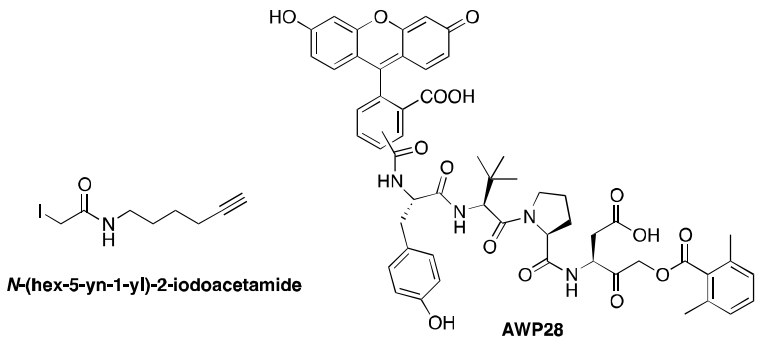


Figure 2.2 Structures of two functionally-different ABPs.
Iodoacetamide-based probe (*left*) and AOMK-based probe, AWP28 (*right*)**.**

In conclusion, though we have a wide variety choice for reactive 'warheads' to date, the design of the 'warhead' still requires caution, thought and novel design when dealing with novel enzyme classes. Further discussion about 'warhead' characteristics in building different types of ABPs will discussed below.

2.2.3 Design of the recognition element

As we discussed above, the affinity of ABPs also affect the degree of labelling efficiency when encountering various target enzymes. Hence, recognition element

design is the key for enhancing the specificity of ABPs to a specific enzyme target. The recognition element can be a small-molecule scaffold which has affinity for the target protein, a substrate peptide sequence, a protein domain, or even a whole substrate protein. An example of the latter is the ubiquitin moiety in DUB ABPs[303,306]. In most of the protease ABPs, the recognition element requires only a peptide sequence corresponding to the proteases substrate. DUB ABPs require the whole ubiquitin protein to interact with the catalytic domains in the DUB due the requirement of engaging the larger Ub-binding surfaces[18]. This is a common challenge when designing ABPs for Ub system enzymes.

Obtaining a highly selective recognition element for ABPs is non-trivial since the substrate protein would normally display some degree of similarity in terms of primary amino acids sequence or similarity in binding mode for closely related enzymes. The high background (off-target labelling) generated by such probes largely decrease their utility in imaging applications and in the development for translating the ABPs into biomarkers for detecting and diagnosing disease states[314,315]. To overcome these issues, related to sub-optimal recognition elements, several different methods have been pursued recently. Peptide inhibitor library screening of fluorescence-labelled substrates has been demonstrated as a good system to identify the highly selective recognition elements. A recognition element with a 30-fold increase in selectivity for caspase-3 has been identified from this platform[316]. The phage display system also allows screening of highly diverse libraries[317,318]. Although this approach was initially applied towards inhibitor development, the principles are similar for ABP recognition elements.

An obvious way to enhance the selectivity of the ABP recognition element towards a protein target would be to utilize the native full-length protein binding partner of the target protein. There are however many caveats inherent to

successfully following such a strategy. Even small, lower molecular proteins, for example, ubiquitin, or recently-developed bicyclic peptides, used as the recognition element for an ABP, cannot overcome cell permeability problems due to their size, electrostatic and steric problems (for example, bulky side chains and ionisable side chains). Examples of the cell-penetrating peptides have been applied to ABPs to address this problem and have been used for proteomic and imaging experiments[319,320]. Other strategies which have been used to enhance the cell permeability for small molecular inhibitors are also the options for solving this problem but with the concerns in disrupting reactivity and selectivity of well-designed ABPs, for example, the liposome-coating method[321]. However, within the recent advances in peptide-based inhibitor development, it is possible to design an ABP with exquisite selectivity even for highly homologues target enzymes.

2.2.4 Design of the reporter tag

The great advantage for using ABPs for profiling a complex cellular proteome is the ability to enrich, isolate and unambiguously identify the ABP-protein adduct from complex biological samples. An affinity handle as reporter tag on an ABP allows retrieval of the ABP-protein complex for further immunoblotting or proteomic analysis, for example, using epitope affinity tag such as the Human influenza hemagglutinin tag (HA-tag), or a polyhistidine-tag (His-tag), Avidin/Streptavidin-Biotin interaction. Bioorthogonal chemical reactivity handles (for example, azides, alkynes, strained alkene and tetrazine functionality[301]) can be used allowing two-step procedures where a sample is treated with ABP and affinity tag (for example, biotin) is subsequently attached by bioorthogonal 'click' reaction. (**Figure 2.1**). The use of epitope tags allow the facile validation of ABP labelling efficiency or selectivity toward panels of recombinant protein and mutants controls[322], and

various types of cell extracts[323]. The mutual reactivity of azides and alkynes is enabled by well-established copper-catalyzed 'click' reaction (Huisgen 1,3-dipolar cycloaddition)[324]. Rapid biorthogonal labeling between strained alkenes and tetrazine functionality, achieved by Diels-Alder cycloaddition reaction, can also be applied to capture and enrich ABP-protein complexes, post-ABP labelling[291,310]. Biotin-tag probes are largely applied in proteomic research since the high affinity of the non-covalent interaction with streptavidin can significantly reduce the non-specific binding of background proteins to the immobilized resins[291,325].

Apart from binding tags for enrichment purposes, fluorescence-tags also allow rapid analysis of ABP-protein target engagement by a simple in-gel fluorescence scanning[323,326]. Such methods allow for rapid analysis of the activity signatures in a particular experiment but offer little in the way of addressing target identification and absolute protein quantitation. The use of fluorescence-tags prevents time-consuming immunoblotting and MS-based analysis. It enables the rapid scanning and analysis of many proteins in the same assay, whether performing experiments with recombinant protein, or with protein lysates from cellular extracts. Furthermore, the high sensitivity and linearity of the fluorescence signal makes the study of labelling kinetics possible. For example, a fluorescence polarization activity-based protein profiling assay could be performed which has been demonstrated by Lahav *et al*[327]*.* A valuable application of application of fluorescently-tagged ABPs is exemplified in imaging assay[328,329]. This allows the tracking of enzyme activity during specific cellular stimuli or the cellular location. For example, tumor cells are normally found in hypoxic or low pH conditions which can quench the fluorescent signal. Therefore, a fluorescence signal is often location specific, and is only observed when the cellular conditions in which ABP is located are appropriate[330].

The choice of the tag utilized in an ABP also requires careful consideration as tag size and solubility affect the ABP utility. Fluorescence tags are useful reporter groups for ABPs but are also frequently bulky and hydrophobic molecules. Similar concerns have been raised for ABPs utilising biotin-tags[288]. Although a huge body of recent literature has demonstrated the utility of these ABPs, another factor to consider is the physiological relevance of the conclusion reached in the vast majority of ABPs were incubated with cell lysates or used *in vitro* and not *in vivo* experiments due to the problems with cell permeability. Also, relative protein concentration and loss of spatiotemporal organization and the biochemical regulation of enzymatic pathways (for example, through interaction with co-factors or regulation of PTMs) will undoubtedly differ between intact cells and cellular extracts.

2.3 How do ABPs detect enzyme activity: The analytical platforms of ABPs

In the field of chemical biology, an important question is how to appropriately apply the tools we have to address and answer important biological questions. The design and synthesis of ABPs has been addressed leaving the question, 'how are ABPs to be used effectively?' The most attractive characteristic about ABPs is their concomitant ability of isolating, enriching and enabling read out (enzyme catalytic activity) target of proteins of interest from a complicated proteome. This allows the systematic profiling of protein enzymatic activity from differing experimental setups (for example, varied cells types within different stimuli, cell-free assays or animal models), termed activity-based protein profiling (ABPP). As a result, choosing the appropriate platform setup for carrying out an ABP-based experiment is essential and several experimental platforms can be utilized with ABPs.

2.3.1 Gel-based platform

The most basic method to validate the labelling ability of ABPs toward target proteins is by gel-based detection. The ABPs can first react with either a recombinant protein mixture or a whole (or fractionated) proteome from cell extracts followed by resolution by using 1D or 2D SDS-PAGE[293,307,331,332]. The labelled enzymes can be visualized by in-gel fluorescence scanning (for fluorescent probes) or immunoblotting (for affinity-tagged ABPs, or Biotinylated-ABPs). The robust fluorescence signal rapidly provides a clear read-out of proteins activity. It allows for even an individual research group to be able to profile hundreds of thousand proteins (different types of proteins or various mutants) in a short period of time. If the probes are constructed based on a full-length protein structure, for example, Ub-VS (-vinyl methyl sulfone) or E2-VME (-vinyl methyl ester) [303,322,333,334], then the corresponding band-shift can also be easily detected by Coomassie staining or by affinity-tag immunoblotting such that a reporter tag is not necessary. Furthermore, biotinylated ABPs together with streptavidin affinity purification, gel separation and tandem mass spectrometry analysis can validate the identity of unknown labelled proteins. Although the 'gel-free' based assays provide a higher-resolution and depth of information towards decrypting a complicated proteome, the facile and simple gel-based assay still provides us with an easily interpretable data set. For example, in *Chapter 4*, I will describe the use of the novel E3-ABPs to dissect the effects of dual-phosphorylation events on the activation mechanism of the RBR E3. Gel-based assays in this work enabled the validation of the importance of p-Ub in a role to prime Parkin activation in cells[322]. Overall, in-gel based analysis is still a valuable format for ABP assays that allow various (but not exhaustive) experimental read-outs to be compared in parallel.

2.3.2 Tandem-mass spectrometry

Gel-based methods for applying ABPs are robust but often result in data which is 'low resolution' and limited in terms of depth of analysis. The 'low resolution' problem often stems from similarities in protein molecular weights (leading to poor resolution by SDS-PAGE analysis) as well as a lack of information regarding protein identity. The latter problem can be addressed by MS/MS analysis[322,334,335]. However, it is clear that data obtained from ABP experiments is limited only when using gel-based experiments exclusively, limiting the utility of ABPs. The introduction of liquid chromatography-mass spectrometry (LC-MS) strategies into ABP work-flows enables direct analysis of complicated proteomes and as such, vastly increased the amount of hitherto inaccessible data obtained from ABP profiling experiments[336].

There are several different ABP work flows that have been developed to address various biological questions. A common approach used is the generation of biotinylated ABPs and their subsequent incubation with a proteome of interest. Streptavidin facilitated capture and enrichment of ABP-protein complexes is followed by on-bead trypsin digestion. The digested-peptides are then analyzed by LC-MS/MS enabling identification of labelled proteins. This strategy has been able to identify more than 50-100 activated proteins from an individual proteome[301,336-339].

Label-free mass spectrometric analysis provides semi-quantitative parameters (for example, peptide counts) for data analysis, whereas peptide labelling methodologies are able to provide quantitative data, allowing in depth analysis of data provided by ABP labelling experiments. Quantitative mass spectrometric methodologies such as SILAC[291,340] (stable isotope labeling by amino acids in cell culture) and TMT (Tandem Mass Tag)[291] are capable of comparing the

relative levels of activated enzymes in parallel from three to ten proteomes simultaneously (SILAC and TMT 10plex, respectively). The state-in-art mass spectrometric technology, MS/MS/MS(MS^3), further provides a platform with a better accuracy in terms of protein quantification[341-343]. Parallel ABPP methodologies enable monitoring of protein activity in response to varied stimuli and form the basis of inhibitor/drug screening platforms in modern drug discovery efforts[302,344,345]. Notably, although multichannel peptide labeling technology, such as iTRAQ (isobaric tags for relative and absolute quantification) or TMT, enable multiple profiling of samples in parallel, labeling efficiency constraints and the resulting peptide signal loss should also be considered in such approaches[346]. Although SILAC technology offers increased efficiencies through isotope labelling, the limited number of parallel channels available (up to three) and the logistics and time employed in establishing and maintaining cultured cells are the major drawbacks of the method. SILAC is also incompatible with primary cells, tissues and biological fluids.

2.3.3 Activity-based imaging platform

The installation of a fluorescent tag on ABPs make useful tools for locating and tracking proteins in the context of live cell[347]. This technology not only provides the cellular location of proteins in context but is a truer representation of intrinsic protein activity as ABP labelling occurs in the context of native protein concentration and spatiotemporal localization. Moreover, protein labelling occurs in the presence of physiologically relevant protein-protein interactions. More recently, Dr. M. Bogyo's group applied this methodology in an animal model to develop a noninvasive imaging system to specifically monitor tumor-associated cysteine proteases, cathepsins, in live mice[330].

2.4 What can ABPs do? ---The application of ABPP

2.4.1 Target discovery

Protein target discovery and identification has been the main experimental application of ABPP[290,348,349]. Application of ABPP in comparing varied biological samples under differing experimental conditions has the ability to uncover protein candidates that are causative of a particular cellular phenotype or disease state. One of the advantages of comparative ABPP is that the data obtained is directly correlated to the intrinsic activity of protein target which might not only be regulated simply by varying protein expression levels but also by posttranslational regulation, for example, via PTMs and/or protein conformational changes. Such methodologies therefore are well suited to address biological questions related to the study of otherwise difficult to study enzymes, for example, protein proteases and E3 ligases. As discussed previously, these enzyme families can be often found in an inhibited form under resting conditions and are only activated in response to specific substrate localization, co-factors engagement or other specific cellular stimuli[81,171,205,206,256,350]. Furthermore, the accuracy and sensitivity of tandem mass spectrometry approaches used in ABPP enables the identification of hitherto undetected low-abundance proteins which have evaded detection by other methods. The application of such methods will undoubtedly provide new perspectives and outlooks on multiple disease-related biology which may translate into novel therapeutic strategies, particularly in relation to novel biomarker development.

2.4.2 Competitive ABPP for inhibitor screening

Competitive ABPP strategies enable the direct application ABPs towards furthering drug discovery. Because ABPs have been designed to specifically target the active

site of particular proteins and protein families, small molecules, developed through tradition medicinal chemistry approaches that target the same protein or protein families can be used in competition with ABPs for proteins labelling therefore allowing elucidation of inhibitor specificity and selectivity[302,307,332] (**Figure 2.3**). In addition, this strategy can also uncover off-target issues associated with small-molecule inhibitor development through comparison of parallel data-sets with and without inhibitor treatment. Furthermore, dose-titration experiments using small-molecule inhibitors in ABP gel-based assays provide a facile way to obtain the inhibitor affinity information toward specific target proteins. Once again, like previously mentioned ABP approaches, the major advantage of competitive ABPP as applied to small-molecule inhibitor development is the ability to interrogate native proteomes whilst minimizing concerns regarding experiments carried out under non-physiologically relevant conditions (for example, effects related to PTMs), apparent in assays using only recombinant protein. Overall, competitive ABPP not only provides a platform for inhibitor development and screening but also generates a useful and rich protein target database, regardless of the target protein or inhibitor under study.

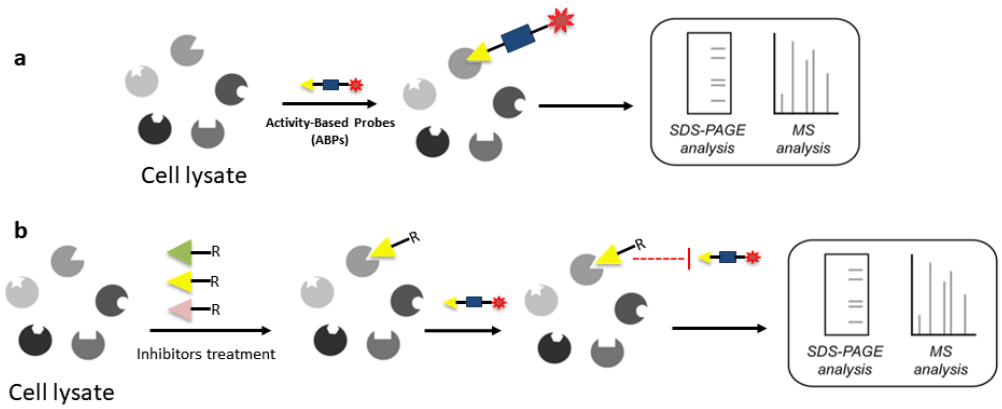


Figure 2.3 The work flow chart of competitive ABPP strategy.

(**a**) The standard ABPP assay. Various types of cellular extracts are incubated with ABPs and the ABP-labeled proteome are subsequently analysed by gel-based assays or comprehensive proteomic strategies. **(b)** In contrast to the standard workflow, the cellular extracts or the intake cells are treated with selected inhibitor prior to the ABPs incubation. The leading inhibitors for targeting specific protein or protein families can be obtained via the comparison of the datasets from two different strategies.

2.5 Case studies of different ABPs

To date, numerous types of ABPs have been developed for the study of various proteins including serine hydrolases[293,307,331,351,352], cysteine proteases[315,330,348,353], DUBs[303,354], metallohydrolases[304,355], kinases[356-358], palmitoyl transferases[359]and other examples[288,302]. The diverse protein families studied by ABP profiling has been beneficial for uncovering fundamental biology related to the proteins in question but has also been a significantly useful strategy for drug development programs[302,360]. Since this thesis is focussed upon on the ubiquitin system, the following discussion will focus on two pioneering ABP families of relevance, serine hydrolases and cysteine proteases ABPs.

2.5.1 Serine hydrolases ABPs

The serine hydrolases family consist of a wide variety of enzymes including proteases, peptidases, lipases, esterases, and amidases which are predicted to make up approximately 1% of mammalian genomes[361]. Serine hydrolases have also

been demonstrated to be important to numerous cellular events, for example, angiogenesis[362], neuronal signaling[363,364], inflammation[365,366], and cancers[367,368]. Although the various structures and sequences among many different serine hydrolases are often varied, this protein family share a similar mechanism of action in order to hydrolyse the appropriate ester or amide bond functionality found in their substrates. The base-activated nucleophilic serine residue targets the substrate ester and amide functional group at the sp^2 carbon. Most commonly, a tetrahedral acyl-enzyme intermediate is formed that is then subsequently hydrolysed by a water molecule[361]. In addition, serine hydrolases have been shown to not only target proteins as substrates, but also polar metabolites, lipids, posttranslational protein modifications and peptides[369]. The high nucleophilicity of the catalytic serine residue present in serine hydrolases make it an tractable chemical target for 'warhead' electrophiles, for example, aryl phosphonates, sulfonyl fluorides, carbamates and fluorophosphonates (FPs)[290,302,368,369] (**Figure 2.1**).

FPs had first been discovered as a mechanism-based inhibitors of serine hydrolases[293,370]. Whilst several other types of other serine hydrolase inhibitors have been demonstrated[371-373], FPs preserve the advantages of broad coverage of the serine hydrolase family and often demonstrate limited cross-reactivity toward other enzymes[361]. The electrophilic FP centre is highly reactive to the serine side chain oxygen atom. Labelling with FPs further stabilize the probe-labelled intermediate by mimicking the enzyme-substrate tetrahedral intermediate as a stable adduct. The above features endow FPs with a broad target selectivity within the serine hydrolase families[361]. In addition, FPs higher reactivity toward activated serine hydrolases over inactive[374]. By demonstrating labelling profiles that distinguish active from non-active enzymes supports the statement that FPs are

bona fide activity based probes for serine hydrolases. A variety of FP-based ABPs have been shown to be able to label more than 80 distinct serine hydrolases in both human and mouse proteomes[290,336,337,375]. Notably, more than one-third of these serine hydrolases had remained uncharacterized before being identified by serine hydrolase ABPP. For example, a serine hydrolase, MAGL, that degrades 2-arachidonoylglycerol (2-AG), an endogenous ligand for cannabinoid receptors, was discovered from mouse brain by FP probes[349]. Cravatt and colleagues have contributed significantly to this field and a huge body of data have been produced by their efforts to date.

Comparative ABPP can be considered a straightforward proteomic method to proceed. Indeed, such FP-based serine hydrolase ABPs have been applied to different biological models for profiling studies. These models include atherosclerosis[376], immune cell activation[377], nervous system signaling[378] and cancers[336,368,375,379]. There are several good cases which exemplify how to apply serine hydrolase ABPs to discover and validate a novel serine hydrolases and to apply findings from those experiments towards small-molecule inhibitor development. Jessani *et al.* first applied fluorescent FP probes to profile a panel of human breast carcinoma and melanoma cell lines[380] from NCI-60. They performed in-gel fluorescent scanning assays to obtain a quantitative comparison of enzyme activities for serine hydrolases. The apparent increased level of some serine hydrolases present in the assay was demonstrated and displayed restricted activity patterns amongst the different cell lines analyzed. Subsequent follow up analysis was performed utilizing biotin-tagged FP ABPP approaches with cellular extract proteomes followed by LC-MS analysis for protein identification. Along with this approach, different groups had discovered more than 50 serine hydrolases and categorized them into subgroups based on their apparent activity signature. As a

result of such studies, it was found that the activity of certain serine hydrolases was consistently increased in aggressive cancer cell lines[366,380]. These enzymes included the known cancer biomarker, urokinase plaminogen activator (uPA), but more interestingly two enzymes, KIAA1363 and MAGL, which had not been previously linked to cancer malignancy[290,360]. Application of quantitative mass activities-based proteomics, further demonstrated that KIAA1363 activity is significantly enhanced in an aggressive cancer cell line (estrogen receptor-negative [ER^-] primary human breast tumors). Chiang *et al.* later developed an inhibitor of KIAA1363 using the competitive ABPP strategy[381]. Surprisingly, by global metabolite profiling of inhibitor-treated cancer cell lines, levels of an unusual class of lipids, monoalkylglycerol ethers (MAGEs), had been found to be decreased in inhibitor-treated cells. It was subsequently determined that KIAA1363 is the hydrolase responsible for hydrolysis of the MAGE precursor, lipid 2-acetyl MAGE, thereby regulating the MAGE levels in cancer cells. The conversion of MAGE into lysophosphatidic acid (LPA) by cancer cells has also been linked to tumor progression[382]. The reduction in tumor cell growth using both KIAA1363 RNA interference (RNAi) or small-molecule inhibitor further indicated the role of KIAA1363 in aggressive cancer cell progression.

The other uncharacterized serine hydrolase that had been identified by FP-probes was monoacylglycerol lipase (MAGL) which has also been found in several aggressive cancer cell lines with increased activity[360,383]. The increased activity of MAGL in cells was further validated by monitoring the hydrolysis level of its substrate, 2-Arachidonoylglycerol (2-AG) combination with MAGL inhibitor (JZL184) treatment [384]. An elevated activity level of MAGL has been shown to be an important feature for aggressive tumor progression and that inhibitor treatment also significantly reduces the tumor growth.

These above examples perfectly demonstrated the utility of FP ABPs applied to the study of seine hydrolases to further progress fundamental and translational biology. The FP probes have progressed from protein identification (within disease or stimuli-specific model systems) to competitive ABPP approaches, aiding inhibitor development through to more sophisticated assay formats which enable discovery and validation of target biology. The success of FP probes in the serine hydrolase field has largely enhanced the research of ABPs application in various biological systems. From both a methodological and a conceptual stand point.

2.5.2 Cysteine protease ABPs

Like serine hydrolases, cysteine proteases are another huge enzyme family found in both prokaryotic and eukaryotic systems. These enzymes have been shown to be involved in regulating multiple cellular functions and having involvement in various diseases, for example, arthritis[385], apoptosis[386-388], tumor progression[389,390], cell cycle regulation[391] and even parasitic invasion[392]. Like serine hydrolases, the distinct catalytic mechanism of cysteine proteases makes them susceptible to labelling by differentially designed electrophiles, for example, epoxides and diazomethyl ketones for cathepsins[315,393], α-haloketones and acyloxymethyl ketone for caspases[348,394] and vinyl sulfones for DUBs[303,305,306] (**Figure 2.1**). The histidine-activated sulfhydryl nucleophile on the catalytic cysteine forms a covalent intermediate with the carboxy terminus of the substrates. The intermediate is then cleaved by the action of a water molecule. Due to the high intrinsic reactivity of the catalytic cysteine of many cysteine proteases, many small-molecule tool compounds and inhibitors (both covalent and non-covalent) have been developed for their study[302,367].

Cysteine proteases can be divided into several different clans based on their evolutionary origins[395,396]. The CA clan subgroup of cysteine proteases has previously been heavily studied due to the many essential roles they occupy in physiological regulation[397,398]. In 1976, a specific cysteine protease inhibitor, E-64, had been discovered from the cultures of *Aspergillus japonicus*[399]. The reactive epoxide warhead present in the E-64 displays high reactivity and specificity towards the papain cysteine protease family[400]. E-64 and its derivatives have been applied widely, for example, DCG-04, a E-64 derivatives which containing a P2 leucine residue, has been used in the studies of tumor progression[389] and prohormone processing[401]. E-64 has significant inhibitory effects upon cathepsins and has minimal cross-reactivity toward other cysteine proteases[402]. Clan CD proteases are the second largest family of cysteine protease, which include legumains, caspases, separases and gingipains. Caspases in particular have been extensively studied due to their essential role in regulating apoptosis[403]. In fact, the cysteine protease ABPs with aldehyde warheads were used to discovered the first caspase, caspase 1[404]. The early caspase ABPs were compromised of an acyloxymethyl ketone (AOMK) as the warhead. In 2005, the AOMK-based probes were demonstrated to have broader coverage of CD clan proteases, including the lysosomal cysteine protease, legumain and cathepsins[348].

A central dilemma in ABP design has always been the balance to retain ABP diversity for broad target discovery but also display selectivity for proteins of interest. By incorporation of optimized peptide sequences into AOMK-based probes, Bogyo *et al.* have previously demonstrated that it is possible to tune the reactivity of AOMK-based probes to selectively target specific proteases [405]. tAB50-Cy5 is an AOMK-based probe which carries a fluorescent cyanine fluorophore, Cy5, reporter tag and also a cell penetrating peptide trans-activating transcriptional

activator (Tat) peptide sequcence[406]. It labels both caspases and legumain proteases. Edgington *et al.* applied this probe to detect and quantify the activity level of caspases and legumain *in vivo*[314,405]. Probe activity was validated in a dexamethasone-induced apoptosis model in $CD4^+CD8^+$ thymocytes of mice. Both caspases and legumain were labelled with the tAB50-Cy5 probe. The quantitative data gathered further indicated that the activity of legumain was low after dexamethasone treatment. However, the activity of caspase-3 was enhanced and reached the highest levels at 12 hour dexamethasone treatment. In addition, subsequent analysis by flow cytometry and comparison with ABP-fluorescence data, indicated that the tAB50-Cy5 was only accumulating within the apoptotic cells. Together, with further in-gel fluorescence scanning and biochemical assay, the study provided a powerful platform to monitor protease activity both *in vivo* and *in vitro*.

Cysteine protease ABPs also contributed to the understanding of tumorigenesis. Combined analyses of the global gene expression in conjugation with ABPP, Joyce *et al.* identified the high activity of cathepsins Z, B/L, and C in tumor progression processes[389]. The application of fluorescently-tagged ABPs allowed the unambiguous localization of the cathepsins within tumor cells. Taken together with data obtained from inhibitor treatment studies of tumor cells, cathepsins have been found to regulate tumor development by affecting invasive proliferation, tumor vascularity and angiogenic switching[290,389]. The other papain family-specific probe, DCG-04, was also applied in ABPP methodologies and identified that the activity of falcipain 1 is enhanced during *P. falciparum* invasion[392].

As previously discussed in the context of serine hydrolase ABPs, cysteine protease ABPs have also been used for inhibitor screening and development.

Cwp84 a papain family protease which regulates *C. difficile* host infection [407]. The inhibition of Cwp84 enhances the host immune response to block the *C. difficile* invasion, making Cwp84 a good target for drug development.

2.5.3 DUB ABPs

Deubiquitinase (DUBs) are another cysteine protease family which have been studied extensively by ABPs[303,408-410]. DUBs play the 'eraser' role in the ubiquitylation system, either trimming or removing the ubiquitin chains from ubiquitylated substrates[43]. DUBs have also been shown to regulate multiple cellular functions[411], for example, membrane trafficking[412], secretory pathway control[413], Wnt signalling, TGF-β signalling and transcription[414] and RNA processing[415]. In addition, the coordination with E3 ligases to further regulate the downstream ubiquitin signalling pathways also serves to highlight the important and multi-faceted role played by DUBs activity. There are nearly ~100 DUBs in the human genome[416] and these can be divided into two main categories based on the catalytic mechanism of substrate hydrolysis; cysteine protease DUBs or metallo-protease DUBs[96]. The cysteine DUBs use a familiar mechanism for (iso)peptide bond cleavage, with similarities to other cysteine proteases where a catalytic triad is employed[417]. Structural studies have shown that the proximal histidine near the catalytic cysteine acts as a general base to deprotonate and activate the thiol group on the cysteine for the nucleophilic attack towards the carbonyl of the (iso)peptide bond[96] (**Figure 2.4**). The conserved catalytic cysteine and chemical environment between cysteine DUBs make this nucleophile suitable for exploitation by ABPs approaches.

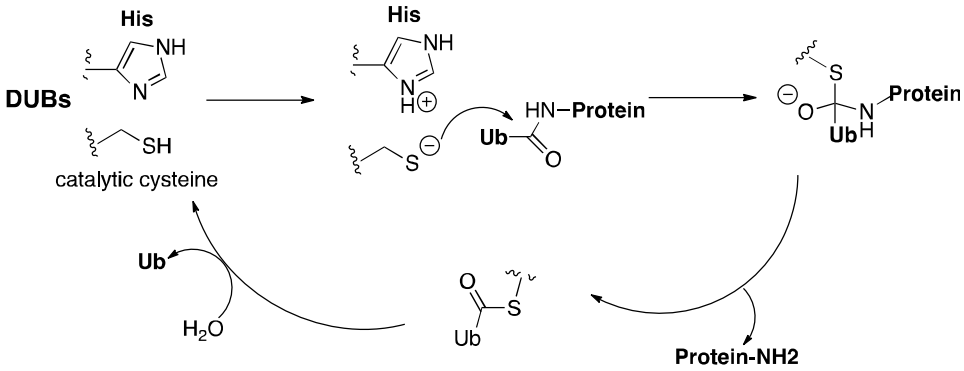


Figure 2.4 The mechanism of DUBs.

The proximal histidine side chain lowers the pKa of the thiol-group of the catalytic cysteine upon DUBs. The nucleophilic attack towards the peptide bond is then performed and further releases the Ub moiety and target proteins. (Figure was reproduced from Komander *et al.*, 2009[96])

Based on the structural folds identified from analysis of DUBs by structural methods, they can be further classified into sevendifferent families[418]. Six of the seven families of DUBs are cysteine protease DUBs there being Ub C-terminal hydrolyses (UCH), Ub-specific proteases (USP), ovarian tumour proteases (OTU), Josephins, MINDY (for MIU-containing novel DUB family) and ZUFSP[418-420]. JAB1/MPN/MOV34 (JAMM) DUBs is belong to metallo-protease DUBs.

The first covalent DUB ABP was developed in 2001 by Borodovsky *et al*. and contains a single full-length ubiquitin as the recognition element and a vinyl sulfone 'warhead' connected to the penultimate C-terminal Gly residue as the electrophile[408]. This first generation of DUB ABP labelled six different putative yeast deubiquitylating enzymes in a recombinant enzyme panel. The authors further identified that the mammalian homolog of yeast Ubp6p, USP14, was bound to the proteasome. To date, a wide variety of DUB ABPs have been developed and applied to different types of ubiquitin research. Ub-aldehyde (Ubal) and Ub-nitrile (Ub-CN) were heavily used in early DUB inhibitor development studies whilst Ubal was also used to solve the first DUB-Ub complex structure[421].

The yeast UCH, Yuh1, was labelled with Ubal and mimics the tetrahedral intermediate during the hydrolysis pathway. The authors also demonstrated that the UCH undergoes a structural rearrangement upon ubiquitin binding which increases its enzymatic activity. However, Ubal and Ub-CN are not formally considered as ABPs since their labelling mechanism is reversible and limits their compatibility for ABP profiling. It was the Ub-vinyl methyl sulfone (Ub-VS) that was initially to demonstrate irreversibly label DUBs in cell lysates[306]. The distinct activity of vinyl methyl sulfone (-VS) has also been applied to the construction of ubiquitin-like protein probes, for example, the NEDD8, ISG-15 and SUMO-1 have all reported to be conjugated with –VS and used as ABPs against their respective (iso)peptidases[422,423]. Following on from the earlier success of Ub-VS, significant effects have been applied to warhead optimization. For example, the Ub-propargyl amide (Ub-Prg)[333,424], Ub-vinyl ethylsulfonate (Ub-OEtVS)[425], Ub-α-amino-β-lactone (Ub-Lac)[425] and more recent Ub-dehydroalanine (Ub-Dha)[426] (**Figure 2.5**). However, vinyl methyl sulfones (VS) and vinyl methyl ester (VME) 'warheads' are still the two most widely used reactive groups utilised in DUB ABPs.

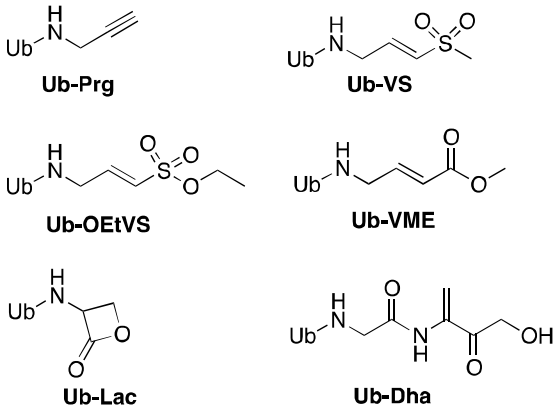


Figure 2.5 Different types of DUB ABPs

Different types of electrophiles are covalently appended to ubiquitin$_{1-75}$ protein (without C-terminal Gly76) for making DUB ABPs.

Another highly reactive warhead, AOMK, is also conjugated with ubiquitin as the DUB ABP, $UbTF_3BOK$. However, it labels not only DUBs but also E1s, E2s and E3s which also carry catalytic residues[248]. Therefore, the profiling data generated by DUB ABPs need to be interpreted with caution. Furthermore, as E2 recognition is a requirement for E1 and E3 activity, activity data inferred from studies on these enzymes should also be treated with caution. A recent report by Hewings *et al.* indicated that multiple non-catalytic cysteines on DUBs were also labelled by DUB ABPs[420]. However, HA-Ub-VS was still capable of validating the activity of USP9X in melanoma cells , a DUB whose activity is regulated by phosphorylation[427]. The enhancement of HA-Ub-VS labelling with USP9X was observed in stimulated $CD4^+$ T cells without an increases in USP9X expression level, emphasising the value of ABPs.

DUB ABPs have been applied toward the discovery of new DUBs, characterization of DUB selectivity, stimuli-regulated DUB activity profiling and DUB inhibitor screening[303]. USP14 was identified by HA-Ub-VS and Ubal[306] was used to identify OTU family DUBs, otubain 1 and otubain 2[428]. Pruneda *et al.* recently applied the Ub-Prg ABP in conjunction with fluorescent substrates to further validate the deubiquitinase and deneddylase activity of proposed bacterial isopeptidases[429]. Non-hydrolyzable DUB ABPs, diUb-AMC (Ub-7-amido-4-methylcoumarin), combined with Ub-Prg provided the data necessary to explain the mechanism behind the K11 chain cleavage preference by OTUD2 and OTUD3[430]. In 2013, Gao *et al.* further applied the HA-Ub-VS probe to demonstrate that USP7-regulated deubiquitylation plays a vital role in early adipogenesis which mediated by histone acetyltransferase, TIP60[431]. As previously discussed with respect to other proteases, ABPP approaches for DUB inhibitor screening and development have also heavily relied upon the use of appropriate DUB ABPs[339,432]. The only DUB

inhibitor currently in clinical trials, VLX1570, has been shown to selectively label USP14 but not UCHL5. The selectivity preference was verified using HA-Ub-VS in a competitive gel-based ABP assay. VLX1570 treatment also induces the accumulation of polyubiquitin conjugates and the apoptotic response in multiple myeloma cells[433]. The DUB ABPs were also applied in two individual research lab for USP7 inhibitor development[345,434]. The above examples demonstrate the necessity and utility of DUB ABPs for the progression of our understanding of DUB biology and inhibiter development.

2.5.4 E1 ABPs

As previously discussed in *Chapter 1*, apart from DUBs, the other enzymes involved in the ubiquitin system (E1s, E2s, E3s) can also transfer ubiquitin via catalytic cysteine residues. Compared to cysteine protease enzymes[288,312], E1, E2, E3 ABP development has only recently been investigated, due in part to the comparatively suppressed reactivity of the catalytic cysteines among these enzymes (compared to cysteine proteases for example)[288]. Furthermore, unlike DUBs that recognize the conserved ubiquitin moiety or distinct ubiquitin chain topology for substrate recognition, E1s/E2s/E3s can display multiple preferred active combinations and have multiple substrates. The lack of a complete mechanistic and biological understanding of these enzymes that on some degree has delayed the development of suitable ABPs. However, recently, several groups have reported the development of ABPs for targeting these enzymes and have discrete demonstrated their application in the ubiquitin field[303].

Though limited in number, the first examples of E1 ABPs have had great impact on ubiquitin research. Olsen *et al.* synthesized and conjugated a non-hydrolysable mimic of the acyl adenylate intermediate (AMSN) onto human

SUMO1 to make SUMO-AMSN[108] (**Figure 2.6**). This conjugate can mimic the intermediate produced during catalytic cysteine attack of the SUMO-AMP intermediate. SUMO1 was conjugated to a 5'-(vinylsulphonylaminodeoxy)adenosine (AVSN) which is able to label the E1 catalytic cysteine like a suicide substrate (**Figure 2.6**). The authors later incubated the SUMO E1, SUMO Activating Enzyme Subunit 1 (SAE1), with SUMO1-AVSN and resolved the conjugation product via SDS-PAGE analysis. A redox stable product was observed and corresponded to the approximate molecular weight of the E1~Ub adduct. Control reactions using catalytically dead E1 further supported the conclusion that the SUMO1-AVSN was covalently labelling the E1 catalytic cysteine residue. The subsequent solving of the E1 structure using both SUMO1-AVSN and SUMO1-AMSN ABPs, uncovered an required E1 conformation change for catalytic cysteine attack of SUMO-AMP. The E1:SUMO-AMSN (E1:SUMO-AMP mimic) adopts an 'open' conformation compared to E1:SUMO-AVSN (E1~SUMO-AMP trapped intermediate). The ATP-binding site is exposed and the catalytic cysteine is buried in the Cys domain. In contrast, E1:SUMO-AVSN adopts a 'closed' conformation and a subsequent 130° rotation of the Cys domain is required to allow successful ubiquitin transfer. Although the SUMO-AVSN labels the SUMO E1 in an activity dependent manner, the lack of reporter tag renders it unsuitable towards profiling E1 activity in ABP-type approaches, although this could be easily introduced. An *et al.* later developed two types of E1 ABPs. Initially, small molecules probe, referred as ABP1 in the literauture[435], was generated and is based on the structure of MLN4924, an inhibitor of the NEDD8 E1, NAE. An analogue of ABP1, ABP3, was also developed and utilized an alkyne handle for reporter tag conjugation. ABP1 has been demonstrated to label a variety of E1s *in vitro* but ABP3 is only able to label Ub and the NEDD8 E1.

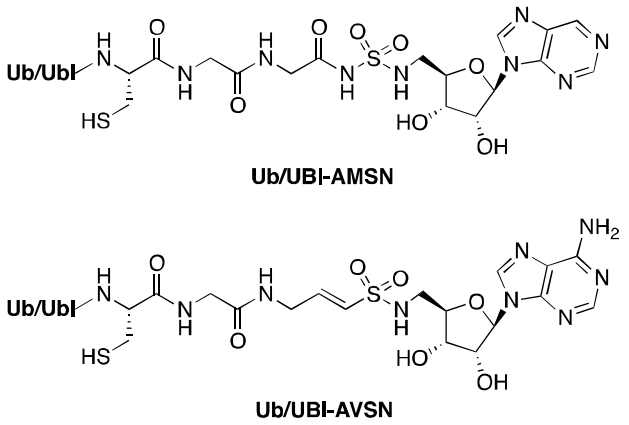


Figure 2.6 The structures of E1-Ub mimic (Ub/Ubl-AMSN) and E1 suicide probe (Ub/Ubl-AVSN).

In 2016, An *et al.* developed a panel of Ubl–AMP electrophilic probes which introduced a dehydroalanine (Dha) functional group as a 'warhead'[436] (**Figure 2.7a**). The facile synthetic methods employed and the native C-terminal sequence of Ub/Ubl allowed these probes to be applied in ABPP experiments. Using a similar construction strategy, the ubiquitin-like protein, microtubule associated protein 1 light-chain 3 (LC3)/ATG8 has also been used for making the LC3-based probe, which has been demonstrated to label autophagy E1-like enzyme, ATG7. The major drawback of these probes is the susceptibility to DUB hydrolysis by IsoT(USP5) which somewhat limits their utility in experiments using cell derived samples. In a different approach, Stanley *et al.* designed E2-based ABPs which were able to efficiently label E1 enzymes[334]. The novel approach developed a small-molecule, termed a tosyl-substituted doubly activated ene (TDAE), that can be conjugated with multiple E2s at their catalytic cysteine generating an electrophilic activated vinylsulfide (AVS) warhead which is then able to react further with a second nucleophile (**Figure 2.7b**). The AVS warhead is thereby able to conjugate two cysteines that come into close proximity. Such probes are useful in studying enzymatic reaction where cysteine-cysteine juxtaposition is a crucial requirement for catalytic activity. Determination of the rate of the second cysteine addition to

the AVS indicated that the -CN functionalised AVS reacts slower than a -VME AVS in the context of model reaction with glutathione (k=1.52x10^{-3}M^{-1}s^{-1} and 7.84x10^{-3} M^{-1}s^{-1}, respectively)[334], therefore showing the warhead reactivity was tunable. In the same study, The UBE2N-CN was demonstrated to be able to label the human E1, UBA1 in an activity-dependent manner that was dependent on the UBA1 catalytic cysteine. UBA1 C632A mutation abolished probe labelling. UBA1 labelling assays were also performed by incubating the UBE2N-CN probe with Human Embryonic Kidney 293 (HEK293) cell extracts and followed by immunoblotting. The E2 AVS-probes labelled greater than 90% of the available UBA1 and demonstrated that probes unique capability of profiling the E1-E2 transthiolation event. This facile installation of the 'warhead' at the E2 catalytic cysteine in this approach enabled the obvious leap to testing beyond E1 enzymes to the more numerous E3 ligase families, for example, HECT and RBR E3 ligases. The failure of HECT/RBR E3 labelling by E2-based ABPs was convincingly attributed to the lack of the ubiquitin moiety upon the E2. However, sub-stoichiometric labelling was observed with non-canonical E3s (for example, bacterial effectors such as NleL), where obviously the strict recognition requirement for the E2¬Ub conjugate are somewhat more relaxed.

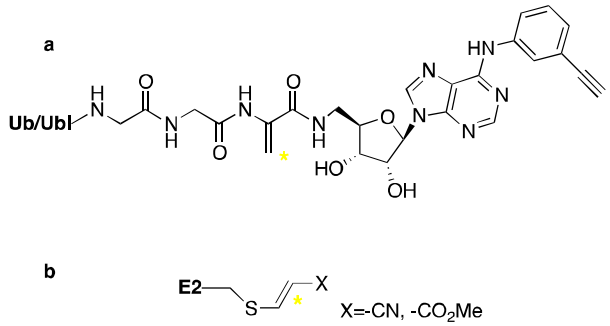


Figure 2.7 The structures of two novel E1 ABPs.

2.5.5 E3 ABPs

Validation of E3 ligases activity has proved challenging as traditional assays are often hampered by deconvolution problems associated with the presence of other critical components of the ubiquitin cascade (E1, E2, Ub, ATP). However, the catalytic cysteines on HECT and RBR E3s, which accept Ub from E2s by transthiolation reactions are potentially suitable sites for labelling by an electrophile-containing probe. Only very recently have the first E2-based E3 ABPs been developed by Pao *et al*.[322]. Part of my research involved optimization of the ABPs where the foundations were reported in Stanley *et al.*[334] to enable labelling of the RBR E3 ligase, Parkin. This research will be discussed in detail in *Chapter 4*. Different types of ABPs and strategies to monitor E3 activity were also reported after the initial publication of research by Virdee lab. For example, Krist *et al.* have demonstrated that the UbFluor probe could be applied to monitor and validate the activity of HECT and RBR E3 ligases[437,438] (**Figure 2.8a**). Although the mechanism of probe labelling is not based on E2-E3 transthiolation, UbFluor can still act as a Ub-thioester mimic to reacting with the catalytic cysteine on HECT/RBR E3s. The fluorescence tag on these probes can be monitored by fluorescence polarization assays which makes high through putscreening of E3 activity using this technique feasible and also potentially accelerating E3 small-molecule inhibitor/activator development via use of competitive ABPP approaches. Of note, this approach is not compatible with cell extracts and does not take into account E2 binding. Another strategy developed by Mulder *et al.* also demonstrated the ability to label E3 ligases both *in vitro* and in cell lysates[439]. The novel ubiquitin-based probe, Ub-Dha, can be activated by E1s and label the downstream E2s and E3s in an ATP-dependent manner (**Figure 2.8b**). The distinct design which features -Dha as a 'warhead' and a native carboxy group at the ubiquitin C-terminus allows E1

enzymatic activations. The activated Ub-Dha is conjugated with E1 via thioester bond and subsequently the Ub-Dha can either form a stable E1-Ub conjugate (by covalent binding between -Dha and the catalytic cysteine) or perform a transthiolation step, transferring the activated Ub-Dha to an E2. Furthermore, the same reaction pathway can occur allowing the potential transfer of Ub-Dha from E2 to E3. In this manner, Ub-Dha is potentially able to detect the activity of all three enzyme families at the same time. However, there are still problems with assay deconvolution, suitable controls and the fact that Dha electrophiles have been shown to be somewhat promiscuous and non-cysteine specific under physiological conditions[440].

Research into the generation and application of E3 ABPs is only in its infancy and there are many technical problems that require addressing for many approaches to enable the maximum impact of these technologies. More comprehensive and novel assays and experimental platforms will undoubtedly be reported in the future.

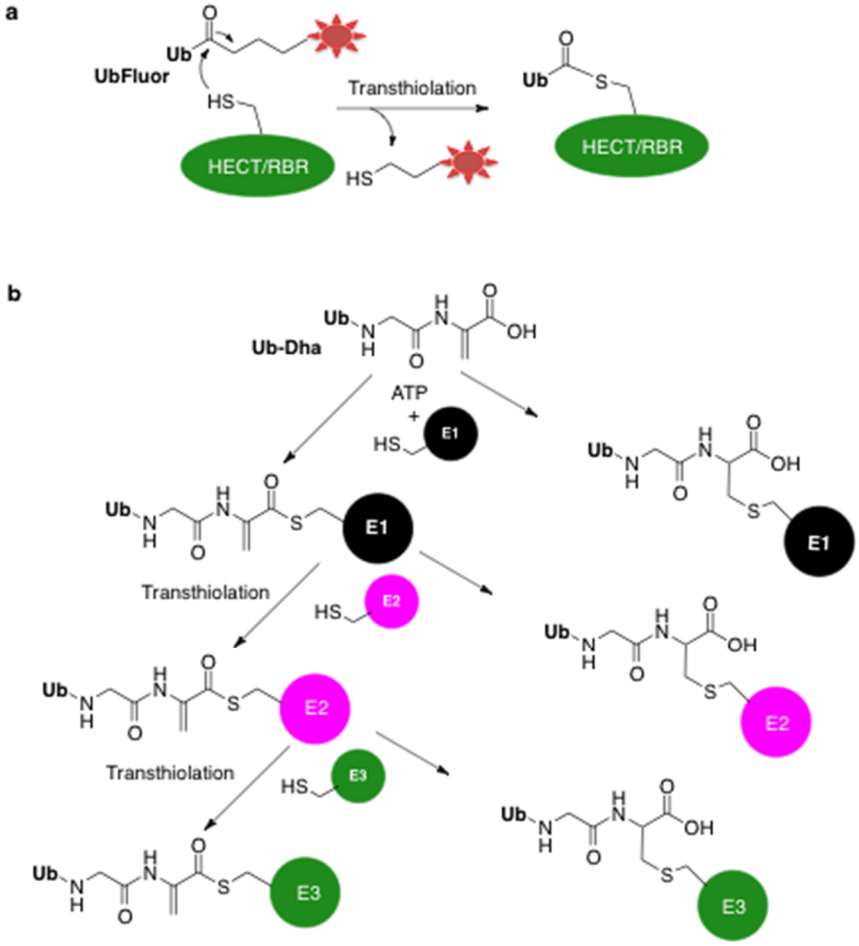


Figure 2.8 The structures and working mechanism scheme of recently developed E3 ABPs.

(**a**) The UbFluor E3 ABP performs a native transthiolation step within the target E3s. It subsequently releases the fluorescent tag from the probe which can be used as the signal to indicate the activity of target E3s. (**b**) Ub-Dha, a 'cascading' ABP, which contains a native Ub C-terminal carboxy group and a Dha electrophile as a 'warhead'. The Ub-Dha can either react within enzymes as a wild type Ub does, that is, transferring itself from E1 through E2 to E3; or to be trapped as the stable conjugates at each step (E1-Ub, E2-Ub, E3-Ub). The Ub-Dha labels the E1s, E2s and E3s in an ATP-dependent manner.

Chapter 3. Material & Method

3.1 General material

3.1.1 Reagents & Chemicals list

Table 3.1 Reagents used in this thesis

Reagents & Chemical	Catalogue number	company
2-mercaptoethanol (β-ME)	444203	Calbiochem
Ac-EGXGN-NH2 (X= K, S or T)	Not provided	Bio-Synthesis Inc
Acetonitrile 'HiPerSolv' for HPLC	20048.32	VWR
Amersham ECL Prime Western Blotting Detection Reagent	GERPN2232	GE healthcare
Apyrase	M0398S	New England BioLabs
Cy3B-maleimide	PA63131	GE Life Sciences
DMEM/F-12 media	21331-046	Thermo Fisher
EZ-link Iodoacetyl-PEG_2-biotin	21334	Thermo Fisher
Isopropyl thio-b-D-galactoside (IPTG)	MB1008	MELFORD
$Na_2S_2O_3$, purum p.a., anhydrous, ≥98.0%	72049-250G	Sigma
Sodium fluoride	S7920-100G	Sigma
(Diacetoxyiodo)benzene	178721-25G	Sigma
2-azidoethan-1-amine hydrochloride	EN300-102102	Enamine
Adenosine 5'-triphosphate disodium salt hydrate	A7699-5G	Sigma
ALBUMIN FROM BOVINE SERUM (Fraction V) BSA	A7906-100G	Sigma
Ampicillin sodium salt	A9518-25G	Sigma
Antimycin A	A8674-50MG	Sigma
Arginine	11009-25G-F	Sigma
Benzamidine hydrochloride	434760-5G	Sigma
Carbonyl cyanide 3-chlorophenylhydrazone (CCCP)	C2759-100MG	Sigma
CentriPure P2 Columns	GEN-CP-0110-50	Generon
Chitin Resin	S6651L	New England Biolabs

Complete, Mini, EDTA-free protease inhibitor cocktail	11873580001	Roche
Coomassie (Bradford) Protein Assay Kit	23200	Thermo Fisher
Copper(II) sulfate anhydrous, powder, ≥99.99% trace metals basis	451657-10G	Sigma
Deoxyribonuclease I from bovine pancreas	DN25-100MG	Sigma
Dimethyl sulfoxide BioReagent, for molecular biology, ≥99.9%	D8418-50ML	Sigma
Dithiothreitol (DTT)	MB1015	Melford
ECL Kit (Reagent detection Western Blotting)	RPN2106	GE healthcare
Ethyl acetate GPR RECTAPUR	23880.368	VWR
EDTA Disodium	EDTA1000	Formedium
Hexane, Laboratory Reagent, ≥95%	208752-2.5L	Sigma
HiMark™ Pre-stained Protein Standard	LC5699	Life Tech
Hygromycin B Gold (solution)	ant-hg-5	Invivogen
Imidazole	I2399-500G	Sigma
Instant*Blue*™	ISB1L	Sigma
iodoacetamide	I1149-5G	Sigma
L-Alanine	A7627-1G	Sigma
L-Ascorbic acid reagent grade, crystalline	A7506-25G	Sigma
Lysine	L5501-5G	Sigma
Lysozyme from chicken egg white	62971-10G-F	Sigma
Magnesium chloride	L6876-5G	Sigma
MG-132	474790-5MG	Merck
Na-glycerophosphate pentahydrate	S/4110/48	Thermo Fisher
Ni-NTA affinity chromatography	30210	Qiagen
Non-Essential Amino Acids Solution (100X)	11140035	Gibco™
NuPAGE LDS loading buffer	NP0007	Invitrogen™
MOPS SDS running buffer 20X	NP0001	NuPage

MES SDS Running Buffer 20X	MES-SDS1000	Formedium
OligomycinA	75351-5MG	Sigma
Phenylmethane sulfonyl fluoride (PMSF)	P7626-25G	Sigma
Pierce™ Streptavidin Plus UltraLink™ Resin	53117	Thermo Fisher
Polyethylenimine, Linear (MW 25,000)	23966-2	Polysciences
Potassium iodide	P2963-100G	Sigma
Propargyl acrylate	543136-5G	Sigma
Propargyloxycarbonyl-L-lysine	HAA2095	Iris Biotech GmbH
Recombinant Human IL-1α 10ug	200-01A	Peprotech
Sequencing Grade Modified Trypsin	V5111	Promega
Serine	S4500-1G	Sigma
Sodium 2-mercaptoethanesulfonate	M1511-10G	Sigma
Sodium bicarbonate	S8875-500G	Sigma
Sodium chloride	10241AP	Merck
Sodium Orthovanadate	S6508-250G	Sigma
Sodium p-toluenesulfinate	89720-100G	Sigma
Sodium pyrophosphate tetrabasic	P8010-500G	Sigma
Sodium Pyruvate (100 mM)	11360070	Thermo Fisher
Sodium sulfate, ACS reagent, ≥99.0%, anhydrous, powder	238597-2.5KG-M	Sigma
Sucrose for molecular biology, ≥99.5% (GC)	S0389-500G	Sigma
TAMRA-azide	CLK-FA008	Jena Bioscience
Tris Acetate Running Buffer (20x)	LA0041	Thermo Fisher
TCEP Hydrochloride	T2650	Melford
Threonine	T8625-10G	Sigma
Triethylamine	471283-500ML	Sigma
Trifluoroacetic acid CHROMASOLV®, for HPLC, ≥99.0%	302031	Sigma
Tween-20	437082Q	VWR
Ubiquitin vinylsulfone (Ub-VS)	U-202	Boston Biochem
QIAprep Spin Miniprep Kit 50	27104	Qiagen
Tris(3-hydroxypropyltriazolylmethyl)amine	762342-100MG	Sigma

3.1.2 Plasmids list

Dr M. Peggie, Dr R. Toth, Dr N. Wood, Mrs M. Wightman and Mr T. Macartney performed the cloning, subcloning and mutagenesis of the constructs described in this thesis. All constructs encoded the human version of the gene unless otherwise indicated.

Table 3.2 Plasmids used in this thesis

Protein	Expressed	Plasmid	DU Number
Ubiquitin	MGCSSG Ubiquitin Q2-L73-Mxe intein/chitin binding domain T3C	pTXB1 MGCSSG Ubiquitin Q2-L73-Mxe intein/chitin binding domain T3C	DU53581
Ubiquitin	MGCSSG Ubiquitin Q2-R74-Mxe intein/chitin binding domain	pTXB1 MGCSSG Ubiquitin Q2-R74-Mxe intein/chitin binding domain	DU53553
Ubiquitin	Ubiquitin 1-73-T3C Mxe intein/chitin binding domain	pTXB1 Ubiquitin (delta R74-G75)-T3C Mxe intein/chitin binding domain	DU24828
Ubiquitin	Ubiquitin I44A (1-73)-T3C Mxe intein/chitin binding domain	pTXB1 Ubiquitin I44A (delta R74-G75)-T3C Mxe intein/chitin binding domain	DU24902
Ubiquitin	Ubiquitin M1-R74-Mxe intein/chitin binding domain	pTXB1 Ubiquitin M1-R74-Mxe intein/chitin binding domain	DU53503
Ubiquitin	FLAG GG Ubiquitin	pET28a FLAG GG Ubiquitin	DU46789
UBCH7	6His UBCH7 C17S C137S TAG3	pET156P UBCH7 C17S C137S TAG3	DU24871
UBCH7	6HIS(TAG3) UBCH7 C17S F63A C137S	pET15 6HIS(TAG3) UBCH7 C17S F63A C137S	DU28262
UBCH7	6His UBCH7 C17S C137S	pET156P UBCH7 C17S C137S	DU39196
UBCH7	His-UbcH7 C17S C137S F63A	pET156P UBCH7 C17S F63A C137S	DU24883

UBCH5c	6His-UBCH5c C21S 107S C111S	pET15b 6His C3 UBCH5c C21S 107S C111S	DU24307
UbcH5c	His-UbcH5c C21S F62A 107S C111S	pET15b 6His C3 UBCH5c C21S F62A 107S C111S	DU28781
UBCH5b	His-UBE2D2 (UBCH5b) C21S C107S C111S	pET28a(+) UBE2D2 (UBCH5b) C21S C107S C111S	DU24825
UBCH5b	His-UBE2D2 (UBCH5b) C21S F62A C107S C111S	pET28a(+) 6His-UBE2D2 (UBCH5b) C21S F62A C107S C111S	DU28325
UBCH5b	His-UBE2D2 (UBCH5b) C21S C107S C111S	pET15b 6His C3- UBE2D2 (UBCH5b) C21S C107S C111S	DU29236
UBCH5b	His-UBE2D2 (UBCH5b) C21S F62A C107S C111S	pET15b 6His C3 -UBE2D2 (UBCH5b) C21S F62A C107S C111S	DU29237
UBCH7	6His UBCH7	pET156P UBCH7	DU12798
UBCH5a	His-UBE2D1 (UBCH5a)	pET28a(+) UBE2D1 (UBCH5a)	DU4315
UBCH5c	His-UBE2D3 (UBCH5C)	pET156P UBCH5c	DU15703
UBE2A	6His-UBE2A	pET28b-UBE2A	DU20178
UBE2B	6His-UBE2B		DU20017
UBE2C	6His UBE2C	pET156P UBE2C	DU32146
UBCH6	6His UBCH6	pET156P UBCH6	DU12803
UBE2G1	6H-UBE2G1	pET28a(+) UBE2G1	DU14055
UBE2H	6His UBE2H	pET156P UBE2H	DU32149
UBE2J2	6His-UBE2J2 1-226	pET28a-UBE2J2 1-226	DU20695
UBC13	6His UBC13	pet156P UBC13	DU15705
UBC1	6His-UBC1	pET156P-UBC1	DU20018
CDC34	6His-CDC34	pET28a(+) CDC34	DU4317
UBE2S	6His-UBE2S	pET28-UBE2S	DU20175
UBE2W	6His-UBE2W	pET28a-UBE2W isoform 1	DU20190
UBE2Q1	6His-UBE2Q1	pET28-UBE2Q1	DU20176
UBE2Q2	6His UBE2Q2	pET156P UBE2Q2	DU12801
UbcH5c	His-UbcH5c N77S	pET15b His UbcH5c N77S	DU28800
UBCH5c	His UBCH5C (UBE2D3) D87A	pET15 6His 3C UBCH5C (UBE2D3) D87A	DU28876

UbcH5c	His-UbcH5c I88A	pET15b 6His C3 UbcH5c I88A	DU28799
UbcH5c	His-UbcH5c L97A	pET15b 6His C3 UbcH5c L97A	DU28798
UbcH5c	His-UbcH5c L104A	pET15b 6His C3 UbcH5c L104A	DU28796
UBCH5c	6His 3C UBCH5C (UBE2D3) S108A	pET15 6His 3C UBCH5C (UBE2D3) S108A	DU28895
UBCH5c	6His 3C UBCH5C (UBE2D3) D117A	pET15 6His 3C UBCH5C (UBE2D3) D117A	DU58041
PINK1 (Pediculus)	GST-PINK1 (Pediculus) 126-end	pGEX6P-PINK1 (Pediculus) 126-end	DU34798
PINK1 (Tribolium)	GST-full PINK1 (Tribolium) kin dead	pGEX-6-full PINK1 (Tribolium) Kdead	DU38037
USP2	GST-USP2 269-605	pGEX6P-2-USP2 269-605	DU15718
USP21	His-USP21 aa196-565	pET156p USP21 aa196-565	DU22483
OTUB2	GST OTUB2	pGEX6P-1-OTUB2	DU32795
UCHL3		pGEX6P-2-UCHL3	DU15696
Parkin	6HIS SUMO Parkin W403A	pET15 6HIS SUMO Parkin W403A	DU51424
Parkin	6His-SUMO Parkin W403A C431S	pET15 6HIS SUMO Parkin W403A C431S	DU28120
Parkin	6His SUMO parkin (223S)	pET156P 6His SUMO parkin (223S)	DU40847
Parkin	6HIS SUMO Parkin C431S	pET 6His SUMO Parkin C431S	DU39796
Parkin	6His SUMO parkin tv3 K27N	pET156P 6His SUMO parkin K27N	DU39837
Parkin	6HIs SUMO PArkin R33Q	pET6HIS SUMO Parkin R33Q	DU39824
Parkin	6HIS SUMO Parkin R42P	pET 6His SUMO Parkin R42P	DU39784
Parkin	6xHIS-SUMO-PARKIN A46P	pET15b SUMO blunt PARKIN A46P	DU44510
Parkin	6His SUMO parkin S65A	pET156P 6His SUMO parkin S65A	DU23314
Parkin	6xHIS-SUMO-PARKIN K161N	pET15b SUMO PARKIN K161N	DU44321

Parkin	6His SUMO parkin K211N	pET156P 6His SUMO parkin K211N	DU23325
Parkin	6His SUMO parkin R275W	pET156P 6His SUMO parkin R275W	DU43127
Parkin	6xHIS-SUMO-PARKIN G328E	pET15b SUMO PARKIN G328E	DU44322
Parkin	6HIS SUMO Parkin T415N	pET 6His SUMO Parkin T415N	DU39786
Parkin	6His SUMO parkin G430D	pET156P 6His SUMO parkin G430D	DU43129
Parkin	6His SUMO parkin C431F	pET 6His SUMO parkin C431F	DU46232
NEDD4L	GST-NEDD4L E576-D955	pGEX_TEV-NEDD4L E576-D955	DU24933
NEDD4L	GST-NEDD4L E576-D955 C922A	pGEX_TEV-NEDD4L E576-D955 C922A	DU24934
NLEL	GST-TEV-NLEL 59- end	pGEX-TEV-NLEL	DU39419
NLEL	GST-TEV-NLEL N59- end C753S	pGEX-TEV-NLEL (N59-end) C753S	DU24819
ARIH1	GST-TEV-ARIH1 aa1-394	pGEX TEV ARIH1 aa1-394	DU24260
UBE3C	GST UBE3C Q641-S1083(end)	pGEX6P UBE3C Q641-S1083(end)	DU45302
MYCBP2	GST-MYCBP2 Ser4378-Phe4640	pGEX6P-1-MYCBP2 Ser4378-Phe4640	DU28681
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4390S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4390S	DU28728
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4393S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4393S	DU28729
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4408S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4408S	DU28730
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4416S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4416S	DU28731
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4417S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4417S	DU28732
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4437S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4437S	DU28733
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4440S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4440S	DU28734

MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4506S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4506S	DU28735
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4509S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4509S	DU28736
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4520A	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4520A	DU28828
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4520A C4572A	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4520A C4572A	DU28897
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4520A C4572S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4520A C4572S	DU28961
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4520K	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4520K	DU29041
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4520S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4520S	DU28737
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4537S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4537S	DU28738
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4540S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4540S	DU28739
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4549S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4549S	DU28716
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4564S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4564S	DU28718
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4565S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4565S	DU28726
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4572A	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4572A	DU28898
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4572A C4520S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4572A C4520S	DU28964
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4572S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4572S	DU28742
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4572S H4583N	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4572S H4583N	DU29097
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4579S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4579S	DU28719

MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4582S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4582S	DU28720
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4600S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4600S	DU28721
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4614S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4614S	DU28727
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4631S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4631S	DU28743
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 C4634S	pGEX6P-1-MYCBP2 Ser4378-Phe4640 C4634S	DU28722
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 E4534A	pGEX6P-1-MYCBP2 Ser4378-Phe4640 E4534A	DU28901
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 E4534Q	pGEX6P-1-MYCBP2 Ser4378-Phe4640 E4534Q	DU28899
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 E4534Q H4583N	pGEX6P-1-MYCBP2 Ser4378-Phe4640 E4534Q H4583N	DU28904
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 F4573A	pGEX6P-1-MYCBP2 Ser4378-Phe4640 F4573A	DU28943
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 F4573A F4586A	pGEX6P-1-MYCBP2 Ser4378-Phe4640 F4573A F4586A	DU28941
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 F4578A	pGEX6P-1-MYCBP2 Ser4378-Phe4640 F4578A	DU28942
MYCBP2	GST-MYCBP2 Ser4378-Phe4640 F4586A	pGEX6P-1-MYCBP2 Ser4378-Phe4640 F4586A	DU28900
NMNAT2	6His SUMO-NMNAT2	pET156His SUMO-NMNAT2	DU28977
MIRO1	6xHIS-SUMO-MIRO1 1-592 K567R K572R	pET15b 6xHIS SUMO Miro1 1-592 K567R K572R	DU44247
Parkin	Parkin E321-V465 C431S	pcDNA5D FRT/TO-Parkin E321-V465 C431S	DU28118
Parkin	Parkin E321-V465	pcDNA5D FRT/TO-Parkin E321-V465	DU28101
Parkin	parkin W403A	pcDNA5 FRT/TO Parkin W403A	DU51408
MYCBP2	MYCBP2 Ser4378-Phe4640	pcDNA5-FRT/TO-DU-MYCBP2 Ser4378-Phe4640	DU29077

MYCBP2	Myc-MYCBP2 Ser4378-Phe4640 C4520A	pcDNA5-FRT/TO-Myc DU-MYCBP2 Ser4378-Phe4640 C4520A	DU28919
Ubiquitin	HA-ubiquitin	pcDNA5D frtTO HA UBIQUITIN	DU48439

3.1.3 Antibody list

Table 3.3 Commercial antibodies used in this thesis

Name	Company	Catalogue Number	Host
6xHis Monoclonal Antibody (Albumin Free)	Clontech	631212	Mouse
Parkin human	SantaCruz	Sc-32282	Mouse
GAPDH (14C10) (HRP Conjugate)	Cell Signaling	3683	Rabbit
Monoclonal Anti-NMNAT2, clone 2E4	Sigma	WH0023057M1-100UG	Mouse
Alpha Tubulin	PROTEINTECH	66031-1-Ig	Mouse

Table 3.4 In-house antibodies used in this thesis

Name	Immunogen	Sheep number	Bleed number
Parkin phospho-Ser65	RDLDQQ**S***IVHIVQR [residues 60 - 72 of human PARKIN]	S210D	2nd
MYCBP2	MYCBP2 (4378 - 4640)	SA357	2nd

3.1.4 Instruments

Bacterial incubators were from Infors (Reigate, UK). Centrifuge tubes, rotors and centrifuges were from Beckmann Coulter (Palo Alto, USA). ÄKTA pure chromatography system, Superdex 200 10/300, Superdex 75 16/600 were from GE Healthcare (Piscataway, NJ). Thermomixer IP shakers and bench-top centrifuge

were purchased from Eppendorf (Cambridge, UK). SpeedVacs were from CHRIST (Osterode, Germany). pH meters and electrodes were from Horiba (Kyoto, Japan). Power packs for electrophoresis and Trans-Blot Cells were from BioRad (Herts, UK). Mini gel tank electrophoresis system was from Atto (Tokyo, Japan). X-Cell SureLock Mini-cell electrophoresis systems and X-Cell II Blot modules were from Invitrogen (Paisley, UK). X-omat autoradiography cassettes, with intensifying screens, were from Kodak (Liverpool, UK). The Konica automatic film processor was from Konica Corporation (Japan). CO_2 incubators were from Mackay and Lynn (Dundee, UK). Tissue culture class II safety cabinets were from Medical Air Technology (Oldham, UK). The Chemidoc Gel Imaging System was from BioRad. The Mosquito Crystal, Dragonfly liquid handling robot and consumables were from TTP Labtech (Herts, UK). Prep-HPLC system components were obtained from Dionex (Camberley, UK). LC-MS system were obtained from Agilent Technologies. Freeze Dryers was obtained from LaboGene (Bjarkesvej, Danmark). The Reveleris® X2 flash chromatography system and The R-210/215 Rotavapor® system were obtained from BÜCHI Labortechnik AG. The ultrasonic cleaning baths was obtained from VWR.

3.1.5 In-house reagents

Primers were obtained by the University of Dundee oligonucleotide synthesis service. Bacterial culture medium Luria Bertani (LB) broth and LB agar plates were provided by the University of Dundee media kitchen facility. The Protein Production Team at Division of Signal Transduction and Therapy (DSTT) provided the affinity chromatography reagents including glutathione sepharose 4b, Ni-NTA agarose. The recombinant E1s, E2s, Flag-Ub, HHARI, UBE3C, USP2, USP21, OTUB2, UCH-L3, NleL, disease mutant Parkin, *ph*PINK1, p-Ub, His-SUMO-Miro1, NMNAT2

proteins expression and purification were carried out by the Protein Production and Assay Development (PPAD) team led by *Dr. Axel Knebel*. The lp-In T-Rex HeLa (Life Technologies) stable cell lines (HeLa cells stably expressing untagged Parkin WT, S65A, H302A or IBR-RING2) were kindly provided by *Dr. Miratul Muqit's lab*.

3.2 Method for Chapter 4

3.2.1 Synthesis of TDAEs (performed by Dr. Mattew Stanely)

3.2.1.1 Synthesis of alkyne-functionalized TDAE *3*

$PhI(OAc)_2$ (966 mg, 2 mmol) was added to a suspension of propargyl acrylate **1** (2 mmol), sodium arenesulfinate (8.0 mmol), and KI (328 mg, 2.0 mmol) in acetonitrile (MeCN) (8 mL). The reaction mixture was stirred vigorously at room temperature for 1 h under an inert atmosphere. The reaction mixture was then quenched by the addition of saturated $Na_2S_2O_3$ aq. (20 mL) followed by a saturated aqueous solution of $NaHCO_3$ aq. (20 mL). Further stirring was followed by extraction with ethyl acetate (EtOAc) (3 x 50 mL). The combined organic phases were washed with saturated NaCl aq. (75 mL) and dried over Na_2SO_4. The organic solvent was removed *in vacuo* and the crude product was directly purified by flash chromatography using a Reveleris® X2 flash chromatography system (Grace) (Reveleris® 12g silica column; EA:Hexane; 15% to 25% to 100% EA gradient elution) to yield **3**. HR-MS: observed $[M-H]^-$ 263.0361 (calculated 263.0373).

3.2.1.2 Synthesis of alkyne-functionalized TDAE *4*

To synthesize **4**, starting material *N*-propargylacrylamide, **2** was first prepared as previously described[441]. Purification was carried out by flash chromatography using a Reveleris® X2 flash chromatography system (Grace) (Reveleris® 12g silica

column; EA:Hexane; 20% to 35% EA gradient elution). Analytical data corresponded to literature values. $PhI(OAc)_2$ (966 mg, 2 mmol) was added to a suspension of *N*-propargylacrylamide, **2** (2 mmol), sodium arenesulfinate (8.0 mmol), and KI (328 mg, 2.0 mmol) in CH_3CN (8 mL). The reaction mixture was stirred vigorously at room temperature for 1 h under an inert atmosphere. The reaction mixture was quenched by the addition of saturated $Na_2S_2O_3$ aq. (20 mL) followed by a saturated aqueous solution of $NaHCO_3$ aq. (20 mL). Further stirring was followed by extraction with EtOAc (3 x 50 mL). The combined organic phases were washed with saturated NaCl aq. (75 mL) and dried over Na_2SO_4. The organic solvent was removed *in vacuo* and the crude product was directly purified by flash chromatography using a Reveleris® X2 flash chromatography system (Grace) (Reveleris® 12 g silica column; Toluene:EA; 0% to 1% to 10% EA gradient elution). HR-MS: observed $[M-H]^-$ 262.0537 (calculated 262.0543).

3.2.2 Preparation of UBE2L3*and UBE2D3*

Full-length E2s were cloned into the pET expression vectors downstream of an N-terminal Hisx6 tag and Precision protease site. *E. coli* strain BL21(DE3, Merck) transformed with E2 clones were cultured at 37 °C in LB medium supplemented with 100 μg mL^{-1} of ampicillin. When the OD_{600} reached 0.6, the culture was induced by addition of 0.2 mM isopropyl thio-β-D-galactoside (IPTG) and incubated at 37 °C for 5.0 hours. The cells were harvested by centrifugation and stored at -80 °C. Cells were resuspended in ice cold lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 25 mM imidazole, 0.5 mg/mL lysozyme, 50 μg/mL DnaseI, Complete, Mini, EDTA-free protease inhibitor cocktail) and incubated on ice for 30 minutes followed by sonication. The clarified lysate was then subjected to Ni-NTA affinity chromatography (Qiagen) and washed with wash buffer (50 mM Tris-HCl

pH 7.5, 150 mM NaCl, 25 mM imidazole). Protein was eluted with elution buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 300 mM imidazole) and dialyzed into storage buffer (20 mM Tris, 150 mM NaCl, 1 mM DTT). The protein was then further purified by size exclusion chromatography using a HiLoad 16/600 Superdex 75 pg column (GE Life Sciences) coupled to an ÄKTA Purifier FPLC system (1.0 mL min^{-1}, running buffer: 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM DTT). Fractions were collected, concentrated and flash frozen for storage at -80 °C.

3.2.3 Preparation of UBE2L3-ProcK

An amber stop codon was introduced at position 3 of the open reading frame in *pET156-UBE2L3** by Quikchange site-directed mutagenesis. This yielded the plasmid *pET156-UBE2L3*-TAG3* (* corresponds to a C17S and C137S mutant). BL21(DE3) cells were transformed with plasmids *pET156-UBE2L3*-TAG3* and *pCDF-PylST* (plasmid harbouring constitutive copies of the *Mb*PylRS/$tRNA_{CUA}$ pair, a kind gift from Dr. Jason Chin) and used to inoculate 200 mL LB media containing ampicillin (100 μg mL^{-1}) and spectinomycin (50 μg mL^{-1}). After overnight incubation at 37 °C, 6 x 1 L of LB medium containing ampicillin (100 μg mL^{-1}) and spectinomycin (25 μg mL^{-1}) were each inoculated with 30 mL overnight culture and incubated at 37 °C, 200 rpm, until cell density reached OD_{600} = 0.8-0.9. Propargyloxycarbonyl-L-lysine (ProcK) (300 mM stock solution in H_2O and pH adjusted to ~7) was then added to a final concentration of 3 mM. Cultures were then incubated for a further 20 min at 37 °C and then induced with IPTG (500 μM). After 4 h cells were harvested and cell pellets were resuspended in 200 mL lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 25 mM imidazole, 0.5 mg mL^{-1} lysozyme, 50 μg mL^{-1} DnaseI, 4 tablets of Complete, EDTA-free protease inhibitor cocktail (Roche) and incubated on ice for 30 min. The suspension was then

sonicated before clarification at 18,000 rpm, 30 min at 4 °C. Ni-NTA resin (600 μl of settled resin, Qiagen) was then added to the lysates and rotated at 4 °C for 1 h. Resin was then washed with wash buffer (50 mM Tris pH 7.5, 150 mM NaCl, 25 mM imidazole) and then eluted with elution buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 300 mM Imidazole). The eluent was then concentrated to 2.5 mL and further purified by size exclusion chromatography using a HiLoad 16/600 Superdex 75 pg column (GE Life Sciences) coupled to an ÄKTA Purifier FPLC system (running buffer: 100 mM Na_2HPO_4 pH 8, 150 mM NaCl, 0.1 mM TCEP). Fractions were collected and concentrated.

3.2.4 TAMRA labelling of UBE2L3-ProcK by Copper-catalyzed Azide-Alkyne Cycloaddition (CuAAC)

UBE2L3-ProcK was diluted in buffer (100 mM Na_2HPO_4 pH 8, 150 mM NaCl, 0.1 mM TCEP) to a final concentration of 50 μM. $CuSO_4$ and Tris(3-hydroxypropyltriazolylmethyl)amine (THPTA)[442], stock solutions were freshly prepared in H_2O and then premixed before being added to the protein at final concentrations of 0.1 mM and 1.0 mM, respectively. TAMRA-azide (Jena Bioscience, #CLK-FA008, M.W. = 512.56 g mol^{-1}) was then added to a final concentration of 150 μM. Ascorbic acid was added to a final concentration of 4 mM. The reaction was incubated at 23 °C for 1-2 h (or until starting material had been consumed as determined by LC-MS). The reaction was then quenched with 5 mM EDTA for 15 min at 23 °C. Small molecule components were then removed by dialysis against 2 x 1L buffer (100 mM Na_2HPO_4 pH 8.0, 150 mM NaCl, 5 mM EDTA, 0.1 mM TCEP) at 4 °C over 48 h. Protein was snap frozen in liquid nitrogen and stored at -80 °C.

3.2.5 Preparation and purification of phospho-ubiqutin (Performed by Dr. Axel Knebel)

23 µM bovine ubiquitin (SIGMA) was phosphorylated for 24 h with 3.7 µM MBP-PINK1 at 22 °C in the presence of 100 µM ATP and 10 mM $MgCl_2$. To replace ADP with ATP, the reaction was dialysed against Mg-ATP solution. Ubiquitin was filtered through a 30 kDa Vivaspin filter to deplete MBP-PINK1, concentrated in a 3 kDa MWCO filter device, washed extensively with water and loaded onto a Mono Q-column, which bound phospho-ubiquitin but not ubiquitin. The former was recovered by washing the column with 50 mM Tris pH 7.5, which was sufficient to elute stoichiometrically phosphorylated ubiquitin

3.2.6 Preparation and purification of Parkin WT, C431S, W403A and p-Parkin (phosphorylation of p-Parkin was performed by Dr. Cong Han)

The wild type human Parkin and C431S mutant clones (pET15-6His-SUMO-Parkin, pET15-6His-SUMO-Parkin C431S and pET15-6His-SUMO-Parkin W403A) were transformed into *Escherichia coli* BL21 (DE3) (Merck) and grown in LB media supplemented with 200 µM $ZnCl_2$ and 100 mg mL^{-1} ampicillin at 37 °C until OD_{600} reached 0.4. The culture was incubated at 16 °C. Upon OD_{600}= 0.8, expression was induced with 25 mM IPTG for 18 h**.** The bacterial cells were then harvested by centrifugation and suspended in lysis buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10 mM imidazole, 1 mM TCEP) supplemented with lysozyme (0.5 mg/mL) and cOmplete™, EDTA-free Protease Inhibitor Cocktail Set (Roche). After sonication treatment on ice, the mixture was centrifuged to yield a clear supernatant, which was loaded onto a column with Ni-NTA resin (Qiagen) pre-equilibrated in lysis buffer. The fraction containing Parkin was eluted, and then dialyzed against buffer A (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM TCEP), then SENP1 was added to

cleave SUMO tag for 12 h at 4 °C (100:1 w/w ratio of Parkin to SENP1). After the digestion, 10 mM imidazole was added to the sample, which was further applied onto a Ni-NTA column. The flow-thru fraction was loaded onto a HiLoad 16/60 Superdex 75 column (GE Healthcare) with buffer B (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM DTT).

The *Pediculus humanus corporis* PINK1 (*Ph*PINK1) clone (pGEX6P-PINK1 126-575) was transformed into *E. coli* BL21 (DE3) and grown in LB media supplemented with 100 mg mL^{-1} ampicillin at 37 °C. Upon OD_{600}= 0.8, expression was induced with 0.1 mM IPTG for 16 h at 26 °C. The bacterial cells were then harvested by centrifugation and suspended in lysis buffer (50 mM Tris-HCl, pH 7.5, 250 mM NaCl, 0.5 mM EGTA, 0.5 mM EDTA, and 1 mM DTT) supplemented with 10 mg/mL Leupeptin and 1 mM Pefabloc. After sonication treatment on ice, the mixture was centrifuged to yield a clear supernatant, which was loaded onto a column with Glutathione Sepharose 4B resin (GE Healthcare) pre-equilibrated in lysis buffer. The *Ph*PINK1 fraction was eluted, and then dialyzed against buffer C (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM DTT, and 5% glycerol).

The Parkin was incubated with *Ph*PINK1 for 8 h at 27 °C in 20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$, 5 mM ATP, and 5 mM DTT (5:1 molar ratio of Parkin to *Ph*PINK1). The reaction mixture was applied to a HiLoad 16/60 Superdex 75 column with buffer B. Subsequently, the Parkin fractions were pooled and concentrated by ultrafiltration with Amicon centrifugal filter device. The phosphorylated Parkin was identified by 8% Zn^{2+}-Phos-tag SDS-PAGE.

3.2.7 Expression and purification of NEDD4L, HOIP and NleL

GST-NEDD4L, GST-NEDD4L C922A, HOIP, HOIP C885S, GST-NleL and GST-NleL C753A, were expressed and purified as previously described[128,242,443]

3.2.8 General method for ΔUb-SR preparation

A plasmid composed of full-length Ub cloned into the pTXB1 vector[444], was taken and the bases coding for the C-terminal Ub residues 74-76 were deleted by Quikchange mutagenesis. To suppress cellular intein activity, a T3C mutation in the GyrA intein was also introduced yielding plasmid *pTXB1-Ub∆74-76-T3C*. ER2566 *E. coli.* cells (NEB) were transformed with *pTXB1-Ub1-74* and recovered with S.O.C. medium (250 µL). The cells were incubated for 1 h at 37 °C and then LB medium (100 mL) containing ampicillin (100 mg mL^{-1}) was inoculated with the recovered cells (200 µL) and the culture was incubated overnight whilst shaking (230rpm) at 37°C. LB medium (2L) containing ampicillin (100 mg mL^{-1}) was inoculated with the overnight culture (50 mL) and incubated whilst shaking (230rpm) at 37°C. At $O.D._{600}$~ 0.4, the cells were transferred to a 25°C incubator and after 30 min the cells were induced with IPTG (0.2 mM). After 5 h the cells were harvested and suspended in 60 ml lysis buffer (20 mM Na_2HPO_4 pH 7.2, 200 mM NaCl, 1 mM EDTA) and frozen. The thawed cells were lysed by sonication on ice and were clarified by centrifugation (39000 x g, 30 min). An empty XK 26/20 column was filled with chitin beads (20 mL) (NEB) and equilibrated with lysis buffer. At 4°C the clarified lysate was loaded (flow rate: 1.0ml min^{-1}) onto the column using an ÄKTA FPLC system. The column was then washed with lysis buffer (~400 mL) and incubated with 50 mL of cleavage buffer (20 mM Na_2HPO_4 pH 6.0, 200 mM NaCl, 100 mM Sodium methanethiolate (MESNA), 1 mM EDTA). The flow was then stopped and the column was incubated for 60 h at 4°C, to allow thiolysis and concomittant S-N transfer. Liberated ΔUb-SH was eluted with elution buffer (20 mM Na_2HPO_4 pH 6, 200 mM NaCl, 1 mM EDTA). The fractions containing ΔUb-SH were pooled and concentrated to ~5 mL using an Amicon Ultra-15 centrifugal filter device (Millipore). The protein was then further purified by semi-preparative RP-HPLC using the

Dionex system. A gradient of 10 % buffer A to 80 % buffer B was applied at a flow rate of 10 mL min^{-1} over 30 min (buffer A=0.1% TFA in H2O, buffer B=0.1% TFA in acetonitrile). After HPLC purification fractions containing ΔUb-SR were pooled and lyophilized yielding ~30 mg.

3.2.9 ΔUb-N_3 aminolysis reaction

ΔUb-SR (25.8 mg) was reconstituted[445], by the addition of DMSO (200 µL). On complete dissolution of ΔUb-SR in DMSO, H_2O (800 µL) was added to give a final DMSO concentration of 20% (v/v). 500 µL of 2-azidoethanamine (Enamine) in 50% (v/v) aqueous DMSO/MQ (8 M) was then added to the ΔUb-SR solution. 60.0 µl of triethylamine (TEA) was then added raising the solution pH to 9 and the mixture was then briefly vortexed. The solution incubated at 30 °C for 16 hours and monitored by LC-MS. The protein (ΔUb-N_3) was then further purified by semi-preparative RP-HPLC (Column: BioBasic-4; Part number: 72305-259270). A gradient of 20 % buffer A to 50 % buffer B was applied at a flow rate of 10 mL min^{-1} over 60 min (buffer A=0.1% TFA in H_2O, buffer B=0.1% TFA in acetonitrile). Fractions containing ΔUb-N_3 were pooled and lyophilized (Yield: 80%).

3.2.10 Conjugation of ΔUb-N_3 with TDAE by CuAAC click reaction

ΔUb-N_3 (3.5 mg) was reconstituted by dissolution into 100 µl DMSO. H_2O (900 µL) was added to give a final DMSO concentration of 10% (v/v) and a final ΔUb-N_3 concentration of 418.6 µM. DMSO stock solutions of TDAEs **3** or **4** were prepared (10 mM). TDAEs **3** or **4** (200 µL) were then mixed with a pre-prepared DMSO/MQ solution (456.2 µL, 20% (v/v)). Phosphate buffer (50.0 µL, 100 mM Na_2HPO_4 pH 7.5, 150 mM NaCl) and ΔUb-N_3 stock solution (238.6 µL, 418.6 µM) were then subsequently added. Tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) in H_2O

(25 µL, 12.5 eq., 50 mM stock solution) was pre-mixed with a freshly prepared solution of $CuSO_4$ (aq.) (5 µL, 2.5 eq., 50 mM stock solution). The THPTA/$CuSO_4$ (aq.) solution (30 µL) was then added to the previously prepared TDAE/ΔUb-N_3 solution. L-ascorbic acid in H_2O (25 µL, 100 mM stock solution) was then added to the previously mixed components. The reaction was incubated at 23 °C with shaking (1400 rpm) for 15 minutes. The reaction was then quenched by the addition of EDTA (2.5 mM final concentration) and the product (**7** or **8**) was purified by semi-preparative RP-HPLC (Column: Thermo Scientific BioBasic-4; Part number: 72305-259270) using the Dionex system. A gradient of 20 % buffer A to 50 % buffer B was applied at a flow rate of 10 mL min^{-1} over 60 minutes (buffer A=0.1% TFA in H_2O, buffer B=0.1% TFA in acetonitrile). Fractions containing ΔUb-TDAE were pooled and lyophilized (Yield: 50-60%).

3.2.11 Synthesis E2~Ub Probes

ΔUb-TDAE (**7**, 7.0 mg) was reconstituted by the addition of 50 µl DMSO. On complete dissolution of ΔUb-TDAE, H_2O (450 µL) was added to give a final DMSO concentration of 10% (v/v) and a final ΔUb-TDAE concentration of 1.6 mM. ΔUb-TDAE was then incubated with 0.5 eq. of the desired E2 in 0.8 mL phosphate buffer (100 mM Na_2HPO_4 pH 7.5, 150 mM NaCl, **reducing agent free,** the desired E2s were performed buffer exchanged before the reaction) at 23 °C for two hours and monitored by LC-MS. For acrylamide probes **10**, **10** F63A, **11**, **12** and **12** F62A the above procedure was carried out with ΔUb-TDAE **8** and 30 °C incubation for 12 h was required to consistently achieve quantitative elimination to the unsaturated AVS species. The products were purified by gel filtration (0.5 mL min^{-1}, 20 mM Tris-HCl pH 7.5, 150 mM NaCl as mobile buffer, HiLoad 16/600 Superdex 75pg).

3.2.12 Molecular modelling of RING- and HECT-probe complexes (Performed by Dr. Virdee)

The RING-E2~Ub model was generated by importing the PDB coordinates 4AP4[446] into the BioLuminate biologics modelling suite (Schrödinger). The AVS-triazole linker present between E2 and Ub in probe **10** was manually built and geometry optimized. Hydrogen atoms were then added using the Protein Preparation tool. The E2 residue Cys85, the Ub residue Leu73, and the linkage between them were selected along with all protein residues within 5 Å. This substructure was then energy minimized using the OPLS2005 forcefield. Minimized coordinates were then exported as mol2 files and imported into Pymol (Schrödinger) for figure generation. The same procedure was repeated for the HECT-E2~Ub model using PDB 3JW0[128].

3.2.13 Parkin activity assay with UBE2L3*

Parkin (3 µg, 760 nM) was incubated with ubiquitylation assay components in a final volume of 50 µl (50 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 240 nM UBA1, 2 µM UBE2L3 variant, 2 mM ATP, 58 µM Ub and 380 nM of GST-*Ph*PINK1 WT. The reaction was incubated at 30 °C for 1 h. The reaction was quenched by the addition of reducing 4X LDS loading buffer, resolved by SDS-PAGE, and visualized by coomassie staining.

3.2.14 *In vitro* probe-labeling assay

The indicated E3 ligases were added into Tris buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl). 5.0 e.q (unless otherwise indicated) of E2~Ub (probe **9** or **10**) was incubated with E3 ligase at 30 °C for the indicated time. Reactions were quenched by the addition of 4X LDS loading buffer (supplemented with ~680 mM 2-

mercaptoethanol) and samples were analyzed by SDS-PAGE (4-12% NuPage gel) followed by coomassie staining or immunoblotting.

3.2.15 Ubiquitylation assay and *in situ* profiling (performed by Kristin Bulk)

Parkin (1.3 μM) was incubated with ubiquitylation assay components in a final volume of 30 µl (50 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 0.12 µM UBA1, 2 µM UBE2L3, 0.83 µM His-SUMO-Miro1, 2 mM ATP, 88 µM Flag-ubiquitin and 0.31 µM *Tc*PINK1 (WT/KD). Ubiquitylation reactions were incubated at 30 °C and 1050 rpm for 60 min. Ubiquitylation reactions were stopped by the addition of 5 U mL^{-1} Apyrase (New England BioLabs) and incubated for 10 min at 30 °C and agitated at 1050 rpm. Probe profiling was carried out by the addition of **9** or **11** (8 µM) and incubated for 10 h at 30 °C and 650 rpm. The reactions were terminated by the addition of LDS sample buffer containing 4 % 2-mercaptoethanol. Reaction mixtures were resolved by SDS-PAGE and immunoblotted with the following antibodies: anti-FLAG (Sigma, 1:10000), anti-Parkin (Santa-Cruz, 1:500) or anti-His (1:5000). For Fluorescence detection, 20 µl of sample were loaded on a 4-12 % gel and subjected to SDS-PAGE. Gels were washed with water for 30 minutes and then visualized by scanning with a FLA-5100 imaging system (Filter: LPG, Laser: 532 nm, Voltage: 400, FUJIFILM Life Science). The gel was subsequently coomassie stained.

3.2.16 DUB resistance assay (performed by Dr. Virdee)

His-$USP2_{259–605}$, His-$USP21_{196-565}$, OTUB2 and GST-UCHL3 were expressed in *E. coli* and purified. The DUB-reactive ubiquitin-based probe, Ub-Alk, was used as a positive control and tested in parallel with probes **9** and **10**. Probes were diluted to 40 μM and DUBs were diluted to 10 μM in buffer (20 mM Tris-HCl pH 7.5, 150

mM NaCl, 1 mM TCEP). Probe (10 μL) and DUB (10 μL) solutions were then mixed and incubated for 2 h at 30 °C. Proteins were then resolved by SDS-PAGE and visualized by coomassie staining.

3.2.17 Tryptic MS/MS sequencing of crosslinked peptides from Parkin and an ABP

The coomassie stained SDS-PAGE band corresponding to the Parkin labeled with ABP **9** was excised, dehydrated and resuspended using standard procedures (In-house protocol from Dr. David Campbell, file name: *In gel digest 170221*). LC-MS/MS analysis was performed on an LTQ Orbitrap Velos instrument (Thermo Scientific) coupled to an Ultimate nanoflow HPLC system (Dionex). A gradient running from 3 % solvent A to 60 % solvent B over 45 min was applied (solvent A = 0.1 % formic acid in H_2O; solvent B = 0.08 % formic acid in 80 % MeCN). Fragment ions were generated by CID and 1+ and 2+ precursor ions were excluded. Thermo .raw data was converted to .mgf format using the MSConvert software (ProteoWizard). Raw data was searched using the pLink software against UBE2L3* and Parkin sequences with trypsin specificity (up to 3 missed cleavages) (Performed by Dr. Virdee). A crosslinker monoisotopic mass of 307.1644 was manually added which accounted for the theoretical mass difference associated with formation of a bisthioether between 2 Cys residues together with the acrylate AVS, the triazole linkage and the tryptic Leu73 remnant from the Ub C-terminus.

3.2.18 Cell culture and lysis protocol

Flp-In T-Rex HeLa (Life Technologies) stable cell lines (HeLa cells stably expressing untagged Parkin WT, S65A, H302A or IBR-RING2) were cultured (37 °C, 10 % CO_2) in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% (v/v) Fetal Bovine Serum (FBS), 2.0 mM L-glutamine, 1X minimum essential medium (MEM,

Life Technologies), 1X non-essential amino acids (NEAA, Life Technologies), 1.0 mM of sodium pyruvate (Life Technologies) and antibiotics (100 units mL^{-1} penicillin, 0.1 mg mL^{-1} streptomycin, 15 μg mL^{-1} blasticidin, 100 μg mL^{-1} hygromycin). At 80 % confluency cells were either untreated or treated with 10 μM CCCP (dissolved in DMSO, 40mM stock solution), and incubated for a further 3 h. Cells were rinsed with ice-cold PBS and extracted in lysis buffer (50 mM Tris-HCl pH 7.5, 5 μM EDTA, 0.27 M sucrose, 10 mM sodium 2-glycerophosphate, 0.2 mM phenylmethane sulfonyl fluoride (PMSF), 1.0 mM benzamidine, 1.0 mM sodium ortho-vanadate, 50 mM sodium fluoride, 5.0 mM sodium pyrophosphate and 10 μM TCEP). Lysis was carried out by sonication (Sonic & Materials INC, VC 100, Jencons Scientific LTD, CT, USA, 55% amplitude, 12 times (1 sec on, 1 sec off)) and then clarified by centrifugation at 4 °C for 30 min at 14,800 rpm. Supernatants were collected (total cell extracts) and protein concentration determined by Bradford assay.

SH-SY5Y (PHE Culture Collection) cells were grown in DMEM/F-12 media supplemented with 15% (v/v) FBS, 2.0 mM L-glutamine and antibiotics (100 units mL^{-1} penicillin, 100 μg mL^{-1} streptomycin). To depolarize mitochondria, cells were treated with 10 μM CCCP, dissolved in DMSO, for the indicated time. Cell extracts were prepared as described for HeLa cells. All cells were cultured at 37 °C in a 5 % CO_2 humidified atmosphere.

HOIP WT and C879S MEF cells were provided by Dr. Philip Cohen lab. The MEF cells was prepared as previously described[447]. The cells were cultured (37 °C, 10 % CO_2) in DMEM supplemented with 10% (v/v) FBS, 2.0 mM L-glutamine, and antibiotics (100 units mL^{-1} penicillin, 0.1 mg mL^{-1} streptomycin, 15 μg mL^{-1}) and treated within 5 ng mL^{-1} IL-1α for indicated time. Cell extracts were prepared as described for HeLa cells. Testing of cell stocks for mycoplasma contamination was routinely carried out in accordance with departmental protocols.

3.2.19 ABPP of total cell extracts

E2~Ub probes were diluted with buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl) and added to total cell extracts at a final concentration of 5.0 µM (final extract concentration 1.6 mg mL^{-1}). Reactions were then incubated at 30 °C for the indicated time. Reactions were quenched by the addition of 4X LDS loading buffer (supplemented with betamercaptoethanol) and samples were analyzed by SDS-PAGE (4-12% NuPage gel) followed by immunoblotting.

3.2.20 Immunoblotting

Samples were resolved by SDS–PAGE (4–12% NuPage gel, Invitrogen) with MOPS or MES running buffer (without boiling) and transferred on to 0.45 µm nitrocellulose membranes (GE Healthcare Life Science). Membranes were blocked with PBS-T buffer (PBS + 0.1% Tween-20) containing 5% (w/v) non-fat dried skimmed milk powder (PBS-TM) at room temperature for 1 h. Membranes were subsequently probed with the indicated antibodies in PBS-T containing 5 % (w/v) Bovine Serum Albumin (BSA) overnight at 4 °C. Detection was performed using HRP-conjugated secondary antibodies in PBS-TM for 1 h at 23 °C. ECL Prime substrate (GE Life Sciences) was used for visualization in accordance with the manufacturers protocol.

3.2.21 Antibodies

His-tagged species were probed with 1:10000 anti-His primary antibody (Clontech, cat number: 631212). Anti-Parkin mouse monoclonal was obtained from Santa Cruz (sc-32282) at 1:2000 dilution (HeLa and SH-SY5Y samples); anti-Parkin phospho-serine 65 rabbit monoclonal antibody was obtained as previously described[258], and used at 1:2500 dilution. GAPDH (14C10) rabbit mAb (HRP

Conjugate, Cell Signaling Technology) was used at 1:5000 dilution. For detecting the endogenous level of phospho-Parkin, an antibody pre-clearing step is required (The pre-blocked nitrocellulose membrane was incubated with anti-Parkin phospho-serine 65 rabbit monoclonal antibody solution (1:2500 dilution) at 4°C for overnight. The pre-clearing procedure was repeated)

3.2.22 Parkin disease mutant *in vitro* profiling

The different recombinant Parkin mutants[448], (1.0 µM, 1.535 µg) were incubated with *Ph*PINK1 or kinase dead *Ph*PINK1 (0.77 µg), TCEP (1.0 mM), ATP (1.0 mM) and Ub (75 µM) in 1X Tris buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5.0 mM $MgCl_2$) at 30 °C for 1.0 h. After 1.0 hour incubation, fluorescent probe **11** was added (5.0 µM) and incubated for another 4 or 8 hours in the dark at 30 °C. Reactions were quenched by the addition of 4X LDS loading buffer (supplemented with betamercaptoethanol) for the indicated time point and samples were resolved by SDS-PAGE (4-12% NuPage gel). Gels were washed with water for 30 minutes and then visualized by scanning with a FLA-5100 imaging system (Filter: LPG, Laser: 532 nm, Voltage: 400, FUJIFILM Life Science). Experiments were carried out in triplicate and quantification was carried out using the Fiji software variant of ImageJ. For each replicate the ratio of the band size relative to the WT signal was taken. The mean of these values were plotted on a Log_2 scale using Prism 6 (GraphPad Software). All the disease mutant Parkin proteins were obtained from Dr. Axel Knebel, DSTT.

3.2.23 Profiling of patient-derived fibroblasts

Low passage fibroblasts from a PD patient carrying a Q456X mutation in the *PARK6* gene and from a healthy age-matched family member, were cultured as previously

described[449]. Fibroblasts were obtained from skin biopsies following routine clinical procedures, underwritten informed consent and approval by a local ethics committee (Comite´ de Protection des Personnes "Ile de France"). Cell lysis, probe profiling and immunoblotting as carried out as described for SH-SY5Y extracts. These fibroblasts sample were kindly provided by Dr. Olga Corti and Dr. Jean-Christophe Corvol (Sorbonne Universités, UPMC Univ Paris 06; and INSERM UMRS_1127, CIC_1422; CNRS UMR_7225; AP-HP and ICM, Hôpital Pitié-Salpêtrière, Department of Neurology, F-75013, Paris, France).

3.2.24 Ubiquitin-based probes

Ubiquitin vinylsulfone (Ub-VS) was sourced from Boston Biochem. The Ub-Alk probe was prepared by Dr. Stanley as follows. Ub1-75 thioester was prepared as previously described except the Gly76 was deleted producing plasmid *pTXB1-Ub1-75.* Pooled fractions were concentrated and adjusted to pH 8 by the careful addition of 1 N NaOH while mixing. Propargyl amine (250 mM) was then added and the reaction left to incubate at 25 °C for 1 h. Protein functionalization was ~90% as determined by LC-MS and was further purified by semi-preparative HPLC and lyophilized. Probe was folded by dissolution in denaturing buffer (200 mM Na_2HPO_4 pH 7.5, 6 M guanidinium chloride) and dialyzed overnight against PBS.

3.3 Method for Chapter 5

3.3.1 Biotin functionalized ABP preparation

Ub with a GCSSG N-terminal extension was expressed from plasmid *pTXB1-UbΔ74-76-T3C* plasmid[7]. An equivalent plasmid encoding Ub residues 1-74 (*pTXB1-UbΔ75-76-T3C*) was also created. Ub thioesters were obtained as described previously

generating cysteine tagged Cys-Ub_{1-73}-SR and Cys-Ub_{1-74}-SR, respectively[7]. The extended Ub_{1-74} was included as this retains Arg74 which forms a favorable electrostatic interaction with the RBR E3 HOIP[31]. Cys-Ub_{1-73}-SR (30 mg) was reconstituted by the addition of DMSO (116 µL) followed by H_2O (456 µL). An aqueous stock solution (48 mM) of EZ-link Iodoacetyl-PEG2-biotin (Thermofisher) was prepared and 200 µL was added to the Cys-Ub_{1-73}-SR solution (580 µl) followed by the addition of 900 µl degased buffer (50 mM Na_2HPO_4 pH 7.5, 150 mM NaCl). The reaction was incubated at 23 °C for 1 hour and monitored by LC-MS. The protein (Biotin-Ub_{1-73}-SR) was then further purified by semi-preparative RP-HPLC (Column: BioBasic-4; Part number: 72305-259270). A gradient of 20 % buffer A to 50 % buffer B was applied at a flow rate of 10 mL min^{-1} over 60 min (buffer A=0.1 % TFA in H_2O, buffer B=0.1 % TFA in acetonitrile). The above procedure was repeated to generate Biotin-Ub_{1-74}-SR. HPLC fractions containing Biotin-Ub_{1-7X}-SR were pooled and lyophilized (Yield: Biotin-Ub_{1-73}-SR 75-85 %, Biotin-Ub_{1-74}-SR 40-50 %). Biotin-tagged ABPs containing thioacrylamide warheads were then prepared as previously described[7] employing the E2 recognition elements UBE2D2*, UBE2D2* F62A, UBE2L3* and UBE2L3* F63A furnishing ABPs **13**, **14**, **15** and **16**. *denotes E2 where non-catalytic Cys residues were mutated to Ser. ABPs based on UBE2D2* and UBE2D3* bearing hexahistidine reporter tags and thioacrylamide warheads, were also prepared yielding ABPs **17** and **18**, respectively.

3.3.2 Activity-based proteomic profiling of SH-SY5Y cells

SH-SY5Y total cell lysate (4.5 mg, 550 µL) was mixed with ABPs **13**, **14**, **15** and **16** (3 µM) and incubated at 30 °C for 4 hours. To induce Parkin activation cells were administered with oligomycin (5 µM) and antimycin A (10 µM) (OA, dissolved in

DMSO) for 3 hours. Control enrichments were also performed where probe was withheld. Extracts were mixed with 100 µL of Pierce™ Streptavidin Plus UltraLink™ Resin (ThermoFisher Scientific) and diluted with 6 % SDS solution (20 µL) to a final concentration of 0.2 % in phosphate buffer. Samples were incubated for 4 hours at 4 °C and resin washed (2 ml 0.2 % SDS/PBS, 2 ml PBS, 1 ml 4 M urea/PBS, 2 ml PBS) and then resuspended in 190 µl Tris buffer (50 mM Tris pH 8, 1.5 M urea). Resin-bound proteins were reduced with TCEP (5 mM) for 30 minutes at 37 °C and then alkylated with iodoacetamide (10 mM) at 23 °C for 20 minutes. DTT (10 mM) was then added followed by washing with buffer (50 mM Tris pH 8, 1.5 M urea) to a final volume of 300 µL. Trypsin (2 µg) was then added and further incubated at 37 °C for 14 hours. Trifluoroacetic acid was added to a final concentration of 0.1 % and samples were desalted with a C^{18} MacroSpin column (The Nest Group Inc). LC-MS/MS analysis was performed on an LTQ Orbitrap Velos instrument (Thermo Scientific) coupled to an Ultimate nanoflow HPLC system (Dionex). A gradient running from 3 % solvent B to 99 % solvent B over 345 min was applied (solvent A = 0.1 % formic acid and 3 % DMSO in H_2O; solvent B = 0.08 % formic acid and 3 % DMSO in 80 % MeCN).

3.3.3 Data processing

Raw files were searched against the Swissprot database and a decoy database using the MASCOT server (Matrix Science). Trypsin specificity with up to three missed cleavages was applied. Cysteine carbamylation was set as a fixed modification and variable modifications were methione oxidation/dioxidation. A PERL script was used to extract the number of rank 1 peptides for each protein from the MASCOT search results and this figure was used as the number of spectral counts which performed by Dr. Virdee. A second PERL script filtered the data by

searching the human swisspfam_v30 database using the E3 domain terms RING, HECT, IBR and zf-UBR. Manual curation was also carried out which involved the addition of E1 enzymes. Any proteins with less than 3 spectral counts and less than 14-fold spectral count enrichment, relative to control experiments where ABP was withheld, were omitted from the list. Pairwise datasets were then plotted as column charts in Prism (Graphpad Software).

3.3.4 Cloning of MYCBP2cat (performed by Dr. Nicola T. Wood)

Human MYCBP2 (NM_015057.4) sequences were amplified from full-length Addgene plasmid #2570. Wild type and mutant fragments were subcloned as BamHI/Not1 inserts into pGEX6P-1 (GE Life Sciences) for bacterial expression, or a modified version of pcDNA TM5/FRT/TO (ThermoFisher) containing an N-terminal Myc tag for mammalian expression.

3.3.5 UBE1 and E2 expression and purification (Performed by Dr. Axel Knebel)

6His-UBE1 was expressed in Sf21 cells and purified via its tag as previously described[334]. Phosphate Buffered Saline was used throughout the purification and hydroxy containing compounds avoided. UBE2D3 was expressed as an N-terminally 6His-tagged protein in BL21 cells and purified over Ni-NTA-agarose and finally dialysed into 50 mM Na_2HPO_4 pH 7.5, 150 mM NaCl, 0.5 mM TCEP. UBE2A was expressed as a GST-fusion in *E. coli* and the GST tag was proteolytically removed. The remaining E2s were expressed as recombinant bacterial proteins and purified via their His-tags and buffer exchanged by size exclusion chromatography into running buffer (50 mM Na_2HPO_4 pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 0.015 % Brij-35) using a Superdex 75 column (GE Life Sciences).

3.3.6 Expression and purification of MYCBP2cat and GST-MYCBP2cat

GST-tagged MYCBP2cat (Ser4378-Phe4640), wt and mutants, were expressed at 16 °C overnight and purified against glutathione resin (Expedeon) using standard procedures. GST-tagged constructs were eluted with glutathione and untagged constructs were obtained by on-resin cleavage with Rhinovirus 3C protease. Proteins were buffer exchanged into 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1.0 mM TCEP buffer and flash frozen for storage at -80 °C.

3.3.7 Expression and purification of NMNAT2 (Performed by Dr. Axel Knebel)

NMNAT2 was expressed with a 6His-SUMO tag in BL21(DE3) cells, induced with 0.1 mM IPTG and incubated for expression at 16 °C. The cells were collected and lysed in buffer (50mM Tris-HCl (pH 7.5), 250 mM NaCl, 0.2 mM EGTA, 20 mM imidazole, 20 mM L-arginine, 0.015% Brij 35, 1 mM Leupeptin, 1 mM Pefabloc, 1 mM DTT using standard protocols and the protein was purified over Ni-NTA-agarose. The eluted protein was incubated with His-SENP1 protease during dialysis against PBS, 20 mM L-arginine, 1 mM DTT. The tag and protease were depeleted against Ni-NTA-agarose and NMNAT2 was concentrated and subjected to chromatography on a Superdex 75 HR 10/30 into buffer (PBS, 20 mM L-arginine).

3.3.8 Activity-based protein profiling of MYCBP2 cysteine mutants

The indicated MYCBP2 mutant was diluted into Tris buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl) to a final concentration of 3 µM. Probe **18** was added (12 µM) and incubated with E3 ligase at 30 °C for four hours. Reactions were quenched by the addition of 4X LDS loading buffer (supplemented with ~680 mM 2-mercaptoethanol) and samples were resolved by SDS-PAGE (4-12 % NuPage gel) followed by Coomassie staining or anti-His immunoblotting.

3.3.9 Tryptic MS/MS sequencing of probe-labelled MYCBP2

Crosslinking MS using ABP 18 was carried out as previously described[322]. In summary Coomassie stained SDS-PAGE band corresponding to ABP-labelled WT MYCBP2 was analyzed by LC-MS/MS using an Orbitrap Fusion™ Tribrid™ mass spectrometer (Thermo Scientific) coupled to an Ultimate nanoflow HPLC system (Dionex). A gradient running from 0 % solvent A to 60 % solvent B over 120 min was applied (solvent A = 0.1 % formic acid in H_2O; solvent B = 0.08 % formic acid in 80 % MeCN). Fragment ions were generated by HCD and 1^+, 2^+ and 3^+ precursor ions excluded. Raw data was searched using the pLink software against UBE2D3* and MYCBP2 sequences with trypsin specificity (up to 2 missed cleavages) (Performed by Dr. Virdee). The error window for MS/MS fragment ion mass values was set to the software default of 20 ppm. A crosslinker monoisotopic mass of 306.1805 Da was manually added which accounted for the theoretical mass difference associated with formation of a bisthioether between 2 Cys residues derived from probe 6, which was based on UBE2D3* and contained a thioacrylamide AVS warhead[322].

3.3.10 Tris/glycerol-mediated E2 discharge assay

Assays were carried out in buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 5 mM $MgCl_2$) containing the indicated MYCBP2 mutant (15 µM), UBE1 (1.5 µM), UBE2D3 (15 µM), Ub (37 µM) and ATP (10 mM). The reactions were incubated at 37 °C for 30 minutes. Reactions were terminated by the addition of 4X LDS loading buffer (with and without ~680 mM 2-mercaptoethanol). A C4572S sample was further incubated with 0.14 N NaOH at 37 °C for 20 minutes and samples were resolved by SDS-PAGE (4-12 % NuPage gel) and visualized by Coomassie staining.

3.3.11 LC-MS analysis of nucleophile discharge assays

Reactions were prepared as described for discharge assay. After 30 minutes, the reaction was analyzed using an Agilent 1200/6130 LC-MS system (Agilent Technologies) using a 10-75 % gradient over 20 minutes (buffer A = H_2O + 0.05 % TFA, buffer B = acetonitrile + 0.04 % TFA).

3.3.12 Preparation of Cy3B-Ub

Ub bearing a TEV protease-cleavable N-terminal hexahistidine tag followed by a ACG motif was expressed in bacteria from a pET plasmid (kindly provide Ronald Hay, University of Dundee). Protein was purified by Ni affinity chromatography, cleaved from the tag with TEV protease then buffer exchanged into reaction buffer (50 mM HEPES, pH 7.5, 0.5 mM TCEP). Protein was concentrated to 2 mg mL^{-1} and 221 μL (50 nmol) was mixed with Cy3B-maleimide (150 nmol, GE Life Sciences) in a final volume of 300 μL and agitated for 2 h at 25 °C. Labelled protein was then further purified with a P2 Centri-Pure desalting column (EMP Biotech) with degassed buffer (50 mM Na_2HPO_4, 150 mM NaCl).

3.3.13 MYCBP2 thioester/ester trapping assay

UBE1 (2 μM) was mixed with Cy3B-Ub (1 μM) in buffer (40 mM Na_2HPO_4-HCl pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 5 mM $MgCl_2$) (Figure 5.4, *Lane 1, 2*). The reaction was then initiated by the addition of ATP (5 mM) and incubated for 10 minutes at 25°C. Samples (*lane 3*, *4)* were taken and combined with UBE2D3 (10 μM). After a further 10 minutes at 25°C, samples (lane 5, 6) were taken and combined with GST-$MYCBP2_{cat}$ (WT, C4520S, C4520A, C4572S, C4572A, C4520A/C4572S or C4520S/C4572A) (15 μM). The reactions were incubated at 25 °C for 30 seconds and terminated by the addition of 4X LDS loading buffer (either non-reducing or

reducing). For Ub~GST-MYCBP2$_{cat}$ C4572S ester bond cleavage, 0.14 N NaOH was added after E3 reaction with E1/E2 mixture for 30 seconds and then further incubated at 37 °C for 20 minutes. The gel was then scanned with a Chemidoc Gel Imaging System (BioRad).

3.3.14 Multiple turnover amino acid and peptide panel discharge assays

Stock solutions (0.5 M) of amino acids were dissolved in MQ water and pH was adjusted to pH ~8. Peptides of the sequence Ac-EGXGN-NH_2 (X= K, S or T) were obtained from Bio-Synthesis Inc. Stock peptide solutions (200 mM) were dissolved in MQ water and pH was adjusted to pH ~8. An E2 (UBE2D3) charging reaction was carried out in buffer (40 mM Na_2HPO_4-HCl pH 8.0, 150 mM NaCl, 0.5 mM TCEP) containing UBE1 (250-500 nM), UBE2D3 (20 μM), Ub (50 μM), or Cy3B-Ub (25 μM), $MgCl_2$ (5 mM) and ATP 10 (mM). The reaction was incubated at 37 °C for 15 minutes and then equilibrated to 23 °C for 3 minutes. An equivalent volume of nucleophile sample containing small molecule/peptide nucleophile (100 mM) and GST-MYCBP2 (10 μM) was then added and incubated at 23 °C. Samples were taken at the specified time points and analyzed as described for Tris/glycerol-mediated E2 discharge assay**.** Cy3B-Ub was visualized using a Chemidoc Gel Imaging System (Biorad). LC-MS was carried out as described for Tris/glycerol discharge but amino acid substrate samples were quenched by the addition of 2:1 parts quenching solution (75 % acetonitrile, 2 % TFA) and peptide substrate samples were quenched by addition of 1:1 parts quenching solution.

3.3.15 Multiple turnover E2 discharge panel

E2s were screened for threonine discharge activity with GST-MYCBP2cat as described for the amino acid panel. E2s were also incubated in the presence of

threonine but in the absence of GST-MYCBP2cat. These samples provided a reference to distinguish between intrinsic E2~Ub instability and E3-dependent discharge. All the E2s were stored in phosphate buffer.

3.3.16 Single turnover E2 mutant discharge by in-gel fluorescence

E2 mutants[126,159,242,450,451] (10 μM) were charged with Cy3B-labelled Ub (12.5 μM) in a final volume of 12 μL at 37 °C for 20 minutes then cooled on ice for 3 minutes. E2 recharging was then blocked by the addition of MLN4924 derivative, Compound 1 (25 μM)[452], which inhibits E1, and then incubated for a further 15 minutes, 23 °C. The mixture was then mixed with 12 μL of GST-MYCBP2$_{cat}$ (5 μM) and threonine (100 mM) and incubated at 23 °C for the specified time. Analysis was carried out as for multiple turnover assays. To account for intrinsic E2~Ub instability the mean % discharge (n=2) calculated against a parallel incubation where E3 was withheld. Data were plotted using Prism (Graphpad).

3.3.17 Expression and purification of ARIH1 and UBE3C (performed by Dr. Axel Knebel)

ARIH1 residues 1-394 (DU24260) was expressed as an N-terminally GST-tagged fusion protein in BL21 cells. UBE3C residues 641-1083 (DU45302) was expressed as a N-terminally GST-tagged fusion protein in Sf21 cells using the baculovirus infection system.

3.3.18 Calculation of observed rate constants for E3-substrate dependent single turnover E2~Ub discharge

UBE2D3 or UBE2L3 (5 μM) were charged with Cy3B-labelled Ub (8 μM) in a final

volume of 30 μL at 37 °C for 25 minutes then incubated at 23 °C for 3 minutes. Single turnover conditions for E2~Ub discharge were achieved by E1 inhibition with MLN4924 derivative, Compound 1 (25 μM) and then incubated for a further 15 minutes. The mixture was then mixed with 30 μL of $MYCBP2_{cat}$ or $ARIH1_{1-394}$ (HHARI) or $UBE3C_{641-1083}$ (1 μM) and threonine (100 mM) and incubated at 23 °C for the specified time. Samples were quenched with non-reducing 4X LDS loading buffer and resolved by SDS-PAGE (Bis-Tris 4-12 %, MES buffer). The gel was then scanned with a Chemidoc Gel Imaging System (BioRad) and subsequently Coomassie stained. E2~Ub signals were quantified using the Fiji software. Observed rate constants were obtained by fitting reaction progressive curves to a single exponential function using Prism (Graphpad Software).

3.3.19 NMNAT2 ubiquitylation assay

NMNAT2 (5 μM) was mixed with E1 (500 nM), UBE2D3 (10 μM), $MYCBP2_{cat}$ (10 μM), Ub (50 μM), ATP (10 mM) and made up with 10X pH 7.5 buffer (40 mM $Na_2H_2PO_4$ pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 0.5 mM TCEP). The reactions were incubated at 37 °C for 1 hour and terminated by the addition of 4X LDS loading buffer (either non-reducing or reducing). For base lability test, reactions were supplemented with 0.14 N NaOH and then further incubated at 37 °C for 20 minutes.

3.3.20 Immunoblotting

The immunoblotting were performed with the same method as described in section 2.2.20. The His-tagged species were probed with 1:10000 anti-His primary antibody (Clontech, #631212). Alpha tubulin (1E4C11) mouse mAb (Proteintech®) was used at 1:10000 dilution. The MYCBP2 antibody was used at 0.5 μg mL^{-1} and

raised by in sheep by MRC PPU Reagents and Services and affinity-purified against the indicated antigen: anti-MYCBP2 (2nd bleed of SA357, residues 4378-4640 of human MYCBP2). Mouse monoclonal NMNAT2 antibody (clone 2E4; Sigma Aldrich) was used at 0.5 μg mL^{-1}.

3.3.21 Cell culture and lysis protocol

SH-SY5Y cells were cultures as previously described in section 2.2.18. HEK293 were cultured (37 °C, 5 % CO_2) in DMEM supplemented with 10 % (v/v) FBS, 2.0 mM L-glutamine, and antibiotics (100 units mL^{-1} penicillin, 0.1 mg mL^{-1} streptomycin). Cell transfections were performed using polyethylenimine (Polysciences) according to the manufacturer's instruction. PEI stock (1 mg/mL) was prepared by dissolving PEI powder in 20 mM Hepes (pH 7). Aliquots were filter sterilized (0.22 μm) and stored at -80 °C. Cells were grown to 50 % confluency on 10 cm dishes; 10 μg of DNA was mixed with 30 μL 1mg/mL PEI and 900 μL serum-free DMEM, vortexed for 5 secs and incubated for 20 - 40 min at room temperature before being added to cells. Cells were harvested 24hrs post transfection. MG-132 (50 μM) was added to cells two hours prior to lysis. Cells were rinsed with ice-cold PBS and extracted in lysis buffer (1 % NP-40, 50 mM Tris-HCl pH 7.5, 1.0 mM EGTA, 1.0 mM EDTA, 0.27 M sucrose, 10 mM sodium 2-glycerophosphate, 0.2 mM PMSF, 1.0 mM benzamidine, 1.0 mM sodium ortho-vanadate, 50 mM sodium fluoride and 5.0 mM sodium pyrophosphate, 50 mM iodoacetamide and cOmplete™, EDTA-free protease inhibitor cocktail (Roche)). Lysates were then clarified by centrifugation at 4 °C for 30 min at 14,800 rpm. Supernatants were collected (total cell extracts) and protein concentration determined by Bradford assay. For the base-lability test, indicated cell lysates were further incubated with 0.5M hydroxylamine, pH 9.0 at 37 °C for 30 minutes.

3.3.22 MYCBP2cat crystallization (collaboration with Dr. Karim Ramfie and Dr. Daan van Aalten)

MYCBP2 was expressed as described for untagged protein. After protease cleavage of the tag the protein was further purified by size exclusion chromatography using an ÄKTA FPLC system and a HiLoad 26/600 Superdex 75 pg column (GE Life Sciences). The running buffer consisted of 20 mM HEPES pH 7.4, 150 mM NaCl, 4 mM DTT. Combined fractions were concentrated to 10.4 mg mL^{-1}. Sparse matrix screening was carried out and Bipyrimidal crystals were obtained from the Morpheus screen condition C1 (Molecular Dimensions). A subsequent optimization screen carried out by Dr. Virdee yielded multiple crystals (Buffer system 1 (MES/imidazole) pH 6.7, 23.3 mM Na_2HPO_4, 23.3 mM $(NH_4)_2SO_4$, 23.3 mM $NaNO_3$, 18 % PEG500 MME, 9 % PEG20000). A single crystal was soaked in mother liquor and further cryoprotected by supplementation with 5 % PEG400 and frozen in liquid N_2. Data were collected to 1.75 Å at the European Synchrotron Radiation Facility at Beamline ID23-1. Energy was set to the peak value of 9.669 keV (1.2823 Å), as determined by an absorption edge energy scan. A total of 360 ° were collected with an oscillation range of Ω = 0.1 °. The phase problem was solved by locating 6 Zn^{2+} sites in the anomalous signal and solvent flattening with the SHELX suite. An initial model was built by ARP/wARP[38] and subsequently optimized by manual building in COOT[39] and refinement with REFMAC5[40] resulting in the final model with statistics as shown in **Extended Data Table 1**. Final Ramachandran statistics were favored: 95.55 %, allowed: 3.24 %, outliers: 1.21 %. Coordinates have been deposited with the Protein Data Bank (PDB ID: 5O6C).

3.3.23 Size exclusion chromatography with multi-angle light scattering (SEC-MALS) (performed by Dr. Ramasubramanian Sundaramoorthy)

SEC–MALS experiments were performed on a Ultimate 3000 HPLC system (Dionex) with an in-line miniDAWN TREOS MALS detector and Optilab T-rEX refractive index detector (Wyatt). In addition, the elution profile of the protein was also monitored by UV absorbance at 280 nm. A Superdex 75 10/300 GL column (GE Life Sciences) was used. Buffer conditions were 50 mM Na_2HPO_4 pH 7.5, NaCl 150 mM, 1.0 mM TCEP and a flow rate of 0.3 mL min^{-1} was applied. Sample (50 µL, 5.5 mg mL^{-1}) was loaded onto the column with a Dionex autosampler. Molar masses spanning elution peaks were calculated using ASTRA software v6.0.0.108 (Wyatt).

3.3.24 Mediator loop modeling (performed by Dr. Peter Mabbitt and Dr. Satpal Virdee)

Mediator loop residues where built and geometry optimized within the Bioluminate Software (Schrödinger). Side chains were modified within COOT[39] and figures were generated with Pymol (Schrödinger). Ramachandran analysis was carried with the RAMPAGE server[453].

3.3.25 Bioinformatic analysis (performed by Dr. Kay Hofmann)

Proteins belonging to the RCR family were identified by generalized profile searches. Overall 671 such sequences were identified. The sequences were aligned by profile-guided alignment using the pftools package. For identifying representative sequences from various taxa, the Belvu program (Sanger Institute) was used to remove sequences with >80 % identity to other sequences. Truncated and misassembled proteins were removed manually, resulting in 130 representative TC domain sequences.

Chapter 4. Development of Novel E3 Activity-Based Probes

4.1 Introduction

The ubiquitylation cascade regulates multiple cellular functions and is involved in numerous diseases. The distinct transfer cascade including E1-E2-E3 enzymes serve as a good target for drug development. Furthermore, how ubiquitin is transferred from E2~Ub to E3 ligases and how substrate specificity is conferred by many E3 ligases are poorly understood. However, to analyze the intrinsic E3 activity has proven to be a challenge task in the field since there is currently no good tool for directly monitoring ligase activity. Furthermore, because of the short reaction time among ubiquitin transfer process and the transient nature of thioester bond, it is difficult to trap the thioester-linked ubiquitin intermediates among all three enzymes. New tools for validating the mechanisms of ubiquitin transfer, the determination of substrate selection, and the identification of E3 substrates are urgently needed. To dissect the regulatory mechanism of HECT/RBR E3 ligases is the main focus of this thesis. Furthermore, the transthiolation ability of these E3s has been poorly addressed even though it is a fundamental step for ubiquitin transfer (**Figure 4.1a**). In addition, currently cellular E3 activity can only be monitored by the autoubiquitination assay which is not typically observed for endogenous E3s.

Activity-based probes (ABPs) have shown their potential for rapidly profiling various enzymes activity in recent years[302,454]. They can also be the powerful structural tools by stabilizing transient enzyme intermediates[108,455]. A key question initially set at to address with a functional ABP was to study the mechanism of ubiquitin transfer for an RBR E3. If we could engineer a stable E2~Ub conjugate this could serve as a recognition element. Careful incorporation of a 'warhead' would

furnish an ABP that covalently and specifically labels the incoming catalytic cysteine in an active E3 (**Figure 4.1b**). To date, our understanding of how an RBR E3 ligase recruits E2~Ub and mediates transthiolation is poorly understood. Additionally, because this ABP is designed with the aim of specifically targeting the catalytic cysteine in any active E3 that functions with an E2 which the ABP was based on, we could simultaneously profile the E3 activities in complex cellular extracts.

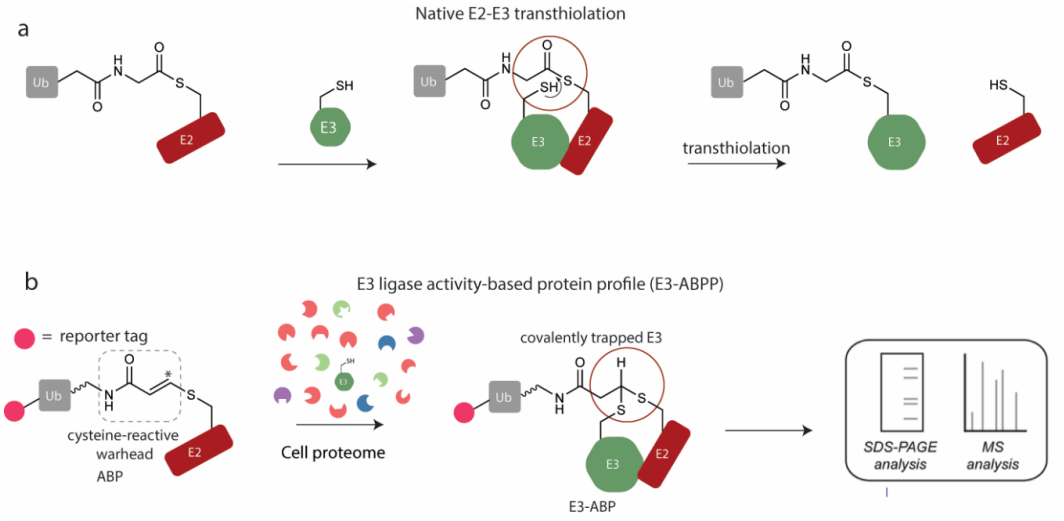


Figure 4.1 Transthiolation between E1s to E3s through juxtaposition of active site cysteine could be abrogated by warhead introduction.

(**a**) Reaction scheme depicting native transthiolation between E2~Ub and HECT/RBR E3 ligases. (**b**) Incorporation of a warhead in E2~Ub makes it as a probe to trap E2-E3 intermediate. Validation of these complexes can be achieved by various methods including mass spectrometry.

Previous research has demonstrated that ABPs consisting of an E2 and a warhead on them can only specifically and efficiently target the catalytic cysteine in E1 and is able to trap an E1-E2 complex in the *in vitro* and *in situ* assays[456]. The chemistry used for probe targeting here ensures the warhead on the probe can only target cysteine residues which come into close proximity to allow crosslinking

to occur. However, the first generation E2-conjugated activity based probes barely labeled any recombinant HOIP RBR E3 ligase *in vitro*, which consists of only the active RBR domain[456]. Ubiquitin itself also seems to play an important role in E2-E3 binding processes and generation of a transthiolation competent conformation. We therefore believed the first generation probes were missing a vital component, the ubiquitin molecule. As a result, developing a E2-ubiquitin-conjugated E3 activity-based probe (E2~Ub*ABP; ubiquitin stably conjugated to an E2 displaying a proximal thiol-reactive warhead) for detecting E3 activity is therefore the first step and the main focus of this project.

4.2 Strategies for preparation of E2-ubiquitin-conjugated E3 Activity-Based Probes

Cysteine−cysteine juxtaposition is the prerequisite for transthiolation activity. The catalytic cysteine that resides in the middle of the E2 polypeptide is the essential residue for thiol-reactive warhead installation and E1/E3 activity-based probe construction. Putting the warhead on E2 ensures the E2 is the recognition element of the probe. A strategy for generating an E3 ABP would be to prepare a stable E2~Ub conjugate with a warhead in place of the thioester bond. We reasoned this could be done by using a trifunctional component that could link E2 to the C-terminus of Ub and also allow warhead attachment.

4.2.1 E3 ABPs synthesis strategy 1: Genetic code expansion technology

In order to introduce the tri-functional molecule linker into ubiquitin and E2s, the first strategy was to use genetic code expansion technology for incorporating an unnatural amino acid containing a *δ*-thiol lysine[457] in place of the catalytic cysteine of E2 enzymes (**Figure 4.3**). In this instance, the genetic incorporation of the

unnatural amino acid into the E2 backbone could be considered the first functionalization step. Mediated by the *δ*-thiol group, the *ε*-amino group could then be reacted with a thioester group on ubiquitin to form a stable ubiquitin-loaded E2 enzyme by formation of an isopeptide bond[125]. This can serve as a stable mimic of a E2~Ub thioester[125]. Subsequent introduction of a thiol-reactive warhead on the proximal *δ*-thiol group by reaction with a TDAE (Tosyl substituted-Doubly Activated Ene, **Figure 4.2**), a bifunctional thiol reactive molecule, could then be carried out[456]. As we discussed in *Chapter 2*, TDAE is a novel compound which allows us to introduce an activated vinylsulfide (AVS) on the E2 catalytic cysteine by an unexpected addition-elimination mechanism[458]. However, the synthetic route for this particular unnatural amino acid bearing a δ-thiol lysine was non-trivial[457]. Herein, I successfully reduced the number of synthetic steps from nine to five for getting the *δ*-thiol-*N*ε-(p-nitrocarboxybenzyloxy)lysine (CBZ-Lys)[445] (**Scheme 4.1**). I successfully incorporated the *Nα*-(*tert*-Butoxycarbonyl)-L-lysine (Boc-Lys) into the position 85 (catalytic cysteine) of UBCH5C; however, I failed with the CBZ-Lys incorporation (data not shown). The incorporation of derivative *δ*-thiol-*N*ε-(((allyloxy)carbonyl)amino)lysine, which is structurally similar to the CBZ-Lys, into UBCH5C was also failed. We reasoned that it was the required *δ*-thiol group in these lysines that was not tolerated at the catalytic cysteine position in the E2 enzyme. Although this approach was unsuccessful, the optimized route for synthesizing CBZ-Lys should prove to be valuable as incorporation of this amino acid allows traceless and site-specific ubiquitylation of recombinant proteins[457]. Because I did not further apply these compounds in my project, the related data are not included.

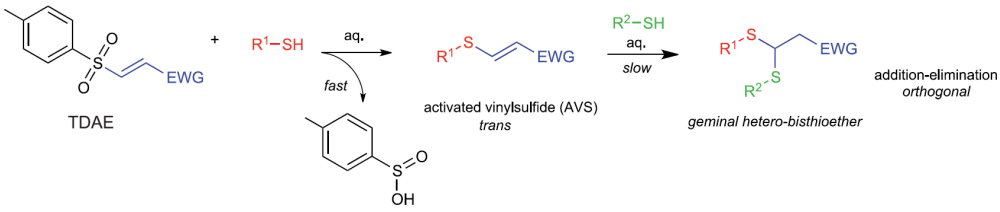


Figure 4.2 TDAE is a powerful molecule for targeting cysteine-containing proteins.

The trans activated vinylsulfide (AVS) is obtained by the reaction of a tosyl-substituted doubly activated ene (TDAE) with a thiol affords. The reactivity-tunable AVS then acts as a warhead to covalently reacts with a cysteine in close proximity. Addition of a second thiol to the AVS is slower than the first reaction which enabling an orthogonal thiol functionalization (OTF).

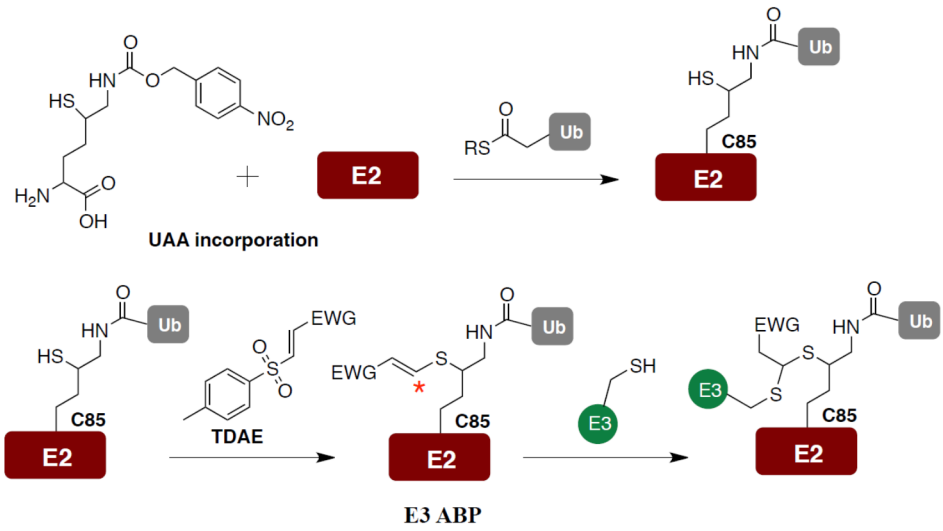
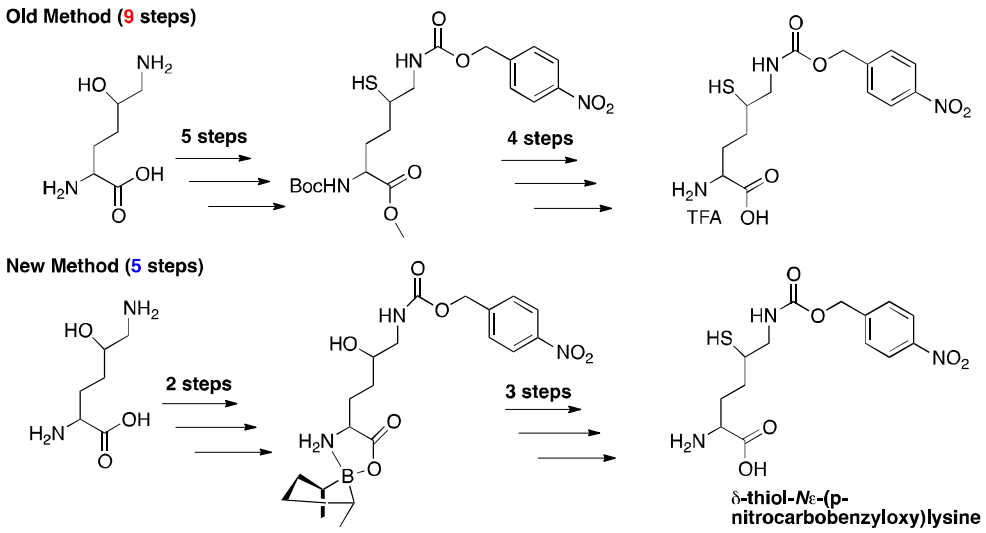


Figure 4.3 Genetically introduction of an unnatural amino acid as a tri-functional linker to E2 and construction of an E3 ABP.

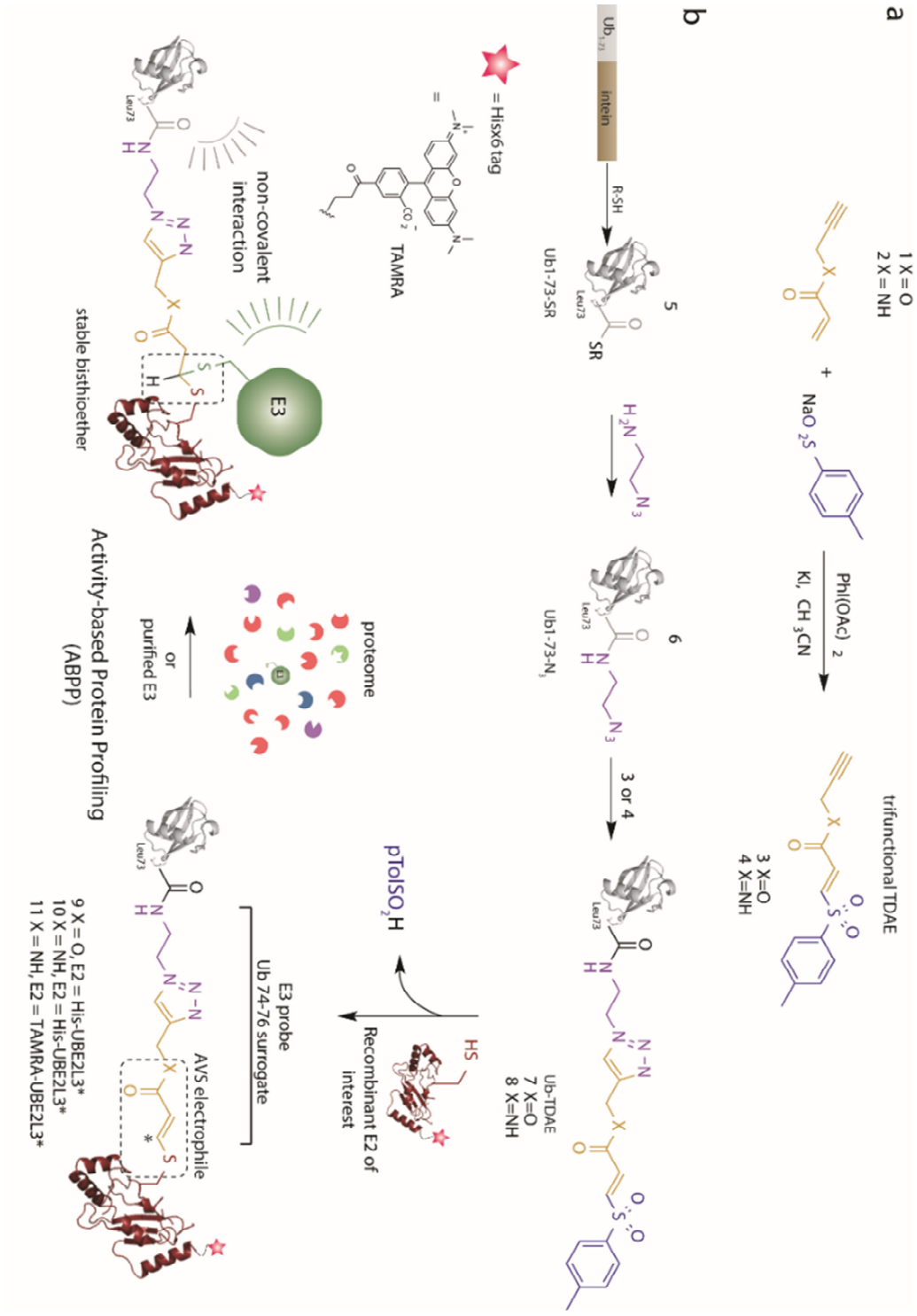
The proposed synthetic route for constructing E3 ABPs by unnatural amino acids incorporation. The incorporation of CBZ-lysine on the position of catalytic cysteine on E2 allows us to generate a stable E2~Ub conjugate. The thiol group on the CBZ-lysine then reacts with TDAE in a mild buffer condition and affords a novel E3 ABP. The asterisk corresponds to the cysteine reactive carbon atom in the ABP.



Scheme 4.1 The application of different protecting group (9-BBN) reduced the synthetic steps for obtaining δ-thiol-*N*ε-(p-nitrocarbobenzyloxy)lysine.

4.2.2 E3 ABPs synthesis strategy 2: Construction of ubiquitin-conjugated E3 ABPs via trifunctional TDAE

Since we failed to incorporate a chemical handle into E2 which should allow us to link the ubiquitin molecule and ligate a cysteine-specific warhead, we designed and synthesized a different trifunctional molecule. We reasoned that appending a tosyl-substituted doubly activated ene (TDAE)[334] to the C-terminus of Ub would generate a highly electrophilic, but thiol-specific, protein conjugate that could undergo addition-elimination with the catalytic cysteine residue of recombinant E2s allowing the production of a stable mimetic of E2~Ub harboring a tried and tested warhead. To achieve this the TDAE would need to be a tri-functional molecule that I would first react with the C-terminus of Ub and then subsequently link to the E2 forming a conjugate containing a mechanistically positioned warhead. We reasoned that an alkyne moiety attached to a TDAE would satisfy this requirement (**Scheme 4.2a**). The Ub-TDAE conjugate could then be prepared by triazole formation between azide-functionalized Ub, that could be prepared using the procedure used for DUB probes, and alkyne-functionalized, TDAEs using Copper-catalyzed Azide-Alkyne Cycloaddition (CuAAC)[459].



Scheme 4.2 Synthesis of alkyne-functionalized TDAEs and their utility for assembling E2~Ub-based probes for profiling E3 transthiolation activity.

(a) Alkyne-functionalized, electron-deficient acrylate **1** and acrylamide **2** were used to prepare trifunctional TDAEs **3** and **4**, respectively. This was achieved by $PhI(OAc)_2$/KI-mediated reaction of **1** and **2** with sodium arenesulfinate. **(b)** Probe construction involved production of Ub truncated to residues 1-73 bearing a C-terminal thioester (Ub1-73-SR). (*Next page*)

Ub1–73-SR was prepared by thiol (R-SH) treatment of a recombinant intein fusion protein. A truncated Ub species was chosen for molecular modeling, based on RING-E2~Ub and HECT-E2~Ub co-crystal structures, indicated that the AVS-triazole linker would satisfactorily mimic the Ub C terminus and serve as a surrogate for residues 74–76 (**Figure 4.4**). Preparation of Ub1–73-SR **5** was achieved by thiolysis of an intein fusion protein. Aminolysis of the Ub1–73-SR with azidoaminoethane afforded azide-functionalized Ub (Ub1–73-N_3) **6**. Copper-catalyzed azide-alkyne cycloaddition (CuAAC) between **6** and TDAEs **3** and **4** yielded TDAE-functionalized Ub molecules (Ub-TDAEs) **7** and **8**. Mild incubation of Ub-TDAEs **7** or **8** with recombinant reporter tagged E2 generated E2~Ub-based probes bearing thioacrylate and thioacrylamide electrophiles poised for actvity-based labeling of catalytic cysteine nucleophiles in E3 ligases. Reporter tags were either a Hisx6 epitope or a 5-carboxytetramethylrhodamine (TAMRA) fluorophore. Ub in E2~Ub makes noncovalent contacts with E3 that are likely to be involved in coordinating a catalytically competent conformation.

We designed and synthesized the trifunctional-TDAE **3** and **4** molecules (**Extended figure 4.1- 4.4**) which incorporated an additional alkyne appendage (TDAE-alkyne) that serves as a vital chemical handle for the click reaction and also harbors latent thioacrylate and thioacrylamide AVS electrophiles, respectively (**Scheme 4.2a**). To obtain the azide-functionalized Ub, different lengths of ubiquitin-thioesters were produced by an intein system[444] with excellent yields after optimization (>10mg protein per Liter LB). Instead of full-length ubiquitin, C-terminal truncations were generated so as to spatially mimic the native E2-Ub conjugate upon incorporation of the warhead and triazole linkage by successful 'click' ligation. A computational model based on Ub-loaded UBCH5B in complex with various E3 ligase (**Figure 4.4**) suggested that the truncated ubiquitin1-73 was the best choice for building the probe. Unfortunately, poor yields of ubiquitin 1-73-SR (thioester) were obtained initially due to the *in vivo* auto-cleavage of the intein fusion that released truncated ubiquitin during protein expression. This was remedied by mutation of a threonine to cysteine at position 3 of the intein protein which prevents the auto-cleavage from occurring[460]. Ubiquitin1-73 thioester, **5**,

could then be successfully generated after modification of the intein system and was purified by preparative reversed phase high-performance liquid chromatography (RP-HPLC) (**Figure 4.5a**).

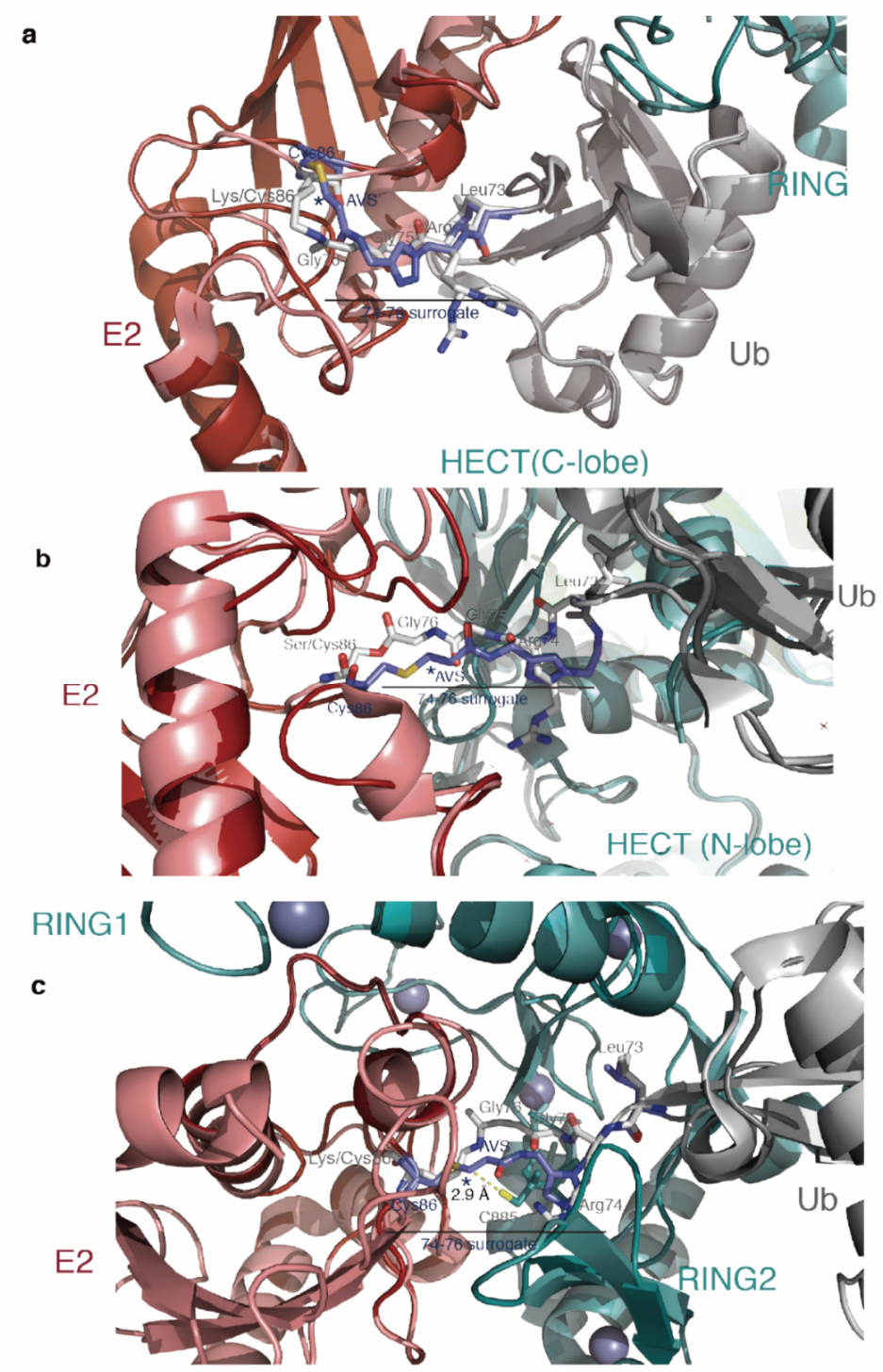


Figure 4.4 Molecular modeling of RING-, HECT- and RBR-E3 ABP complexes. (*Next page*)

(a) Crystal structure of RING E3 ligase in complex with E2~Ub (PDB ID 4AP4, thioester linkage was stabilized by mutation of E2 catalytic cysteine to lysine and an isopeptide was formed)[125]. E2 (salmon), RING (light teal) and Ub (light gray) are in cartoon format. Ub residues 74-76 are in gray stick format. The computational model, where residues 74-76 have been substituted for the AVS-triazole linkage (blue stick notation) in probe **10**, is superposed. In the computational model E2 (firebrick), RING (dark teal) and Ub (dark gray) are in cartoon format. The modeling suggests that the AVS-triazole satisfies *(next page)* the conformational requirements of E2~Ub for RING engagement. **(b)** Crystal structure of HECT E3 ligase in complex with E2~Ub (PDB ID 3JW0, thioester linkage was stabilized by mutation of E2 catalytic cysteine to serine and an ester was formed[128]). E2 (salmon), HECT (light teal) and Ub (light gray) are in cartoon format. Ub residues 74-76 are in gray stick format. In the computational model E2 (firebrick), HECT (dark teal) and Ub (dark gray) are in cartoon format. The modeling suggests that the AVS-triazole also satisfies the conformational requirements of E2~Ub for HECT engagement. **(c)** Crystal structure of RBR E3 ligase (HOIP) in complex with E2~Ub (PDB ID 3EDV, thioester linkage was stabilized by mutation of E2 catalytic cysteine to lysine and an isopeptide was formed[129]). E2 (salmon), RBR (light teal) and Ub (light gray) are in cartoon format. Ub residues 74-76 are in gray stick format. In the computational model E2 (firebrick), RBR (dark teal) and Ub (dark gray) are in cartoon format. The modeling suggests that the AVS triazole also satisfies the conformational requirements of E2~Ub for HOIP engagement. The incoming cysteine nucleophile in RING 2 (C885) is 2.9 Å from the reactive center in the AVS electrophile. (*The figures are provided by Dr. Virdee*)

To obtain the necessary azide-functionalized ubiquitin for click reaction, ubiquitin 1-73 thioester **5**, underwent on aminolysis reaction with 2-azidoethanamine in the presence of triethylamine at 30°C for 16 hours yielding ubiquitin-azide **6** (**Scheme 4.2b** & **Figure 4.5b**). Conjugation of ubiquitin-azide **6** with TDAE-alkyne **3** and **4** was successfully achieved by developing optimized conditions for CuAAC yielding ubiquitin-TDAE **7**, **8** (**Scheme 4.2b** & **Figure 4.5c**). E2~Ub*ABP (E3 ABP) was then made by treating ubiquitin-TDAE with different E2s under native but non-reducing condition (**Figure 4.6a-c**). To ensure that the TDAE only reacts with the catalytic cysteine, C17S and C137S on E2 UBE2L3* (UBCH7) were introduced. These mutations on E2s did not compromise activity as the

mutant supported Parkin activity the same as wild-type (**Figure 4.7**). UBE2L3* was then successfully conjugated with ubiquitin-TDAE **7**, **8** and became novel E3 activity-based probes **9** and **10** (**Figure 4.6a-c**). In conclusion, we have developed a facile, modular and scalable method for synthesizing ubiquitin-conjugated E2 as E3 activity-based probes. Herein, I would also like to highlight the user-friendly method we had developed for probe construction since it largely enhances the possibility for large scale production for future applications. It is extremely important for method development since we would like to apply the ABP in crystal structure studies which generally require large milligram amounts of proteins for condition-screening.

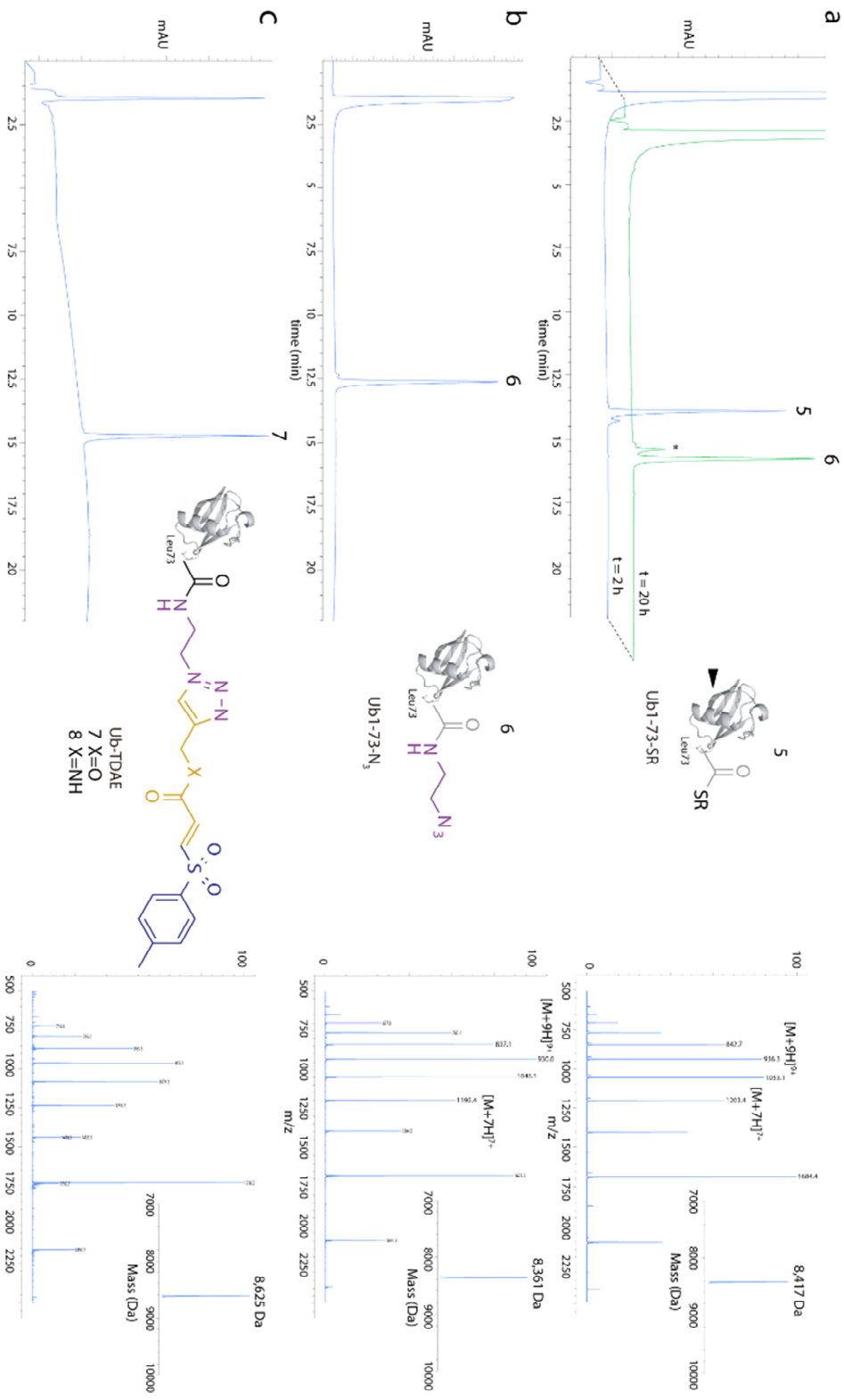


Figure 4.5 Characterization data for the preparation of intermediates (5), (6) and (7).

(**a**) Left panel, Ub thioester **5** (blue) underwent near quantitative aminolysis with azidoaminoethane (green), in aqueous DMSO buffer containing 4 % triethylamine, after 16 hincubation at 30 °C. The peak denoted with an asterisk corresponds to hydrolysis product of **5**. Reaction was monitored by HPLC at 214 nm. (*Next page*)

Right panel, ESI-MS mass spectrum for starting thioester **5**. Inset, deconvoluted mass spectrum (observed mass = 8,417 Da; calculated mass = 8,418 Da). **(b)** Left panel, analytical HPLC analysis of purified aminolysis product **6**. Right panel, ESI-MS mass spectrum of **6**. Inset, deconvoluted mass spectrum (observed mass = 8,361 Da; calculated mass = 8,362 Da). **(c)** HPLC analysis of the purified Ub-TDAE conjugate **7** obtained by Copper-catalyzed Azide-Alkyne Cycloaddition (CuAAC) between azide-functionalized Ub **6** and TDAE **3** in pH 7.5 phosphate buffer, $CuSO_4$ and Tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) after 15 min incubation at 23 °C. Right panel, ESI-MS mass spectrum of **7**. Inset, deconvoluted mass spectrum (observed mass = 8,625 Da; calculated mass = 8,626 Da).

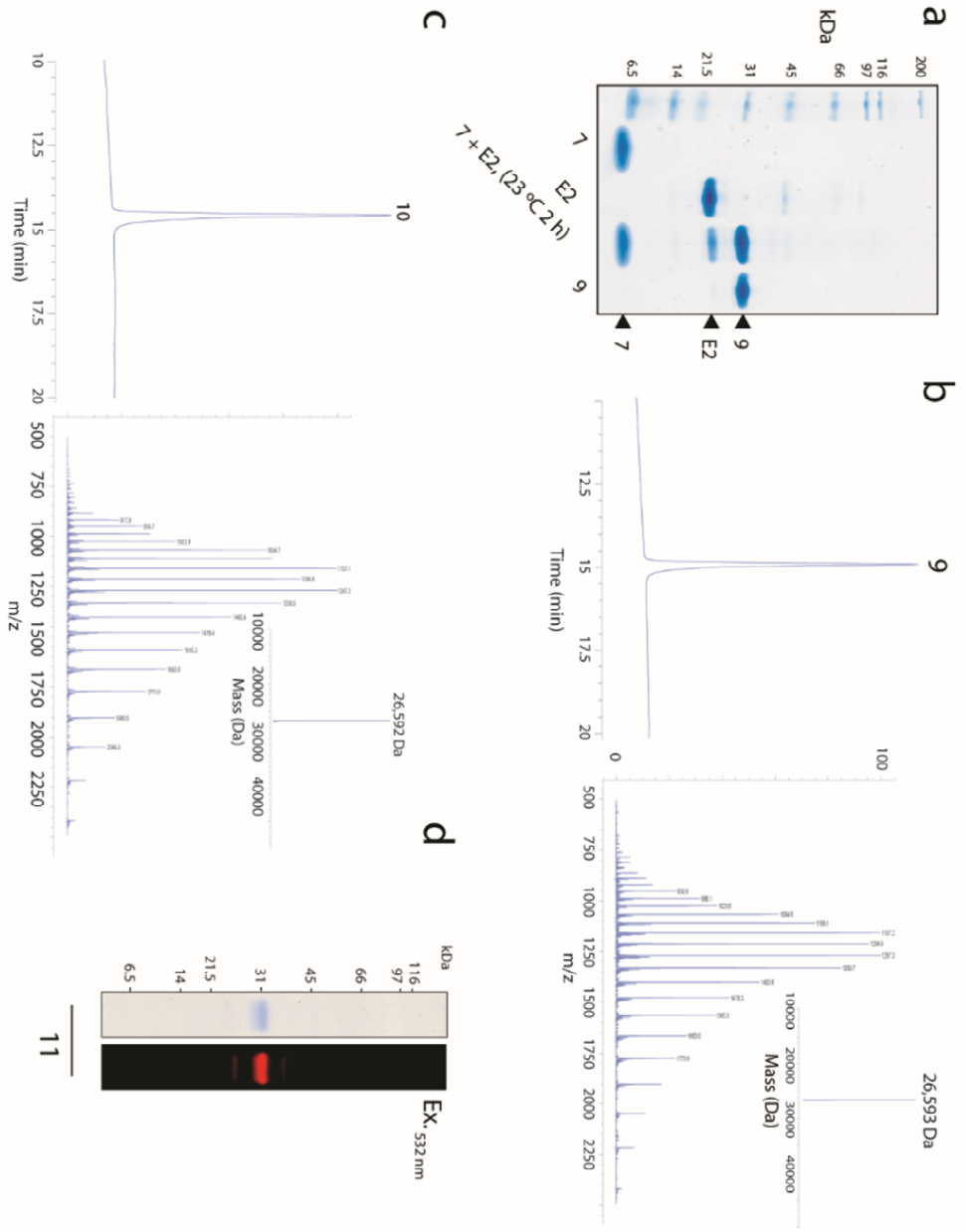


Figure 4.6 Characterization data for the preparation of ABP (9), (10) and (11).

(a) Efficient addition-elimination reaction between **7** (lane 2) and E2 (His-UBE2L3*) (lane3) to produce E2~Ub conjugate, probe **9** (lane 4). **5** (2 eq.) was added to His-UBE2L3* in phosphate buffered saline and incubated for 2 h at 23 °C). Probe **9** was then purified by size exclusion chromatography (lane 5). (**b**) Left panel, HPLC analysis of purified probe **9**. Right panel, ESI-MS spectrum of **9** (observed mass = 26,593 Da; calculated mass = 26,598 Da). The N-terminal Hisx6 tag has been proteolytically cleaved for clarity. (**c**) Left panel, LC-MS characterization of probe **10** prepared by reaction between 6 and His-UBE2L3*. Observed mass = 26,592 Da; theoretical mass = 26,597 Da (N-terminal Hisx6 tag cleaved). (**d**) SDS-PAGE analysis of fluorescent TAMRA-conjugated probe **11**. Left lane is a coomassie stain and the right lane is a scan of in-gel fluorescence using an imaging system (excitation λ = 532nm).

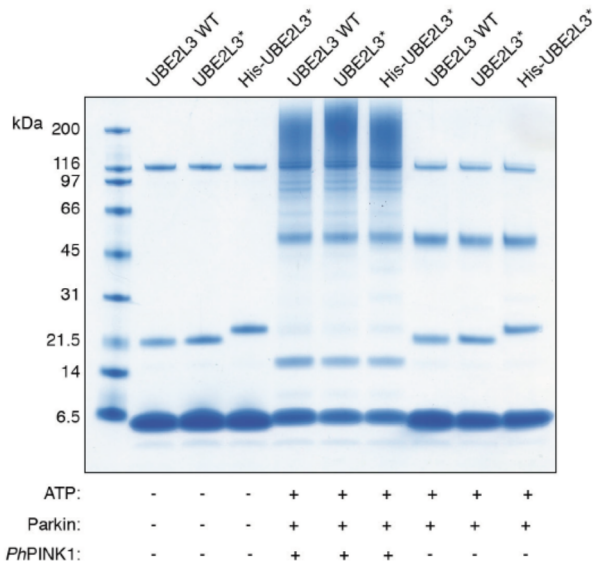


Figure 4.7 Tagged and untagged forms of double mutant C17S C137S UBE2L3 (UBE2L3*) support Parkin activity.

The His-tagged form of UBE2L3* that probes **9**, **10** and **11** were based on was tested for its ability to support Parkin activity (N.B. probe **11** contained a Ser to ProcK mutation N-terminal of the His-tag that was labelled with a TAMRA fluorophore). Both His-UBE2L3* and UBE2L3* (where the N-terminal Hisx6 tag was removed with Rhinovirus C3 protease) demonstrated activity comparable to WT protein in a qualitative SDS-PAGE gel-based activity assay. Unless otherwise labeled, 50 µL reactions consisted of 50 mM Tris-HCl pH7.5, MgCl2 (5 mM), ATP (2 mM), UBA1 (240 nM), UBE2L3 variant (2 µM), Ub (58 µM), Parkin (1.2 µM), GST-*Ph*PINK1 (383 nM). Reactions were incubated at 30 °C for 1 h, resolved by SDS-PAGE, and visualized by coomassie staining.

4.2.3 Construction of fluorophore-tagged E3 ABP for systematically profiling

In order to expand the application of our ABPs, I focused on the reporter tag replacement which provide us choices for different readout based on the native of the assay. Furthermore, by simply introducing a biotin-tag it could make this probe a very powerful tool for activity-based proteomics. A versatile handle on probes also allows us to swap various tags via facile reactions. We therefore genetically incorporated a bio-orthogonal alkyne handle into UBE2L3* through the unnatural amino acid propargyloxycarbonyl-L-lysine (ProcK). This was positioned N-terminal to the hexahistidine (Hisx6) tag (UBE2L3*-ProcK). The alkyne group allows us to

use well-developed CuAAC click reaction for tag conjugation. Azide-functionalized 5-carboxytetramethylrhodamine (TAMRA) was then used to label UBE2L3*-ProcK, yielding UBE2L3*-TAMRA. The fluorescent probe **11** was then generated via conjugated Ub-TDAE **6** with UBE2L3*-TAMRA by the same method I used in probe **9** and **10** construction. (**Figure 4.6d, Extended Figure 4.5: the LC-MS data for ABP 11**).

4.3 Validation of novel synthesized E3 Activity-Based Probes with the RBR E3 Parkin

4.3.1 RBR E3 ligase, Parkin, is labeled with novel E3 ABPs in an activity-dependent manner

In order to validate the functionality of our probes, we attempted a labeling assay with probe **9** and **10** with recombinant Parkin. Since the dual phosphorylation on both Parkin and ubiquitin might be required for optimal activity, we generated phosphorylated-Parkin (p-Parkin) and phosphorylated-ubiquitin (p-Ub) via incubation with PINK1 and ATP. We obtained nearly 60% p-Parkin and quantitative p-Ub. Purification was by size exclusion chromatography and we confirmedthe purity by LC-MS and Phos-tag SDS-PAGE. The activity of p-Parkin was then further tested by auto-ubiquitylation assay (data not shown). We then incubated probe **9** and **10** with either non phosphorylated-Parkin or p-Parkin at 30°C for four hours in the presence of p-Ub. The reaction was resolved by SDS-PAGE and validated by Coomasie staining or immunoblotting (**Figure 4.8a**). The only protein bearing Hisx6 tag is our probes which allowed me to use highly specific Hisx6 antibody to analyze the labeling. In **Figure 4.8a**, probe **9** and **10** obtained efficient labeling with p-Parkin but not with non- phosphorylated Parkin (lane 2,3 versus lane 4,5). The probe **10** likely gained more efficient labeling with p-Parkin due to the greater electrophilicity of the acrylamide electrophile, to it being a closer mimic of the Ub

C terminus, or due to a combination of both. Moreover, to dissect whether probe labeling was attributable to E2 engagement of the RING1 domain on Parkin, we also generated a mutant version of probe **9** (**9 F63A**) containing a Phe63Ala mutation in the UBE2L3* component of the ABP (**Extended Figure 4.6**). Aforementioned in *chapter 1*, E2 Phe63 is important for RING, HECT and RBR E3 ligase binding[129,461]. The loss of p-Parkin labeling within probe **9 F63A** indicated that the mechanism for our probe labeling is based on a canonical E2-E3 interaction. Although we expected the decrease labeling in non-phosphorylated Parkin group, it was not consistent with the findings that the p-Ub alone is able to sterically activate Parkin[82]. We doubted that the weak activation came from the low concentration of p-Ub since Parkin is activated when a molar excess of p-Ub was presented[60]. Indeed, I observed the probe labeling with non-phosphorylated Parkin in a dose dependent manner while increasing p-Ub up to 0.2 and 1mM concentration (**Figure 4.8a**).

Since we observed the probes label Parkin in an activity dependent manner, we wanted to know whether the probe is labeling to catalytic cysteine in Parkin. We applied probe **10** to WT p-Parkin or catalytic mutant C431S p-Parkin accompanied with tryptic MS/MS sequencing of the corsslinked WT p-Parkin to confirm the labeling position. As shown in **Figure 4.8b** and **4.8c**, probe **10** only labeled WT p-Parkin but not the C431S mutant. The MS/MS peptide sequence data and F63A experiment further indicated that our probes label Parkin via a mechanism consistent with the proposed E2-RBR E3 mechanism where the catalytic Cys85 in UBE2L3 and Cys431 in Parkin are in juxtaposed to perform transthiolation.

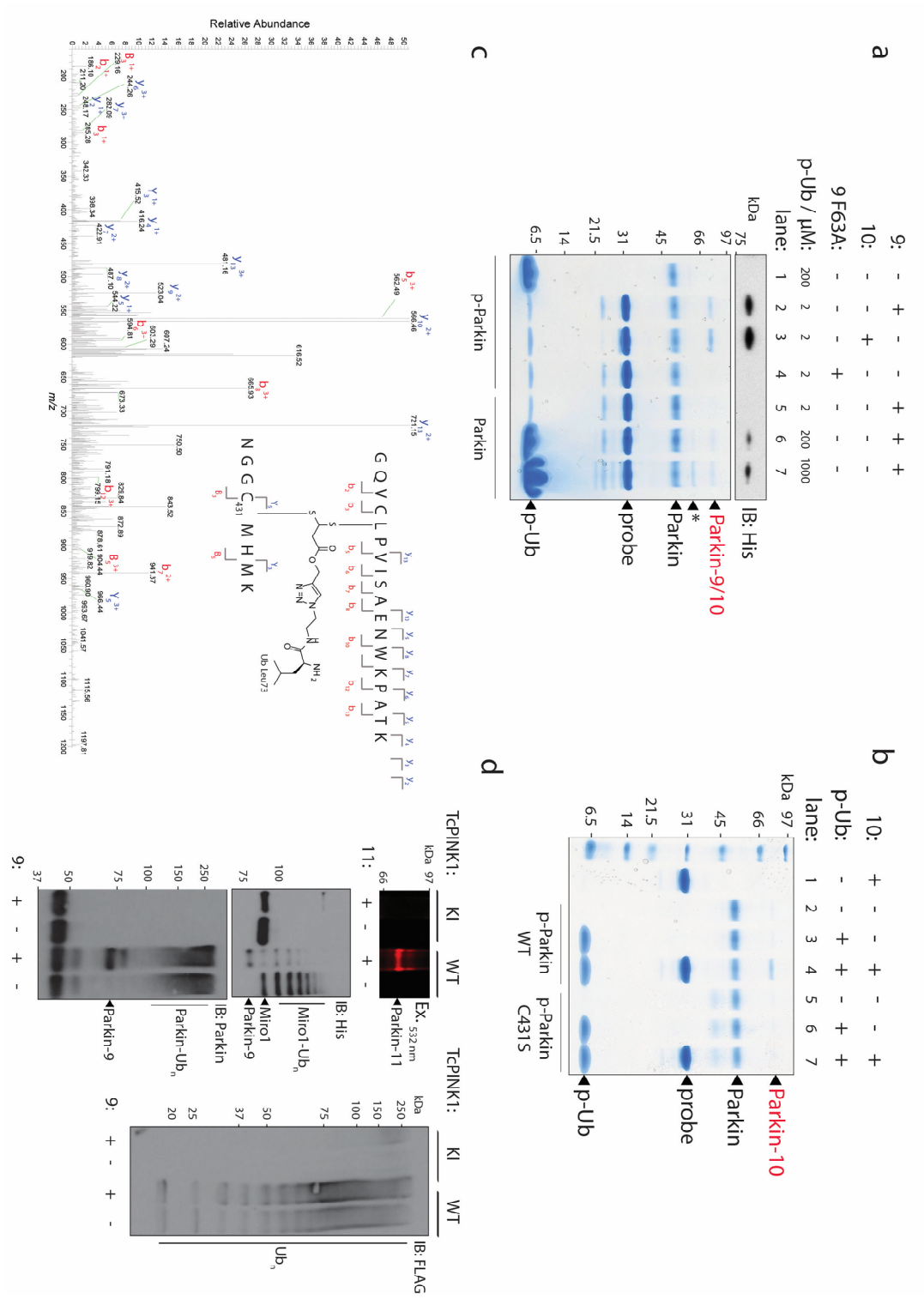


Figure 4.8 E2~Ub-based probes label the RBR E3 ligase Parkin in an activity-dependent manner.

(a) Coomassie stained reducing SDS-PAGE and anti-His immunoblotting reveal that **9** and **10** (10 µM) form a covalent adduct with p-Parkin (2 µM) in the presence of p-Ub (2 µM) (lanes 2 and 3). Probe **9 F63A** (predicted to abolish E3 binding) failed to label Parkin under the same conditions (lane 4). (*Nex page*)

Non-phosphorylated Parkin failed to undergo labeling with probe **9** in the presence of p-Ub (2 μM) (lane 5). Labeling could be affected by the inclusion of molar excess levels of p-Ub (lanes 6 and 7). * corresponds to contaminating band from p-Ub preparation. **(b)** Probe **10** does not label p-Parkin C431S in the presence of p-Ub (lane 8 versus lane 5). All Parkin species and p-Ub were prephosphorylated by treatment with *Ph*PINK1. **(c)** Annotated tryptic MS/MS spectrum for a tryptic 5+ charged precursor ion of the crosslinked peptide derived from labeling of Parkin with **9** (observed m/z = 625.7106; expected m/z = 625.7126) further confirms probe labeling of Parkin Cys431. **(d)** In situ probe labeling of reconstituted substrate ubiquitylation assays. Parkin and FLAG-Ub in the reactions were phosphorylated by pre-incubation with *Tc*PINK1. Parkin labeling with fluorescent probe **11** and probe **9** was strictly consistent with the Parkin activity readouts of His-SUMO-Miro1 substrate ubiquitylation, parkin autoubiquitylation and free polyubiquitin chains formation. In all cases, activity was strictly dependent on *Tc*PINK1 activity. (Performed by Kristin Balk) Consistent results were obtained over three replicate experiments.

4.3.2 *In vitro* Parkin labelling by probes is dependent on PINK1 kinase activity

To further confirm the probe labelling with Parkin is dependent on Parkin activity, we performed post-probe labeling of a Parkin ubiquitylation assay reaction mixture which provided me with these independent readouts for Parkin activity. In **Figure 4.8d**, the Parkin was incubated with either WT *Tribolium castaneum* PINK1 (*Tc*PINK1 WT) or kinase inactivated PINK1 (*Tc*PINK1 KI) and supplied with E1, UBE2L3, FLAG-tagged Ub, ATP and the substrate for Parkin, Miro1^{249,462}. As shown in **Figure 4.8d**, the ubiquitylation on Miro1 was detected and depended on PINK1 kinase. After the kinase reaction, probe **10** or **11** was further reacted with mixtures and observed by either His-tag blot or in-gel fluorescent scan. Again, we observed probe labelling with activated Parkin consistent with the kinase activity of the respective mixtures. Fluorescent-tag containing probe, **11**, provided another rapid and independent readout since the His-tag was both present in probe **10** and Miro1 which we used in the assay which might lead to confusion during data interpretation.

4.4 Applying novel E3-ABPs for RBR E3 activity studies

RBR E3 ligase, Parkin, has been proven to be involved in Parkinson's disease. The huge body of data suggested that Parkin has 100's of substrates and its activity could associated with neurodegeneration[463-465]. However, the transthiolation activity of Parkin itself in cells is less well addressed, herein, I provide a tool that can monitor Parkin activity directly from its transthiolation ability.

4.4.1 Dual phosphorylation events on Parkin and ubiquitin are both essential for Parkin activity

We had demonstrated that our probes are compatible to label Parkin E3 in an activity-dependent manner, we next would like to gain more biological insight about Parkin activation mechanism. Since we were unable to isolate each phosphorylation step of Parkin and Ub under cellular level to dictate their effect toward Parkin activity, we then incubated the recombinant p-Ub and p-Parkin or non-phosphorylated Parkin individually with our probes. Intriguingly, appreciable labeling was observed only with probe **10** and p-Parkin in the presence of p-Ub, suggesting that both phosphorylation events are critical for optimal Parkin transthiolation activity **(Figure 4.9**). Probe **10** still labeled p-Parkin in low level without p-Ub supplement, however, the labeling greatly enhanced when p-Ub was present. Consistent with the previous data that probe **10** labeled non-phosphorylated Parkin when p-Ub was supplied. Another interesting finding from **Figure 4.9** was that wild-type ubiquitin can also increase p-Parkin activity based on probe labeling. One explanation is that the non-phospho Ub still interacts with the p-Ub binding helix (pUBH)[84] but with less affinity. The lack of phospho-Ser65 maybe makes non-phospho Ub unable to position its Ile44 patch correctly to react with extended helix in the RING1 domain, or, the Ub structure is not as dynamic as p-Ub which allow its extended C-terminal tail to interact with the IBR domain. The

other possibility might be that the excess of ubiquitin acts a substrate (as an acceptor ubiquitin for chain formation) for Parkin and further primes Parkin activity although there is no similar data to support this hypothesis. This result suggests that p-Ub is not redundant to Parkin activation after Parkin phosphorylation. Therefore, maintaining p-Ub dependence on sustained Parkin activity could continuously ensure that Parkin activity correlates closely with the levels of the p-Ub primary signal.

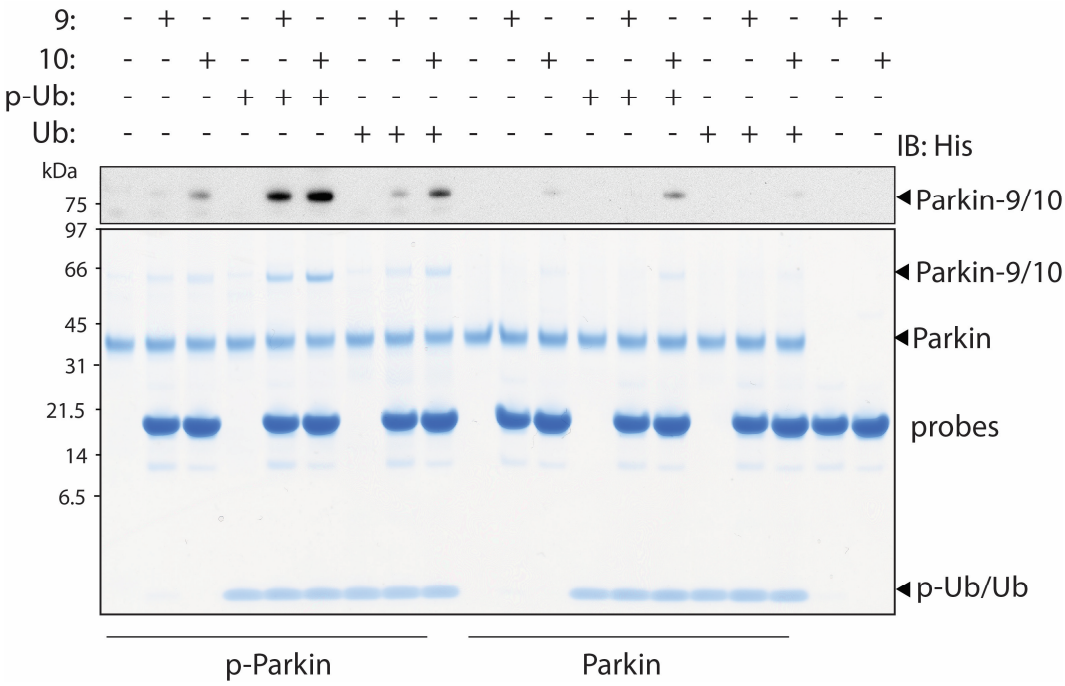


Figure 4.9 Optimized parkin transthiolation activity requires phosphorylation of both p-Parkin and p-Ub.

Prephosphorylated p-Parkin (3 µM) underwent significant labeling with probe **9** only in the presence of p-Ub (6 µM). Substantially reduced labeling was observed in the presence of Ub. Parallel profiling with probe **10** demonstrated that **10** was more sensitive than **9**. Labeling was consistently more efficient with probe **10**, enabling detectable parkin labeling in the presence of individual phosphorylation sites.

4.4.2 Determinants of Parkin activity in the context of PINK1-Parkin signaling in cells

As mentioned, I next wanted to validate ABP functionality within cells by profiling the activity of Parkin using probes **9** and **10** via analysis of HeLa cells (that do not express endogenous Parkin) stably expressing untagged Parkin WT, Parkin S65A and Parkin H302A. Parkin S65A cannot be phosphorylated by PINK1 and the H302A mutation significantly impairs p-Ub binding[84,250,258]. Therefore, in combination with S65A, the effects of Ub and Parkin phosphorylation in the context of PINK1-Parkin signaling could be assessed. The cells were either treated with or without CCCP (carbonyl cyanide m-cholorophenylhydrazone) to impair the mitochondrial membrane potential which activates Parkin. The cells were then lysed under a mild conditions using sonication and further treated with probe **9** or **10** (5μM) for 4 hours at 30°C. In parallel, the E2 only probe (no ubiquitin) containing vinylmethyl ester (VME) (**UBE2L3-VME**) as a warhead was also added in the same assay as a control. According to total Parkin immunoblotting, it revealed a fraction of Parkin WT (~40%) was activated, as judged by the presence of new bands corresponding to Parkin WT labeled with probes **9** or **10** in the presence of CCCP (**Figure 4.10**). Furthermore, according to pSer65-Parkin specific immunoblotting, the phosphorylated pool of WT Parkin was labeled with probe **10** nearly 100%.

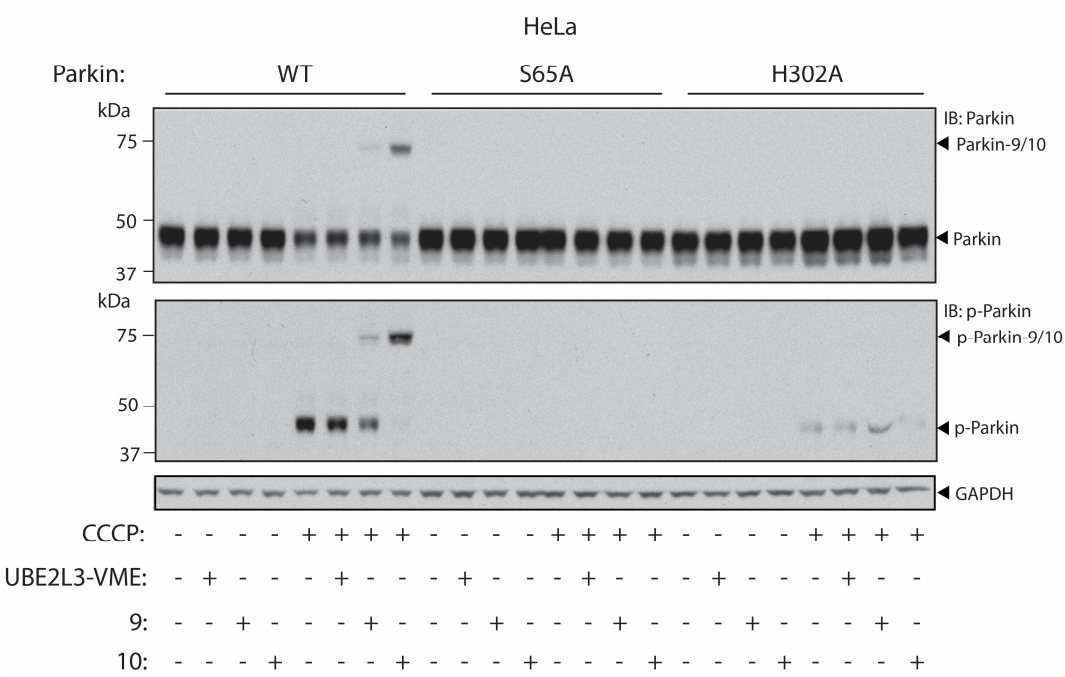


Figure 4.10 Activity-based profiling of cellular Parkin provides insights into the hierachy of Parkin and Ub phosphorylation in the context of PINK1-Parkin signaling.
HeLa cells stably expressing Parkin WT, Parkin S65A and Parkin H302A were untreated or treated with the CCCP (10 µM) for 3 h and labeling of Parkin WT was only observed with probes **9** and **10** in response to CCCP treatment. Probe labeling correlated with Parkin phosphorylation and labeling of the phosphorylated pool was more efficient than that of the total Parkin pool. Non-phosphorylatable Parkin S65A did not undergo labeling with probe **9** or **10**. Parkin H302A with significantly impaired p-Ub binding ability exhibited drastically reduced phosphorylation and no detectable labeling with **9** or **10**.

In contrast, we did not see any labeling of S65A highlighting the importance of Parkin phosphorylation in cells. Furthermore, no labeling was observed with the H302A mutant, and, p-Parkin levels were substantially reduced (**Figure 4.10**). The fact that the total pool of H302A Parkin is unchanged after CCCP treatment and p-Parkin is drastically reduced suggesting that phosphorylation of Ub by endogenous PINK1 contributes to Parkin activation (**Figure 4.10**). Moreover, this shows that p-Ub and p-Parkin are required for optimal activity, this result supports a model whereby the primary mechanism of Parkin activation in response to mitochondrial depolarization in cells involves initial p-Ub binding to Parkin that primes it for

PINK1 phosphorylation, by disrupt the auto-inhibition between Ubl domain and RING1 domain, and optimal activation. To further explore the possibility that Parkin phosphorylation was also associated with RING0 displacement, we carried out profiling of HeLa cells stably expressing Parkin WT, Parkin S65A and a truncated Parkin IBR-RING2 construct (residues 321-465, which missing the Ubl, RING0 and RING1 domain) with our probes (**Figure 4.11**). The IBR-RING2 construct enabled further verification that Ub access to C431 is inhibited by RING0 in cells. Contrary to Parkin WT, Parkin S65A did not undergo CCCP-dependent labelling with Ub-VS which can measure C431 accessibility[466]. Profiling of cells stably expressing the IBR-RING2 construct with Ub-VS revealed significantly enhanced labelling efficiency relative to full-length Parkin, that was CCCP independent. Taken together, this indicates that RING0 does inhibit Ub accessibility to C431 and suggests that Parkin phosphorylation mediates Ub accession to C431, most likely by allosteric displacement of RING0. This latter point requires further validation as we cannot decouple the feed-forward aspect that is activity, and therefore S65 dependent. Furthermore, IBR-RING2 did not undergo labelling with probe **10** and is therefore not proficient to undergo transthiolation with E2~Ub. In conclusion, these results support a model where the primary mechanism of Parkin activation in response to mitochondrial depolarization involves initial p-Ub binding to Parkin that primes it for PINK1 phosphorylation that is essential for its sustained activation[84,250,258].

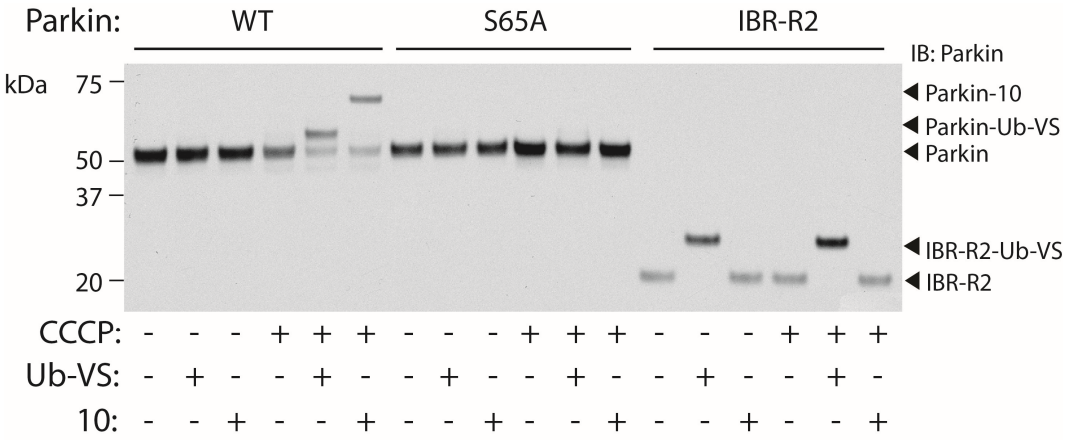


Figure 4.11 Parkin phosphorylation may also mediate RING0 displacement.
Profiling HeLa cells stably expressing Parkin WT, Parkin S65A and a truncated Parkin IBR-RING2 construct (residues 321-465) by our probes was performed. The cells were either treated with CCCP or not and lysed by the same method we used in **Figure 4.9**. No labelling from probe **10** with IBR-RING2 Parkin was observed under CCCP treatment but Ub-VS did. Ub-VS underwent a catalytic Cys431 dependent labelling with Parkin. The result further supported a model whereby parkin phosphorylation allosterically mediates Cys431 accession in addition to displacement of the REP element. These findings also demonstrate that Ub-VS can only be used as a tool to assess C431 accessibility but not native E3 transthiolation activity.

4.4.3 Profiling of endogenous Parkin activity in domaminergic SH-SY5Y cells

As we discussed in the *Chapter 1* that to date the majority of data for Parkin research is based on overexpression systems (the N-terminal tag might even aberrantly activate it[234]). Furthermore, since optimal activity of Pakin needs both phosphorylation on Parkin and ubiquitin, it also raised the question whether the concentration of PINK1 and ubiquitin is sufficient to activate Parkin in the overexpression system. Therefore, we would like to use probe **10** to determine the activation status of endogenous parkin in response to mitochondrial uncoupling in undifferentiated dopaminergic SH-SY5Y neuroblastoma cells. The cells were untreated or treated with CCCP (10 μM) for 3, 6 and 9 h and then lysed as with the above method. Profiling was then carried out with probe **10** and probe

10 F63A as a further control, and validated by immunoblotting against total-Parkin and p-Parkin. Again, I did not observe probe labeling and Parkin phosphorylation in untreated group. However, the labeling by probe **10** was detected after CCCP treatment and peaked at the 3 hour time point (**Figure 4.12**). Notably, the probe could label nearly 95% of p-Parkin in repeated assays and while the F63A control probe failed to label any p-Parkin. Our data is consistent with the finding that Parkin proceeds to undergo autoubiquitylation and destabilization since detected Parkin decreases in a time-dependent way. Moreover, based on the fact that probes nearly label the total pool of p-Parkin, the fraction of probe-labeled total-Parkin can be considered as the pool of fully-activated Parkin in SH-HY5Y cells. Hence 75% of total Parkin is activated in SH-HY5Y cells in response to CCCP treatment. The data also suggests that the parkin activation signal is attenuated through degradation of parkin in the activated state rather than through its deactivation as probe labelling stoichiometry remains approximately constant.

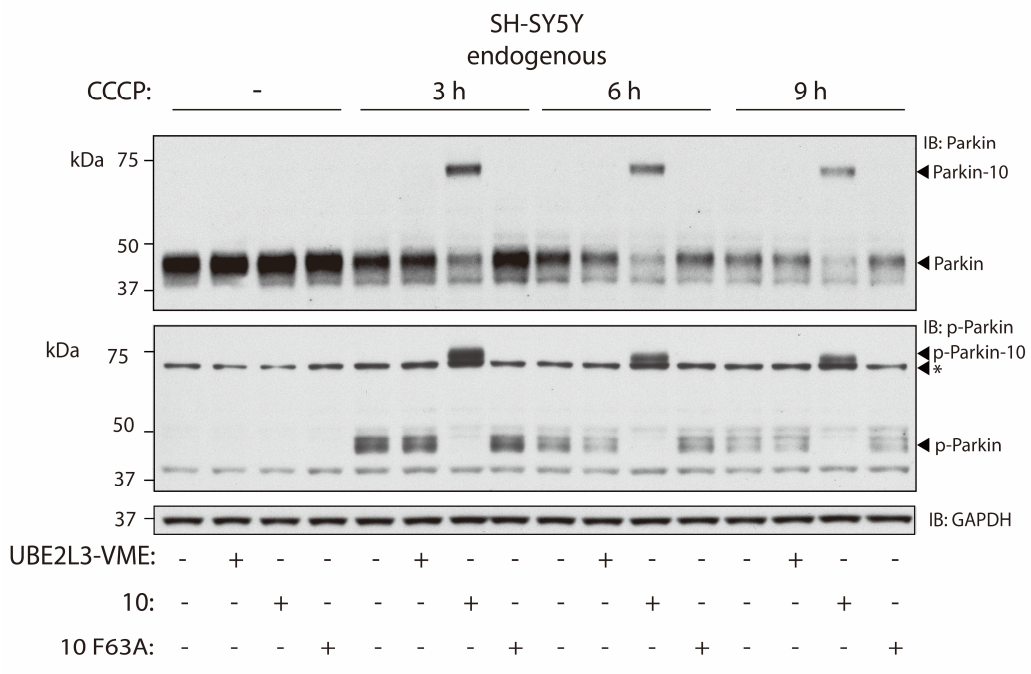


Figure 4.12 Activity-based profiling of endogenous Parkin in SH-SY5Y cells directly reveals Parkin phosphorylation and activation of transthiolation.

In response to CCCP treatment, labeling of the total parkin pool is more efficient than with overexpressed Parkin. Labeling efficiency of the p-Parkin pool approached nearly 100%. Labeling was not observed with the control probes UBE2L3*-VME or **10 F63A**. Over the time course, probe labeling efficiency of parkin remained unchanged while parkin levels declined, suggestive of Parkin clearance in the activated state. Asterisk (*) corresponds to non-specific bands. Consistent results were obtained over three replicate experiments.

The failure to label Parkin with UBE2L3-VME (**Figure 4.12**) further indicated that the contact between Ub moiety on E2 and RBR E3s is essential for ubiquitin transfer. We hypothesized that the interaction between E3 and Ub triggers a Parkin conformational change required for transthiolation. To confirm this was the case and that E2~Ub engineering did not inadvertently alter the warhead properties, I made an E2~Ub probe with an IIe44Ala mutation in the Ub. Consistent with our hypothesis, probe **10 (Ub-I44A)** failed to label recombinant p-Parkin in the presence of p-Ub nor endogenous Parkin in CCCP-treated SH-HY5Y cells (**Figure 4.13a** and **4.13b**).

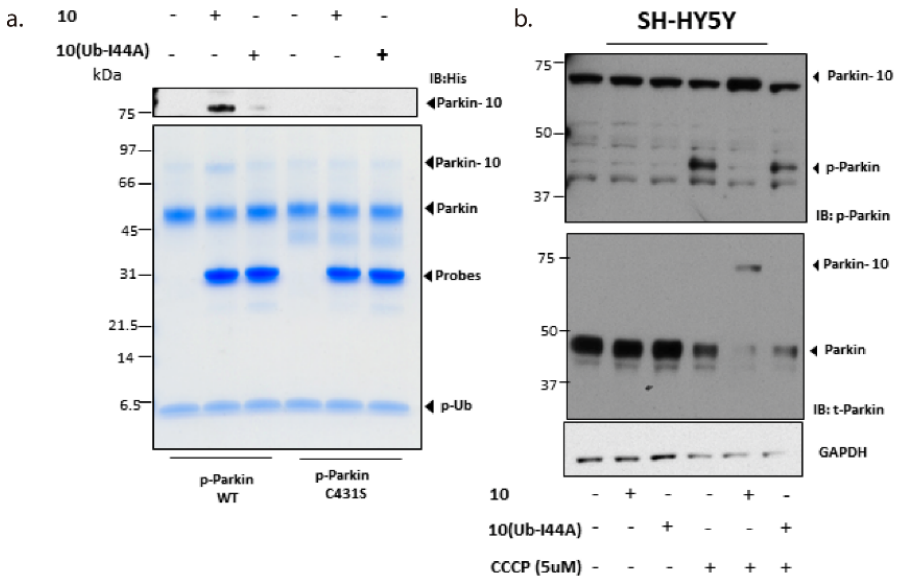


Figure 4.13 The ubiquitin Ile44A mutation on probes disrupts the binding ability of E3 ABPs toward Parkin.

(a) The probe **10** or **10** (Ub-I44A) was treated with either WT or C431S p-parkin. The WT probe **10** as a positive control labeled p-parkin while **10** (**Ub-I44A**) did not. (**b**) Ubiquitin Ile44A mutation on E3 ABP also abolish its ability to label endogenous Parkin in SH-HY5Y cell lysate.

4.4.4 Profiling of endogenous Parkin activity in primary fibroblasts derived from PD patients

To explore the translational potential of our probes we used probe **10** to profile primary fibroblasts derived from a skin biopsy from PD patients carrying a homozygous Q456X mutation in the *PARK6* gene and an exon 5 deletion/C441R (Δex5/C441R) compound heterozygous mutation in the *PARK2* gene. The Q456X mutation disrupts PINK1 activity and the PARK2 compound heterozygous mutation would be predicted to result in aberrant Parkin expression and/or stability[449]. As expected, I did observe significant probe labeling in healthy control cells in response to CCCP treatment whereas probe **10** failed to label Parkin in PINK1 Q456X mutation cells (**Figure 4.14**). Parkin was undetectable in the Δex5/C441R sample and as expected, no probe labelling signals were observed.

These data provide the first quantitative assessment of Parkin activity in patient-derived cells and indicate that our probes can be deployed to determine the general functionality of the PINK1-Parkin signaling pathway in PD patients. We are excited about the potential of our probes for identifying individuals who might be predisposed to PD, and stratifying patients suffering from idiopathic PD for treatment with future therapeutics.

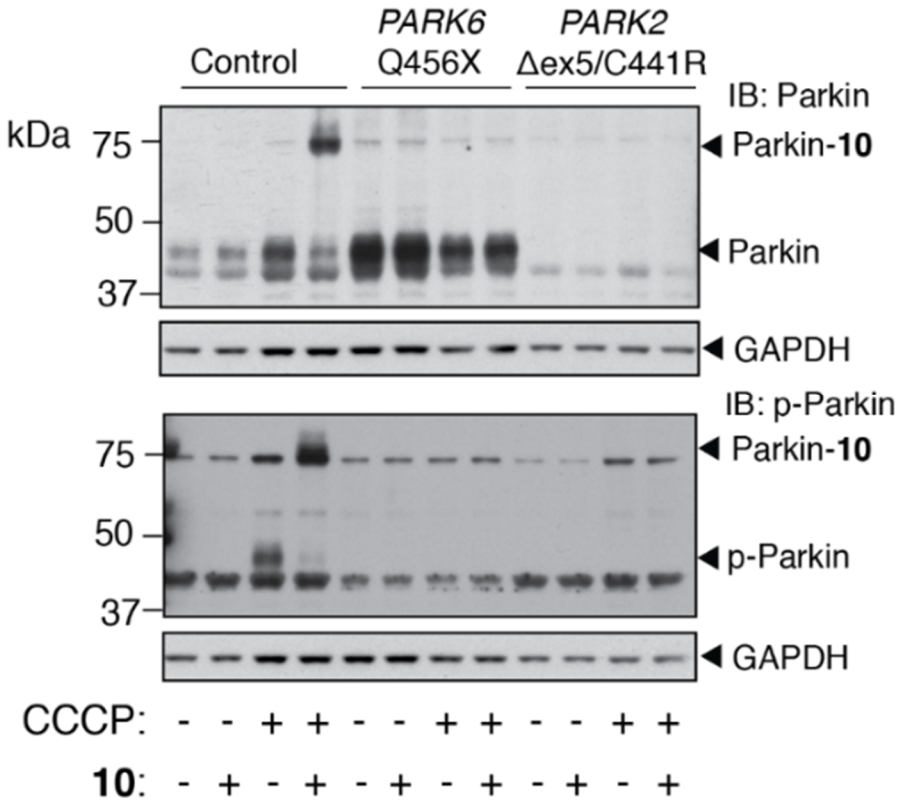


Figure 4.14 Activity-based profiling of PD patient-derived fibroblasts reveals biomarker potential.

Low passage primary fibroblasts derived from skin biopsies from a healthy age-matched family member, a PD patient harbouring a Q456X homozygous mutation in the *PARK6* gene, and a PD patient harboring a Δex5/C441R compound heterozygous mutation in the *PARK2* gene were profiled with probe **10**. Control cells undergo robust Parkin phosphorylation and activation. In contrast, Parkin in either *PARK6* or *PARK2* mutant PD patient cells cannot be activated in response to CCCP treatment. With expectation, probe **10** only labeled Parkin in the positive control group. Both mutation on *PARK6* or *PARK2* abolishes Parkin activity.

4.4.5 Parkin phosphorylation displaces the REP element

The repressor element of Parkin (REP) that blocks the RING1 domain for recruiting E2~Ub is one of the Parkin autoinhibitory mechanisms[236,245]. The side chain of Trp403, located in this element of Parkin, inserts into a pocket in RING1 and masks the predicted E2 docking site on Parkin[245]. A W403A mutation in the REP region disrupt this inhibition and results in increased Parkin autoubiquitylation activity and accelerated Parkin recruitment to damaged mitochondria upon activation of PINK1 kinase activity[245]. However, the mechanism behind the relief of REP element by PINK1 phosphorylation is not clear. We then explored whether relief of the autoinhibitory REP element was associated with phosphorylation of Parkin at Ser65 by performing a probe labelling assay with the non-phosphorylated Parkin W403A mutant in parallel with Parkin and p-Parkin, in the presence and absence of p-Ub (**Figure 4.15a**). Probe **10** only labelled non-phosphorylated Parkin W403A in the presence of p-Ub but not Parkin W403A alone. This finding was consistent with the data that Parkin W403A activity was still dependent on PINK1 and CCCP treatment in cells but retained activity in vitro when supplied with p-Ub. Apparently, without activating PINK1 to phosphorylate ubiquitin in cell, Parkin is unable to maintain in full-activated form. However, according to the obtained data, we concluded that Parkin phosphorylation was only partially recapitulated by the W403A mutation upon consideration of the stoichiometric nature of the W403A mutation and the substoichiometry of Parkin phosphorylation (~60%). In addition, it has been shown that the enhanced affinity of Parkin toward E2 enzyme via phosphorylation on Ubl domain was only emulated when Parkin W403A mutation and Ubl domain deletion coexist[250]. For further confirm the probe was being labelled with the correct position, we carried out probe labelling with Parkin W403A/C431S double mutant (**Figure 4.15b**). These data suggest that parkin

phosphorylation is directly associated with relief of the autoinhibitory constraints imposed by the REP element through an allosteric mechanism, thereby facilitating E2 binding, and REP displacement is independent of p-Ub binding.

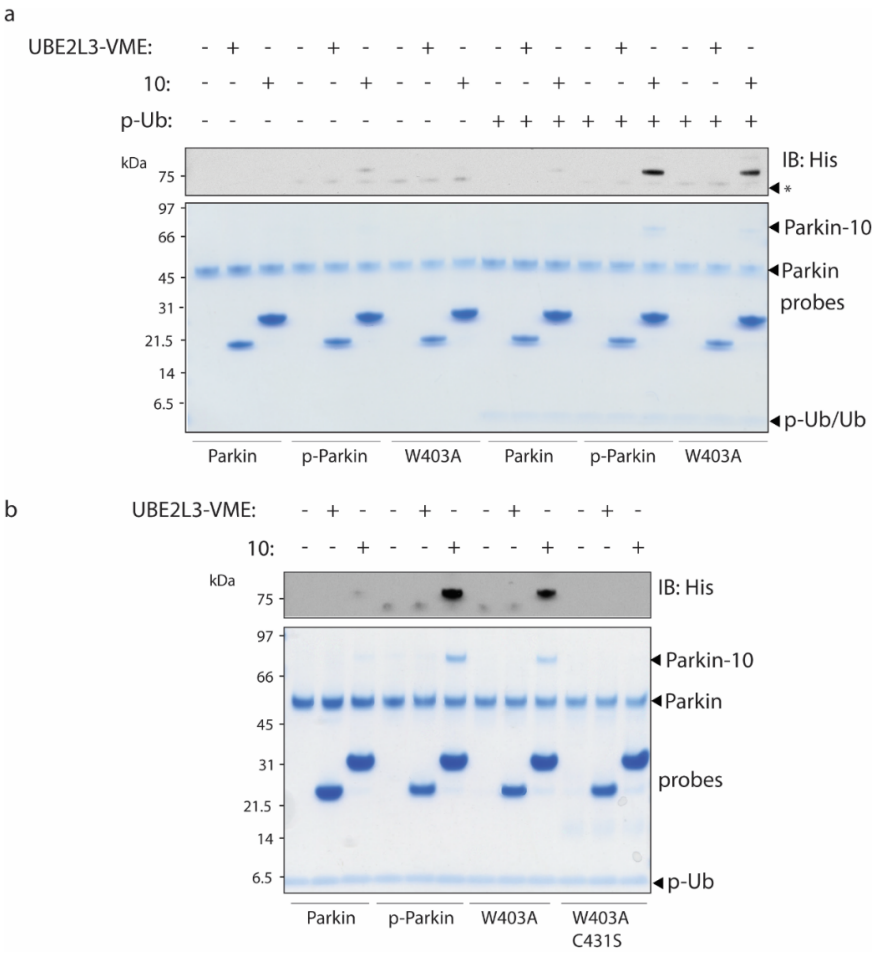


Figure 4.15 REP domain blocks the E2~Ub contact with Parkin and attenuates Parkin activity.

(a) ABP **10** was used to determine whether a distinct phosphorylation event contributed to displacement of the autoinhibitory REP element. Non-phosphorylated Parkin W403A, in the presence of p-Ub, underwent robust labeling (lane 18). The first-generation E2-based probe UBE2L3*-VME, which contains a chemically analogous electrophile to **9** but lacks the Ub component, failed to label Parkin under any of the conditions tested. Consistent results were obtained over two replicate experiments. **(b)** Profiling of Parkin mutants was carried in the presence of p-Ub with control probe UBE2L3*-VME and ABP **10**. Importantly, the Parkin W403A C431S double mutant, in the presence of p-Ub, does not undergo labeling. This confirms that labeling remains activity dependent in the context of the W403A mutation. Asterisk (*) indicates a non-specific band.

4.4.6 Rapid profiling of Parkin disease mutants

In order to translate our probes into clinical application, we built fluorescent probe **11** by conjugation of the fluorophore TAMRA with E3 ABP (**Scheme 4.2**, **Figure 4.6d**) and profiled a panel of recombinant Parkin disease mutants (K27N, R33Q, R42P, A46P, K161N, K211N, R275W, G328E, T415N, G430D and C431F, **Figure 4.16a**) in parallel, with Parkin WT and the phosphorylation defective S65A mutant. This was achieved by incubating probe **11** with the panel of mutants that had been phosphorylated with WT *Pediculus humanus corporis* PINK1 (*Ph*PINK1) in the presence of Ub (**Figure 4.16 b-c**). Among these Parkin mutants, 8 of 12 were inactivate. The K27N mutant was reduced 0.7-fold and the R33Q mutant was found to activate Parkin 2.2-fold. As these disease mutants are located through out the Parkin polypeptide, this suggests that transthiolation activity is dependent on all of the domains within Parkin and leads towards a unifying pathogenic basis. However, we cannot formally exclude the possibility that the mutations may also affect the latter step of Ub transfer to substrate[249].

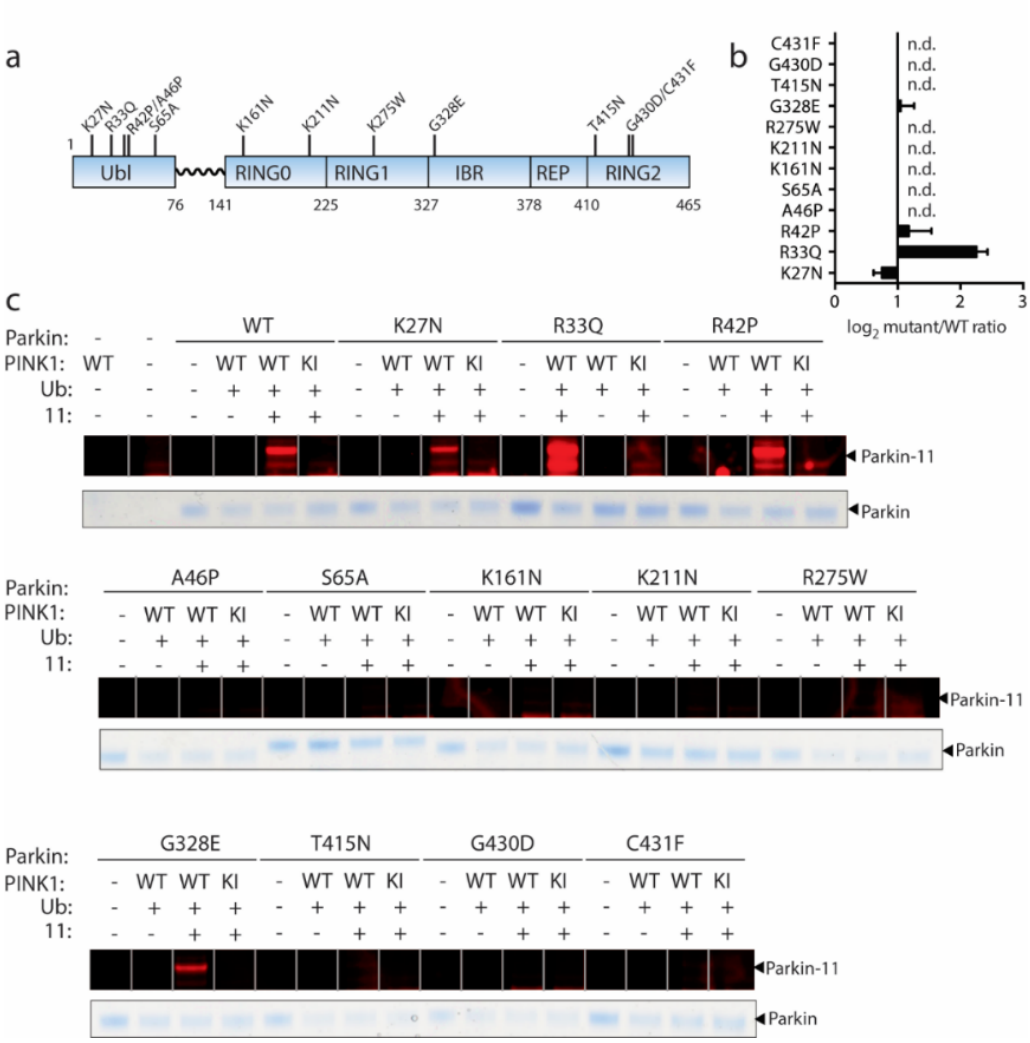


Figure 4.16 Quantitative and direct activity-based Protein Profiling of transthiolation activity of Parkin patient mutations reveals distinct activity signature.

(a) Amino acid boundaries of the multidomain architecture of parkin. **(b)** Recombinant Parkin mutants were incubated with *Ph*PINK1 in presence of Ub and ATP. Mutations resided throughout the multi-domain architecture of Parkin. Incubations were then directly profiled for Parkin transthiolation activity with ABP **11**. Mutations mildly perturbed, activated or abolished transthiolation activity as determined by fluorescence of labeled Parkin. **(c)** Fold-change of mutant Parkin transthiolation activity relative to WT represented on a Log_2 scale (n.d. means no detectable labeling. Data represent mean values and error bars correspond to ± s.d, n = 3). (* the doublet bands in K27N and R33Q are the impurity band from the ABP 11.)

4.5 Further validation of E3 ABP specificity and compatibilty

Since we established the activity-based probe for RBR E3 ligases by profiling Parkin, to build a platform for E3s activity-based proteomic is an important next step. In addition, the details of activity regulation for most of E3s are still poorly understood, our probes are able to provide more biological insight into this field. However, before we carried out more sophisticated assays for E3 activity profiling, we need to confirm that the E3s are the major target for our probes and also to demonstrate that more E3s can be labelled by our probes.

4.5.1 DUBs profiling with E3 ABPs

DUBs play an important role in ubiquitin system by editing ubiquitin chain level in cells. They also cooperate with E3s to regulate various cellular functions. Intriguingly, DUBs can also have a catalytic cysteine as the catalytic reside for their activity. As a result, we next tested whether probe **9** and **10** cross reacted with DUBs and whether they were susceptible to DUB mediated hydrolytic degradation, either of which would restrict the utility of the probes activity-based proteomics. We found that although the potential control DUB ABP propargylated-Ub (Ub-Alk)[333] efficiently labeled DUBs from three different subfamilies (USP, OTU and UCH), probes **9** and **10** exhibited no labeling, nor were they degraded by DUB isopeptidase activity (**Figure 4.17a**). These data further indicated the selectivity of our probes toward E3s.

4.5.2 Profiling HECT E3 ligase with E3 ABPs

In order to expand the utility of our probes, we next tested the compatibility of probes against another catalytic cysteine-containing E3 ligase family, HECT E3s. Again, labelling of probe **9** and **10** with activated forms of the HECT E3 NEDD4[128]

and the bacterial HECT-like E3 NleL[443] was observed in all cases and the efficiency was dependent on the catalytic cysteine nucleophiles (**Figure 4.17b-c**). It demonstrated that our probes are able to target a different E3 family and indicated the potential of our probes for systematic profiling.

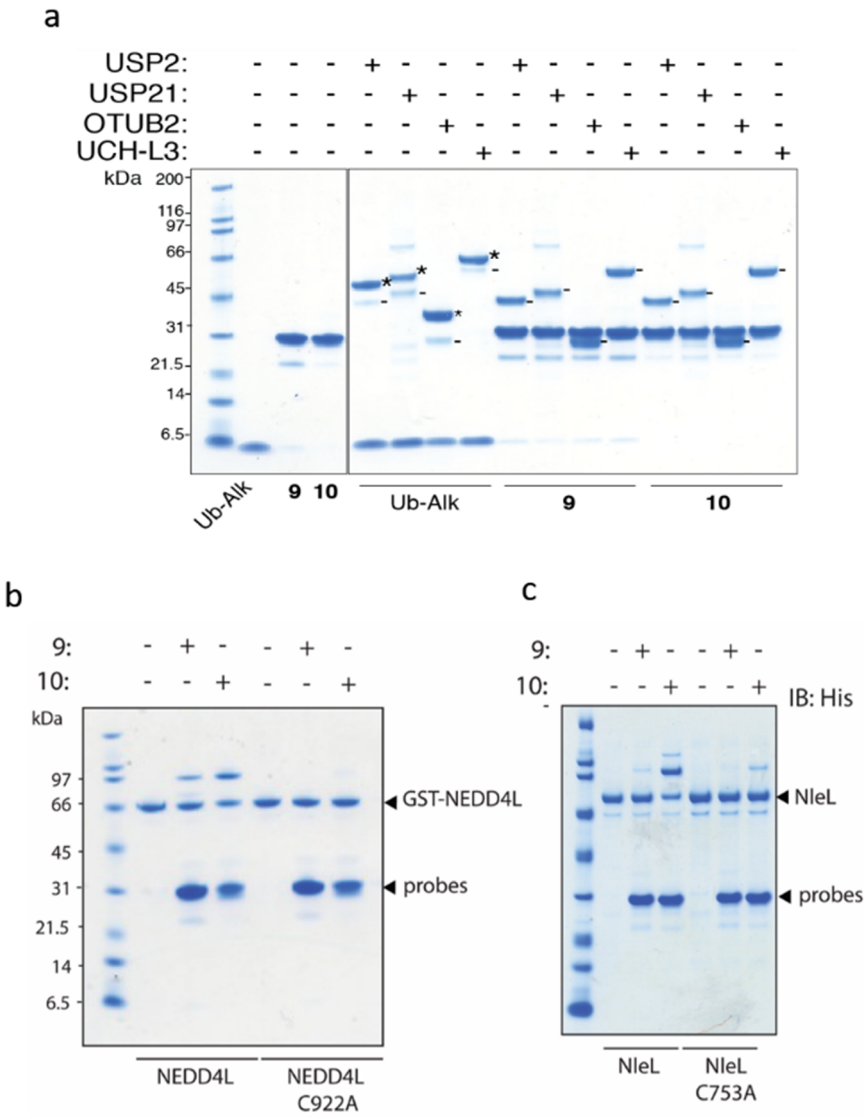


Figure 4.17 E3 ABPs have no affinity toward DUBs but retain activity for HECT E3s.
(**a**) DUB ABP, Ub-Alk, showed great labelling efficiency with USP2, USP21, OTUB2 and UCH-L3. In contrast, the labelling of these DUBs by E3 ABPs **9** and **10** was missing further indicated the specificity of E3 ABPs is strict with E3 only. **(b-c)** Probe **9** and **10** labelled HECT and HECT-like E3, NEDD4L and NleL, in an activity dependent manner.

4.5.3 In situ profiling of HOIP in complex proteomes of mouse embryonic fibroblasts (MEFs) cells

HOIL-1L (haem-oxidized iron-regulatory protein 2 ubiquitin ligase-1) and HOIP (HOIL- 1L-interacting protein) are important RBR E3 ligases that form a complex with SHARPIN called LUBAC (The linear ubiquitin chain assembly complex)[242]. Although the LUBAC complex is involved in regulating several cell signaling pathways[467 447,468], and binding of either HOIL-1L or SHARPIN to HOIP is required for HOIP E3 activity, the regulation of HOIL-1 or HOIP RBR E3 ligase activity in response to dynamic signaling is still not well understood. Recently, the co-crystal structure of UBCH5B with HOIP was solved[129]. It showed that the Arg 74 on ubiquitin might be important for ubiquitin transfer as it forms a salt bridge with E976 on HOIP. This interaction was also observed in a truncated HOIP structure published by Stieglitz *et al.*[239]. By building probes on alternative E2s and including Ub Arg74, we built a new type of ABP, based on Ub 1-74 and UBCH5B (Probe **12**). We also built a control version of this probe containing an F62A mutation in the E2 component that would be predicted to abolish RBR E3 recognition. In parallel, we profiled mouse embryonic fibroblasts (MEFs) and catalytically inactive HOIP 879S knock-in MEFs (provided by Dr. Philip Cohen's lab) with probe **12** (**Figure 4.18**). Labeling of HOIP was only observed in WT MEFs and F62A mutant probe did not undergo HOIP labeling in either cell line. As the probe signal was dependent on the catalytic cysteine and the F62 residue in E2, this suggests that our probe labeling signal provides a physiological measure of endogenous HOIP transthiolation activity.

In a previous study, the linear ubiquitin chain made by the LUBAC complex was shown to be dependent on IL-1α stimulation[447]. However, the same group

also claimed that HOIP *activity* is not correlated, suggesting that perhaps an opposing activity to HOIP (e.g. the linear Ub chain-specific DUB, OTULIN) is deactivated in response to IL-1α treatment. This latter argument is consistent with our results. Our data suggested that HOIP might be constitutively active in mouse embryonic fibroblasts. However, probe labeling indicated that activation of the endogenous HOIP pool was substoichiometric. Therefore our probe should be a valuable tool for determining whether HOIP can be further activated, or deactivated, under certain conditions. A hypothesis we could test is whether HOIP activity is not binary but rather behaves like a rheostat where different levels of activity illicit distinct cellular processes.

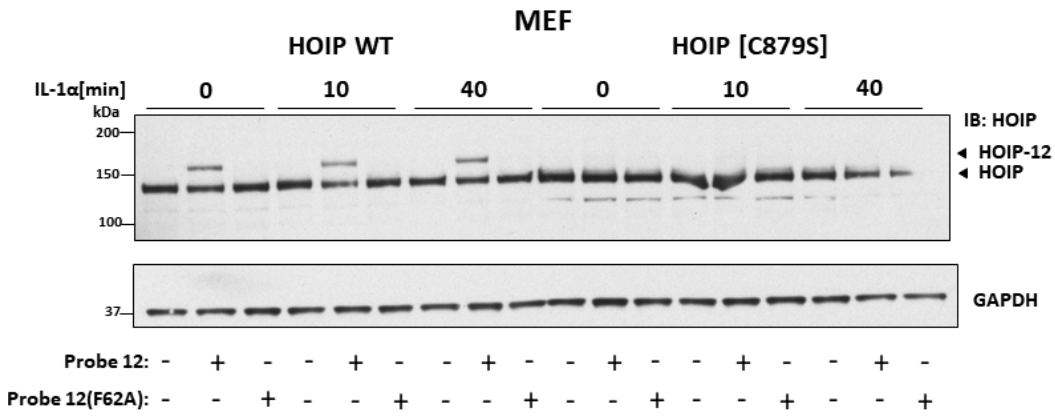


Figure 4.18. HOIP E3 is constitutively active in MEFs.
HOIL-1Land HOIP are important RBR E3 ligases which form a complex with SHARPIN called LUBAC. Wild-type (WT) HOIP and catalytic dead mutant HOIP [C879S] knock-in MEFs were treated with 5 ng/mL IL-1α for certain time points. Cellular extracts were incubated with either probe **12** or control probe **12 (F62A)**. HOIP labeling was only observed in WT MEFs with probe **12** and was dependent on Cys879. However, labeling of HOIP was independent of IL-1α stimulation which suggested that HOIP was constitutive active.

4.6 Conclusion & Discussion

A useful tool to study Ub transfer from E2s to E3s has remained elusive in the field. Herein, we present methodology to synthesize novel ABPs that profile HECT/RBR E3 transthiolation activity. The utility of these ABPs was demonstrated by profiling the transthiolation activity of the PD-associated RBR E3 ligase Parkin both *in vitro* and in cellular extracts. Our study revealed that Parkin phosphorylation displaces the inhibitory REP element and suggests that p-Ub binding to Parkin occurs upstream of Parkin phosphorylation in cells. Furthermore, our probes enabled direct and quantitative measurement of endogenous Parkin activation in neuronal cell lines as well as in PD-derived patient fibroblasts. We also demonstrated the compatibility of our ABPs with another RBR E3 HOIP, and HECT E3s. In conclusion, our findings show the potential of this new technology as a robust readout of PINK1-Parkin signaling, as a novel biomarker of PD patient-derived samples, and a powerful tool for profiling the transthiolation activity of E3s in general.

4.6.1 The Parkin mechanism

The mechanism for regulating Parkin activity has been heavily studied in recent years. I have already discussed the model based on the current discoveries in *chapter 1*. Our data supports the model that initial binding of p-Ub to Parkin renders Parkin a more effective PINK1 substrate in cells, and both phosphorylation events are then absolutely required for sustained Parkin activity, as judged by transthiolation. However, Parkin is located in cytoplasm when cells are in the resting stage, how does Parkin get recruited to mitochondria is still in debate[80,469]. Some data support a model where Parkin will be recruited to mitochondria by p-Ub but how the very first p-Ub or p-Ub chains are made that recruit Parkin is still unknown. Although people in the field are now more aware of the fact that the

large body of data about Parkin research are based on overexpression systems, there is still little attention on the reagents we used. CCCP or O/A (oligomycin A/Antimycin A) treatment is a facile and fast method to depolarize mitochondria. The consequences provide us a good model to monitor mitochondria damage and the PINK1/Parkin signaling pathway. However, since there was a subtle phenotype from either PINK1 or Parkin KO mice[248,470], I doubt that neurons will encounter the same level of damage imparted by CCCP treatment. Based on this hypothesis, does the PINK1/Parkin pathway still play a major role in neuron cell protein quality control? The recent study by McWilliams *et al.* has demonstrated that the mice basal mitophagy is regulated in a PINK1 independent manner[471].

Although our probes might not be able to answer all of the above questions, we still can follow up one more important issue in Parkin research. What is the role of Parkin in the cytoplasm? Since the major pool of Parkin is in the cytoplasm, does Parkin retain its ligase activity in this location? Activation of Parkin by Ubl domain phosphorylation and p-Ub interaction is a well-accepted model to date. If Parkin can become activated in the cytoplasm, what is the kinase responsible for Parkin phosphorylation or is there an alternative activation mechanism? We can first use our probes to detect Parkin activity differences in mitochondria and cytoplasm since our probes only labels fully activated Parkin. It can also allow us to screen various mild stimuli to activate Parkin in cells. I believe our probes can bring more insight into Parkin research.

4.6.2 Probes can do more than measure Parkin activity

There is a longstanding interest in the conformational changes and mechanistic aspects of Ub transfer throughout the Ub conjugation cascade. Although several different groups solved the Parkin structures recently, none of them are in complex

with E2s. Furthermore, albeit we know the phosphorylation events that activate Parkin *in vitro* and *in vivo*, there is still no data to explain how conformational changes brings the RING2 domain on Parkin in close proximity of the catalytic Cys in E2s. Overall, it is the transient nature of the E2~Ub~E3 complex that keeps us in dark. However, a working theory of our probes is that they can stabilize the transient E2~Ub~E3 complex which could be used to facilitate its structural characterization. Moreover, our probes can also serve as a E2~Ub intermediate which also has been proven to be unstable or difficult to make. As a result, we can measure the affinity differences between E2 or E2~Ub toward different E3s. Moreover, the methodology for making the probes allows us to replace the E2 component without difficulty, we can easily build a library to screen different E2~Ub affinities towards a wide variety of E3s.

4.6.3 E3 ABPs in activity-based protein profiling (ABPP)

It is certainly possible to profile more RBR or HECT E3 ligase activities by our probes. Dr. Julien Licchesi's group have demonstrated that our probes can also label multiple HECT E3 ligases *in vitro* and in cell lysate, for example NEDD4, UBE3C, and HECTD1[248]. However, to validate all the RBR and HECT E3s one by one is not realistic and time consuming. As a result, we applied our probes in the activity-based protein profiling (ABPP) system within mass spectrometry[302]. We modified our probes by adding a biotin-tag on to them which allow us to immobilize the probe-E3 complexes with streptavidin beads. The detail will be discussed in next chapter (*Chapter 5*).

4.6.4 The future optimization of E3 ABP construction

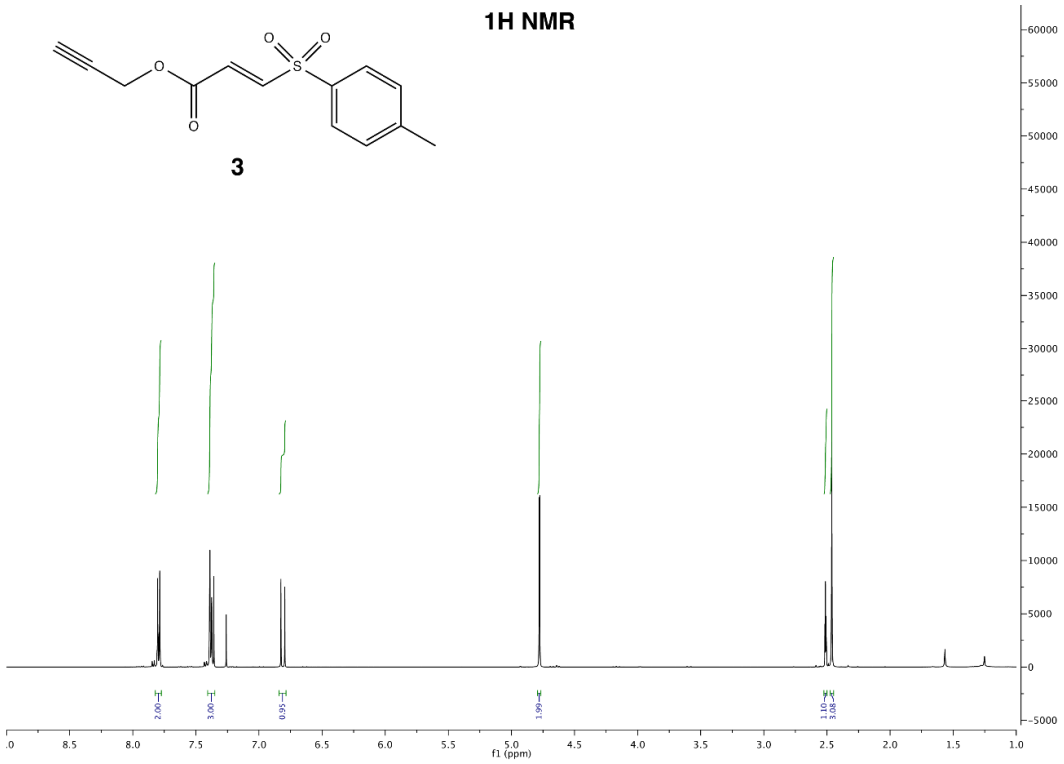
In this chapter, I already demonstrated that the ability and potential of our novel E3 ABPs. They can detect the endogenous Parkin in an activity-dependent manner

and distinguish the activity of WT Parkin among Parkin mutants. However, the probes are not in perfection yet. First, the linker which connects ubiquitin and E2 does not mimic the native E2~ubiquitin thioester linkage. Although the distance between them is similar to the native conjugate according to modeling software (**Figure 4.4**), we still cannot ignore the impact from the differences between triazole linker and wild type peptide bond. In addition, the E2~Ub adopt an 'open' conformation when they encounter HECT/RBR E3s but a 'closed' conformation with RING E3s. Whether the linkage between Ub and E2s in our probe permits this conformational change is still unknown.

The truncation on the ubiquitin might be tolerated by some E3s but certainly may interfere with the interaction with others (discussed already in section *4.5.3*[129]). In fact, the probes containing Ub_{1-73} did not have good selectivity and labeling efficiency toward HOIP E3 ligase in the early test. As a result, how to optimize the tri-functional TDAE to preserve the native ubiquitin C-terminus is a question we might need to address in the future.

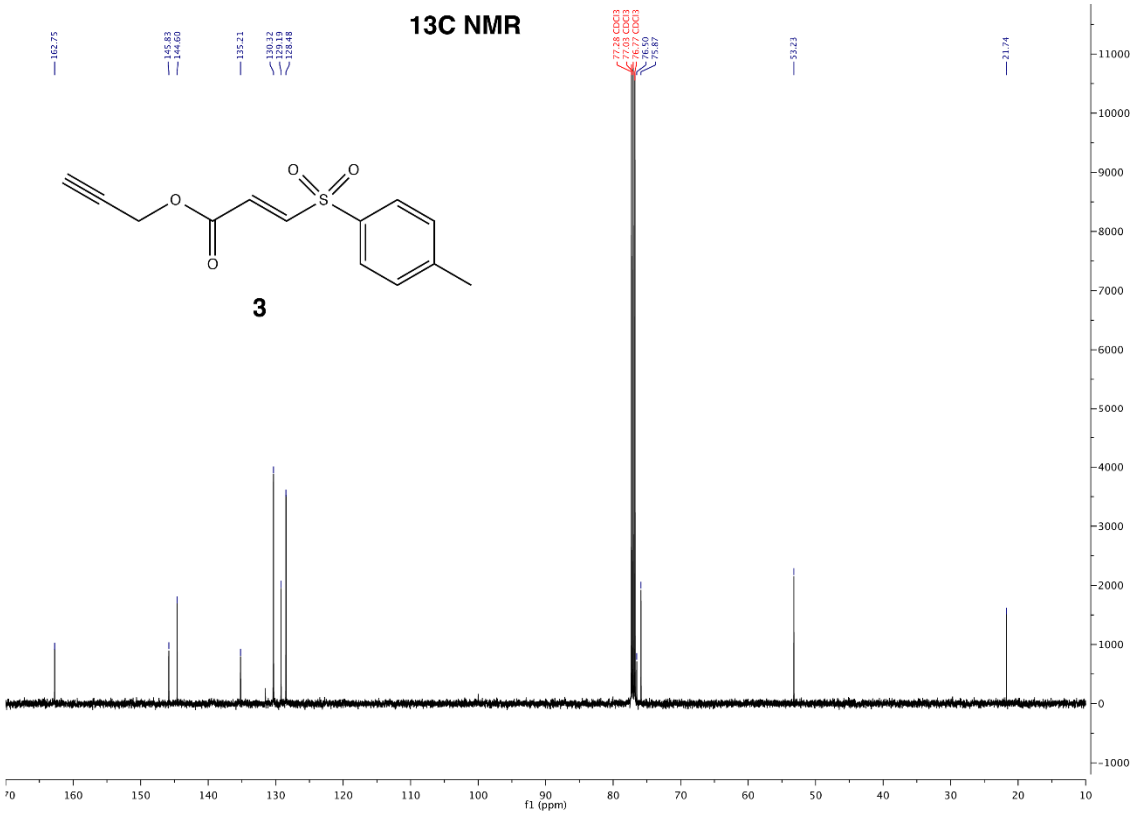
The other drawback of the probes I demonstrated in this section is that they work only with RBR/HECT type E3 ligases which contain catalytic cysteine on them. The lack of reactivity toward RING E3s limits the coverage of possible E3 activity profiling since there are nearly 600 RING E3s in the human genome. There is no catalytic nucleophile on RING E3s, we cannot apply the same chemistry to label RING E3s in an activity dependent manner. Introducing a photocrosslinker into a stabilized E2~Ub conjugate may be one strategy for making RING E3 ABPs.

4.7 Extend Figures

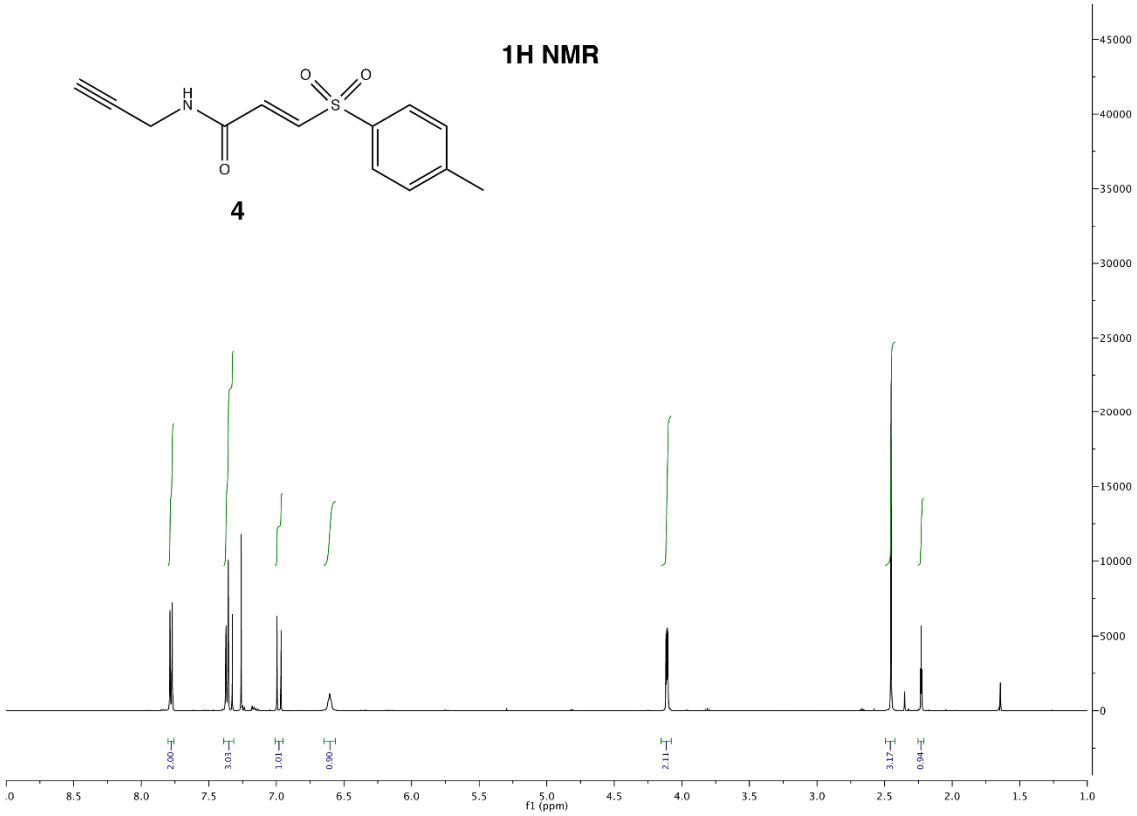


Extended Figure 4.1 1H NMR spectrum for TDAE 3.

1H NMR (500 Mhz, $CDCl_3$) δ 7.79 (2H, d, 8.4 Hz), 7.35-7.40 (3H, m), 6.81 (1H, d, 15.2 Hz), 4.78 (2H, d, 2.5 Hz), 2.51 (1H, t, 2.5 Hz), 2.46 (3H, s).

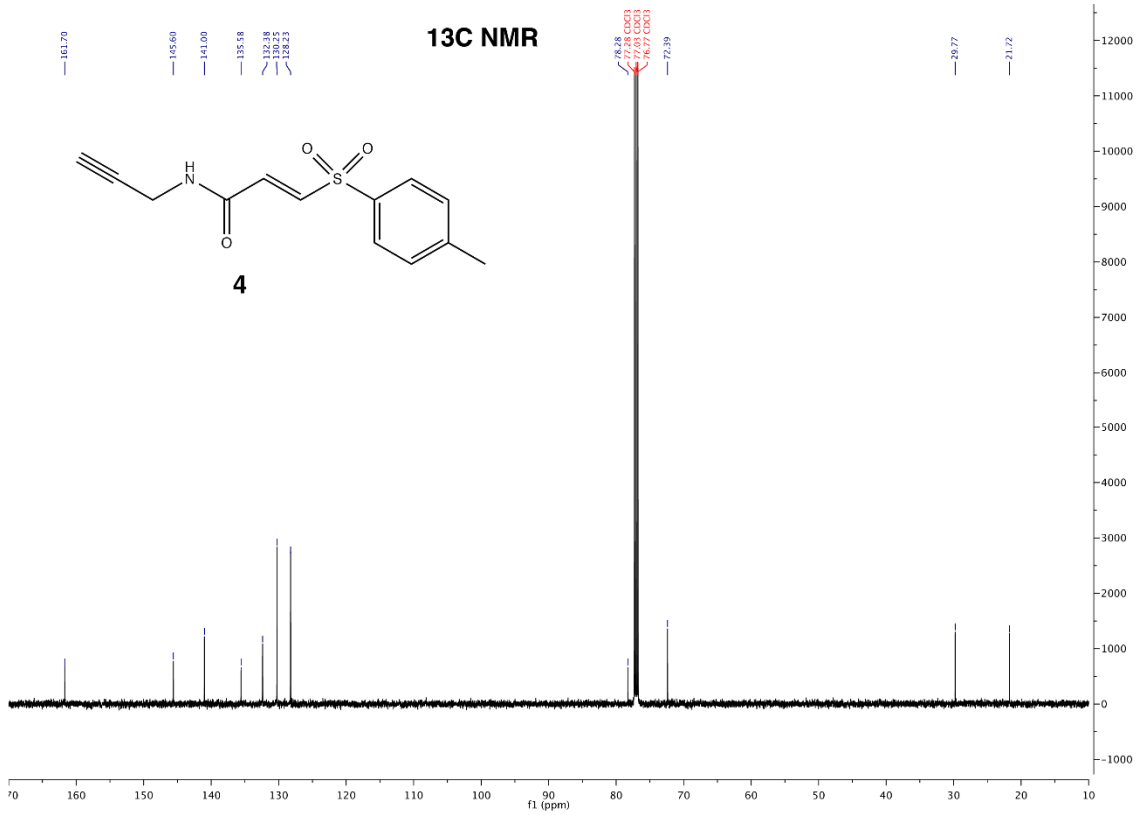


Extended Figure 4.2 13C NMR spectrum for TDAE 3. (Provided by Dr. Mathew Stanley)
13C NMR (126 MHz, $CDCl_3$) δ 162.75, 145.83, 144.60, 135.21, 130.32, 129.19, 128.48, 76.50, 75.87, 53.23, 21.74.

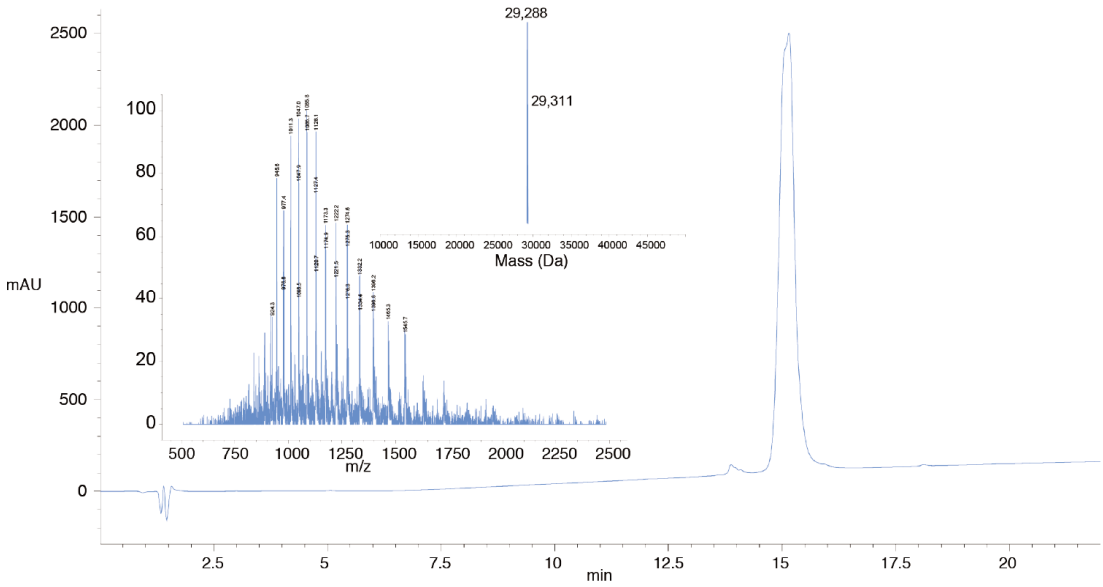


Extended Figure 4.3 1H NMR spectrum for TDAE 4.

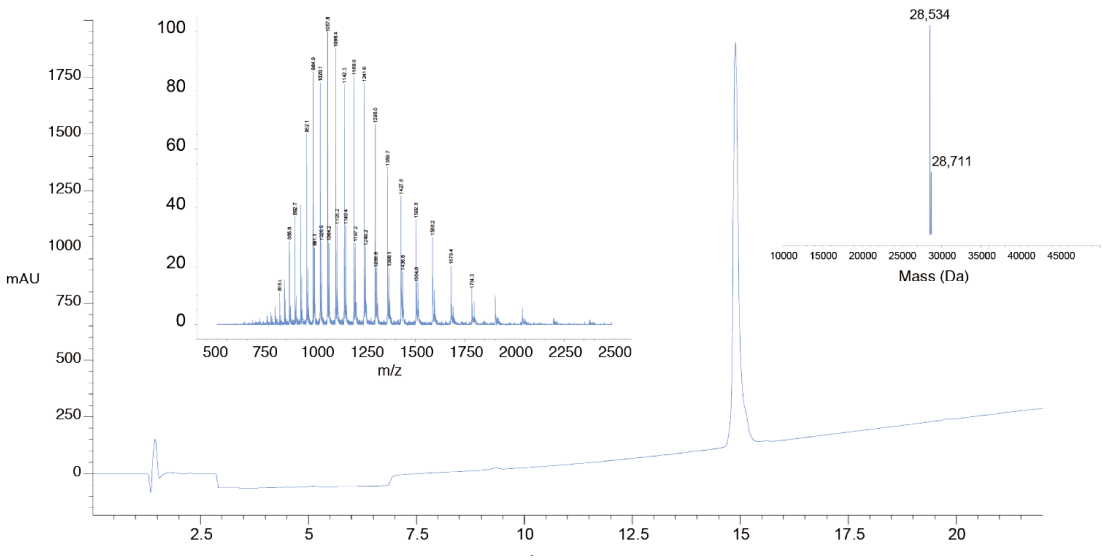
1H NMR (500 Mhz, $CDCl_3$) δ 7.78 (2H, d, 8.3 Hz), 7.37 (2H, d, 8.5 Hz), 7.34 (1H, d, 14.9 Hz), 6.98 (1H, d, 14.8 Hz), 6.60 (1H, t, 5.1 Hz), 4.10-4.12 (2H, dd, 5.3, 2.6 Hz), 2.45 (3H, s), 2.23 (1H, t, 2.6 Hz).



Extended Figure 4.4 13C NMR spectrum of TDAE 4. (Provided by Dr. Mathew Stanley)
13C NMR (126 MHz, CDCl3) δ 161.70, 145.60, 141.00, 135.58, 132.38, 130.25, 128.23, 78.28, 72.39, 29.77, 21.72.



Extended Figure 4.5 LC-MS characterization of ABP 11.
Observed mass = 29,288 Da; theoretical mass = 29,290.3 (+Ac) Da. Mass of 29,311 Da is a putative Na+ adduct.



Extended Figure 4.6 LC-MS characterization of ABP 9 F63A.

Observed mass = 28,534 Da; theoretical mass = 28,537.5 Da. Additional species with an observed mass of 28,711 Da (+177) corresponds to gluconoylated form of **9 F63A**.

Chapter 5. Activity-based proteomics and discovery of a new class of E3 ligase

Activity-based protein profiling (ABPP) has been proven to be a powerful platform for studying protein activity. In addition, as I discussed in the *Chapter 2 &4*, it has also enabled the discovery of new proteins and drug screening[288,302,360]. However, activity-based E3 ligase profiling has not yet be reported. In *chapter 4*, I discussed and demonstrated a facile way to generate RBR/HECT E3 ABPs. These probes have been shown their ability to dissect the activity of comprehensive Parkin E3 system and are able to detect the endogenous level of Parkin E3[322]. Moreover, our probes can also label endogenous HOIP from cell extract or HECT-domain only Nedd4L, UBE3C, HECTD1[472] recombinant proteins in an activity-dependent manner. On the other hand, the low reactivity toward DUBs further indicated the selectivity of our E3 ABPs. As a result, we reasoned that our probes can be applied in E3s ABPP. There are nearly 700 different E3s in the human genome which can be further subgroup into three families: RING, HECT and RBR E3 ligases. HECT and RBR E3s possess a catalytic cysteine nucleophile that undergoes transthiolation reactions with E2~Ub forming a covalent intermediate that subsequently transfers Ub to substrate[92,473,474]. Alternatively, RING E3s recruit Ub~E2s and position them proximal to substrates and mediate direct ubiquitin transfer via RING domains[125,126,159]. Interestingly, the presence of the RING domain does not indicate the conserved RING E3-like function. The RBR E3s also contain the canonical RING domain but have characteristics of the HECT family which transfer ubiquitin to substrates via a covalently-linked E3~Ub intermediate[121,242]. RING E3s are the biggest family among E3s with more than 600 predicted members and the detailed mechanism behind E3s possessing a RING domain is not fully understood.

5.1 Activity-based proteomics of E3 ligases

5.1.1 Development of biotin-conjugated E3 ABPs

The mechanism of E3 labeling by our probes involves trapping of the native ubiquitin E2-E3 transthiolation intermediate (**Figure 5.1a-b**). In order to retrieve ABP-E3 complexes from cell extracts and to perform the proteomic analysis, I prepared biotinylated variants of our recently developed ABPs. In the *Chapter 4*, I described the attachment of a TAMRA reporter tag by genetic encoding of an unnatural amino acid at E2 N-terminus. Here we simplified reporter tag attachment by introducing a GCSSG sequence between the Met1 and Gln2 of C-terminally truncated Ub plasmid, I obtained the N-terminal cysteine tagged Ub thioester Cys-Ub_{1-73}-SR and Cys-Ub_{1-74}-SR[322] (**Figure 5.1b**). These Cys-tagged Ub thioesters were then conjugated with iodoacetyl-PEG_2-biotin and obtained biotin-Ub_{1-7x}-SR (**Extended Figure 5.1 and 5.4**). It has been demonstrated that different E3s have distinct E2 preference[20,116,244]. In order to obtained broad coverage of E3s, we generated the biotin tagged ABPs based on the promiscuous E2, UBE2D2, (UBE2D2-ABP **13**) (**Extended Figure 5.2**), and the HECT/RBR specific E2, UBE2L3, (UBE2L3-ABP **15**) (**Extended Figure 5.5**), by the same method I described in *chapter 4*. As an E2-E3 labeling control, I also prepared mutant biotin tagged ABPs predicted to disrupt or impair E3 ligase binding (UBE2D2-F62A-ABP **14** and UBE2L3-F63A-ABP **16**, respectively) (**Extended Figure 5.3 and 5.6**). The biotin tagged ABPs have been widely used in protein proteomic profiling[340,475,476] due to their extremely high binding affinity and selectivity with streptavidin. The high tolerance of this interaction toward extreme conditions allowed me to perform harsh washing methods to eliminate the off-target binding with the immobilized beads. The newly developed biotin tagged ABPs (ABP **13** and **15**) both labeled Parkin E3 with the same efficiency as His tagged ABPs (ABP **10**) (data not shown).

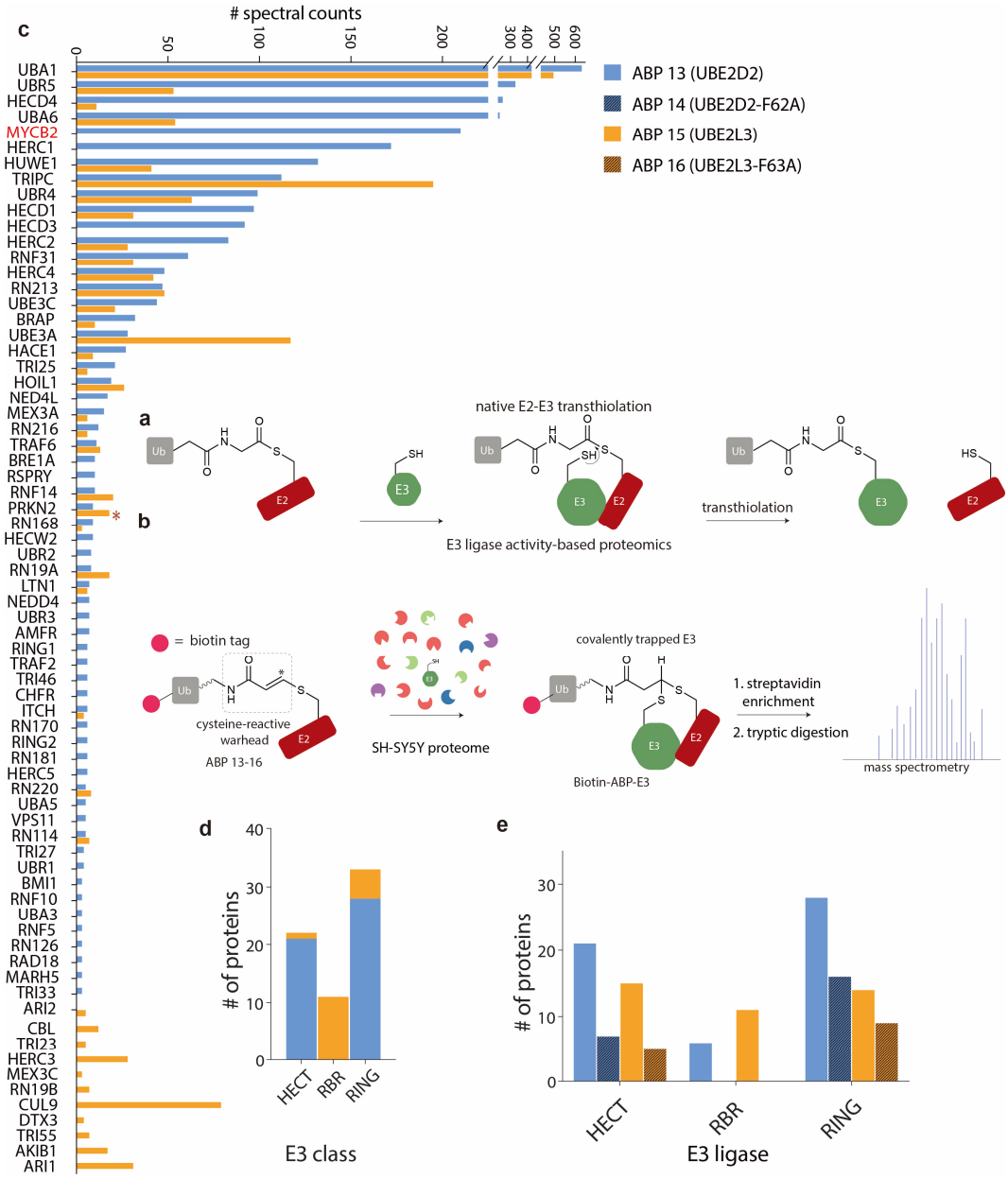


Figure 5.1 Activity-based Proteomics of E3 ligases.

(**a**) Native transthiolation between thioester-linked E2~Ub and E3. (**b**) The ABP acts like a suicide substrate and covalently traps E3s that demonstrate transthiolation activity. Biotin-tagged ABPs enable streptavidin enrichment of labelled proteins from extracted proteomes. Proteins can then be identified by semi-quantitative mass spectrometry-based spectral counting. The wavy bond represents triazole linkage to a truncated ubiquitin molecule (as described[322]; **Extended Figure 5.1 and 5.4**). (**c**) Number of spectral counts obtained from activity-based proteomic profiling of OA-treated SH-SY5Y neuroblastoma cells for proteins currently annotated as E3 ligases and E1s (E1s are labelled because they also demonstrate transthiolation activity with E2s). Proteins with less than 3 spectral counts were rejected. (*Next page*)

(**d**) Total number of HECT, RBR and RING E3 ligases detected from the different probes. As the mutations (E2 F62/63A mutation) introduced into control ABPs **14** and **16** are likely to impair rather than abolish E3 labelling, proteins with less than half the number of spectral counts of those obtained with their parent ABP (**13** and **15**, respectively) were not plotted. The asterisk highlights the RBR E3 ligase Parkin (PRKN2) which is only detected when OA is added to cells, further suggestive of activity-dependent labelling. (**e**) Numbers of E3 ligases detected using ABPs **13** and **15**. 22 HECT E3s were detected, 21 of which could be recovered with ABP **13** alone (HERC3 was only detected with ABP **15**). 11 RBRs were detected with ABP **15** alone. In aggregate 33 RING E3 ligases were detected, 28 of which could be recovered with ABP **13** alone (CBL, TRI23, MEX3C, DTX3 and TRI55 were only detected with ABP **15, Extended Figure 5.8 & 5.10**).

5.1.2 Profiling of endogenous E3 ligase activity by MS-based proteomics

Since our probe, ABP**10**, has been demonstrated to label endogenous RBR E3 Parkin in an activity-based manner, I applied the same model system to validate my biotin tagged ABPs. Neuroblastoma SH-SY5Y cells were left untreated or treated with the inhibitors of oxidative phosphorylation, oligomycin and antimycin A (OA) which leads to ABP-detectable activation of the RBR E3 Parkin by depolarizing mitochondria[60,477]. Parallel profiling of SH-SY5Y cell extracts was performed with our biotin tagged ABPs (ABP **13**-**16**) and labelled proteins were enriched against streptavidin resin. Stringent washing conditions were applied followed by on-bead trypsin digestion (**Figure 5.1b**). ABP-enriched protein peptides were analyzed by an LC-MS/MS platform which based on an LTQ Orbitrap Velos instrument coupled to an Ultimate nanflow HPLC system. Data was processed by the MRC proteomics Facility, Dr. Virdee and me. The E3s were filtered against PFAM domain term searches combined with manual curation. E3s that showed less than 14-fold spectral count enrichment compared to control purifications, where biotin tagged ABP was not supplied, were excluded (**Figure 5.1c**). As expected, we observed > 50 % reduction of the number of spectral counts for the majority of HECT/RBR E3s when the E2 binding-defective control ABPs **14**

and **16** were supplied in the reactions (**Figure 5.1d** and **Extended Figure 5.7 & 5.8**). These data demonstrate that members from all E3 subtypes can be sensitive to the E2 F62/F63A mutation, especially for the RBR family (almost all of them are sensitive to the mutation), suggestive of activity-dependent labelling. We cannot formally exclude whether the detection of mutation insensitive E3s is activity-independent or whether these E3s are simply permissive to the F62A/F63A mutation. The retaining E3 reactivity of E2 binding-defective ABP **14** and **16** may due to unknown E2-E3 binding mechanisms or off-target labelling by incorrect positioning of the probes on the E3s (false E2-E3 pairs). The exact probe labeling sites should be further identified in the future by crosslinking MS. Notably, Parkin peptides were only be observed from the cells treated with OA (**Extended Figure 5.9 & 5.10**). This was consistent with the immunoblotting data (data not shown) and previous research. Taken together, this demonstrated that for at least a subset of detected E3s, spectral counts correlated with E3 activity.

ABP **13** and **15** are constructed with two different E2 (promiscuous UBE2D2 and HECT/RBR favored UBE2L3 respectively), as expected, they retrieved distinct and overlapping E3 subsets which was reflective of their different E3 preferences (**Figure 5.1c**). The aggregate number of recovered HECT/RBR E3s from probes **13** and **15** was 33 (22 HECT and 11 RBR) thus representing ~80 % of the currently annotated HECT/RBR E3s (**Figure 5.1e**). However, we also observed in total 33 RING E3s from our dataset, which was unexpected. RING E3s do not contain a catalytic cysteine so in theory, should not be labeled by our probes. One explanation is that mechanistically off-target labeling was taking place and due to the high sensitivity of MS-based detection, these aberrantly labelled E3 were being recorded. Alternatively, these RING E3s may display an unknown mechanism to interact with E2 enzymes for ubiquitin transfer. Notably, some of the labelled RING E3s are

poorly characterized and all that is known about them is the presence of a predicted RING domain, for example, RNF181, DTX3, ZN598 and MYCBP2. The RING E3 with the highest number of spectral counts was MYCBP2 (myc-binding protein 2, also termed Phr-1[478]) which prompted us to explore this putative RING E3 further (**Figure 5.1c**).

5.2 Characterization of the neuron-associated protein, MYCBP2

Among various batches of sample we tested, a giant 0.5 MDa neuron-associated E3 MYCBP2, which has a C-terminal RING domain, was unexpectedly labeled by our ABPs. As MYCBP2 does not contain any domains known to be involved in transthiolation (its C-terminal RING domain is devoid of HECT or IBR/RING2 domains), we hypothesized that it was a novel class of E3 ligase. MYCBP2 and its orthologues are collectively termed PHR (PAM-Highwire-Rpm-1) proteins (PAM in Human, Phr1 in mouse, Esrom/Phr1 in zebrafish, Highwire in Drosophila and RPM-1 in *C. elegans*[478]). Although there are quite a few papers describing the functions of MYBCP2, research on the function of the isolated domains on MYBCP2 is missing. The large size of MYCBP2 is likely to have made its study challenging.

MYCBP2 was first isolated by a phage screen for binding proteins of the protooncogene *MYC*[479]. The authors further demonstrated that MYCBP2 is widely expressed in human tissue samples but with particularly high levels in brain. The RCC-1 domain and Myc protein binding site were also identified in the same report[479]. However, it was later demonstrated as a neuron-associated protein by studies in Drosophila[480-482]. It has also been shown to regulate synapse formation and axon termination in mice and fish[483-485]. The embryonic lethal effect of Phr1 knock-out mice, due to respiratory distress, further heighted its importance[486]. Although MYCBP2 is less studied in human models, it has been shown that the loss

of Highwire exclusively impairs the motor neurons in Drosophila[481]. The *RPM1* mutants in *C. elegans* was also observed to have impairment of the neuron-neuron synaptic connection[480]. Moreover, this mutant also causes severe axon termination defects in *C. elegans*[480,487,488]. Consistent with their functions in regulating neuron development, PHR proteins are often found in presynaptic terminals[487,489,490].

PHR proteins have been shown to regulate several signaling pathways[478]. It has been shown that MYCBP2 can regulate mTOR signaling pathway by regulating TSC2 (Tuberous Sclerosis Complex (TSC) gene) ubiquitylation levels in an over-expression system [491]. MYCBP2 has also been indicated to form a non-canonical SCF ubiquitin ligase complex with Fbxo45, a putative E3 substrate receptor, and Skp1[492,493]. The orthologue of this complex is found in Drosophila[*490*]. In addition, MYCBP2 has also been found to mediate the activity of p38 MAPK via regulation of transient receptor potential vanilloid receptor 1 (TRPV1)[494].

MYCBP2 is also a positive regulator of Wallerian axon degeneration where it promotes the degradation of the axonal neuroprotectant, Nicotinamide Mononucleotide Adenyl Transferase (NMNAT2)[493,495,496]. Consequently, MYCBP2 mutant animals demonstrate potent neuroprotection in response to axonal injury[493,495]. Above reports all demonstrate the importance of MYBCP2 in cells; these synaptic and neuroprotective phenotypes are observed with RING domain mutants and those pertaining to a structurally uncharacterized cysteine-rich region immediately C-terminal to the RING domain[482,495]. Since there is no elegant biological report on C-terminal region containing the RING domain, and our probes labelled this protein, we suspected an unknown mechanism for MYCBP2 E3 activity.

5.2.1 Profiling of neuron-associated protein, MYCBP2, by E3 ABPs

Since we obtained the highest number of spectral counts for a RING E3 was MYCBP2, we carried out an *in vitro* probe labelling assay with recombinant MYCBP2. We expressed a C-terminal version of MYCBP2 encompassing the RING domain in *E. coli* (residues 4378-4640; MYCBP2cat; **Figure 5.2b**). MYCBP2cat underwent robust ABP labelling (**Figure 5.2a**) with an efficiency comparable to that observed with E3 ligases known to demonstrate transthiolation activity[322,472]. This implied that MYCBP2 might be a novel class of RING-linked E3 ligase which used a different mechanism to transfer ubiquitin to the substrates.

We speculated that if a novel domain was present that contained an active site cysteine it would be C-terminal to the RING domain. We carried out structure prediction using the Phyre2 and I-TASSER servers and both predicted low confidence similarity to RBR E3 ligases. Immediately after the RING domain is an unannotated region that contains 16 cysteine residues, many of which reside in Cys-X-X-Cys motifs, characteristic of structural Zn-binding sites[497]. In order to clarify the category of MYCBP2 without any structural information on this region, I next set out to identify the possible catalytic cysteine in this region. Since the ABP undergoes mechanism-based labelling of the active site cysteine sulfhydryl group in HECT/RBR E3s, identifying a Cys mutant that abolished ABP labelling would allow me to figure out the catalytic cysteine (**Figure 5.1b**)[121,168,242]. I systematically generated a panel of cysteine to serine MYCBP2cat mutants and reacted with the ABP **18** (**Figure 5.2c**). We validated the labelling efficiency by both Coomassie staining and immunoblotting against the primary His-tag antibody. Three cysteine residues (C4506S, C4520S and C4537S) in MYCBP2cat were found to abolish ABP labeling (**Figure 5.2c**). Since C4506 and C4537 were located in a Cys-X-X-Cys motif and were likely to have a structure role, C4520 became the most promising

candidate as the catalytic cysteine. To further confirm our hypothesis, we subsequently carried out tandem mass spectrometry[322] of ABP-labelled wild type MYCBP2cat and the data supported the role of C4520 as an active site nucleophile (**Figure 5.2d**). We observed 38 spectral matches corresponding to probe labelling sites from MS result and 36 of which were assigned to C4520 (**Extended Figure 5.11**). This further indicated that C4520 might be a catalytic cysteine in a new class of E3.

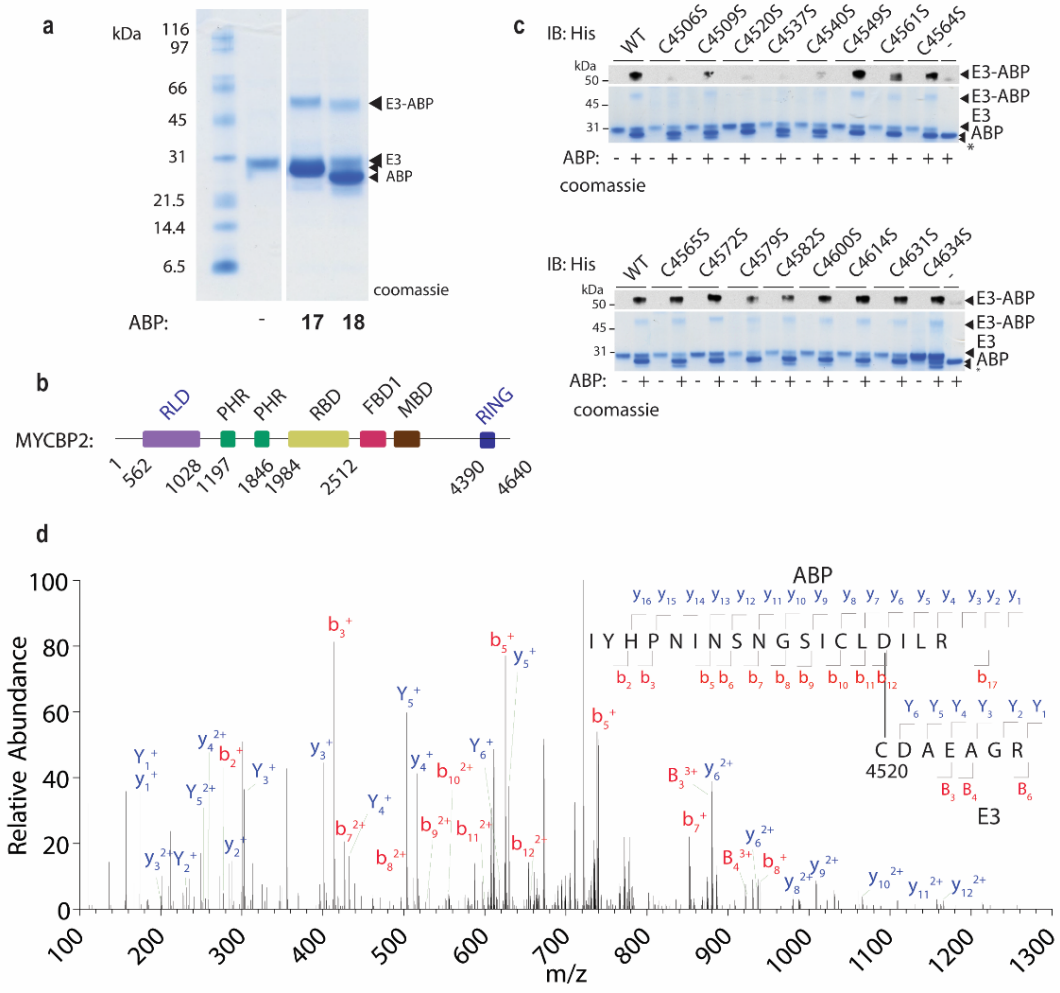


Figure 5.2 MYCBP2cat is labelled by activity-based probes that profile transthiolaiton activity.

(**a**) A C-terminal version of MYCBP2 (residues 4378-4640) is labelled by ABPs based on the E2s UBE2D2 (**17**) and UBE2D3 (**18**)[322]. (*Next page*)

(**b**) Domain architecture of MYCBP2. Abbreviations: RCC1-like GEF domain (RLD), two PHR family-specific (PHR) domains, a RAE1 binding domain (RBD), a F-box binding domain 1 (FBD1), a Myc binding domain (MBD), and a C-terminal RING domain[478]. (**c**) Putative active site cysteines in MYCBP2cat were determined by ABP profiling of a panel of cysteine to serine mutants. MYCBP2cat mutant (3 μM) was incubated with ABP **18** (12 μM) at 30 °C for 4 hours. ABP-treated samples were resolved by SDS-PAGE and visualized by Coomassie staining and immunoblotting against the hexahistidine reporter tag on the ABP. (**d**) C4520 was underscored as the active site cysteine residue following tryptic digestion of the ABP-labelled protein and MS/MS sequencing. The spectrum is for a 5^+ precursor ion that was within 1 ppm of its theoretical *m/z* value (observed *m/z* = 614.5088; expected *m/z* = 614.5094).

5.2.2 Catalytic cysteine-dependent esterification activity of MYCBP2

Based on the above data and the software structure prediction, we speculated that MYCBPcat displays RBR E3 characteristics. Furthermore, as RBR E3s were first considered as RING E3s, it seemed plausible that MYCBP2 was another new class of RING-linked E3. One of the features of RBR E3s is that they undergo auto-ubiquitylation Ub chain formation dependent on their catalytic cysteine[121,159,242]. In order to further dissect the role of C4520 in MYCBP2cat, I established an in-vitro E3 activity assay by using E1, UBE2D3 and Ub[121]. Surprisingly, I could not detect any ubiquitin chain formation or even auto-ubiquitylation of MYCBP2cat (**Figure 5.3a** and **Extended Figure 5.12**). However, rapid MYCBP2cat-dependent ubiquitin discharge from E1 and E2 were observed and were dependent on C4520 (**Figure 5.3a** and **Extended Figure 5.12**). Since there was no nucleophile present that is known to become modified by E3 ligases, such as lysine and cysteine[121], the high level of ubiquitin discharge was unexpected. We therefore analyzed the whole reaction by liquid chromatography-mass spectrometry (LC-MS). This demonstrated that the ubiquitin was been quantitatively modified by tris(hydroxymethyl)aminomethane (Tris) and glycerol (8639 and 8668 Da), both of which were present in our assay buffer at concentrations of 50 mM and ~65 mM,

respectively (**Figure 5.3c-d**). I assayed activity for all of the generated cysteine mutants and found it to be dependent on the active site cysteine (**Extended Figure 5.12**), C4520S (a characteristic of E3s that form a thioester-linked E3~Ub intermediate) thereby underscoring its catalytic role. Consistent with the result I found earlier, the C4520S mutation abolished the ubiquitin ester adducts which further emphasized the catalytic role for C4520 (**Figure 5.3a**).

In the E2 discharge assay, I found that another cysteine mutation, C4572S, had slightly reduced E3 activity but formed a mono-Ub MYCBP2 adduct (**Figure 5.3a**). In addition, the mono-Ub adduct was resistant to reducing agent (2-mercaptoethanol, BME) but was labile to base environment after treatment with 0.14 N NaOH (**Figure 5.3b**). This demonstrated that the mono-Ub adduct was probably linked via an oxyester bond to S4572. The striking discovery from these data suggested to us that C4572 may also play a catalytic role in ubiquitin transfer.

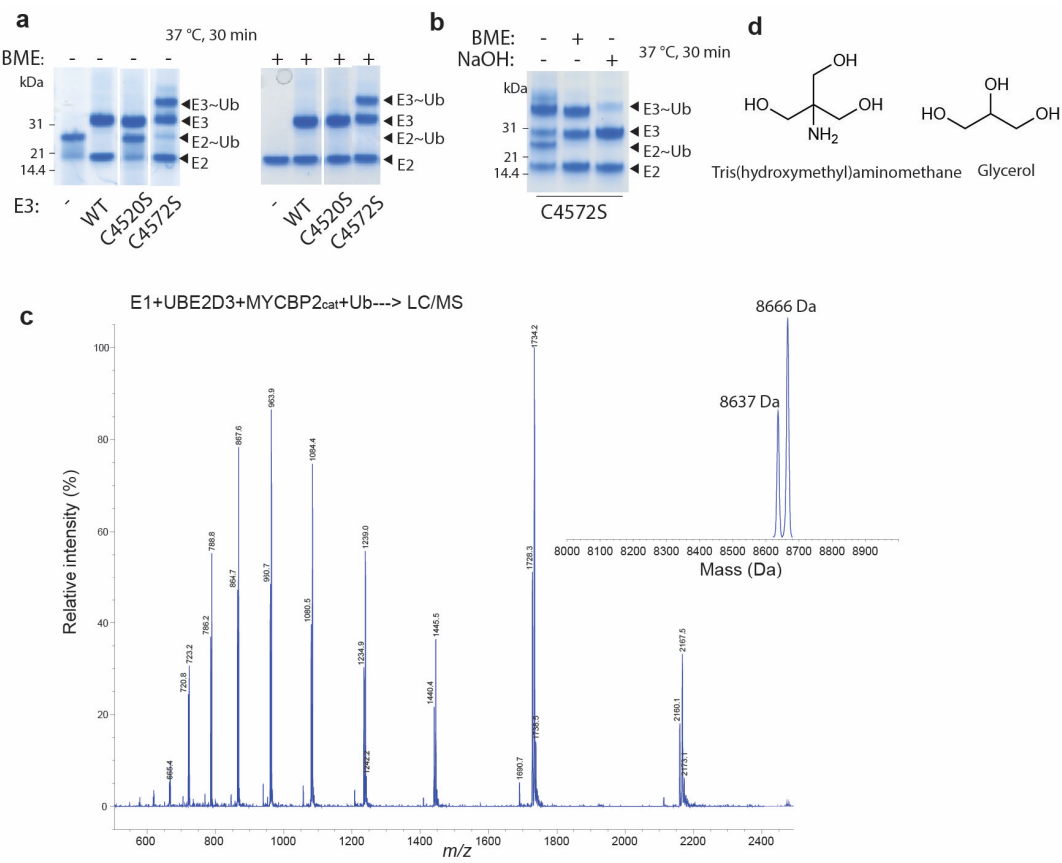


Figure 5.3 MYCBP2cat C4520 plays an essential role in E2~Ub discharge reaction.

(**a**) Multiple turnover E2 discharge assay onto Tris (50 mM)/glycerol (~65 mM). In the absence of the thiol 2-mercaptoethanol (BME; left panel), E2 (15 μM) is loaded with Ub. In the presence of untagged MYCBP2cat wild type (WT) (15 μM), E2~Ub undergoes quantitative discharge due to progressive covalent attachment to threonine, which depletes the activatable pool of Ub. A MYCBP2cat C4520S mutant abolishes E3-mediated discharge of Ub from E2~Ub whereas a MYCBP2cat C4572S mutant demonstrates a modest reduction in activity and forms a Ub adduct. In the presence of BME (right panel), E2~Ub undergoes thiolysis whereas the Ub adduct formed on the C4572S mutant is thiol resistant. Discharge reactions were incubated at 37 °C for 30 minutes. (**b**) The Ub adduct on MYCBP2 C4572S is base labile (0.14 N NaOH) indicative of the formation of an engineered (oxy)ester linkage between Ub and the S4572 residue. (**c**) Mass spectrum of condensation productions between Ub and glycerol and, Ub and Tris. Expected mass for Ub condensation with glycerol is 8,639 Da (found = 8,637 Da). Expected mass for Ub condensation with Tris is 8,668 Da (found = 8,666 Da). (**d**) Chemical structure of tris(hydroxymethyl)aminomethane (Tris) and glycerol.

5.2.3 MYCBP2 operates via a cysteine relay mechanism

BME-resistant but base-labile MYCBP2cat C4572S Ub-adducts could be obtained in repeated assays. An explanation for this observation is that the serine mutant

was contributing to catalysis via formation of a less transient (oxy)ester-linked intermediate between Ub that was still functionally active. This suggested that in addition to C4520, C4572 may have a catalytic contribution to E3 activity. We hypothesized that C4520 and C4572 were both catalytic residues that functioned in tandem and relayed Ub from one cysteine to the other by undergoing an intramolecular transthiolation reaction. To further test this hypothesis, I established a gel-based thioester/ester trapping assay[242] using fluorescently labelled Ub, (Cy3B-Ub), and GST tagged MYCBP2cat (**Figure 5.4**). After removal of all traces of hydroxyl components in my buffers, I managed to trap a ubiquitin thioester on wild type MYCBP2cat (**Figure 5.4**, lane 7) and it was labile to BME treatment (**Figure 5.4**, lane 8). In contrast, MYCBP2 C4520S did not form a detectable thioester adduct (**Figure 5.4**, lane 9). Within my expectation, the C4520A mutant was not able to form a mono-ub adduct (**Figure 5.4**, lane 11). Consistent with earlier experiments (**Figure 5.3b**), I obtained a C4572S adduct which was thiol-resistant but base labile (**Figure 5.4,** lanes 13 - 15). Moreover, if a transient thioester intermediate was being formed between Ub and C4520, then a proposed unreactive C4572A mutant should trap it and enhance the MYCBP2~Ub level. The data was consistent with this hypothesis; compared with wild type MYCBP2cat we observed more Ub adduct formation on the C4572A mutant and the adduct was thiol-labile (**Figure 5.4**, lane 16-17). I did not observe any adduct on the C4520A/C4572S mutant further supporting the idea that the intermediate adduct was indeed formed on the C4520 residue.

In light of the above data, we proposed a novel E3 ubiquitin transfer mechanism with dual catalytic cysteines, C4520 and C4572. A plausible model was that they were both catalytic residues that functioned in tandem and relayed Ub from one cysteine to the other by undergoing an unprecedented intramolecular

transthiolation reaction. The inability to render S4520 catalytic was likely to be reflective of C4520 and C4572 residing in distinct chemical environments. The intrinsic nucleophilicity reactivity of C4520, that cannot be copied with serine, might be essential at this position.

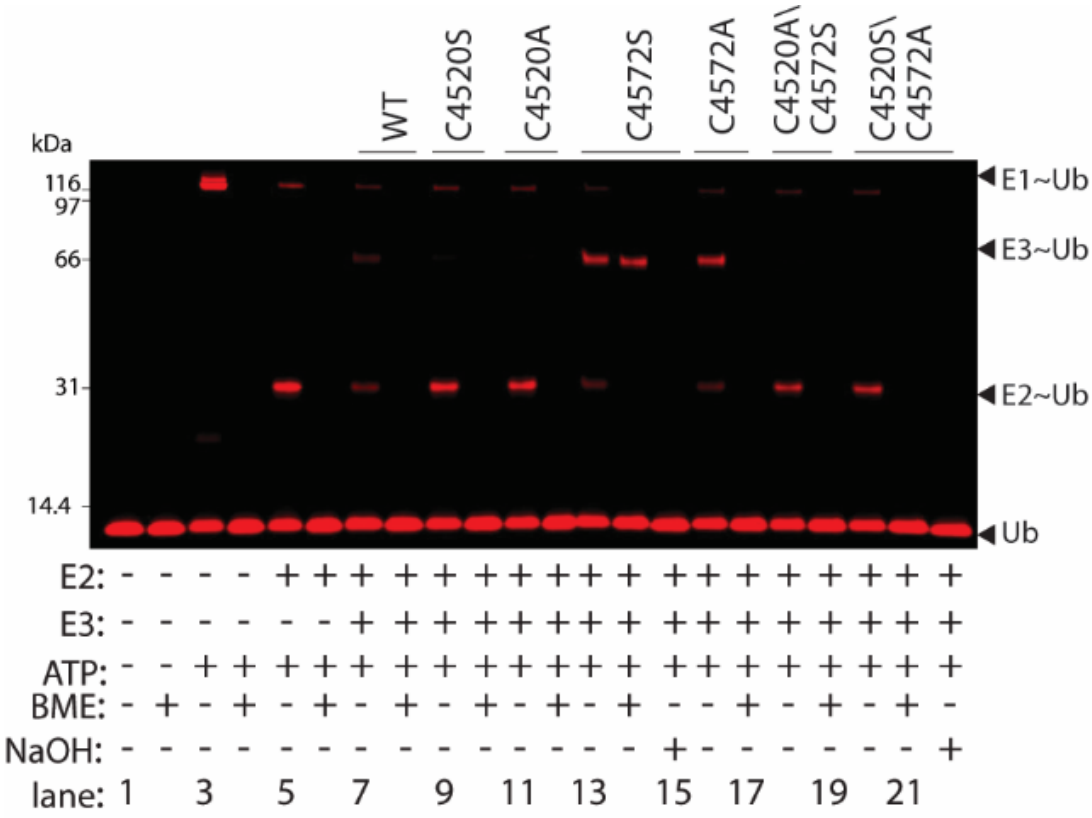


Figure 5.4 MYCBP2 E3 ligase activity is dependent on a cysteine relay mechanism.
In the absence of hydroxy substrates, a BME-sensitive Ub adduct can be detected on MYCBP2cat WT indicative of it being thioester-linked (lanes 7 and 8). Mutation of C4520 to serine (lanes 9 and 10), or alanine (lanes 11 and 12), abolishes thioester adduct formation. MYCBP2cat C4572S forms a stable adduct that is resistant to BME but labile in the presence of 0.14 N NaOH (lanes 13-15). The MYCBP2cat C4572A mutant forms a BME labile adduct that is more stable than that formed with MYCBP2cat WT (lanes 16 and 17 vs. lanes 7 and 8). A C4520A/C4572S double mutant abolishes adduct formation entirely indicating that the intermediate adduct is formed on C4520. (Commasie stained-gel is in **Extended Figure 5.13**)

5.2.4 MYCBP2 ubiquitylates threonine with high selectivity

Since we discovered that MYCBP2 has high ubiquitylation activity towards hydroxyl groups, we speculated that it may not ubiquitylate lysine residues like conventional

E3 ligases[498]. We followed up with a screen with MYCBP2cat mediated ubiquitin discharge onto a panel of amino acids including alanine, lysine, arginine, serine and threonine[121]. Glycerol was applied in the screen as a positive control. We first loaded E2 with Ub via a standard Ub charging assay by incubation with E1, E2, Ub and ATP. MYCBP2cat and amino acids mixtures were then supplied into the reaction. With surprise, the discharge activity of MYCBP2 was high toward threonine rather than the canonical ubiquitin substrate lysine (**Figure 5.5a**). LC-MS data further confirmed the adduct of ubiquitin-threonine and the product formation was also dependent on catalytic cysteine C4520 (**Figure 5.5b**). In order to further validate the selectivity of MYCBP2 towards these amino acids, I performed the discharge assay under long time course contains and quantified the modified and unmodified species by LC-MS. I observed nearly 10-fold selectivity for the amino acid threonine over serine based on the time point when 50% ($t_{1/2}$) of ubiquitin was modified by amino acid (**Figure 5.5c**, ~5 mins for Thr and ~50 mins for Ser). The lysine-modified Ub was observed by LC-MS but it was independent of MYCBP2cat. It implied that the modification may most likely be due to direct transfer from E2[121].

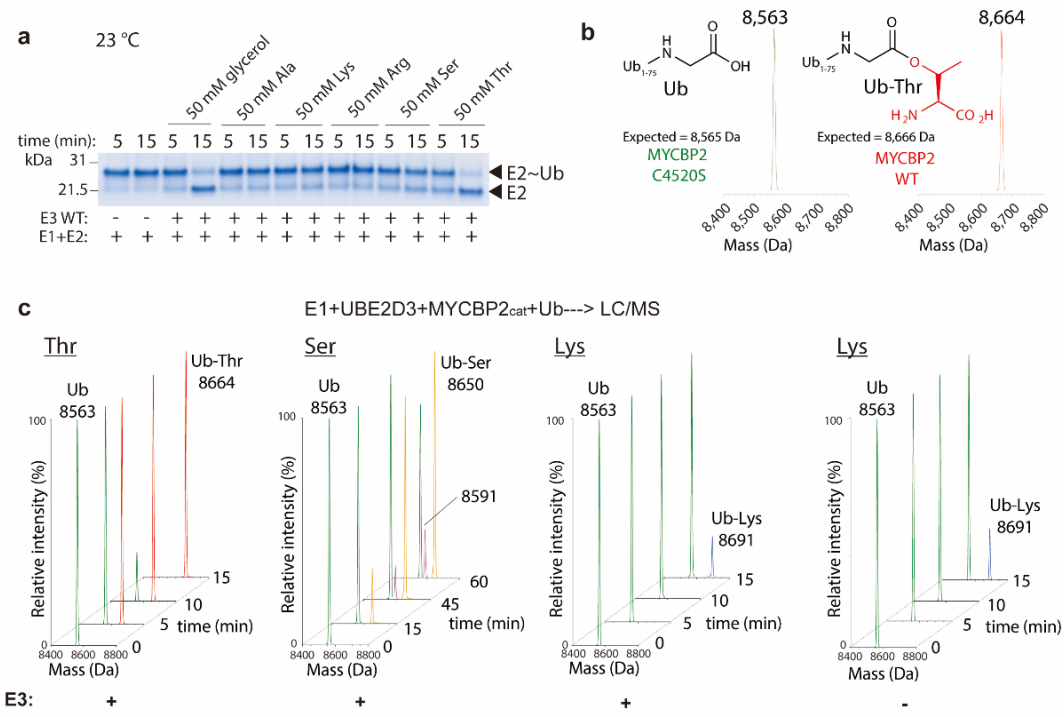


Figure 5.5 MYCBP2 is a threonine selective E3 ligase.

(**a**) E3-mediated multiple turnover discharge reaction with UBE2D3 (10 μM) and GST-MYCBP2$_{cat}$ (5 μM) against glycerol (50 mM) as a positive control, and a panel of amino acids (50 mM). Discharge was carried out at 23 °C for the indicated time. (**b**) The discharge assay was analyzed by LC-MS revealing that Ub was being converted to a product with a molecular weight corresponding to the condensation product with threonine (observed molecular weight=8,664 Da; theoretical= 8,666 Da). With MYCBP2$_{cat}$ C4520S, Ub remained unmodified (observed molecular weight = 8,563; theoretical = 8,565 Da). (**c**) Deconvoluted mass spectra for ubiquitin species in the presence of amino acid (50 mM). The intensities of the Ub reactant and product are reflective of their relative abundance. Ubiquitinated serine (Ub-Ser) observed molecular weight = 8,650 Da; theoretical = 8,652 Da. Observed mass at 8,591 Da corresponds to a side product that is only observed after extended incubation. Ubiquitinated lysine (Ub-Lys) observed molecular weight = 8,691 Da; theoretical = 8,693 Da. Assuming exponential Ub consumption, $t_{1/2}$ is ~5min for threonine. For serine, $t_{1/2}$ is 10-fold slower. Lysine ubiquitylation is E3-independent as a similar degree of modification is observed in the absence of E3.

In order to validate whether the threonine activity and selectivity were retained in a peptide context or not, we firstly carried out time course analysis for ubiquitylation of a panel of model peptides of the sequence Ac-EGXGN-NH2 (X=Lys, Ser or Thr). The random single peptide sequences were chosen due to the lack of

knowledge for the proper MYCBP2 substrates (papers published to date did not indicate the ubiquitylation sites within substrates for MYCBP2). By supplying fluorescent Ub, Cy3B-Ub, into the reaction, I was able to quantify the activity of MYCBP2 toward peptides. The MYCBP2cat threonine selectivity over serine for the peptide context was retained approximately 3-fold (**Figure 5.6**, higher panel **Extended Figure 5.14**). Although the ubiquitylation of the Lys-peptide was low as expected, we observed its ubiquitylation was partially inhibited in the presence of MYCBP2cat. Since free Lys-ubiquitylation was found to be E3-independent in our previous LC-MS data, similarly I presumed the inhibition was due to MYCBP2 transthiolation activity competing with E2 lysine aminolysis activity. This underscores the absence of MYCBP2cat lysine reactivity (**Figure 5.6**, lower panel). Taken together, the above experimental data revealed that MYCBP2 was not only a novel class of E3 enzyme that operates via a cysteine relay mechanism, it specifically esterifies threonine with Ub. Because of its novel mechanism we termed it a RING Cys-Relay (RCR) E3 ligase.

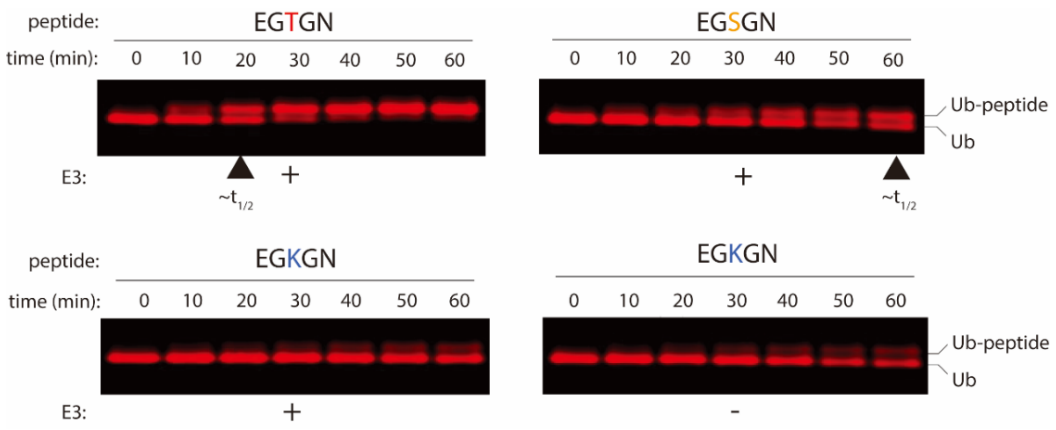


Figure 5.6 MYCBP2 retains threonine selectivity in a peptide context.

In-gel fluorescence scan of peptide modification with fluorescently labeled Ub. The $t_{1/2}$ for Ub modification with the threonine and serine peptides is ~20 min and ~60 min, respectively. Observed mass of Cy3B-Ub = 9,534 Da; theoretical mass = 9,537 Da. Observed mass of Cy3B-Ub modified threonine peptide = 10,033 Da; theoretical mass = 10,036 Da. Observed mass of Cy3B-Ub modified serine peptide = 10,019 Da; theoretical mass = 10,022 Da (**Extended Figure 5.14**). Inefficient modification of the lysine peptide is observed which is moderately enhanced in the absence of E3.

5.2.5 MYCBP2 ubiquitylates NMNAT2 by esterification

MYCBP2 has been demonstrated to promote Wallerian degeneration through destabilization of NMNAT2 protein[493,495]. Furthermore, it has also been demonstrated that the level of NMNAT2 was increased in the Phr1 (MYCBP2 orthologue in mice) KO primary neurons[493]. I suspected the NMNAT2 is the downstream substrate for MYCBP2. In order to validate whether MYCBP2 retains the same esterification reactivity towards NMNAT2 ubiquitylation, we expressed and purified recombinant NMNAT2 and carried out ubiquitylation assays by incubating NMNAT2 within E1, E2 (UBE2D3) and MYCBP2cat. Mono-ubiquitinated NMNAT2 was observed and it was performed in an E3-dependent manner. Moreover, the mono-ubiquitylated NMNAT2 adduct was resistant to reducing agent but was hydroxide labile (**Figure 5.7a**). It was further demonstrated that MYCBP2 had non-lysine ubiquitylation activity because there are 13 lysine residues in NMNAT2.

To recapitulate the ubiquitinated-NMNAT2 under cellular conditions was challenging since NMNAT2 ubiquitylation is mediated by a Skp1/Fbox45 substrate receptor co-complex that binds to a site ~1940 residues N-terminal to the MYCBP2cat region[493]. The activation mechanism for the endogenous MYCBP2 is also poorly understood. In addition, NMNAT2 undergoes palmitoylation and rapid axonal transport making reconstitution and cellular study of its ubiquitylation extremely difficult. Considering the entirely different pathways non-lysine ubiquitylated proteins may take, the inhibition of proteasome activity which is the standard procedure to facilitate the detection of ubiquitylated species in cells, may not contribute the same effects to non-lysine ubiquitylated proteins. However, in order to confirm that the non-lysine esterification activity by MYCBP2 occurs in cells, the overexpression of myc-tagged MYCBP2cat in HEK (human embryonic kidney) 293 cell either with or without HA-Ub were performed. A mono-ubiquitylated form of MYCBP2 was observed and that was dependent on C4520. Furthermore, the mono-Ub adduct was also resistant to thiolysis after BME treatment but labile under basic treatment, suggestive of the Ub being ester-linked to the protein (**Figure 5.7b)**. These data demonstrate that MYCBP2 can retain specificity for hydroxy amino acids in mammalian cells.

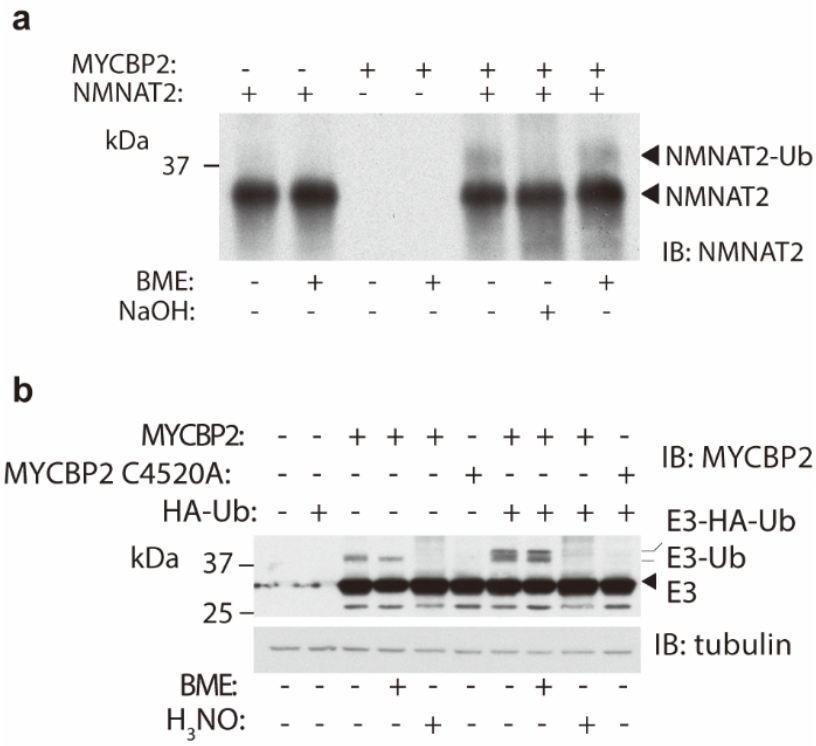


Figure 5.7 Non-lysine ubiquitylation specificity of MYBCP2 is retained in mammalian cells.

(**a**) Recombinant NMNAT2 undergoes base-labile ubiquitylation in a MYCBP2cat dependent manner. The NMNAT2 protein was incubated with E1, UBE2D3, MYCBP2cat, Ub and ATP at 37°C for 30mins. The mixtures were resolved by SDS-PAGE and performed immunoblotting with NMNAT2 primary antibody. (**b**) MYCBP2cat transiently transfected into HEK293 cells undergoes hydroxy ubiquitylation that is dependent on C4520. The discrete ester-linked species that corresponded to the molecular weight of MYCBP2cat plus the addition of Ub (or HA-Ub) were observed. These species were also resistant to reducing agent but labile to hydroxide. The mono H_3NO corresponds to treatment with 0.5 M hydroxylamine at pH 9.0.

5.2.6 MYCBP2cat performs as a monomer for ubiquitin transfer

As I discussed in *Chapter 1*, some RING E3s have been demonstrated as a dimer when they are activated[248]. In order to validate whether C4520 and C4572 in MYCBP2 relay Ub intramolecularly (*cis*) or via a dimeric intermolecular (*trans*) mechanism, I used untagged MYCBP2cat and initially demonstrated that a C4572A mutation abolished threonine discharge activity (**Figure 5.8a**). If the dimeric intermolecular mechanism was taking place, I should be able rescue MYCBP2

C4572A activity by supplying MYBCP2 C4520S which I previously showed was inactive in solution (**Figure 5.8a**). However, the threonine discharge activity was not restored when two mutants were incubated together (**Figure 5.8a**). The same result was obtained even after the concentration of MYBCP2 mutants were doubled. Moreover, the data from size exclusion chromatography-multi angle light scattering (SEC-MALS) of MYCBP2cat revealed a single species consistent with the mass of the monomeric protein which further supported the intramolecular monomeric relay mechanism (**Figure 5.8b**). The full-length MYBCP2 protein is nearly 500 kDa which contains multiple different domains with different roles. As a result, it is reasonable for MYBCP2 to retain as a monomer under physiological conditions. However, we still cannot formally rule out other possible mechanisms but we cannot think of any other models that are consistent with the data.

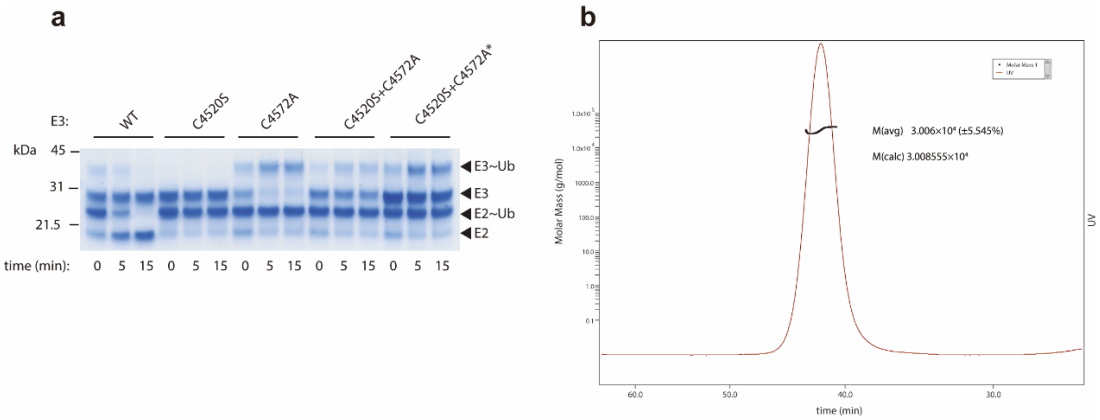


Figure 5.8 MYCBP2cat performs threonine discharge activity via intramolecular Ub realy transthiolation.

(**a**) The RCR E3 ligase activity is dependent on both C4520 and C4572. These residues relay Ub via an intramolecular transthiolation reaction rather than an intermolecular reaction as negligible threonine discharge activity is observed when the MYCBP2cat C4520S and MYCBP2cat C4572A mutants are combined (*corresponds to elevated concentrations of E3 mutants). (**b**) SEC-MALS analysis for untagged $MYCBP2_{cat}$. $MYCBP2_{cat}$ ran as a monodisperse species with a calculated molecular weight of 30.06 (±6) kDa (theoretical = 30.08 kDa). For the note, the SEC-MALS assay was performed by Dr. Ramasubramanian Sundaramoorthy.

5.2.7 The catalytic efficiency of MYCBP2cat

Since MYBCP2 adopts an unprecedented mechanism and a distinct preference towards substrate, I next wanted to further test the efficiency of this newly discovered machinery. In order to assess activity towards threonine I compared it to established HECT/RBR E3 lysine aminolysis activity. To do this, I developed a single-turnover E2~Ub discharge assay using fluorescently tagged Ub. The E2 UBE2D3 was preloaded with Cy3B-Ub by E1 enzyme. To prevent E1 reloading the reaction was treated with a highly specific E1 inhibitor[452]. The reaction was then incubated with indicated E3s (5:1 molar ratio for E2~Ub:E3) and amino acids. The observed rate constants (k_{obs}) for E3-substrate dependent single turnover E2~Ub discharge for UBE3C-mediated discharge was too low to be quantified. However, MYCBP2cat activity towards threonine was 22-fold lower than that of HHARI (**Figure 5.9**). Notably, the HHARI RBR E3 displays the auto-ubiquitylation activity which further enhanced the observed rate constants for HHARI. Since MYCBP2 does not contain the intrinsic auto-ubiquitylation activity, the comparison of threonine activity between lysine activity is not perfectly fair. In order to obtain the accurate MYCBP2 threonine reactivity, the direct protein substrates are needed.

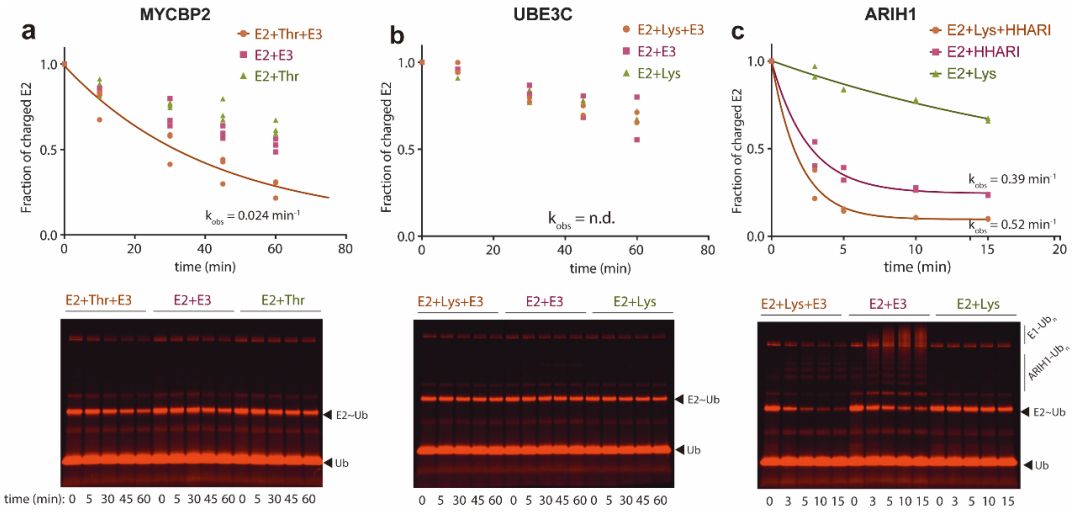


Figure 5.9 Observed rate constants (k_{obs}) for E3-substrate dependent single turnover E2~Ub discharge.

(**a**) Observed rate constant (0.024 min^{-1}) for MYCBP2cat-threonine mediated single turnover E2~Ub discharge determined by in-gel fluorescence of Cy3B labelled Ub. E2 was UBE2D3 and threonine (50 mM) was the substrate. The reaction was terminated on the indicated time points by supplied with non-reducing loading buffer. (**b**) Observed rate constant for UBE3C-lysine mediated single turnover E2~Ub discharge was too slow to measure. E2 was UBE2L3 and lysine (50 mM) was the substrate. (**c**) Observed rate constant (0.52 min^{-1}) for HHARI-lysine mediated single turnover E2~Ub discharge. E2 was UBE2L3 and lysine (50 mM) was the substrate. The major component of this rate is attributable to auto-ubiquitylation of lysine residues within HHARI because when lysine is withheld, k_{obs} HHARI mediated E2~Ub discharge is 0.39 min^{-1} and this is only partially outcompeted by the addition of lysine.

5.2.8 MYCBP2 displays distinct E2 requirements

In order to verify whether MYCBP2cat E3 ligase activity is exclusively attributed to E3~Ub thioester intermediate formation, or if it can perform direct Ub transfer from E2~Ub like RING E3s[91], I tested MYCBP2cat E3 activity with the HECT/RBR specific E2 UBE2L3[121] and a number of UBE2D3 mutants. Surprisingly, we observed negligible activity of the HECT/RBR specific E2, UBE2L3, with MYCBP2 cat (**Figure 5.11**). Considering the UBE2L3 is involved in mediating HECT/RBR~Ub thioester formation, the weak activity with MYCBP2cat RCR E3 may imply a different E2 requirement for MYCBP2cat E3 ligase. We then carried out the analysis with a

panel of UBE2D3 mutants (N77S, D87A, I88A, L97A, L104A, S108A and D117A)[125,126,159] to gain more insight about the E2 requirements for MYCBP2cat (**Figure 5.10**). The N77 and D117 mutants have been found to disrupt the amino acids in E2 required for suppressing the *pKa* of the acceptor nucleophile which are the essential residues for RING E3. In contrast to the requirement for in RING E3 activity, D117A showed negligible effect on MYCBP2 transthiolation while the N77S mutant activity was massively reduced. In addition, the E2 residues D87, I88, L97, L104 and S108 are involved in adopting the 'closed' conformation of E2~Ub which is required for RING E3 mediated ubiquitin transfer[125,126,159]. Unlike the canonical E2 requirement for RING E3s, S108A mutant retained the transthiolation activity with MYCBP2cat whereas I88A showed reduced activity. D87A, L97A and L104A mutants demonstrated the low transthiolation activity with MYCBP2cat. Moreover, the L104A mutant has been found to be permissive to RBR E3 activity[129] which was not the case for MYCBP2cat activity transthiolation (**Figure 5.10**). Taken together, the above data show that MYCBP2cat has evolved distinct E2 requirements that are neither consistent with a HECT/RBR-like mechanism nor a RING-like mechanism. We certainly cannot formally exclude the possibility that MYCBP2cat induces a closed E2~Ub conformation during contact with MYCBP2cat. The structure of the E2:MYCBP2cat complex is required for the answer.

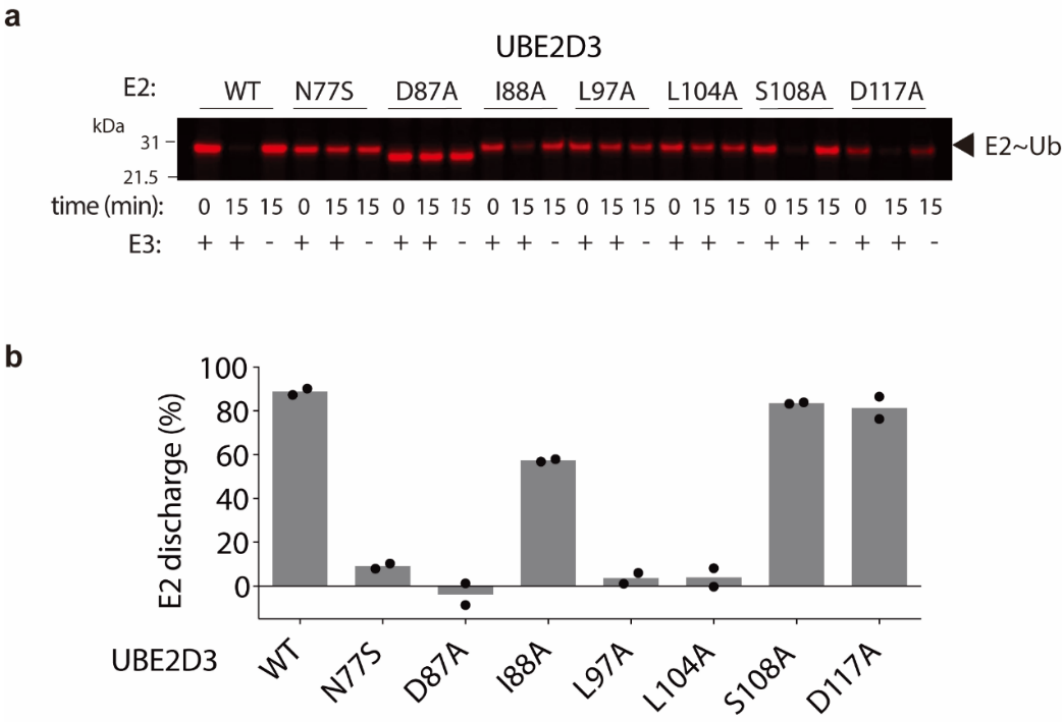


Figure 5.10 MYCBP2cat has a distinct E2 requirement.

(**a**) SDS-PAGE analysis of single turn over E2~Ub discharge assays employing Cy3B labelled Ub demonstrate that MYCBP2 has E2 requirements that are neither consistent with a HECT/RBR nor a RING mechanism. (b) Quantification of the different E2 mutant activities (n=2).

We would also like to discover the other E2 partners for MYCBP2 as many E3s have been found to work with several different E2 enzymes[499,500]. We carried out threonine discharge assays with a further 15 E2 conjugating enzymes. We observed the highly homologous UBE2D1 isoform had similar activity to UBE2D3. In addition, UBE2E1 (UbcH6) was the only other that was showed activity with MYCBP2cat (**Figure 5.11**).

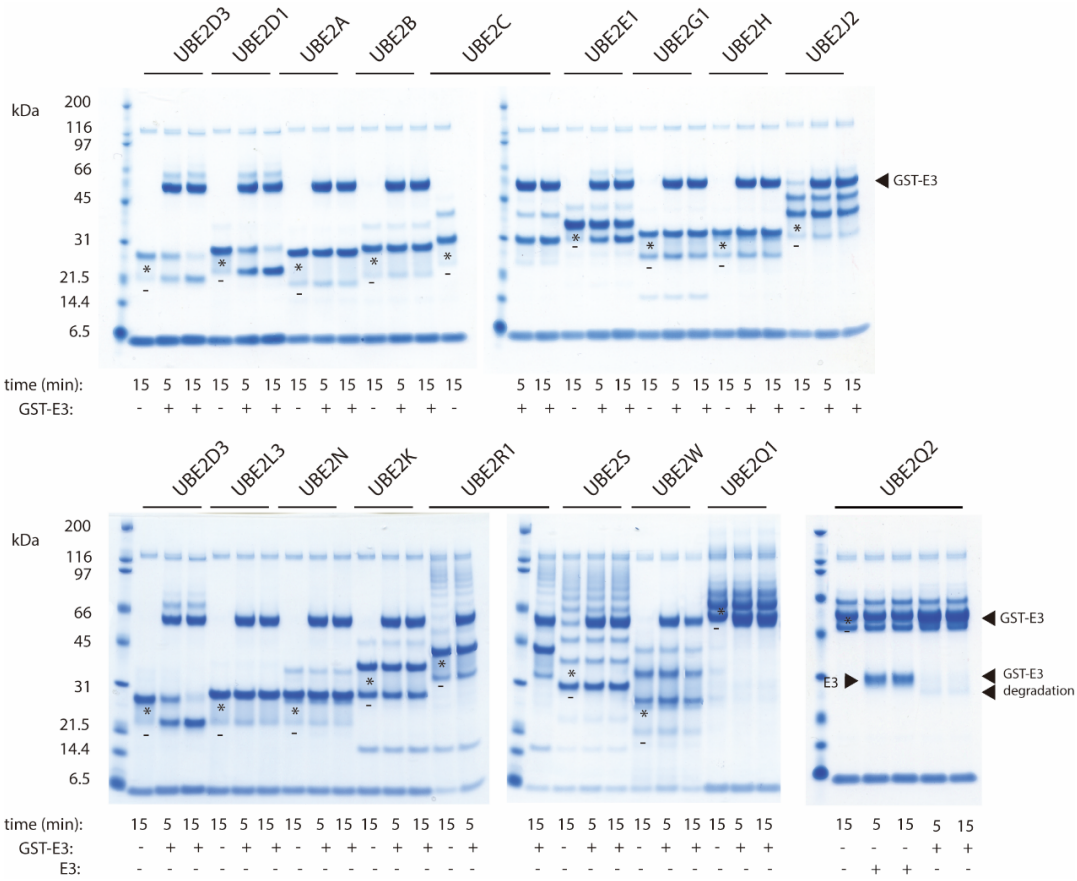


Figure 5.11 MYCBP2cat only works with UBE2D1, UBE2D3 and UBE2E1 enzymes.

17 E2s were tested for threonine discharge activity with GST-MYCBP2cat by Dr. Virdee. UBE2D1, UBE2D3 and UBE2E1 were the only E2s that demonstrated detectable activity. -, position of unmodified E2; *ubiquitin-charged E2. Unexpectedly, the HECT/RBR-specific E2 UBE2L3 shows negligible activity with MYCBP2. Certain E2s undergo E3-independent polyubiquitin chain formation and/or autoubiquitylation. In the presence of UBE2Q2, GST-MYCBP2 undergoes minor degradation resulting in the appearance of two lower molecular weight species. Consequently, we also carried out the assay with untagged MYCBP2 that did not undergo degradation; this produced similar results, showing that UBE2Q2 does not support MYCBP2 activity.

5.3 The crystal structure of MYCBPcat explains its distinct activity for Ub transfer

5.3.1 Structural characterization provides molecular basis of RCR E3 ligase activity

In order to understand the detailed molecular mechanism of the RING-Cys-relay machinery, we crystallized MYCBPcat (residues 4378-4640). We collected

synchrotron diffraction data to 1.75 Å. To clarify the contribution for the crystal-relevant works, the protein preparation and crystal-plate screening was performed by me. Dr. Virdee optimized crystals and Karim Rafie mounted crystals and collected synchrotron diffraction data. Daan van Aalten solved the MYCBP2cat crystal structure (PDB: 5O6C).

The 6 highly ordered Zn ions in the MYCBP2cat structure addressed the phasing issue during structure solving (data collection and refinement statistics shown in **Extended Table 5.1**). MYCBP2cat is composed of four distinct regions which was observed in a single MYCBP2cat molecule present in the asymmetric unit (**Figure 5.12a** and **Extended Figure 5.15**). The predicted cross-brace C3H2C3 RING domain was found at the N-terminal area spanning residues 4388-4441 (**Figure 5.12c**). A long α-helix (residues 4447-4474) follows the RING domain that leads into small helix-turn-helix motif (residues 4475-4500) (**Figure 5.12d**). A structurally unprecedented globular domain that binds four ions (residues 4501-4638) was found at the C-terminal region (**Figure 5.12b**). Dr. Virdee superposed known RING domains with the RING domain in MYCBP2 and observed high similarity among them (RING E3 RNF4[125] and RBR E3 HOIP[129]). Notably, MYCBP2cat resembles the extended architecture adopted by HOIP RING1 domain (**Figure 5.12c**). This region may play an important role for the distinct E2 requirement of MYCBP2cat. The following Zn-binding region adopts a unique unprecedented architecture which does not resemble any known protein folds (as determined using the EBI ePDB server). We termed this domain the TC (Tandem Cysteine) domain since the two catalytic cysteines, C4520 and C4572, reside at this region (**Figure 5.12b**). Notably, the predicted Zn binding residues C4506 and C4537 were found to have structural roles as they coordinate Zn ions. Mutations of the residues in earlier experiments must have caused a structural defect explaining

the poor probe labelling result (**Figure 5.12e-f** and **Figure 5.2c**). An unstructured region (4519-4526) was observed between the βA2 strand and helix 3_{10}A. The first proposed catalytic cysteine C4520, which undergoes transthiolation with E2~Ub, existed in this region. With the flanking residues around it, this region forms a mobile loop which we termed the mediator loop. Considering the high quality and high resolution of the crystal structure we obtained, we suspected that this loop is dynamic in nature. This dynamic nature could facilitate relay of the Ub from C4520 to C4572. The Zn coordination configuration (C5HC7HC2) of the TC domain is semi-contiguous and does not adopt cross-brace architecture (**Figure 5.12e**). Moreover, the residues E4534 (located at the end of the 3_{10}A helix) and H4583 (located in the αA1 helix) together with C4572 were be observed as an apparent catalytic triad. We referred this region as the esterification site (**Figure 5.12b**). Since histidine has been found to act as a general base to suppress the *pKa* of the sulfydryl group on cysteine[239,361], we suspected that H4583 could also play a similar role.

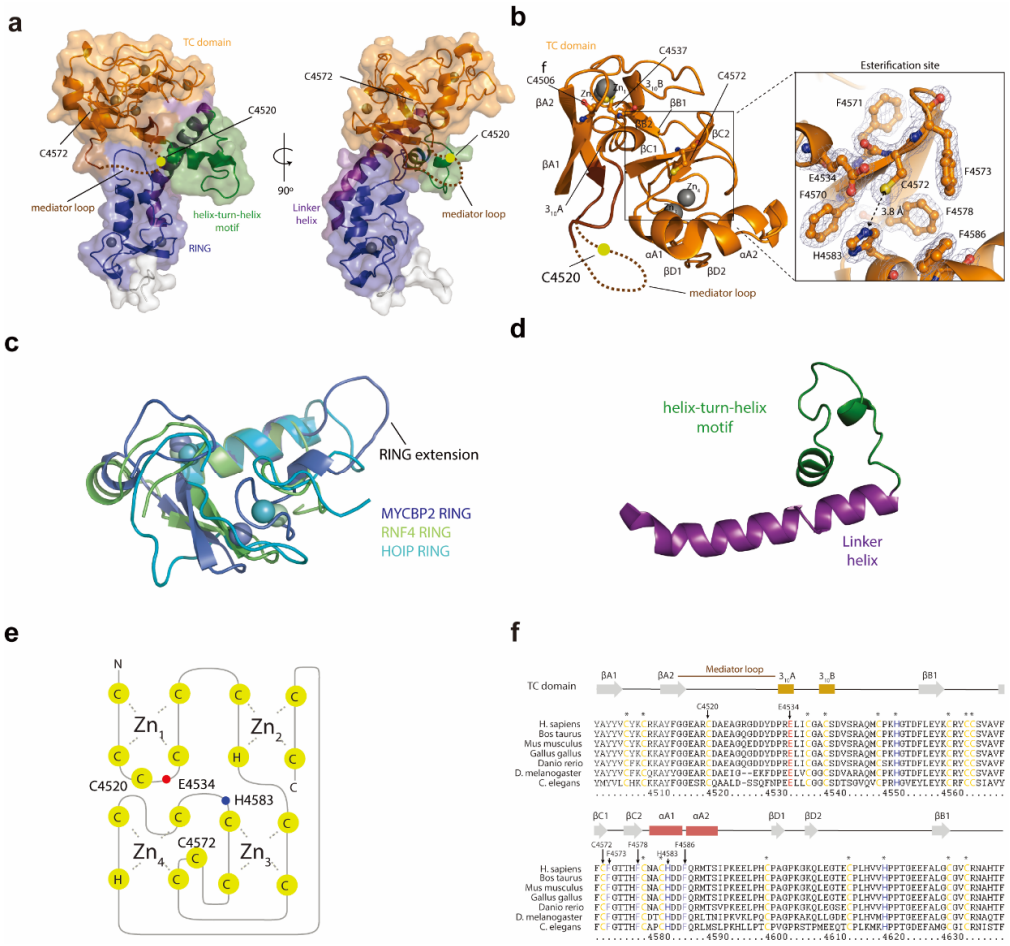


Figure 5.12 Molecular basis for the Ub relay mechanism utilized by the RCR E3 ligase.

(**a**) Surface and cartoon representation of the MYCBP2cat structure refined to 1.75 Å. C4520 begins Ub relay by undergoing transthiolation with E2~Ub and resides within a disordered region (residues 4519-4526) which form part of a region termed the mediator loop. (**b**) The tandem cysteine (TC) domain contains four short two-stranded β-sheets (βA1-βA2, βB1-697 βB2, βC1-βC2 and βD1-βD2), two short 310-helices (310A and 310B) and extensive loop regions. Inserted directly within the domain between the βA2 strand and helix 310A is the mediator loop which harbors the first catalytic cysteine residue, C4520. The second catalytic cysteine residue, C4572, is located within the esterification site at the end of the βC1 strand. The sulfhydryl group of C4572 is positioned between α1 and βC1 and is encircled by 5 Phe residues which flank C4572 (F4570, F4571, and F4573), emanate from the end of the βC1 strand (F4578), or emanate from the turn between αA1 and αA2 (F4586). Inset is a close up of the esterification site where C4572S forms an apparent catalytic triad-like arrangement with E4534 and H4583. (**c**) Superposition of MYCBP2 RING domain with the RING domain from the canonical RING E3 ligase RNF4[125], and from the RBR E3 ligase HOIP[129]. (**d**) The linker helix and helix-turn-helix motif that connect the RING domain to the Tandem Cysteine (TC) domain. (**e**) Diagram depicting Zn coordination network for the TC domain. Catalytic residues (numbered) (*nex page*)

are distributed throughout the TC polypeptide. (**f**) The Tandem Cysteine (TC) domain that confers threonine specificity is present in all MYCBP2 orthologues. All residues shown to be required for threonine esterification activity are conserved. Asterisks correspond to Zn-binding residues, grey arrows correspond to β-strands, gold rectangles correspond to 3_{10}-helices, and red cylinder corresponds to α-helix. (The figures are provided by Dr. Virdee)

5.3.2 Structural basis for threonine selectivity

The obtained MYCBP2cat crystal structure not only uncovers the unprecedented TC domain providing an insightful glimpse of the Ub-relay system, but also provide a hint to explain its preference for threonine. Crystal packing revealed that the N-terminus of a symmetry-related MYCBP2cat molecule was packed against the esterification site of the molecule in the asymmetric unit. A threonine in this terminus (T4380sym) forms a number of substrate-like interactions (**Figure 5.13a-b**). The β-hydroxy group of T4380sym, like the sulfhydryl group of C4572, forms an apparent triad with E4534 and H4583 (**Figure 5.13a**). However, we observed that the β-hydroxy group of T4380sym is closer to the imidazole nitrogen atom than the sulfur atom of C4572 (2.7 Å vs. 3.8 Å) (**Figure 5.13a**). It indicated that the H4583 can act as a general base to deprotonate the β-oxygen atom of T4380sym for the nucleophilic attack and the triad centred on T4380sym has more ideal geometry than C4572. Since the Ub relayed toward C4572 would form a thioester-linked intermediate, the Ub C-terminus is a productive electrophile for nucleophilic attack from threonine. Although the ubiquitin moiety is absent in our structure, the sulfur atom on C4572 is only 3.8 Å away form β-oxygen atom of T4380sym. This structural feature supports the specific esterification reactivity from MYCBP2cat. Moreover, the nearby residues (F4573, F4578 and F4586) are also found to form a sub-cluster which is proximal to the β-methyl group of T4380sym (**Figure 5.13b**). These three continuous Phe residues perform a well-defined hydrophobic pocket which the T4380sym β-methyl group docks into

(**Figure 5.13b**). The sub-cluster pocket stabilizes the threonine residue via a hydrophobic interaction with the methyl-group side chain on Thr. This proposed model may explain MYCBP2 high selectivity toward threonine rather than serine.

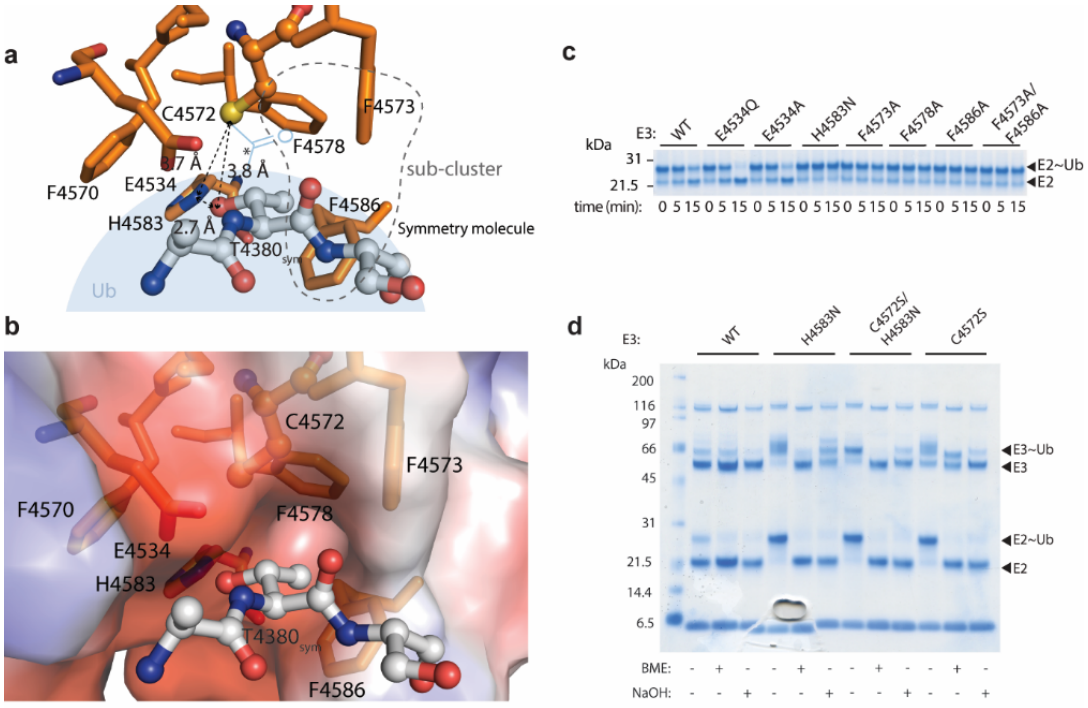


Figure 5.13 Structural basis for threonine selectivity.

(**a**) A threonine residue at the N-terminus of a symmetry related molecule (T4380sym) in an Ala-Thr-Ser sequence motif (grey ball and stick representation) is docked into the esterification site (orange stick). The β-oxygen atom of T4380sym is 2.7 Å from the imidazole nitrogen atom of H4583. The β-oxygen atom is also 3.8 Å away from the C4572 sulfur atom. In light blue is a schematic of Ub thioester-linked to C4572. The electrophilic center of the Ub thioester carbonyl carbon atom (*) is a single bond-length away from the C4572 sulfur atom. (**b**) Electrostatic potential of the esterification site (blue represents positive potential, red represents negative potential, and grey represents neutral potential). Residues F4573, F4578 and F4586 form a hydrophobic pocket which the Thr β-methyl group docks into. (*The structral figures are provided by Dr. Virdee*) (**c**) Residues deemed important for catalysis were mutated and tested in threonine discharge assays. (**d**) The H4583N mutant undergoes near quantitative Ub adduct formation. The adduct is largely removed after thiol treatment indicating that Ub is thioester-linked to the E3. The diffuse nature of the upper band might be due to the presence of a trapped thioester-linked Ub on C4520, and C4572, as the H4583N mutation prevents substrate deprotonation. In addition, S4572/H4583N double mutant does not form an stabilized ester adduct, rather increased formation of thiol-sensitive adduct.

In order to test the proposed model we observed from the crystal structure, I expressed a panel of MYCBP2cat mutants and compared their threonine activity with WT MYCBP2cat via the E2~Ub discharge assay. Consistent with the proposed role for H4583 to act as a general base to deprotonate threonine, the H4583N mutants abolished the MYCBP2cat threonine activity and ceased the ubiquitin transfer cascade (**Figure 5.13c**). As a consequence, the ubiquitin should stay upon two catalytic cysteines as the thioester adducts. The increased level of thiol-sensitive MYCBP2~Ub was observed with the H4583N mutant (**Figure 5.13d**). On the other hand, the C4572S/H4583N mutant only forms a discrete Ub adduct that is thioeser-linked. The H4583 residue is crucial for suppressing the pK_a of S4572 and thereby rendering it catalytic. In its absence, Ub remains trapped as a thioester-linked species on the upstream C4520 residue. To test the roles of the three Phe residues in the sub-cluster, the threonine activity of the individual Phe mutants was also tested. Strikingly, the F4573A, F4578A and F4586A mutants all abolished the MYCBP2cat threonine activity (**Figure 5.13c**). The role of E4534 was uncertain since the E4534 mutants did not decrease the MYCBP2cat activity. The findings show that E4534 residue is not part of a catalytic triad and establishing its precise role needs further study.

5.3.3 The proposed RCR E3 mechanism

Although the structure of MYCBP2cat was obtained and solved, how MYCBP2 binds E2 was unknown. However, the high similarity of the RING domain structures among MYCBP2cat and RING E3 and RBR E3, together with their largely conserved mode of binding to E2[123,125,126,159,501,502] allow us to perform the modelling of an E2-RCR E3 ligase complex. Dr.Virdee carried out the modeling via superposition of the MYCBP2cat RING domain with the RING1 domain from the RBR E3 HOIP bound

to the E2 complex structure (PDB ID: 5EDV)[129]. It allowed us to place the E2 with our currently solved MYCBP2cat structure. The modeling results showed that the E2 catalytic Cys residue falling immediately proximal to the mediator loop, which contains the catalytic C4520 residue (**Figure 5.14a-b**). Notably, it would appear that no significant conformational change is needed for MYCBP2cat to perform the E2-E3 transthiolation based on the structure model we obtained (**Figure 5.14a)**. The proposed dynamic mediator loop should be able to bring the C4520 in the close proximity of Cys85 in E2.

The missing structure of the mediator loop prevents us from conclusively explaining the Ub-relay mechanism. As a result, to simulate the conformation required for Ub relay from C4520 to C4572, Dr. Virdee and Dr. Mabbitt modelled the missing mediator loop residues with a GlyGly dipeptide thioester linked to C4520, representative of the Ub C-terminus that would be transferred via transthiolation with E2~Ub (**Figure 5.14b**). The carbonyl C atom of the Ub thioester on C4520 was observed in a position ~3.3 Å of the C4572 sulfhydryl sulfur atom. This modeling result supported the mechanism of the ubiquitin relay from C4520 to C4572 although the clashes were observed between mediator loop residues further C-terminal. The most notable were R4533, E4534, N4580, H4583 and D4584. However, these events could be largely relieved by rotations of their sidechains into available space (**Figure 5.14b**). In addition, the modeling result further supported the dynamic nature of the mediator loop since the large positional movement was required to bring the C4520 in close proximity with C4572. This may also explain the failure to trap a C4520S ester intermediate in my earlier experiments. Furthermore, the lack of a general base around the C4520 was also consistent with the result for the inability to transfer ubiquitin onward C4572S in C4572S/H4583N mutant (**Figure 5.13d**). We believe that catalytic

activity at the C4520 position is strictly dependent on the chemical properties of cysteine.

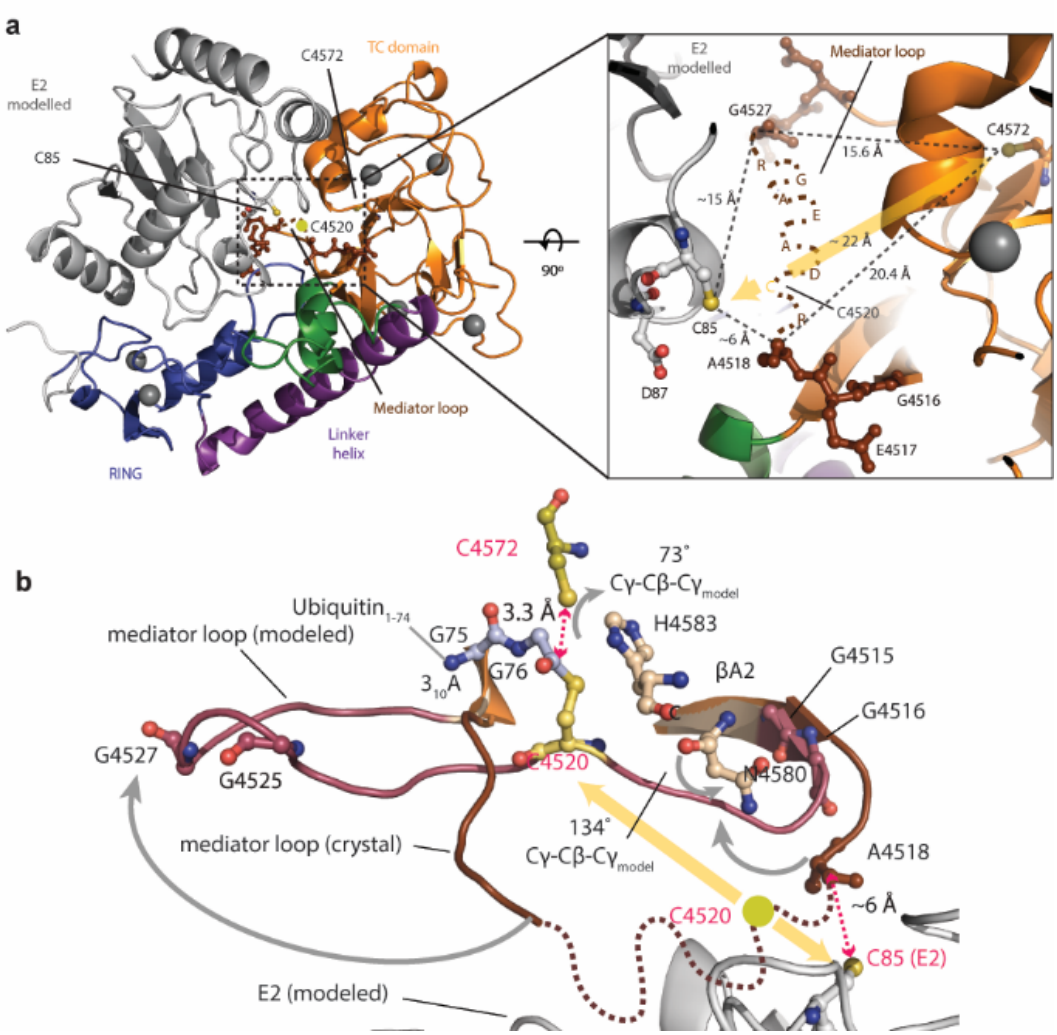


Figure 5.14 The modelling of an E2-MYBCP2cat intermediate and Ub relay.

(**a**) Superposition of the RING domain from the RBR E3 ligase HOIP in complex with E2 (Ub linked to E2 has been omitted due to direct clash with the TC domain; PDB ID: 5EDV[129]) allows modelling of the E2 into our structure (grey cartoon representation). The catalytic C85 residue in E2 (mutated in silico from Lys to Cys43) is proximal to C4520 which undergoes transthiolation with E2~Ub. The right-hand panel is a top-down close up of the mediator loop region. The eight missing residues that form the mediator loop are schematically shown in brown text. (**b**) Modeling of the Ub relay intermediate. For the experimental structure, TC domain residues are in orange and mediator loop residues are in dark brown. For the model, TC domain residues are in light orange and mediator loop residues are in mauve. The modeled E2, based on the superposition in (**a**), is in grey cartoon. Upstream (C4520) and downstream (C4572) catalytic residues are in yellow and colored by atom type. (Nex page)

Ub residues (G75-G76) are in blue ball and stick notation and are colored by atom type. Gly residues in the mediator loop that are likely to be important for loop mobility are displayed in mauve ball and stick and colored by atom type. N4570 and H4583 side chains have been rotated by the specified angles to relieve steric clash (**Extended Figure 5.16**). (*The figures are provided by Dr. Virdee*)

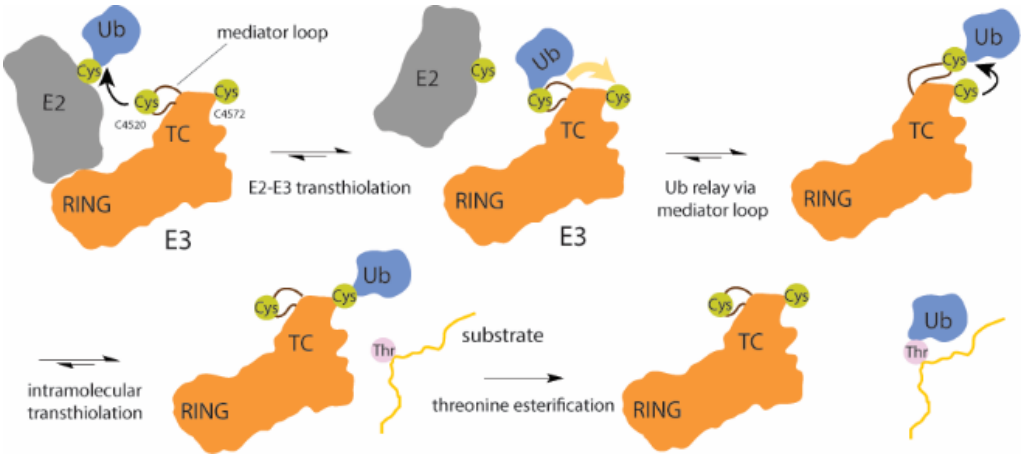


Figure 5.15 Schematic representation of the proposed model of the RCR E3 ligase mechanism.

Cartoon illustrating the proposed cysteine-relay and threonine esterification mechanism of RCR E3 ligases.

5.4 Conclusion & Discussion

Herein, I applied novel E3 activity-based probes to profile E3 activity in cell extracts (Activity-based protein profiling, ABPP). Combining two different E3 ABPs in this powerful platform, the system has been demonstrated to be able to label 80% of HECT and RBR E3 ligases. In addition, by profiling different types of cells in parallel (only SH-HY5Y data shown), we also identified many RING E3s or RING domain-containing proteins which are poorly characterized. MYCBP2 (PHR protein) was one of these proteins and yielded with the highest number of spectral counts in the mass spectrometry assay. Unexpectedly, we identified MYCBP2 as a novel subtype of E3 ligase which adopts a distinct catalytic cysteine relay mechanism (**Figure 5.15**) for ubiquitin transfer. It also displays different amino acids substrate preference with high selectivity for threonine esterification. I successfully

demonstrated the utility of our E3 ABPs by revealing a novel ubiquitin transfer mechanism by a poorly characterized E3, MYCBP2. This is note the first time the power of ABPs has been applied to discover new enzymes[288,290,302]. The discovery of this new E3 category not only expands the knowledge of the ubiquitin system but also provides a potential new target for drug development and the treatment of various neurological disorders.

5.4.1 What can E3 ABPP achieve?

In this thesis, I proved the concept that the E3 ABPs invented by the Virdee lab could be used in a comprehensive ABPP platform. The next question is how can I apply them for further research? There are several assays which are currently under investigation. Firstly, performing the E3s activity profiling among different biological models. The activity of serine hydrolases has been shown to display differences between tumor and non-tumor cells[368]. The systematic profiling of E3 activity among such models is still missing. Although the profiling platform has not yet been optimized for high throughput, the broad coverage of E3 (nearly all HECT/RBR E3s) labelling by our probes, combined with the multi-plex labelling tags (TMT-sixplex or TMT-tenplex), it is feasible to potentially compare up to 50 E3 activities across 10 different samples in a single LC-MS/MS run. For example, the cell lysates which treated with 10 different reagents individually can be reacted with biotin-probes and then the complexes are further be immobilized by streptavidin resins. After the on-beads trypin digestion, the trypsinized peptides can be labelled by TMT-tenplex tags. The 10 different tag-labelled samples are then pooled together which can further analyze by one single LC-MS/MS run. Various agents such as CCCP, Oligomycin A, Antimycin A, Rotenone, Valinomycin or deferiprone (DFP)[503] , have been shown to interrupt mitochondrial functions.

However, are these reagents all capable of inducing the activity of Parkin in mitochondrial quality control? Or, does the disruption of mitochondria regulate the activity of other E3s via an unknown signaling pathway? Our probes may be able to address these questions. Understanding the aging process is key for treating neurodegeneration diseases. Tau protein, β-amyloid and α-synuclein have been proven to be associated with Parkinson's disease[504,505]. The misregulation of the level of these proteins has been believed to be one of the key causal reasons for PD. However, the responsible E3s for these proteins or how the E3s regulate theses proteins are not fully understood. Systematic profiling of E3 activities for immature cells, mature cells and aging cells should help us understand more about this disease.

Kinases have been valuable drug targets for treating cancers for a long time[506,507]. In contrast, there is not a lot of research targeting ubiquitin cascade enzymes for drug development[508,509]. E3 ligases should be good targets for drug development, but the lack of tools to systematically profile the pattern for E3 activity among various cancer types, and other models of diseases, prevent the access to this research. Furthermore, the difficulties for obtaining good antibodies against the protein of interest also hinder studies. The sensitivity of our probes, which has been demonstrated as being able to detect endogenous Parkin from PD patients[322], should provide the inventory of E3s which are associated with tumorigenesis. For example, HECT E3 Nedd4L has been shown to mediate prostate cancer[510], lung cancer[511] and melanoma[512]. RBR E3 Parkin has also been found to be associated with glioma[513], colorectal cancer[514], ovarian cancer and breast cancer[515]. In addition, competitive ABPP can be performed for inhibitor screening once we know the target E3s. An alternative strategy is to develop an E3-focused inhibitor library and use this for phenotype-screening within the tumor models

and then perform ABPP. This should uncover the potential targets for the chosen inhibitors and also reveal the possible off-target proteins. One of the biggest advantages for using the ABPP is that there should be no bias on the protein hits. Discovering the inhibitors for E3s is not the only thing which ABPs can achieve. By applying fluorescent-E3 ABPs and the competitive ABPP strategy, quantitative inhibitory activity toward the target E3 by the selected inhibitors should be obtained by dose-response experiments.

In addition to the application in drug screening, the discovery of potential E3 pairs for the E2s can also be studied by our E3 ABPs. There are nearly 700 E3s in the human genome but only 40 E2s have been identified to date. As we discussed in the *chapter 1*, it is possible for one E2 to work with several E3s or one E3 to work with several E2s. Since the E3 ABPs are composed of an E2s backbone that serves as a recognition element, the application of various E3 ABPs, built on different E2 backbones, should uncover the unappreciated E2-E3 pairs.

By performing the ABPP with different proteomes, our E3 ABPs unexpectedly labelled multiple RING E3s as identified by mass spectrometry analysis. MYCBP2 is the only one of these RING E3s we have studied to date. Is there another RING E3 that adopts the same ubiquitin relay mechanism as MYBCP2? Or do these poorly characterized RING E3s belong undiscovered classed of RING-linked E3? More comprehensive studies on these E3s need to be performed since many E3s that contain a poorly characterized cysteine-rich region after their conserved RING domain.

Multiple Ubl systems also use a homologous E2-E3 cascade and use transthiolation for transferring Ubl protein toward substrates[516], but little is known about their function or even which enzymes are responsible for these events. Since the delicate methodology for making the E3 ABP allows us to replace not only back

bone E2 but also the ubiquitin, the various Ubl E3 ABPs should be able to build. Notably, there are only few SUMO, NEDD8, and ISG15 E3s discovered; only one E3 for ATG8 and UFM1 and no E3 has been identified to date for FAT10, URM1, and ATG12[516]. Combined with the ABPP strategy, the Ubl E3 ABPs which based on Ubl¬E2s, should be able to identify the Ubl E3s which are currently missing to date. Together with the Ub E3 ABPs, we should be able to draw an elegant map for the Ub/Ubl system.

5.4.2 The limitation of E3 ABPs

ABPP is definitely a powerful platform for protein activity study, but certain issues are must be currently considered when interpreting data. First, the concentration of the probe we use in the assay. How do we determine the 'correct' concentration for the target proteomes is always a dilemma since the level of proteins vary among different samples. If the concentration was too low, we may not be able to detect differences in protein activity. For example, if there were only 100 probe molecules in the reactions but the activated protein molecules went from 100 to 200 after stimuli treatment, the activity of this protein would remain the same from the data interpretation. Furthermore, the affinity of the protein-based ABPs toward their target protein, changes in which might be a regulatory mechanism, usually affect their reactivity, the ABPs may not be able to label the target proteins at the low concentration. However, supplying too much probe in the reaction may also cause the off-target labelling since it may produce a physiologically irrelevant environment for the ABPs and target proteins. Moreover, the presence of too much protein-based ABPs in the samples for mass spectrometry analysis may also generate the high abundant peaks during the assay which may mask the signal from low abundance proteins.

The other obvious drawback for our E3 ABPs is their inability to interact with the intact cells. The E2-Ub backbone is certainly a benefit for the selectivity but the size of the protein also decreases the cell permeability of E3 ABPs. It has already been proven that the ABPs can distinguish differences in protein activity in cell extracts[368,410,439]. However, it could also be argued that the local E3 concentration in the intact cells would be dramatically different than in the cell extracts. Apparently, the relative concentration between ABPs and target protein would affect the labelling efficiency which may mislead the research focus. Furthermore, whole cell extracts may also increase the relative concentration of DUBs, kinases, phosphatases and proteases toward the target E3s which may eliminate a regulatory PTM and inactivate the E3s of interest. As a result, optimizing the cell permeability of our E3 ABPs should be addressed in the near future. For example, Mulder *et al.* reported the introduction of Cy5–Ub-Dha into HeLa cells via electroporation[439]. The electroporation method should be tested with our E3 ABPs. In addition, by conjugating the cell-penetrating peptides (CPPs) upon our probes[517], for example, *trans*-activator of transcription (TAT) protein peptide of HIV-1, might be able to deliver our probes into the intact cells[518]. The only concern of this strategy is whether the size and properties of the additional peptide sequences would affect the probe labeling or not.

5.4.3 Why the ubiquitin relay?

The newly discovered cysteine relay mechanism for ubiquitin transfer indeed expands the knowledge of the field. It is not immediately clear why this mechanism has evolved. In fact, MYCBP2 is not the only protein which regulates its activity by two cysteines wihtin the same protein. The two cysteine residues with catalytic properties, Cys46 and Cys165, on peroxidatic AhpC protein regulate the protein

activity by forming a disulfide bond[519]. However, individual mutation of cysteine in AhpC protein, only C165S abolished its activity. Apparently, the cysteine relay mechanism is not the case for this protein. We did not observe the disulfide bond formation in the MYCBP2 crystal structure. Furthermore, the transition of the ubiquitin from the first catalytic cysteine to the second one is extremely fast since I did not observe any doubly-loaded MYCBP2cat, not using wild type or the C4572S mutant (**Figure 5.4**). The only doubly-loaded MYCBP2cat was observed with the H4583N mutant where the ubiquitin was unable to be transferred to threonine due to the lack of a general base H4583 (**Figure 5.13**). These results indicated that the intramolecular transthiolation of MYBCP2 is the process which provides a facile means to shuttle Ub throughout the ubiquitin system.

The other explanation for the cysteine relay mechanism is the esterification activity of MYBCP2. The special structural features of the esterification site and the TC domain, which are distinct in MYCBP2 differing greatly from conventional E3s that undergo transthiolation with E2~Ub. In order to allow Ub transthiolation with E2 *and* esterification ubiquitin towards substrates, we propose the necessity of the cysteine relay mechanism. As we discussed above, in order to keep the esterification reactivity, the C4572 residue and Phe sub-cluster together adopt a distinct esterification site (**Figure 5.12b**). I speculate that the rigid and highly conserved E2 ubiquitin conjugating domain (Ubc)[520] is not able to perform the E2-E3 transthiolation reaction at this site for ubiquitin transfer. Therefore, a region which is capable of accepting ubiquitin from E2 via the conserved E2-E3 transthiolation is needed. According to our modeling result (**Figure 5.14a**), the first catalytic cysteine C4520 can feasibly accept the ubiquitin from E2 via the contact between E2 and the RING domain in MYCBP2. The biochemical experimental data also support this hypothesis. Apparently, the remaining question is why the

ubiquitin is transferred to the non-lysine residue[521].

From our collaboration with Dr. Hofmann, we observed two other cysteine-rich domains N-terminal to the RCR domain. These two domains may adopt Zn-binding regions which are currently under investigation. MYCBP2 is a huge protein which also contains multiple additional domains, for example, the RCC-1 domain and Myc protein binding domain[478]. Although the individual function of these domains have been demonstrated already[478], how these domains might affect E3 activity of MYCBP2 is still unknown. It has been shown that MYCBP2 forms a complex with Fbxo45/Skp1 for substrate ubiquitylation[478,492,493]. Since we discovered the ability for MYCBP2 to transfer the ubiquitin to the substrates directly via its catalytic cysteine, it raised the question that who is the E3 ligase machinery in the MYCBP2/Fbxo45/Skp1 complex responsible for ubiquitin transfer? In addition, is the ubiquitylation event from the MYCBP2/Fbxo45/Skp1 complex also non-lysine ubiquitylation? Based on the data from biochemistry experiments, the mono-Ub adducts are the only product from MYCBP2cat E3 ligase. I propose that the complex formation may be essential for physiological substrate ubiquitylation.

5.4.4 Non-lysine ubiquitylation

The ubiquitin esterification reactivity of MYCBP2 is not the first example of non-lysine ubiquitylation. It has already been reported by different groups for cysteine, serine and threonine[521,522]. The early evidence for non-lysine ubiquitylation was discovered with viral E3 ligases MARCH (membrane-associated RING-CH) [523]. One of these MARCH E3, the viral ligase kK3, has been found to ubiquitinate cysteine, a residue on the major histocompatibility complex (MHC) class I. An engineered Lys-less cytoplasmic tail of MHCI heavy chain (HC) was unexpectedly ubiquitylated

and downregulated by kK3 E3[524,525]. Although the ubiquitin-thioester linkage has already been demonstrated by E2~Ub conjugates, the kK3-mediated substrate cysteine ubiquitylation was still the first evidence for non-lysine ubiquitylation in a substrate protein. Furthermore, the homologue of kK3 E3 in murine virus, mK3, has later been demonstrated to have ubiquitin esterification reactivity towards serine and threonine in substrate proteins[498]. Furthermore, the serine residue on theses substrates has been shown to be the primary residue for ubiquitylation even with lysine in proximity. It further indicated that non-lysine ubiquitylation is not simply a compensation event for the shortage of lysine residue on substrate proteins[498].

The mono-ubiquitylation on the conserved Cys residue upon Pex5p has been found to be a key regulators for releasing Pex5p from the peroxisome membrane[119]. The Pex4p and Pex10 have also been found as the E2 and E3 enzymes for this Cys-mono-ubiquitylation events[119,526]. Surprisingly, Platta *et al.* observed that two conserved lysine residues in Pex5p were ubiquitylated by E2 Ubc4 working with the E3 Pex2p when the releasing cascade was blocked. [527]. In addition, the cysteine residue in the proneural transcription factor, Ngn2, has also been demonstrated to be ubiquitylated as well[528]. In mammalian cells, the N-terminal region (tBid-N) of the Bid protein has been found to be ubiquitylated for freeing the C-terminal proapoptotic Bcl-2 homology 3 (BH3) domain on Bid protein when sensing the apoptotic signal[529]. However, no lysine residue has been identified on tBid-N domain. In 2007, Tait *et al.* demonstrated that the multiple mutations on Cys, Ser and Thr residues in tBid-N abolish the ubiquitylation of this domain[530]. The tail of the T-cell antigen receptor α-chain (TCRα) has also been shown to undergo non-lysine ubiquitylation. The mutation on the two evolutionarily conserved Ser residues to Arg massively reduces the ubiquitylation and the rate of degradation

of TCRα[531]. Interestingly, the mutation of these two Ser residues to either Cys, Thr, or Lys had no effect on the ubiquitylation level. NS-1 immunoglobulin (NS-1 lg) has also been found to be ubiquitylated via Ser and Thr residues by HRD1 E3 (a major ERAD-associated E3 ligase in yeast and mammals)[532]. Unexpectedly, HRD1 E3 was found to be able to ubiquitylate the Lys residue on NS-1 when all Ser and Thr residues were mutated[533].

The above examples indicated that non-lysine ubiquitylation is not just an 'exception' to the canonical ubiquitylation events but also an important regulation mechanism. However, most of these E3s perform the ubiquitylation by a RING E3 mechanism implying that the E2 is the enzymes which mediate ubiquitin esterification[522]. As a result, MYCBP2 is the first discovered human E3 which displays autonomous non-lysine ubiquitylation activity. Furthermore, from the bioinformatics analysis, the RCR machinery which gives E3 activity to MYCBP2 has been well-conserved in higher eukaryotes (from *C. elegans*, *Drosophila* to mouse and human).

In conclusion, several features of non-lysine ubiquitylation have been discovered[521,522]. Although MYBCP2 shows negligible activity towards free lysine amino acids, other non-lysine ubiquitylation examples indicated the possibility for the coexistence of the non-lysine and lysine ubiquitylation events. This intrinsic activity toward Lys, Cys, Ser, and Thr makes the identification of the specific ubiquitylation site extremely difficult. Interestingly, apart from MYBCP2, the non-lysine ubiquitylation activity is not determined by a single E2 or E3. mK3 E3 has been found to collaborate with the E2 Ube2j2 for both Lys and Ser/Thr ubiquitylation. It means that the Ube2j2 is not the E2 specific for Ser or Thr residue. In other words, mK3 E3 should play a role in determining the substrate preference of Ube2j2. It has been shown that the context of the substrate sequence is not the

determinant for the ubiquitin conjugation[534]. It is likely the three-dimensional location of the targeting residues determines the outcome of ubiquitylation[534]. The targeting residues have to be in close proximity of the E2 active site which is positioned by a paired E3[154,155]. Although the proximity between E2 and substrates may play an important role in selecting the substrate amino acid residues, it is not the only factor to affect the reaction. It has been found that the mK3/Ube2j2 can enhance Lys or Thr/Ser ubiquitylation in both cases, however, mK3/Ube2d1 can only target Lys residues in the substrate[122]. Taken together, these proposed mechanisms further implied the importance of the cysteine relay mechanism and the structural distinct esterification site for MYCBP2 threonine esterification. The cysteine relay mechanism can make sure that the ubiquitin is correctly being transferred to the MYCBP2 C4572 residue which is the catalytic cysteine responsible for stringent esterification (**Figure 5.12b** and **5.13a**).

5.4.5 What is the role for threonine ubiquitylation?

To date, non-lysine ubiquitylation has only been discovered on the substrate residues but not upon ubiquitin itself[122]. It is consistent with the biochemical data we obtained from MYCBP2 assays. We did not observe any free-ubiquitin chain formation even under hydroxyl-group free conditions. It is not clear why the non-lysine polyubiquitin chains have not been found yet. The isopeptide (amide) bond on the ubiquitin conjugates has been demonstrated as the most stable linkage compared to oxyester or thioester linkages[521,535]. The instability of these non-lysine linked ubiquitin conjugates may provide the explanation for the missing non-lysine DUBs. Validation of the redox environment of cells and the change of the Ser/Thr *pKa* all affect the linkage stability of these Ub conjugates. It is plausible for the deubiquitylation occurs via a non-enzymatic pathway. In other words, it also

provides a quick response mechanism for stabilizing non-lysine ubiquitin conjugates. For example, MYCBP2 has been demonstrated to mediate the internalization of transient receptor potential vanilloid receptor 1 (TRPV1)[494,536]. The loss of MYCBP2 caused the constitutive p38 MAPK activation and prevented the internalization of TRPV1. Furthermore, p38 MAPK has been shown to play an important role in the sensibilization procedures of nociceptive neurons (sensory neurons for damage respons)[537]. It is possible that the MYCBP2-dependant threonine ubiquitylation is also involved in the fast stimuli-induced responses of nociceptive neurons. MAPK kinase kinase 12 (MAP3K12, also known as dual leucine zipper bearing kinase, DLK) has been indicated as the MYCBP2 substrate for regulating this p38 MAPK-TRPV1 pathway[538].

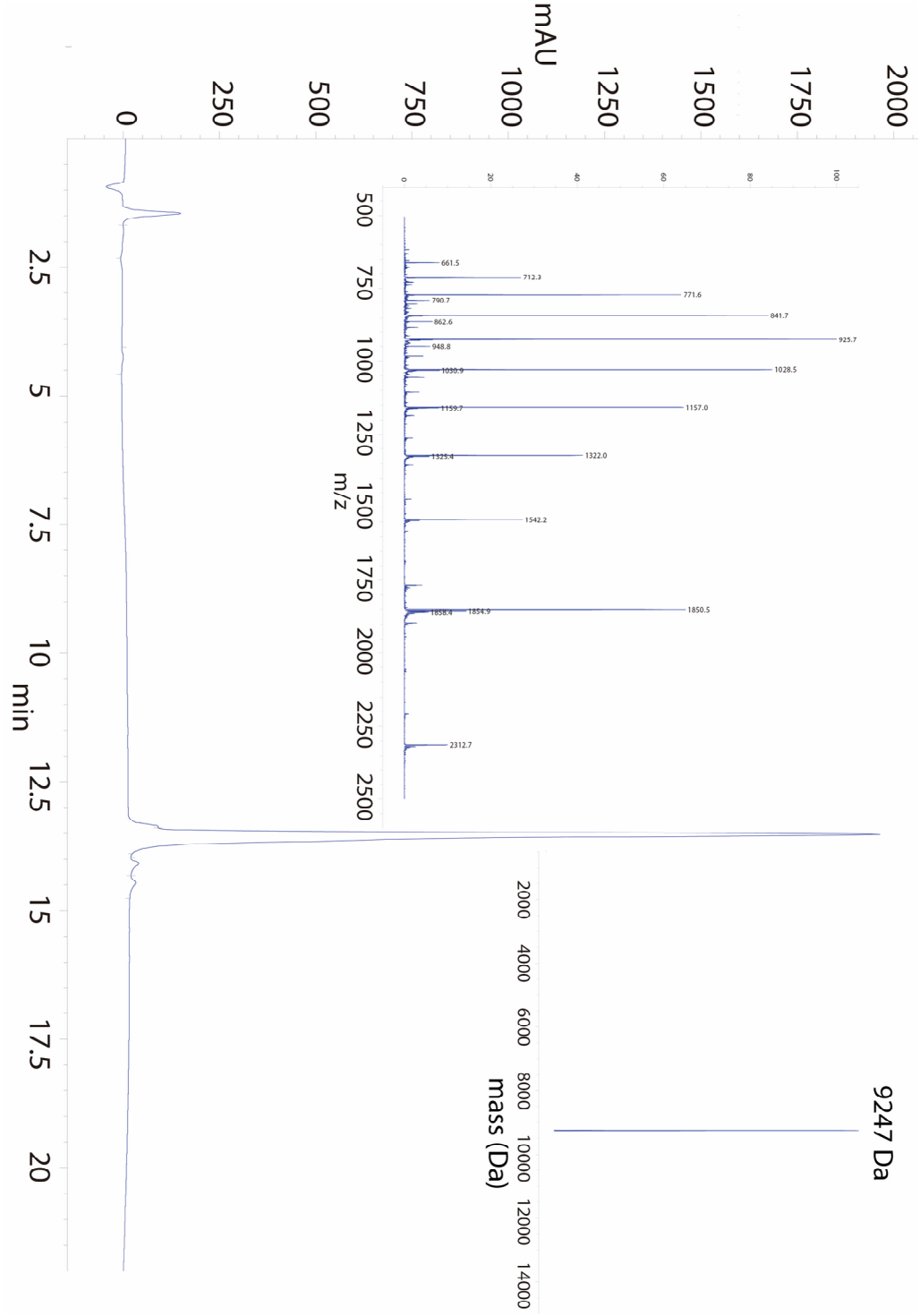
The actual physiological role for the MYCBP2-meidated ubiquitin esterification is still unknown. It is plausible that the MYCBP2-mediated Thr-ubiquitylation performs as a competitor with Thr-phosphorylation, considering the high selectivity of MYCBP2 toward the threonine amino acid. Furthermore, the reported substrate candidates for MYCBP2, for example, DLK1[539] and TSC2[540], have also been shown to be phosphorylated by kinases. Whether MYCBP2 can regulate the activity of these proteins via ubiquitylation requires further study[541].

It has proven challenging to study non-lysine ubiquitylation due to instability of the oxy-ester or thio-ester linkage[521,522]. The supplement of thiol-containing agents in typical biochemistry experiments simply eliminates the observation of cysteine-based ubiquitylation. The combination of heating and reducing environment may also break the oxyester-linked ubiquitin conjugates. Mass spectrometry analysis has been frequently used in discovering ubiquitinated proteins[542,543]. However, the frequently used ‘trypsin digestion’ step has to be performed under basic conditions for long incubation which may hinder the

detection of the oxyester-linked diGlyGly peptides[544]. Although the oxyester-bond should be reasonable stable under pH 8.0 buffer condition, the appropriate buffer condition for performing oxyester-linked diGlyGly peptides digestion should be further studied in very near future. Furthermore, we cannot rule out the possibility that the ubiquitin esterification is independent of the proteasome degradation machinery. It makes proteasome inhibitor treatment, which is frequently applied for stabilizing ubiquitin-conjugates, unable to trap the ester-linked ubiquitin conjugates.

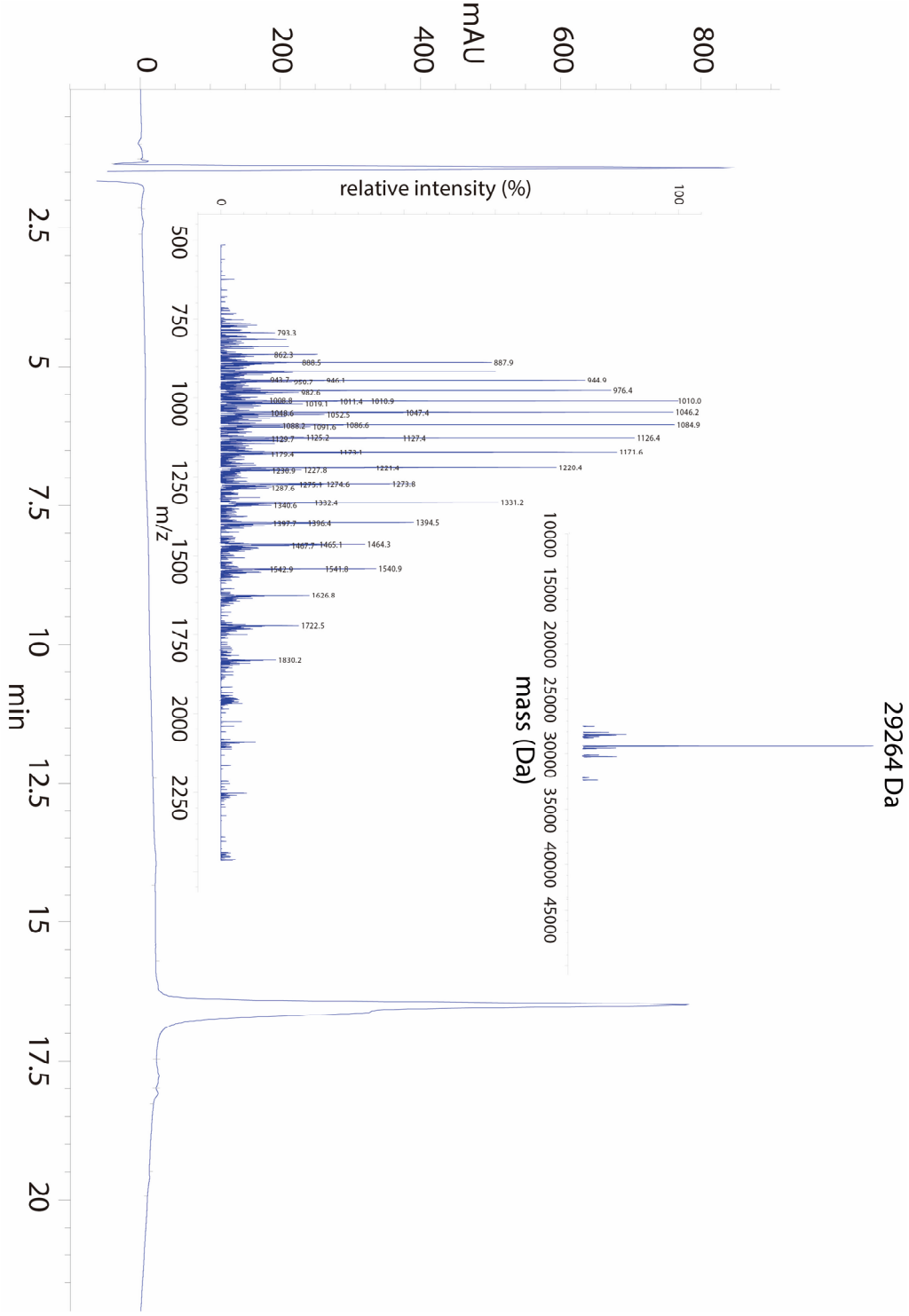
To date, ubiquitylation has only been discovered in a polypeptide context. Since the high esterification reactivity of MYBCP2 has been demonstrated in this thesis, the possibility of lipid or carbohydrate ubiquitylation exist. Furthermore, the ubiquitylation on lipoproteins[545] or *N*-glycosylated proteins[546] have been demonstrated already. It is plausible that we may have missed some critical data due to limitations of the detection method. The development of a delicate method to study the non-lysine ubiquitylation is therefore urgently needed.

5.5 Extended Figures



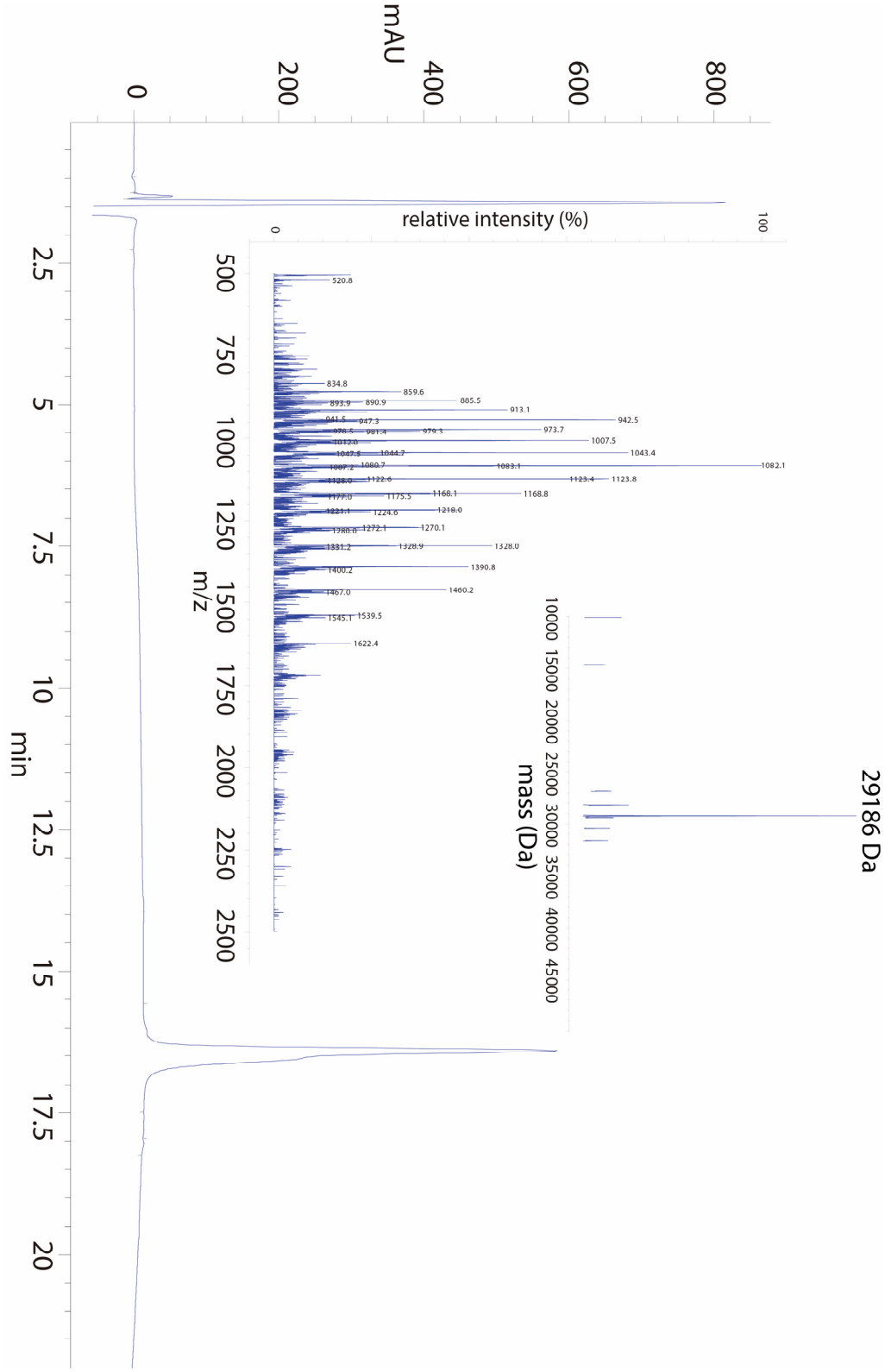
Extended Figure 5.1 LC-MS characterization of biotin-Ub$_{1-74}$-SR.

Biotin-Ub$_{1-74}$-SR. HPLC chromatogram monitoring UV absorbance at 214 nm. ESI mass spectrum (inset left) and deconvoluted mass spectrum (inset right). Expected mass = 9249.2 (-Met); found = 9,247 Da.



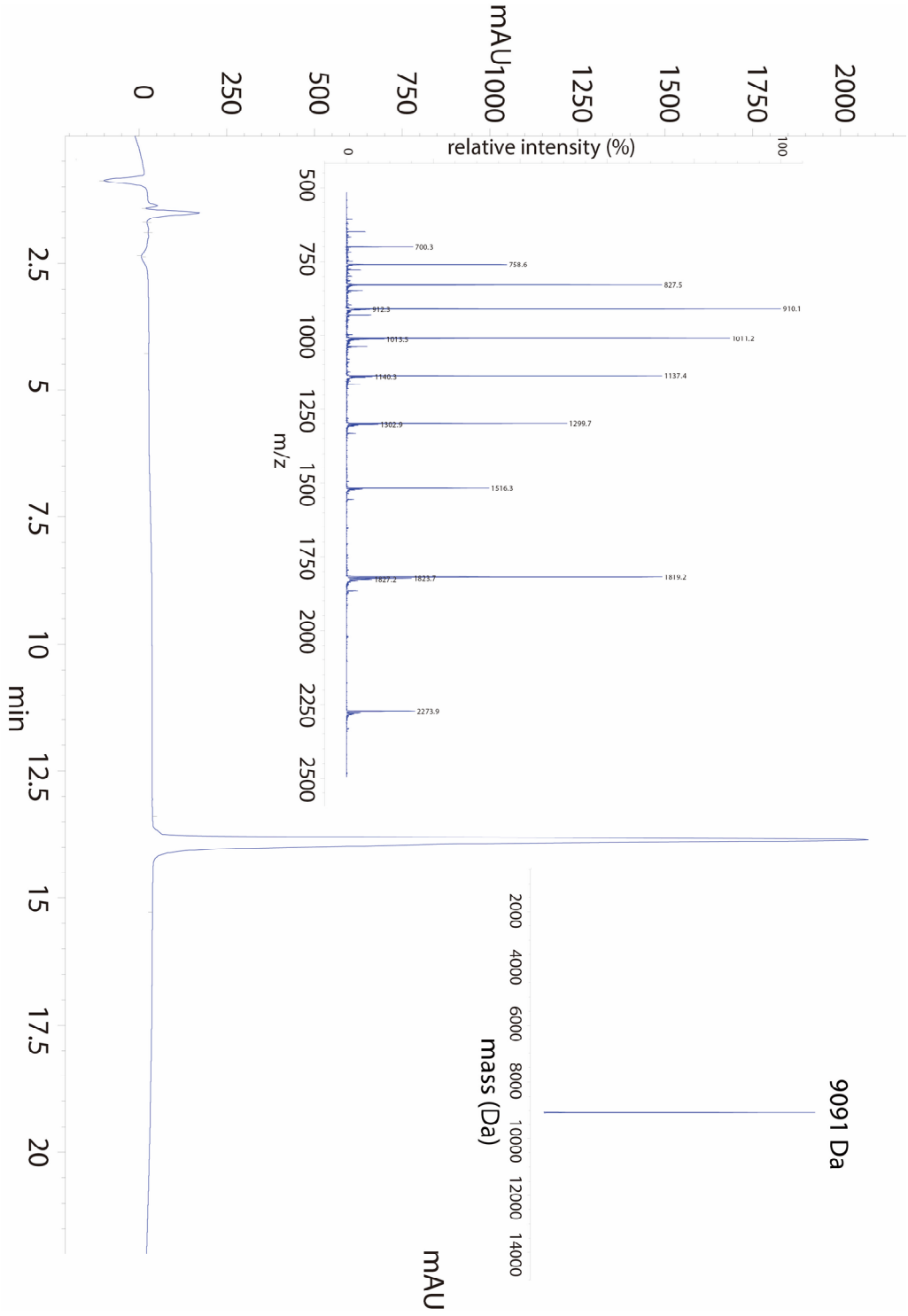
Extended Figure 5.2 LC-MS characterization of biotin-ABP 13.

UBE2D2 ABP **13**. HPLC chromatogram monitoring UV absorbance at 214 nm. ESI mass spectrum (inset left) and deconvoluted mass spectrum (inset right). Expected mass = 29,268 Da (-Met); found = 29,264 Da



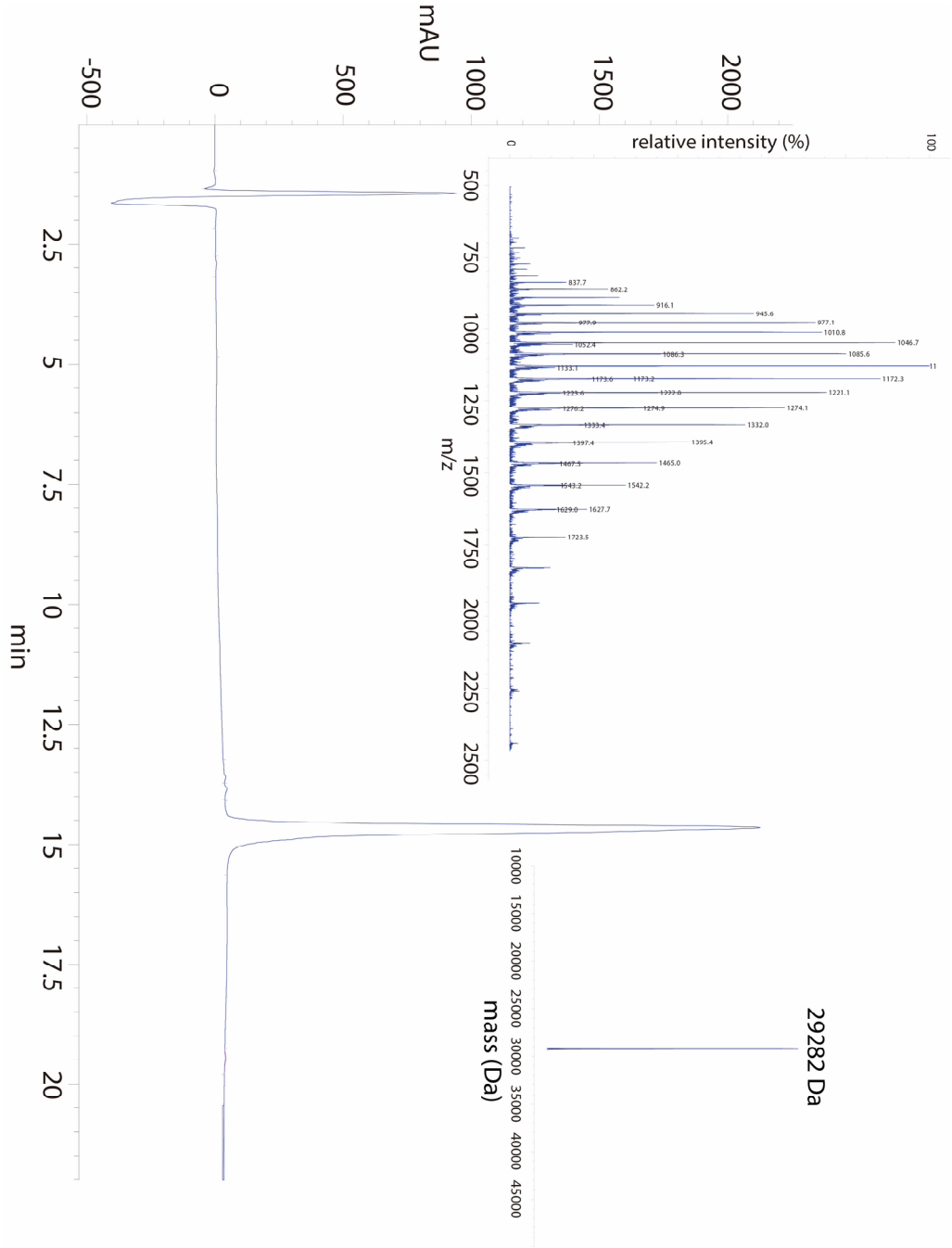
Extended Figure 5.3 LC-MS characterization of biotin-ABP 14.

UBE2D2 F62A ABP **14.** HPLC chromatogram monitoring UV absorbance at 214 nm. ESI mass spectrum (inset left) and deconvoluted mass spectrum (inset right). Expected mass = 29,192 Da (-Met); found = 29,186 Da.



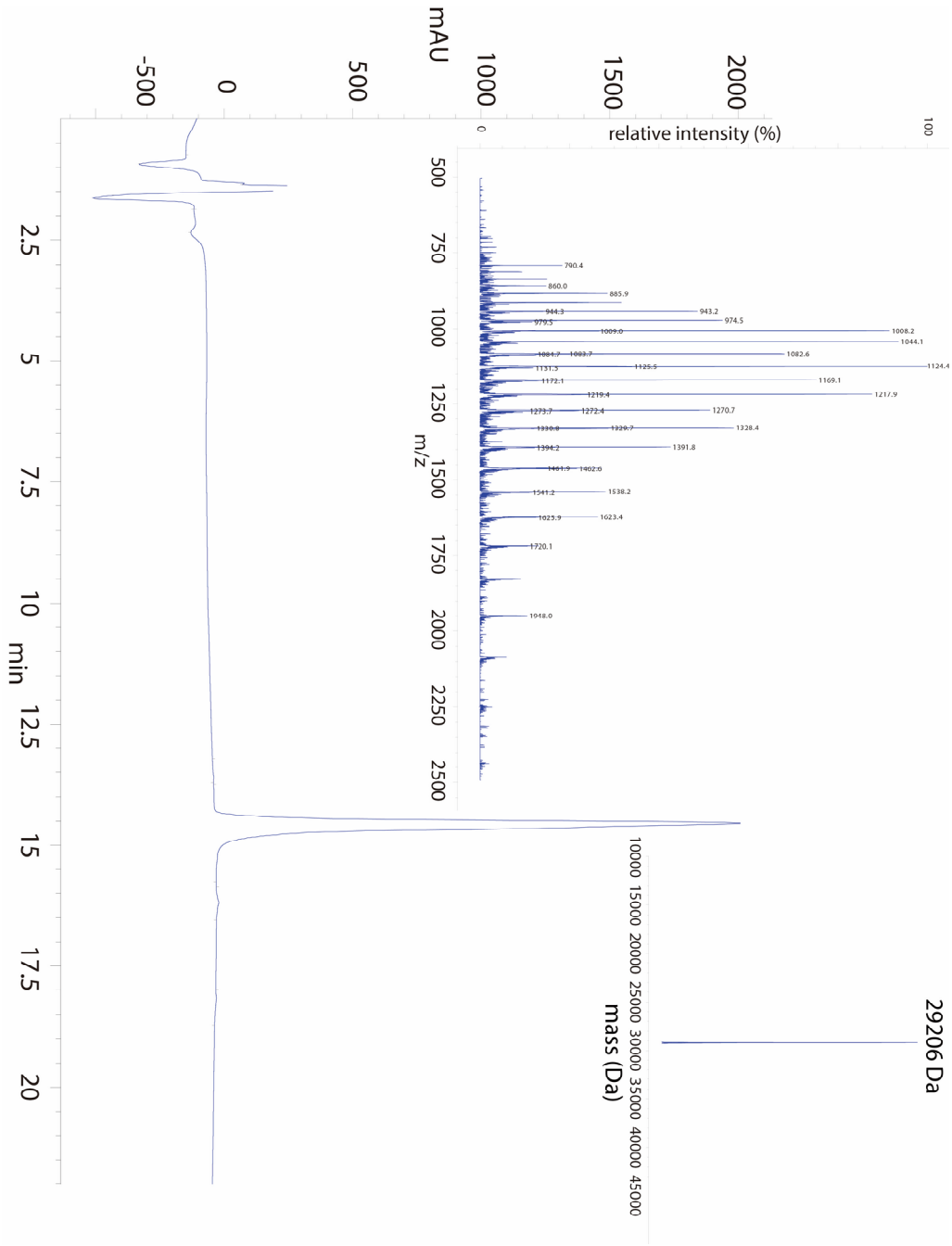
Extended Figure 5.4 LC-MS characterization of biotin-Ub$_{1-73}$-SR.

Biotin-Ub$_{1-73}$-SR. HPLC chromatogram monitoring UV absorbance at 214 nm. ESI mass spectrum (inset left) and deconvoluted mass spectrum (inset right). Expected mass = 9093 (-Met); found = 9091 Da.



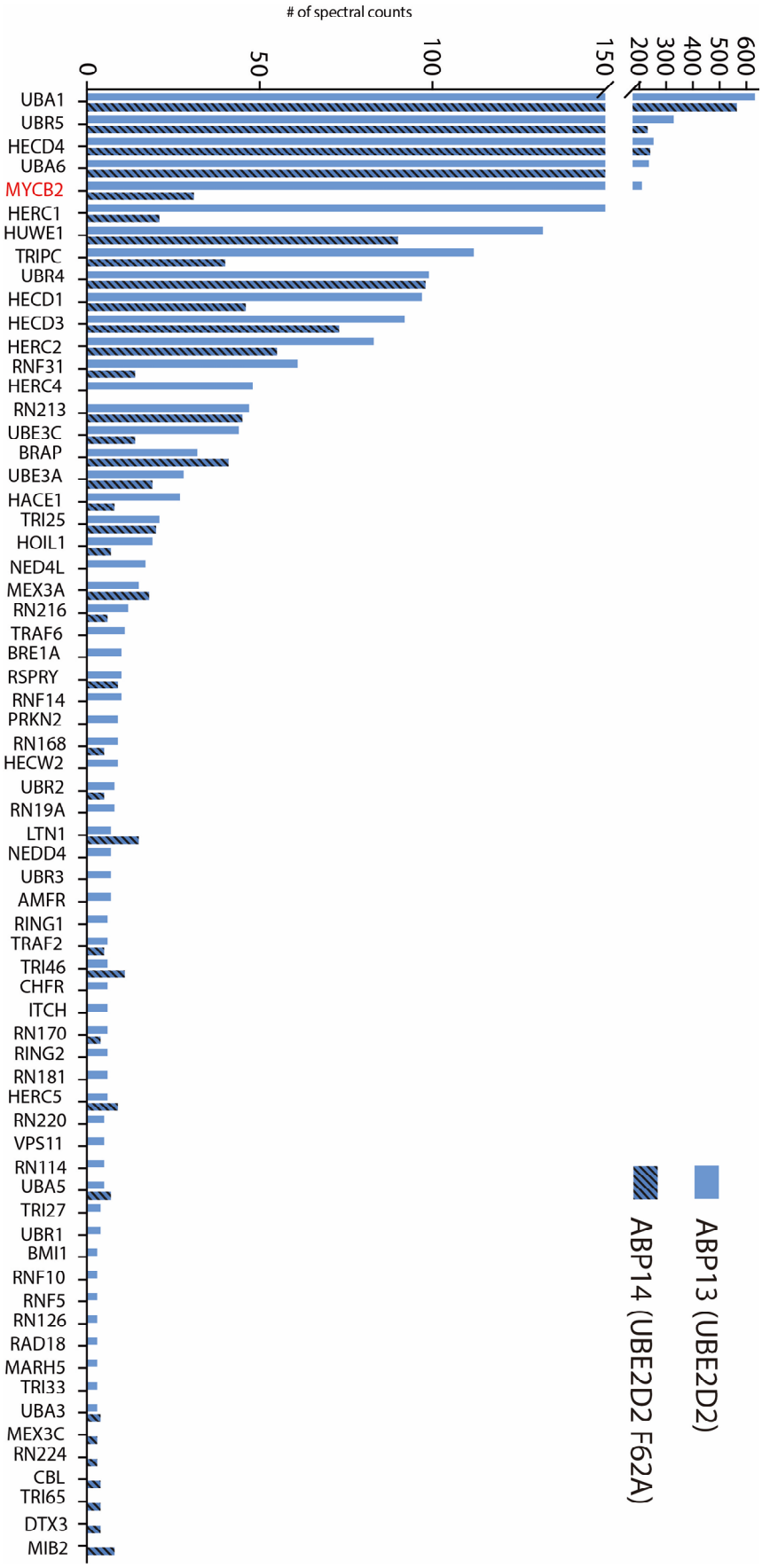
Extended Figure 5.5 LC-MS characterization of biotin ABP 15.

UBE2L3 ABP **15**. HPLC chromatogram monitoring UV absorbance at 214 nm. ESI mass spectrum (inset left) and deconvoluted mass spectrum (inset right). Expected mass = 29286.9 (-Met); found = 29282 Da.



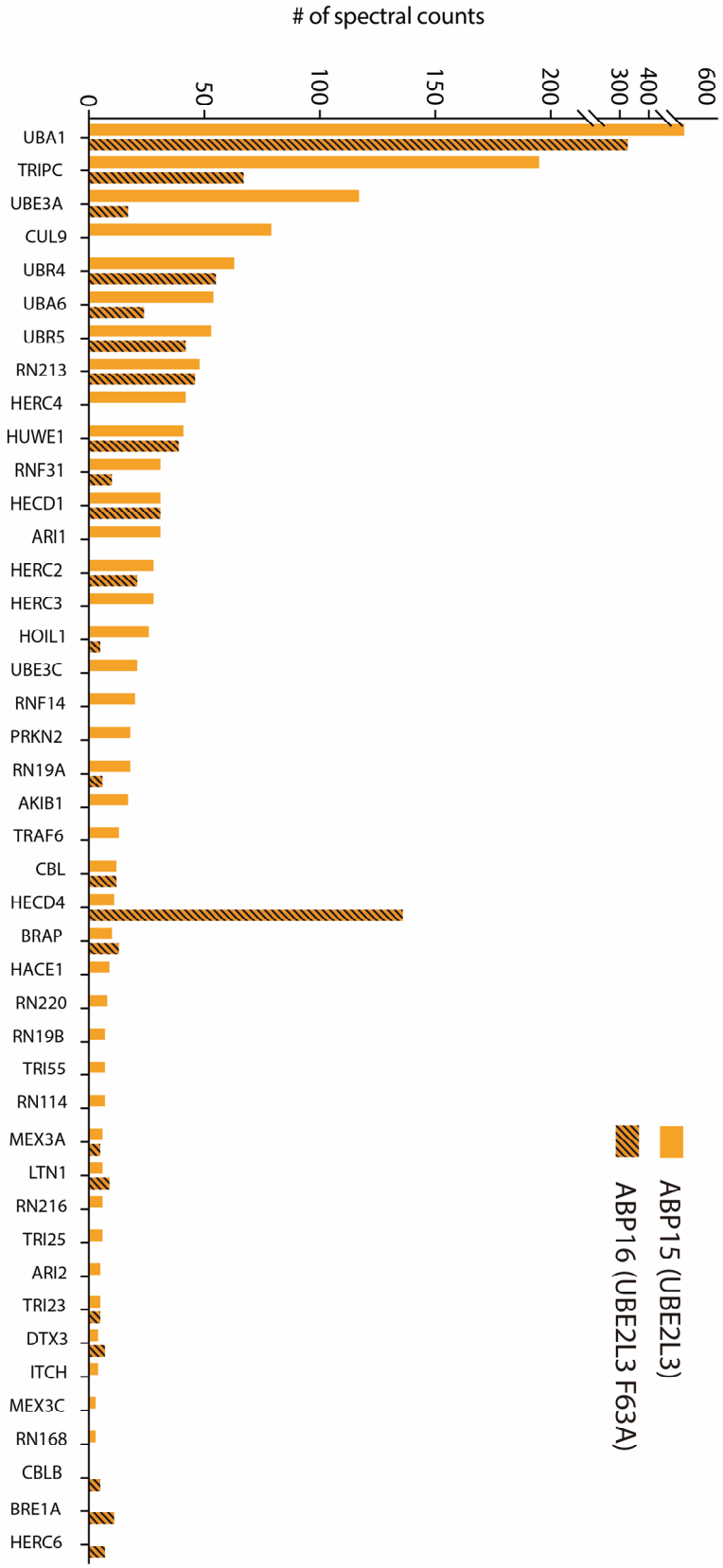
Extended Figure 5.6 LC-MS characterization of biotin ABP 16.

UBE2L3 F63A ABP **16**. HPLC chromatogram monitoring UV absorbance at 214 nm. ESI mass spectrum (inset left) and deconvoluted mass spectrum (inset right). Expected mass = 29210.8 (-Met); found = 29206 Da.



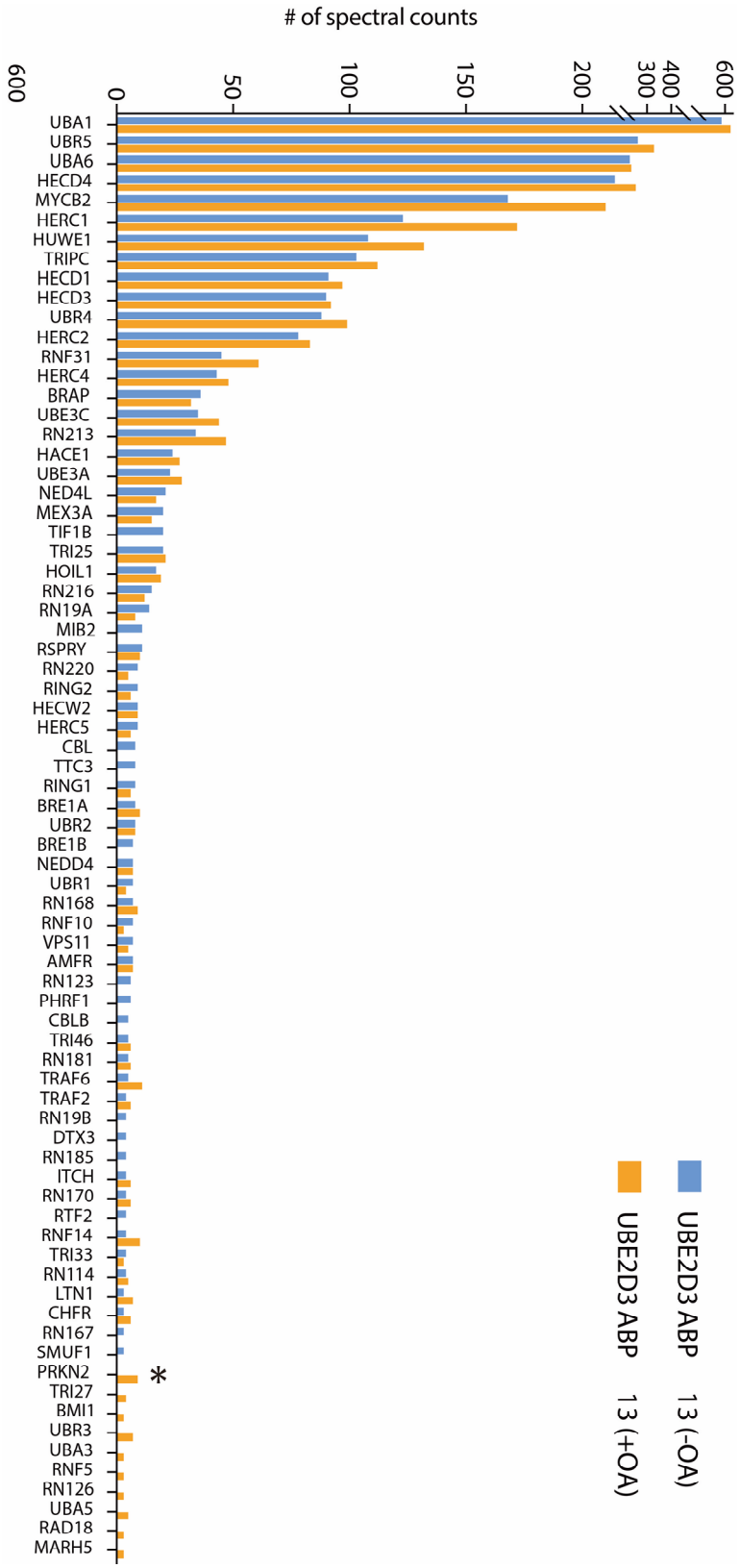
Extended Figure 5.7 The E3 spectral counts by LC-MS/MS analysis. (*next page*)

Control probes containing F62A mutation recover fewer E3s which is consistent with activity-based labelling. To induce Parkin activation SH-SY5Y cells were administered with oligomycin (5µM) and antimycin A (10 µM) (OA) for 3 hours. Extracted proteomes were then incubated with UBE2D2 ABP **13** and the control probe, UBE2D2 F62A **14** and ABP-labelled proteins were enriched against streptavidin resin followed by on-resin tryptic digestion. Obtained peptides were analyzed by data dependent LC-MS/MS. Recovered proteins were filtered against the pfam domain terms (RING, HECT, IBR, zf-UBR) and proteins with <3 spectral counts were excluded. E1s were also included as these undergo transthiolation and are highly enriched by our ABPs. The number of spectral counts for the recovered E1/E3 proteins plotted against protein ID for **13** and **14**.



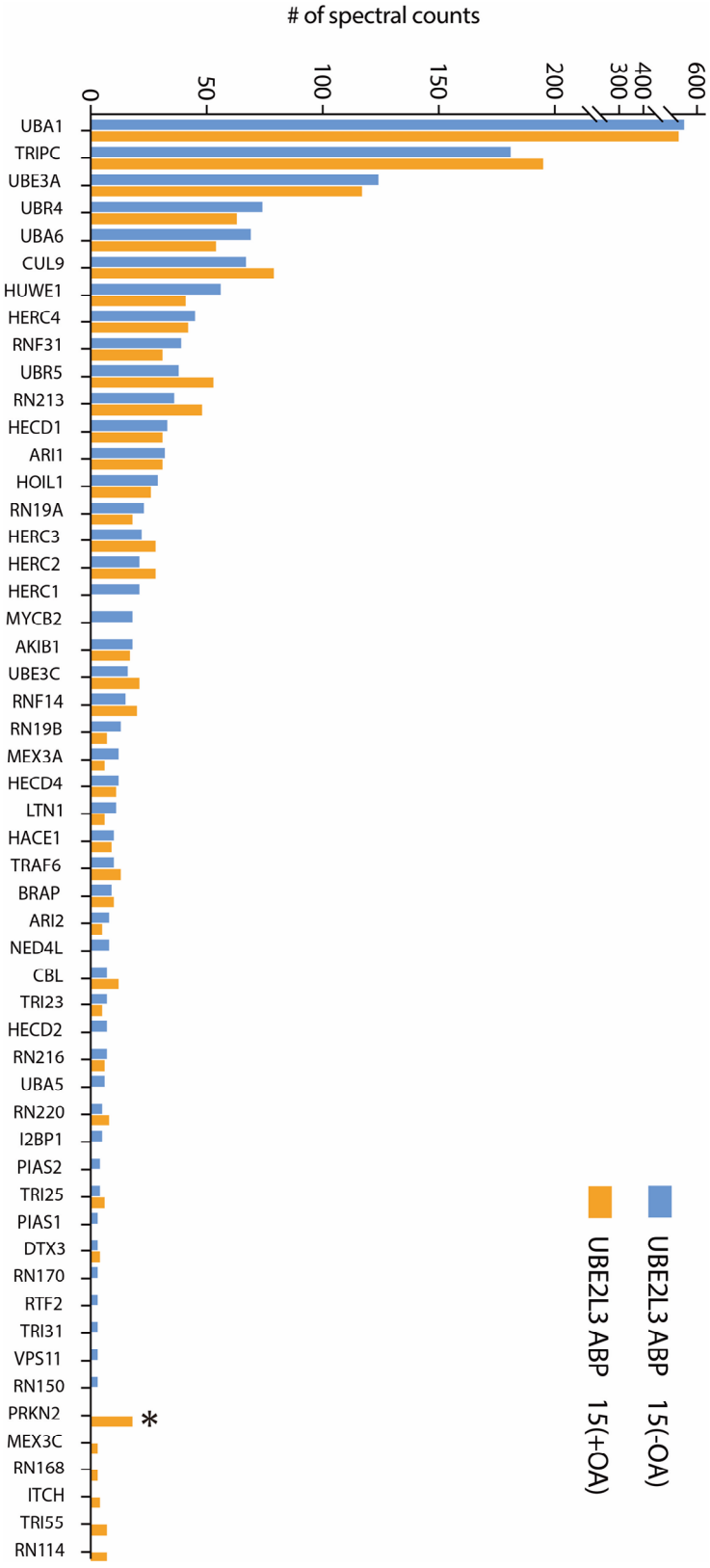
Extended Figure 5.8 The E3 spectral counts by LC-MS/MS analysis.

The same analysis as in the **extended figure 5.7** was carried out using UBE2L3 ABP **15** and is respective control ABP, UBE2L3 F63A **16.**



Extended Figure 5.9 The E3 spectral counts by LC-MS/MS analysis.

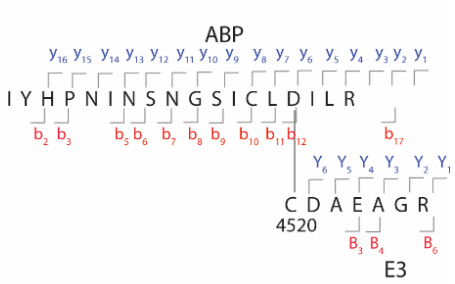
SH-SY5Y cells were untreated or administered with oligomycin (5 μM) and antimycin A (10 μM) (OA) for 3 hours. Extracted proteomes were then incubated with UBE2D2 ABP **13**. The data was analyzed by the same method used in **extended figure 5.7**.



Extended Figure 5.10 The E3 spectral counts by LC-MS/MS analysis.

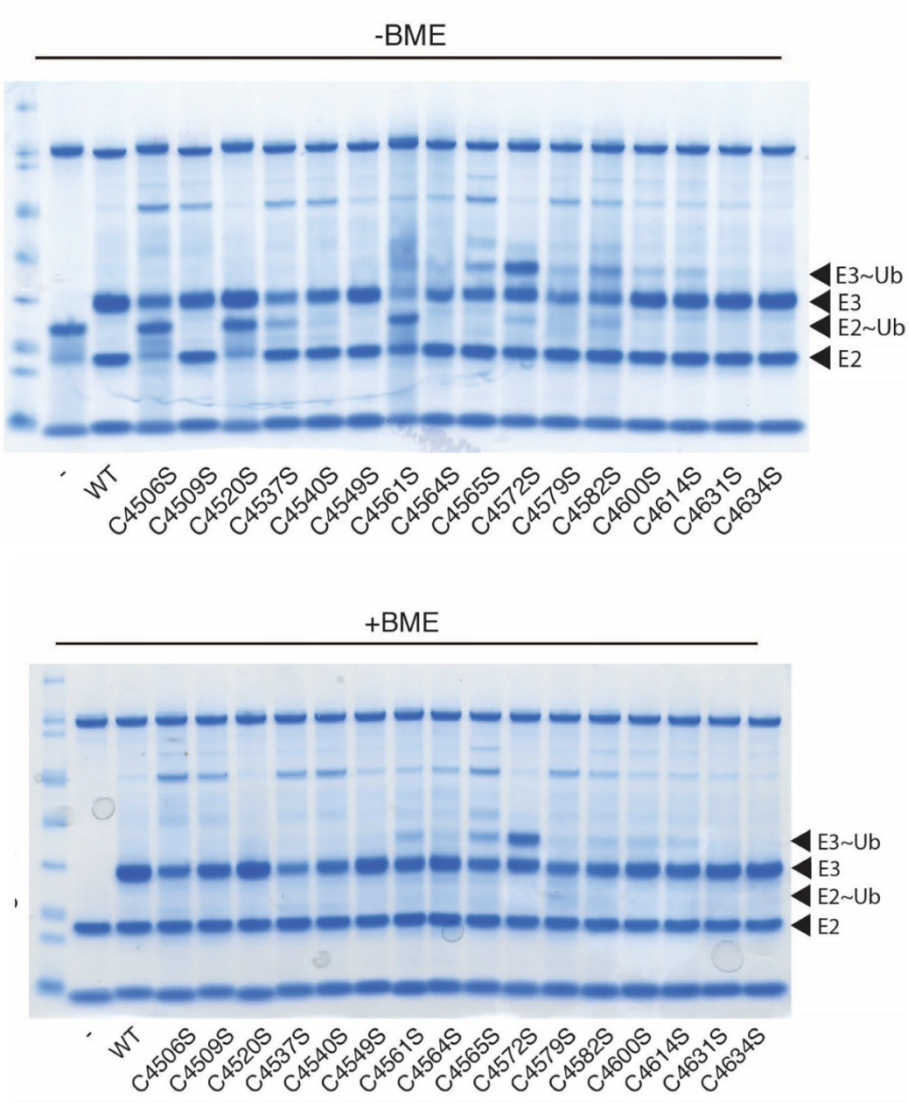
SH-SY5Y cells were untreated or administered with oligomycin (5 μM) and antimycin A (10 μM) (OA) for 3 hours. Extracted proteomes were then incubated with UBE2D2 ABP **15**. The data was analyzed by the same method used in **extended figure 5.7**.

	+	++	+++
b1	114.0913	57.5493	38.702
b2	277.1547	139.081	93.0564
b3	414.2136	207.6104	138.7427
b5	625.3093	313.1583	209.1079
b6	738.3933	369.7003	246.8026
b7	852.4363	426.7218	284.8169
b8	939.4683	470.2378	313.8276
b9	1053.5112	527.2592	351.8419
b10	1110.5327	555.77	370.8491
b11	1197.5647	599.286	399.8598
b12	1310.6488	655.828	437.5544
b17	2894.4037	1447.7055	965.4727
B3	2637.3025	1319.1549	879.7723
B4	2766.3451	1383.6762	922.7865
B6	2894.4037	1447.7055	965.4727
y1	175.119	88.0631	59.0445
y2	288.2030	144.6051	96.7392
y3	401.2871	201.1472	134.4339
y4	516.3140	258.6606	172.7762
y5	629.3981	315.2027	210.4709
y6	1758.8738	879.9406	586.9628
y7	1871.9579	936.4826	624.6575
y8	1958.9899	979.9986	653.6682
y9	2016.0114	1008.5093	672.6753
y10	2130.0543	1065.5308	710.6896
y11	2217.0863	1109.0468	739.7003
y12	2331.1293	1166.0683	777.7146
y13	2444.2133	1222.6103	815.4093
y14	2558.2563	1279.6318	853.4236
y15	2655.3090	1328.1582	885.7745
y16	2792.3679	1396.6876	931.4608
Y1	175.1190	88.0631	59.0445
Y2	232.1404	116.5738	78.0517
Y3	303.1775	152.0924	101.7307
Y4	432.2201	216.6137	144.7449
Y5	503.2572	252.1322	168.4239
Y6	618.2842	309.6457	206.7662



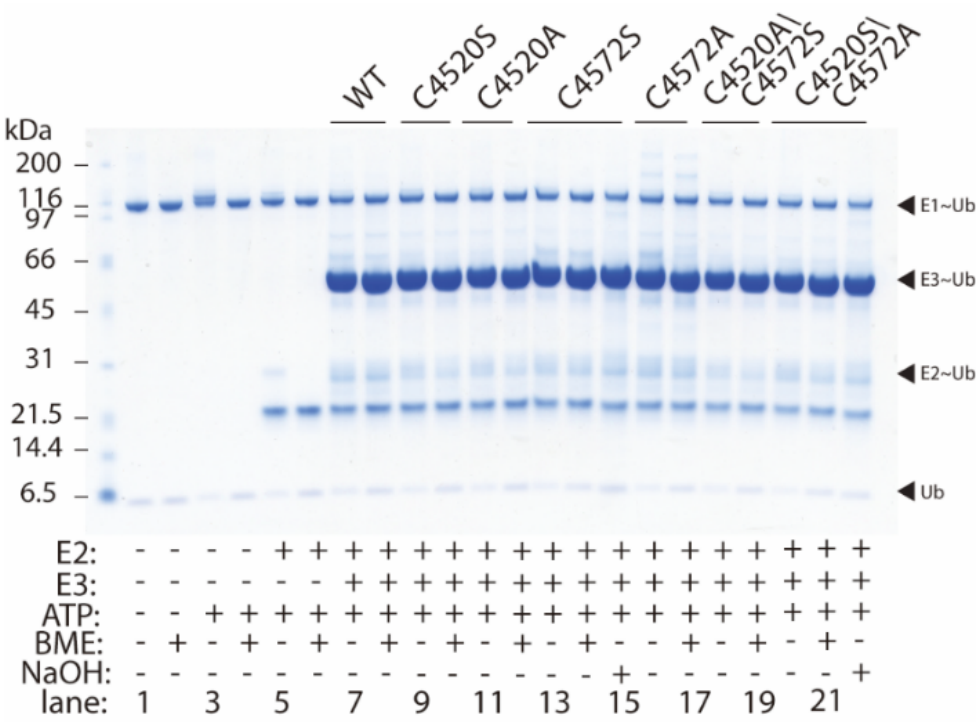
Extended Figure 5.11 Expected and found fragment ions in the MS^2 spectrum for the predominant cysteine residue labelled with the ABP.

Using the pLink software, 38 spectral matches corresponding to cysteine labelling sites in MYCBP2cat were identified. 36 of these corresponded to C4520. **Figure 5.2d** is a representative MS^2 spectrum. Above, is the predicted and found fragment ions for the representative spectrum. The spectrum is for a 5$^+$ precursor ion (observed *m/z* = 614.5088; expected *m/z* = 614.5094). A mass tolerance of 20 ppm was applied for fragment ion assignment.



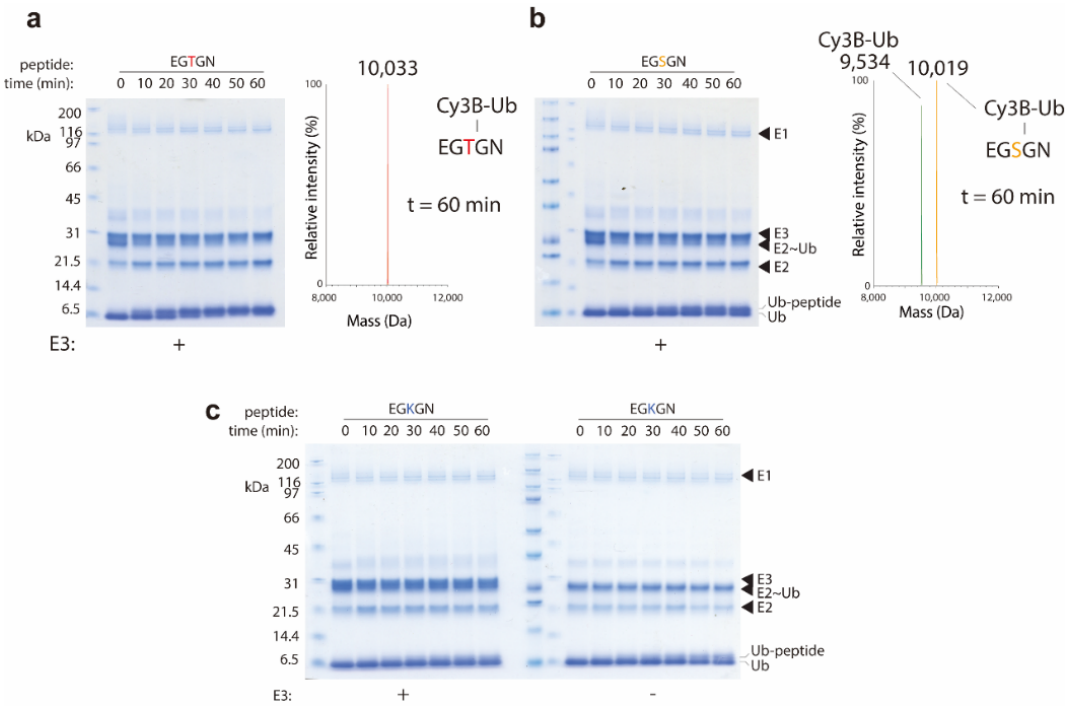
Extended Figure 5.12 Discharge activity towards Tris/glycerol for all of the tested MYCBP2cat cysteine to serine mutants.

The C4506S mutation abolishes discharge activity but because C4506S resides within a Cys-X-X-Cys Zn-binding motif, this was assumed to be due to disruption of the protein fold and not a catalytic defect. The C4561S mutation undergoes aberrant thioester adduct formation. An explanation is that C4561S mutation (also in a structurally important Cys-X-X-Cys Zn-binding motif) unfolds the protein and liberates Cys residues, that would otherwise be occupied as Zn ligands, that undergo aberrant transthiolation.



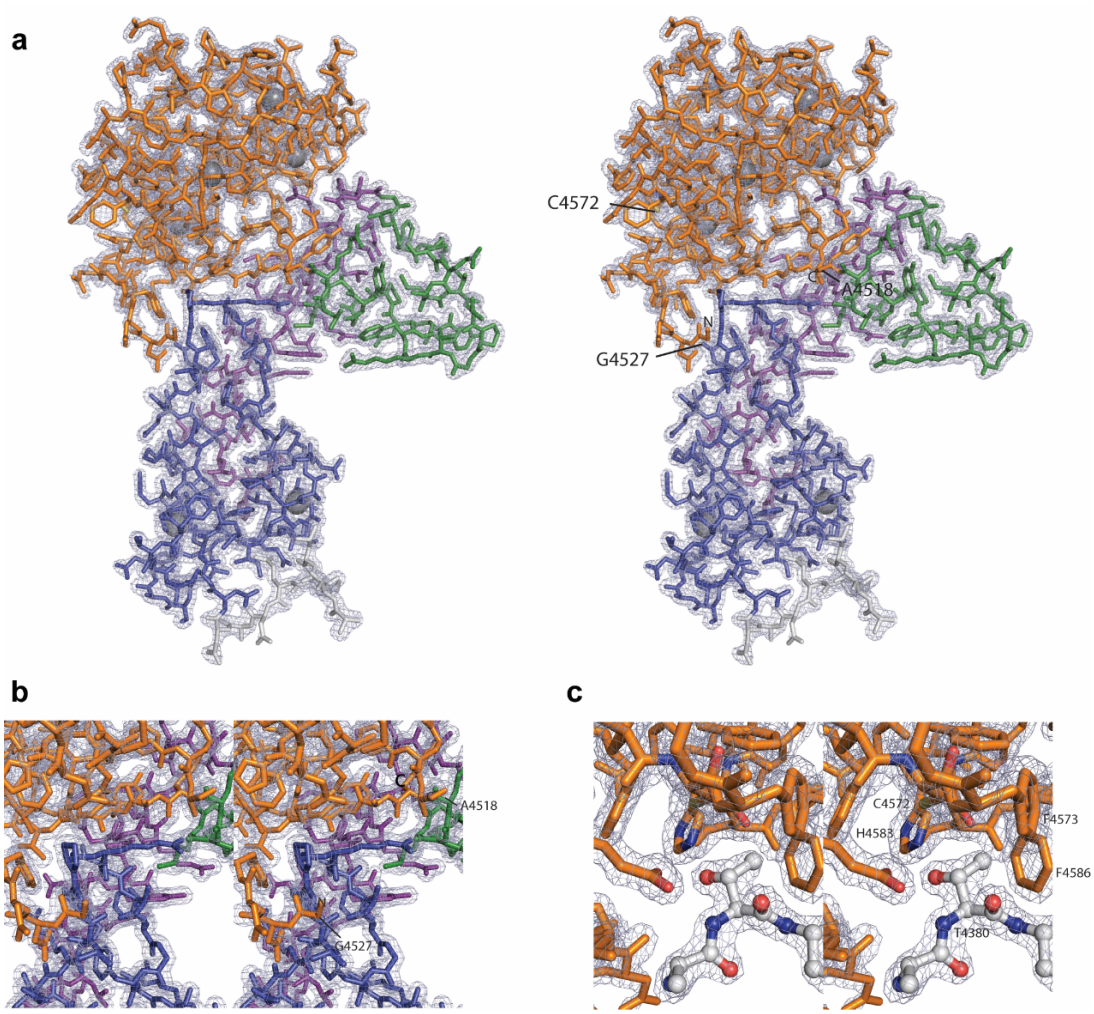
Extended Figure 5.13 Coomassie stain of the thioester/ester trapping assay with GST-MYCBP2cat.

After in-gel fluorescence scan, as shown in **Figure 5.4**, the gel was Coomassie stained.



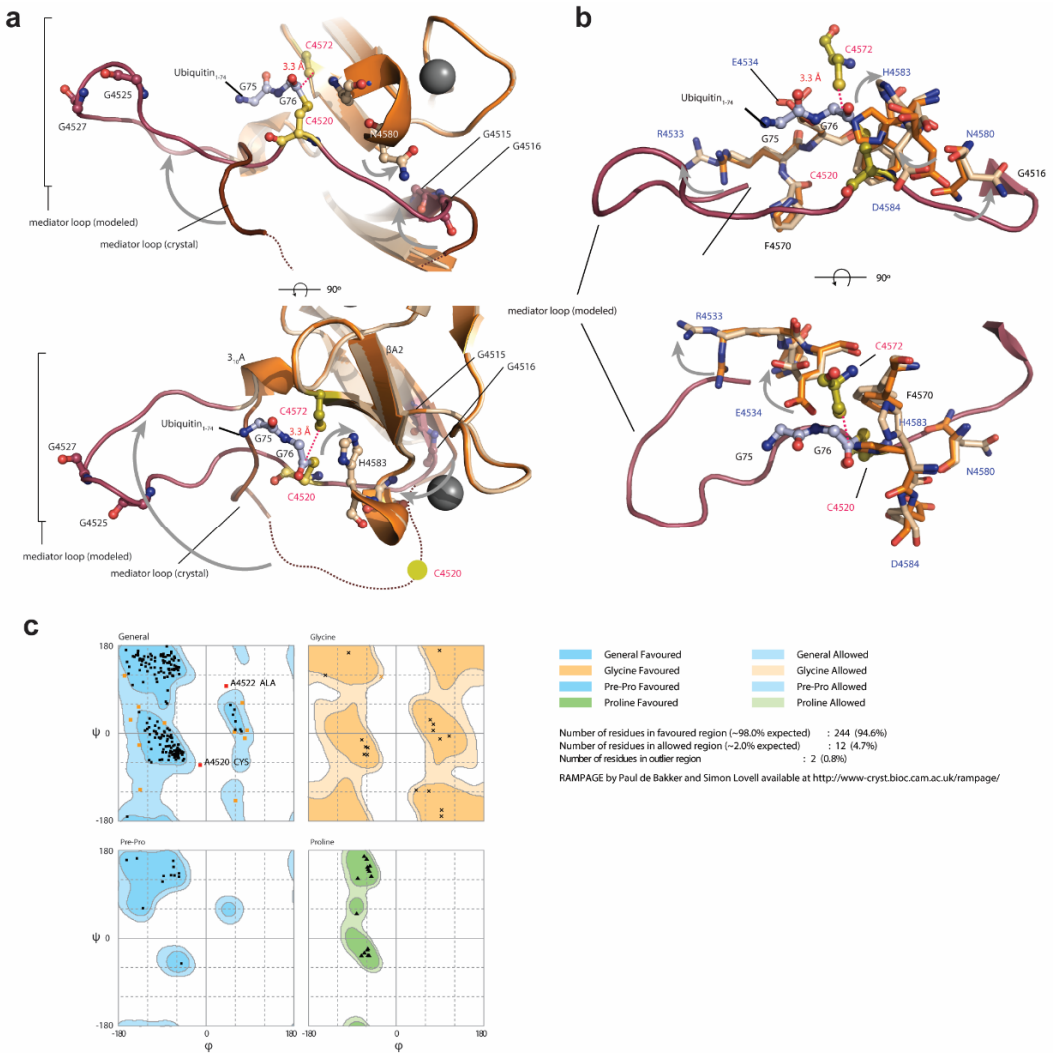
Extended Figure 5.14 Coomassie stain of peptide-assay gel.

(a) Coomassie stain of threonine gel presented in **Figure. 5.6**. Also shown is the deconvoluted mass spectrum representative of all ubiquitin species at the 60 min time point. Observed mass of Cy3B-Ub modified threonine peptide = 10,033 Da; theoretical mass = 10,036 Da. (**b**) Coomassie stain of the serine gel presented in **Figure 5.6**. Also shown is the deconvoluted mass spectrum representative of all ubiquitin species at the 60-min time point. Observed mass of Cy3B-Ub = 9,534 Da; theoretical mass = 9,537 Da. Observed mass of Cy3B-Ub modified serine peptide = 10,019 Da; theoretical mass = 10,022 Da. (**c**) Coomassie stain of lysine gels, in the presence and absence of E3, presented in **Figure 5.6**. Inefficient modification of the lysine peptide is observed, which is moderately enhanced in the absence of E3. Experiments shown in **a**-**c** were repeated more than three times.



Extended Figure 5.15 Representative stereo views of the MYCBP2cat crystallographic model.

The distinct regions are colored in stick representation: RING domain (blue), linker helix (purple), helix-turn-helix motif (green) and tandem cysteine (TC) domain (orange). The mesh represents the experimental $2|F_{obs}|-|F_{calc}|$ electron density map contoured at 1.5 σ. **(a)** Wide field view. C4572 is the downstream active cysteine residue in the esterification site. The mediator loop region is formed between A4518 and G4527 and is disordered in the structure. (**b**) Close up of the mediator loop region. (**c**) Close up of the esterification site. Threonine 4380 motif from the symmetry related molecule ($T4380_{sym}$) is shown and represented in grey ball and stick. (*The figures are provided by Dr. Virdee*)



Extended Figure 5.16 Modelling of mediator loop.

(**a**) Model of the proposed ubiquitin relay intermediate as shown in **Figure 5.14** but from aalternative perspective. In the experimental structure, tandem-cysteinedomain residues are shown in orange and mediator-loop residues are in dark brown. In the model, tandem-cysteine-domain residues are in light orange and mediator-loop residues are in mauve. The modelled E2, based on the superposition in **a**, is in grey cartoon. Essential cysteines C4520 and C4572 are in yellow and coloured by atom type. Ubiquitin residues G75-G76 are in blue ball-and-stick representation and are coloured by atom type. Gly residues in the mediator loop that are likely to be important for loop mobility are displayed in mauve ball and stick and coloured by atom type. N4570 and H4583 side chains have been rotated by the specified angles to relieve steric clash. (**b**) As in **a**, but amino acid side chains that have been flipped to relieve steric clash with the modelled mediator loop are labelled in blue (**c**) All phi and psi angles in the modelled structure fall within accepted values as determined by Ramachandran analysis with the RAMPAGE server. (*The figures are provided by Dr. Virdee and Dr. Mabbitt*)

	$MYCBP2_{4378-4640}$
Data collection	
Space group	$P\,6_1$
Cell dimensions	
a, b, c (Å)	82.58, 82.58, 103.30
α, γ, β (°)	90.00, 90.00, 120.00
Resolution (Å)	40 -1.75 (1.78 -1.75)
R_{merge}	0.051 (0.404)
I / σI	18.3 (2.4)
Completeness (%)	99.3 (93.3)
Redundancy	6.9 (3.0)
Refinement	
Resolution (Å)	40.0-1.75
No. reflections	40007
R_{work} / R_{free}	0.172/0.196
No. atoms	
Protein	1988
Ligand/ion	6
Water	257
B-factors	
Protein	28.2
Ligand/ion	26.6
Water	35.5
R.m.s. deviations	
Bond lengths (Å)	0.019
Bond angles (°)	1.9

Data were collected from a single crystal. Values in parentheses are for highest-resolution shell.

Extended Table 5.1 Data collection and refinement statistics of MYBCP2cat (*The table is provided by Dr. Virdee)*

Perspectives

In this thesis, I demonstrated a sophiscated method to generate E3 ABPs[322]. I further applied the ABPs using the activity-based prteomic strategy to uncover a novel E3 mechanism, RING-Cysteine-Relay, for ubiquitin transfer[547]. Furthermore, poorly characterized E3, MYBCP2, contains distinct ubiquitin esterification reactivity to selectively transfer ubiquitin upon the threonine amino acid. Taken together, this thesis provides the first example for discovering and profiling E3 activity via novel E3 activity-based probes. However, in order to broadly apply this platform across multiple research areas, there were a few issues that need to be addressed.

Although the methodology for producing the E2-based E3 ABPs is straight forward, the hydrolysis of the truncated Ub_{1-74}-SR during the aminolysis step resulted in a low yield for making this specific type of ABP (the E3 ABPs with a Ub_{1-74} moiety). The CuAAC click reaction though is a quick route for conjugating TDAE compounds with the ubiquitin-azide moiety, the 40-50% product loss during the reaction and purification step is a major drawback for producing ABPs on a large scale. Moreover, the last two or three residues missing (-GlyGly or –ArgGlyGly) in the ubiquitin moiety (Ub_{1-74} ABP or Ub_{1-73} ABP, respectively) of the E3 ABPs may still affect the activity or selectivity of our ABPs. Therefore, optimization of the TDAE compounds for making the E3 ABPs is required. In addition, despite the molar concentration of the ABPs being used at a 4:1 ratio towards target E3s, the labeling efficiency of E3 ABPs toward some E3s in the *in vitro* activity profiling assay is still low.

The generally used strategies to determine E3 activity are either by observing E3 autoubiquitylation, ubiquitin chain formation, or by measuring the E2~Ub

discharge rate[82,121,239]. However, the above methods may not accyrately reflect the 'genuine' E3 activity since these reactions are normally supplied with excess E1, E2 and ubiquitin. It is possible that the small amount of activated E3 enzyme is sufficient to make siginificant amoniuts of ubiquitin chains (that might give a disproportionate refecltion of E3 activity). Furthermore, the highly sensitive immunoblotting and silver-staining method that are broadly applied in these assays may also overestimate the activity of the E3s. However, the labeling mechanism of our E3 ABPs is by detecting the transthiolation activity of the target E3[322,334]. The chemistry associated with native E2-E3 transthiolaiton and E3 ABs labelling are different. As a result, it is not appropriate to claim that specific E3 is 'fully-active' when it catalyzed the ubiquitin chains formation under *in vitro* assays.

The conditions for our E3 ABP labelling assay is also an issue that needs to be considered. The availability of the catalytic cysteine on HECT/RBR/RCR E3s is an imporatant criterion for E3 ABP labelling. Oxidation of the cysteine or formation of a disulfide may prohibit the ABP labeling. The supply reducing agents in the reactions can be performed but side reactions of the reducing agents toward E3 ABPs were observed. The excess reducing agents (DTT or TCEP) are able to react with the 'warhead' on our ABPs which could abolish the activity of the E3 ABPs (data not shown). In addition, how much reducing agent is defined as a 'sufficient level' for keeping the catalytic cysteine available for E3 activity? To estimate the required reducing agent concentration for cellular extracts is especially challenging.

The incubation time for the in *situ* probe labeling assay is considerably long (4.0 hours) compared to the autoubiquitylation assay (0.5-1.0 hours). Albeit TCEP is relatively stable, the degradation of these reducing agents can still happen during the long incubation time, especially at 30°C or 37°C conditions. Furthermore, insufficient reducing agent in the reaction buffer may disrupt the

integrity of the Zn-binding domain on E3s since zinc ions are coordinated by cysteine residues in E3 proteins (the RING and IBR domains in RBRs[236] , or RCR E3 and the TC domain in RCR E3[547]). This may decrease the stability of the target E3s and abolish ABP labelling.

How to correctly apply E3 ABP to E3 research is also another important issue for follow-up work. Despite the ABPP strategy and the crystallography which are discussed previously, what kinds of assays can we design and perform to obtain more insight about E3 activity regulation? Furthermore, it has been demonstrated that the interactions between cofactors and E3s can activate or enhance certain types of E3s. How to dissect such complicated components behind such E3 activation mechanisms that coulde be detected by our E3 ABPs needs to be addressed.

The other part of this thesis is the discovery of a novel E3 mechanism perform by a giant E3, MYCBP2. Potentially, this is an exciting discovery which could open a new dimension to the ubiquitin field. However, the most important concern for this project is the missing MYCBP2 ubiquitin esterification substrate in cells. We clearly observed base labile Ub-conjugated adducts in cellular extract samples but it was largely attributed to auto-ubiquitylated MYCBP2cat. The identity of the genuine protein substrates of MYCBP2 which are conjugated within the ubiquitin via an oxyester-bond is under investigating. Obviously, the main question is that how should we perform the substrate search. Optimized biochemical conditions for assays are required due to the unstable nature of the oxyester-bond[521]. For example, maintaining the correct pH for the cell lysis buffer, protein digestion buffer for mass spectrometry, or even the sample loading buffer for SDS-PAGE; preparing the lysis buffer with suitable protease/phosphatase inhibitors (the inhibitors which have been proven to be unable to react with oxyester-bond).

Moreover, the algorithm of the mass spectrometry software for searching the peptides PTM sites also requires modification. Based on the current settings in the software, the predicated molecular weight (typically +144 Da) of the non-lysine ubiquitination modification is often not included in the search software. As a result, we may still be unable to detect the peptides which contain the ubiquitylated Thr/Ser residues even we perform the assay under a idealised biochemical conditions.

Finding the direct substrate for a protein has been proven to be challenging in many fields. Performing MYCBP2 immunoprecipitation with mass spectrometry analysis could provide the candidate substrates. However, a good antibody for MYCBP2 with high selectivity is required for this strategy. Furthermore, choosing a correct biological model for MYCBP2 substrate research is required since MYCBP2 is highly expressed in brain and cerebellum[478]. It has also been proved to regulate axon and synapse development[484,488]. As a result, performing the immunoprecipitation in physiologically-irrelevant but frequently used cell lines such as HEK293, U2OS or HeLa cells may not provide the correct information on MYCBP2 substrates. MYCBP2 KO or ligase-dead KI primary neuronal cells, or differentiated neuroblastoma cells, should be studied in future research. Another strategy which has already been applied in searching for substrates is by performing immunoprecipitation and mass spectrometry using 'ubiquitin remnant motif antibody K-ε-GG'[27,548]. The analogous di-GlyGly-threonine based antibody is also being developed by our lab. This tool should provide an informative dataset for threonine ubiquitination.

References:

1 Kerscher, O., Felberbaum, R. & Hochstrasser, M. Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annual review of cell and developmental biology* **22**, 159-180, doi:10.1146/annurev.cellbio.22.010605.093503 (2006).

2 Goldstein, G. *et al.* Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proc Natl Acad Sci U S A* **72**, 11-15 (1975).

3 Ciehanover, A., Hod, Y. & Hershko, A. A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. *Biochem Biophys Res Commun* **81**, 1100-1105 (1978).

4 Ciechanover, A., Heller, H., Elias, S., Haas, A. L. & Hershko, A. ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. *Proc Natl Acad Sci U S A* **77**, 1365-1368 (1980).

5 Hershko, A., Ciechanover, A., Heller, H., Haas, A. L. & Rose, I. A. Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. *Proc Natl Acad Sci U S A* **77**, 1783-1786 (1980).

6 Harrison, J. S., Jacobs, T. M., Houlihan, K., Van Doorslaer, K. & Kuhlman, B. UbSRD: The Ubiquitin Structural Relational Database. *J Mol Biol* **428**, 679-687, doi:10.1016/j.jmb.2015.09.011 (2016).

7 Lenkinski, R. E., Chen, D. M., Glickson, J. D. & Goldstein, G. Nuclear magnetic resonance studies of the denaturation of ubiquitin. *Biochim Biophys Acta* **494**, 126-130 (1977).

8 Perica, T. & Chothia, C. Ubiquitin--molecular mechanisms for recognition of different structures. *Curr Opin Struct Biol* **20**, 367-376, doi:10.1016/j.sbi.2010.03.007 (2010).

9 Vijay-Kumar, S., Bugg, C. E. & Cook, W. J. Structure of ubiquitin refined at 1.8 A resolution. *J Mol Biol* **194**, 531-544 (1987).

10 Winget, J. M. & Mayor, T. The diversity of ubiquitin recognition: hot spots and varied specificity. *Mol Cell* **38**, 627-635, doi:10.1016/j.molcel.2010.05.003 (2010).

11 Hochstrasser, M. Origin and function of ubiquitin-like proteins. *Nature* **458**, 422-429, doi:10.1038/nature07958 (2009).

12 Gareau, J. R. & Lima, C. D. The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nat Rev Mol Cell Biol*

11, 861-871, doi:10.1038/nrm3011 (2010).

13 Rabut, G. & Peter, M. Function and regulation of protein neddylation. 'Protein modifications: beyond the usual suspects' review series. *EMBO Rep* **9**, 969-976, doi:10.1038/embor.2008.183 (2008).

14 Geng, J. & Klionsky, D. J. The Atg8 and Atg12 ubiquitin-like conjugation systems in macroautophagy. 'Protein modifications: beyond the usual suspects' review series. *EMBO Rep* **9**, 859-864, doi:10.1038/embor.2008.163 (2008).

15 Xie, Z., Nair, U. & Klionsky, D. J. Atg8 controls phagophore expansion during autophagosome formation. *Mol Biol Cell* **19**, 3290-3298, doi:10.1091/mbc.E07-12-1292 (2008).

16 Stolz, A. *et al.* Fluorescence-based ATG8 sensors monitor localization and function of LC3/GABARAP proteins. *EMBO J* **36**, 549-564, doi:10.15252/embj.201695063 (2017).

17 Mueller, T. D. & Feigon, J. Structural determinants for the binding of ubiquitin-like domains to the proteasome. *EMBO J* **22**, 4634-4645, doi:10.1093/emboj/cdg467 (2003).

18 Ronau, J. A., Beckmann, J. F. & Hochstrasser, M. Substrate specificity of the ubiquitin and Ubl proteases. *Cell Res* **26**, 441-456, doi:10.1038/cr.2016.38 (2016).

19 Komander, D. & Rape, M. The ubiquitin code. *Annu Rev Biochem* **81**, 203-229, doi:10.1146/annurev-biochem-060310-170328 (2012).

20 Yau, R. & Rape, M. The increasing complexity of the ubiquitin code. *Nat Cell Biol* **18**, 579-586, doi:10.1038/ncb3358 (2016).

21 Hicke, L. Protein regulation by monoubiquitin. *Nat Rev Mol Cell Biol* **2**, 195-201, doi:10.1038/35056583 (2001).

22 Ulrich, H. D. & Walden, H. Ubiquitin signalling in DNA replication and repair. *Nat Rev Mol Cell Biol* **11**, 479-489, doi:10.1038/nrm2921 (2010).

23 Goldknopf, I. L., French, M. F., Musso, R. & Busch, H. Presence of protein A24 in rat liver nucleosomes. *Proc Natl Acad Sci U S A* **74**, 5492-5495 (1977).

24 Hoege, C., Pfander, B., Moldovan, G. L., Pyrowolakis, G. & Jentsch, S. RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* **419**, 135-141, doi:10.1038/nature00991 (2002).

25 Dupont, S. *et al.* FAM/USP9x, a deubiquitinating enzyme essential for TGFbeta signaling, controls Smad4 monoubiquitination. *Cell* **136**, 123-135, doi:10.1016/j.cell.2008.10.051 (2009).

26 Haglund, K. *et al.* Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation. *Nat Cell Biol* **5**, 461-466, doi:10.1038/ncb983 (2003).

27 Kim, W. *et al.* Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol Cell* **44**, 325-340, doi:10.1016/j.molcel.2011.08.025 (2011).

28 Kaiser, S. E. *et al.* Protein standard absolute quantification (PSAQ) method for the measurement of cellular ubiquitin pools. *Nat Methods* **8**, 691-696, doi:10.1038/nmeth.1649 (2011).

29 Xu, P. *et al.* Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation. *Cell* **137**, 133-145, doi:10.1016/j.cell.2009.01.041 (2009).

30 Saeki, Y. Ubiquitin recognition by the proteasome. *J Biochem* **161**, 113-124, doi:10.1093/jb/mvw091 (2017).

31 Thrower, J. S., Hoffman, L., Rechsteiner, M. & Pickart, C. M. Recognition of the polyubiquitin proteolytic signal. *EMBO J* **19**, 94-102, doi:10.1093/emboj/19.1.94 (2000).

32 Petroski, M. D. & Deshaies, R. J. Mechanism of lysine 48-linked ubiquitin-chain synthesis by the cullin-RING ubiquitin-ligase complex SCF-Cdc34. *Cell* **123**, 1107-1120, doi:10.1016/j.cell.2005.09.033 (2005).

33 Lu, Y., Lee, B. H., King, R. W., Finley, D. & Kirschner, M. W. Substrate degradation by the proteasome: a single-molecule kinetic analysis. *Science* **348**, 1250834, doi:10.1126/science.1250834 (2015).

34 Deng, L. *et al.* Activation of the IkappaB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. *Cell* **103**, 351-361 (2000).

35 Doil, C. *et al.* RNF168 binds and amplifies ubiquitin conjugates on damaged chromosomes to allow accumulation of repair proteins. *Cell* **136**, 435-446, doi:10.1016/j.cell.2008.12.041 (2009).

36 Lee, B. L., Singh, A., Mark Glover, J. N., Hendzel, M. J. & Spyracopoulos, L. Molecular Basis for K63-Linked Ubiquitination Processes in Double-Strand DNA Break Repair: A Focus on Kinetics and Dynamics. *J Mol Biol* **429**, 3409-3429, doi:10.1016/j.jmb.2017.05.029 (2017).

37 Silva, G. M., Finley, D. & Vogel, C. K63 polyubiquitination is a new modulator of the oxidative stress response. *Nat Struct Mol Biol* **22**, 116-123, doi:10.1038/nsmb.2955 (2015).

38 Gack, M. U. *et al.* TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. *Nature* **446**, 916-920,

doi:10.1038/nature05732 (2007).

39 Martin-Vicente, M., Medrano, L. M., Resino, S., Garcia-Sastre, A. & Martinez, I. TRIM25 in the Regulation of the Antiviral Innate Immunity. *Front Immunol* **8**, 1187, doi:10.3389/fimmu.2017.01187 (2017).

40 Song, E. J. *et al.* The Prp19 complex and the Usp4Sart3 deubiquitinating enzyme control reversible ubiquitination at the spliceosome. *Genes Dev* **24**, 1434-1447, doi:10.1101/gad.1925010 (2010).

41 Spence, J. *et al.* Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain. *Cell* **102**, 67-76 (2000).

42 Emanuele, M. J. *et al.* Global identification of modular cullin-RING ligase substrates. *Cell* **147**, 459-474, doi:10.1016/j.cell.2011.09.019 (2011).

43 Komander, D. *et al.* Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains. *EMBO Rep* **10**, 466-473, doi:10.1038/embor.2009.55 (2009).

44 Kirisako, T. *et al.* A ubiquitin ligase complex assembles linear polyubiquitin chains. *EMBO J* **25**, 4877-4887, doi:10.1038/sj.emboj.7601360 (2006).

45 Ikeda, F. *et al.* SHARPIN forms a linear ubiquitin ligase complex regulating NF-kappaB activity and apoptosis. *Nature* **471**, 637-641, doi:10.1038/nature09814 (2011).

46 Tokunaga, F. *et al.* SHARPIN is a component of the NF-kappaB-activating linear ubiquitin chain assembly complex. *Nature* **471**, 633-636, doi:10.1038/nature09815 (2011).

47 Tokunaga, F. *et al.* Involvement of linear polyubiquitylation of NEMO in NF-kappaB activation. *Nat Cell Biol* **11**, 123-132, doi:10.1038/ncb1821 (2009).

48 Gerlach, B. *et al.* Linear ubiquitination prevents inflammation and regulates immune signalling. *Nature* **471**, 591-596, doi:10.1038/nature09816 (2011).

49 Rahighi, S. *et al.* Specific recognition of linear ubiquitin chains by NEMO is important for NF-kappaB activation. *Cell* **136**, 1098-1109, doi:10.1016/j.cell.2009.03.007 (2009).

50 Jing, H. *et al.* Porcine Reproductive and Respiratory Syndrome Virus nsp1alpha Inhibits NF-kappaB Activation by Targeting the Linear Ubiquitin Chain Assembly Complex. *J Virol* **91**, doi:10.1128/JVI.01911-16 (2017).

51 Wang, L. *et al.* The Linear Ubiquitin Assembly Complex Modulates Latent Membrane Protein 1 Activation of NF-kappaB and Interferon Regulatory Factor 7. *J Virol* **91**, doi:10.1128/JVI.01138-16 (2017).

52 Lafont, E. *et al.* The linear ubiquitin chain assembly complex regulates

TRAIL-induced gene activation and cell death. *EMBO J* **36**, 1147-1166, doi:10.15252/embj.201695699 (2017).

53 Matsumoto, M. L. *et al.* K11-linked polyubiquitination in cell cycle control revealed by a K11 linkage-specific antibody. *Mol Cell* **39**, 477-484, doi:10.1016/j.molcel.2010.07.001 (2010).

54 Meyer, H. J. & Rape, M. Enhanced protein degradation by branched ubiquitin chains. *Cell* **157**, 910-921, doi:10.1016/j.cell.2014.03.037 (2014).

55 Bremm, A., Moniz, S., Mader, J., Rocha, S. & Komander, D. Cezanne (OTUD7B) regulates HIF-1alpha homeostasis in a proteasome-independent manner. *EMBO Rep* **15**, 1268-1277, doi:10.15252/embr.201438850 (2014).

56 Qin, Y. *et al.* RNF26 temporally regulates virus-triggered type I interferon induction by two distinct mechanisms. *PLoS Pathog* **10**, e1004358, doi:10.1371/journal.ppat.1004358 (2014).

57 Elia, A. E. *et al.* Quantitative Proteomic Atlas of Ubiquitination and Acetylation in the DNA Damage Response. *Mol Cell* **59**, 867-881, doi:10.1016/j.molcel.2015.05.006 (2015).

58 Cripps, D. *et al.* Alzheimer disease-specific conformation of hyperphosphorylated paired helical filament-Tau is polyubiquitinated through Lys-48, Lys-11, and Lys-6 ubiquitin conjugation. *J Biol Chem* **281**, 10825-10838, doi:10.1074/jbc.M512786200 (2006).

59 Shang, F. *et al.* Lys6-modified ubiquitin inhibits ubiquitin-dependent protein degradation. *J Biol Chem* **280**, 20365-20374, doi:10.1074/jbc.M414356200 (2005).

60 Ordureau, A. *et al.* Defining roles of PARKIN and ubiquitin phosphorylation by PINK1 in mitochondrial quality control using a ubiquitin replacement strategy. *Proc Natl Acad Sci U S A* **112**, 6637-6642, doi:10.1073/pnas.1506593112 (2015).

61 Durcan, T. M. *et al.* USP8 regulates mitophagy by removing K6-linked ubiquitin conjugates from parkin. *EMBO J* **33**, 2473-2491, doi:10.15252/embj.201489729 (2014).

62 Michel, M. A., Swatek, K. N., Hospenthal, M. K. & Komander, D. Ubiquitin Linkage-Specific Affimers Reveal Insights into K6-Linked Ubiquitin Signaling. *Mol Cell* **68**, 233-246 e235, doi:10.1016/j.molcel.2017.08.020 (2017).

63 Wang, Q. *et al.* The E3 ubiquitin ligase RNF185 facilitates the cGAS-mediated innate immune response. *PLoS Pathog* **13**, e1006264, doi:10.1371/journal.ppat.1006264 (2017).

64 Gatti, M. *et al.* RNF168 promotes noncanonical K27 ubiquitination to signal DNA damage. *Cell Rep* **10**, 226-238, doi:10.1016/j.celrep.2014.12.021 (2015).

65 Nucifora, F. C., Jr. *et al.* Ubiqutination via K27 and K29 chains signals aggregation and neuronal protection of LRRK2 by WSB1. *Nat Commun* **7**, 11792, doi:10.1038/ncomms11792 (2016).

66 Liu, J. *et al.* Rhbdd3 controls autoimmunity by suppressing the production of IL-6 by dendritic cells via K27-linked ubiquitination of the regulator NEMO. *Nat Immunol* **15**, 612-622, doi:10.1038/ni.2898 (2014).

67 Johnson, E. S., Ma, P. C., Ota, I. M. & Varshavsky, A. A proteolytic pathway that recognizes ubiquitin as a degradation signal. *J Biol Chem* **270**, 17442-17456 (1995).

68 Koegl, M. *et al.* A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. *Cell* **96**, 635-644 (1999).

69 Michel, M. A. *et al.* Assembly and specific recognition of k29- and k33-linked polyubiquitin. *Mol Cell* **58**, 95-109, doi:10.1016/j.molcel.2015.01.042 (2015).

70 Kristariyanto, Y. A. *et al.* K29-selective ubiquitin binding domain reveals structural basis of specificity and heterotypic nature of k29 polyubiquitin. *Mol Cell* **58**, 83-94, doi:10.1016/j.molcel.2015.01.041 (2015).

71 Chastagner, P., Israel, A. & Brou, C. Itch/AIP4 mediates Deltex degradation through the formation of K29-linked polyubiquitin chains. *EMBO Rep* **7**, 1147-1153, doi:10.1038/sj.embor.7400822 (2006).

72 Zotti, T. *et al.* TRAF7 protein promotes Lys-29-linked polyubiquitination of IkappaB kinase (IKKgamma)/NF-kappaB essential modulator (NEMO) and p65/RelA protein and represses NF-kappaB activation. *J Biol Chem* **286**, 22924-22933, doi:10.1074/jbc.M110.215426 (2011).

73 Al-Hakim, A. K. *et al.* Control of AMPK-related kinases by USP9X and atypical Lys(29)/Lys(33)-linked polyubiquitin chains. *Biochem J* **411**, 249-260, doi:10.1042/BJ20080067 (2008).

74 Yuan, W. C. *et al.* K33-Linked Polyubiquitination of Coronin 7 by Cul3-KLHL20 Ubiquitin E3 Ligase Regulates Protein Trafficking. *Mol Cell* **54**, 586-600, doi:10.1016/j.molcel.2014.03.035 (2014).

75 Huang, H. *et al.* K33-linked polyubiquitination of T cell receptor-zeta regulates proteolysis-independent T cell signaling. *Immunity* **33**, 60-70, doi:10.1016/j.immuni.2010.07.002 (2010).

76 Yang, M. *et al.* K33-linked polyubiquitination of Zap70 by Nrdp1 controls CD8(+) T cell activation. *Nat Immunol* **16**, 1253-1262, doi:10.1038/ni.3258

(2015).

77 Lin, M. *et al.* USP38 Inhibits Type I Interferon Signaling by Editing TBK1 Ubiquitination through NLRP4 Signalosome. *Mol Cell* **64**, 267-281, doi:10.1016/j.molcel.2016.08.029 (2016).

78 Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Res* **26**, 399-422, doi:10.1038/cr.2016.39 (2016).

79 Herhaus, L. & Dikic, I. Expanding the ubiquitin code through post-translational modification. *EMBO Rep* **16**, 1071-1083, doi:10.15252/embr.201540891 (2015).

80 Kane, L. A. *et al.* PINK1 phosphorylates ubiquitin to activate Parkin E3 ubiquitin ligase activity. *J Cell Biol* **205**, 143-153, doi:10.1083/jcb.201402104 (2014).

81 Koyano, F. *et al.* Ubiquitin is phosphorylated by PINK1 to activate parkin. *Nature* **510**, 162-166, doi:10.1038/nature13392 (2014).

82 Kazlauskaite, A. *et al.* Parkin is activated by PINK1-dependent phosphorylation of ubiquitin at Ser65. *Biochem J* **460**, 127-139, doi:10.1042/BJ20140334 (2014).

83 Okatsu, K. *et al.* Phosphorylated ubiquitin chain is the genuine Parkin receptor. *J Cell Biol* **209**, 111-128, doi:10.1083/jcb.201410050 (2015).

84 Wauer, T., Simicek, M., Schubert, A. & Komander, D. Mechanism of phospho-ubiquitin-induced PARKIN activation. *Nature* **524**, 370-374, doi:10.1038/nature14879 (2015).

85 Wauer, T. *et al.* Ubiquitin Ser65 phosphorylation affects ubiquitin structure, chain assembly and hydrolysis. *EMBO J* **34**, 307-325, doi:10.15252/embj.201489847 (2015).

86 Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu Rev Biochem* **67**, 425-479, doi:10.1146/annurev.biochem.67.1.425 (1998).

87 Smit, J. J. & Sixma, T. K. RBR E3-ligases at work. *EMBO Rep* **15**, 142-154, doi:10.1002/embr.201338166 (2014).

88 Schulman, B. A. & Harper, J. W. Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat Rev Mol Cell Biol* **10**, 319-331, doi:10.1038/nrm2673 (2009).

89 Pickart, C. M. & Rose, I. A. Functional heterogeneity of ubiquitin carrier proteins. *J Biol Chem* **260**, 1573-1581 (1985).

90 Berndsen, C. E. & Wolberger, C. New insights into ubiquitin E3 ligase mechanism. *Nat Struct Mol Biol* **21**, 301-307, doi:10.1038/nsmb.2780 (2014).

91 Deshaies, R. J. & Joazeiro, C. A. RING domain E3 ubiquitin ligases. *Annu*

Rev Biochem **78**, 399-434, doi:10.1146/annurev.biochem.78.101807.093809 (2009).

92 Scheffner, M. & Kumar, S. Mammalian HECT ubiquitin-protein ligases: biological and pathophysiological aspects. *Biochim Biophys Acta* **1843**, 61-74, doi:10.1016/j.bbamcr.2013.03.024 (2014).

93 Dikic, I., Wakatsuki, S. & Walters, K. J. Ubiquitin-binding domains - from structures to functions. *Nat Rev Mol Cell Biol* **10**, 659-671, doi:10.1038/nrm2767 (2009).

94 Scott, D., Oldham, N. J., Strachan, J., Searle, M. S. & Layfield, R. Ubiquitin-binding domains: mechanisms of ubiquitin recognition and use as tools to investigate ubiquitin-modified proteomes. *Proteomics* **15**, 844-861, doi:10.1002/pmic.201400341 (2015).

95 Heideker, J. & Wertz, I. E. DUBs, the regulation of cell identity and disease. *Biochem J* **467**, 191 (2015).

96 Komander, D., Clague, M. J. & Urbe, S. Breaking the chains: structure and function of the deubiquitinases. *Nat Rev Mol Cell Biol* **10**, 550-563, doi:10.1038/nrm2731 (2009).

97 Popovic, D., Vucic, D. & Dikic, I. Ubiquitination in disease pathogenesis and treatment. *Nat Med* **20**, 1242-1253, doi:10.1038/nm.3739 (2014).

98 Cohen, P. Immune diseases caused by mutations in kinases and components of the ubiquitin system. *Nat Immunol* **15**, 521-529, doi:10.1038/ni.2892 (2014).

99 Senft, D., Qi, J. & Ronai, Z. A. Ubiquitin ligases in oncogenic transformation and cancer therapy. *Nat Rev Cancer* **18**, 69-88, doi:10.1038/nrc.2017.105 (2018).

100 Atkin, G. & Paulson, H. Ubiquitin pathways in neurodegenerative disease. *Front Mol Neurosci* **7**, 63, doi:10.3389/fnmol.2014.00063 (2014).

101 Ciechanover, A., Heller, H., Katz-Etzion, R. & Hershko, A. Activation of the heat-stable polypeptide of the ATP-dependent proteolytic system. *Proc Natl Acad Sci U S A* **78**, 761-765 (1981).

102 Jin, J., Li, X., Gygi, S. P. & Harper, J. W. Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging. *Nature* **447**, 1135-1138, doi:10.1038/nature05902 (2007).

103 Chiu, Y. H., Sun, Q. & Chen, Z. J. E1-L2 activates both ubiquitin and FAT10. *Mol Cell* **27**, 1014-1023, doi:10.1016/j.molcel.2007.08.020 (2007).

104 Olsen, S. K. & Lima, C. D. Structure of a ubiquitin E1-E2 complex: insights to E1-E2 thioester transfer. *Mol Cell* **49**, 884-896, doi:10.1016/j.molcel.2013.01.013 (2013).

105 Walden, H., Podgorski, M. S. & Schulman, B. A. Insights into the ubiquitin transfer cascade from the structure of the activating enzyme for NEDD8. *Nature* **422**, 330-334, doi:10.1038/nature01456 (2003).

106 Lois, L. M. & Lima, C. D. Structures of the SUMO E1 provide mechanistic insights into SUMO activation and E2 recruitment to E1. *EMBO J* **24**, 439-451, doi:10.1038/sj.emboj.7600552 (2005).

107 Lee, I. & Schindelin, H. Structural insights into E1-catalyzed ubiquitin activation and transfer to conjugating enzymes. *Cell* **134**, 268-278, doi:10.1016/j.cell.2008.05.046 (2008).

108 Olsen, S. K., Capili, A. D., Lu, X., Tan, D. S. & Lima, C. D. Active site remodelling accompanies thioester bond formation in the SUMO E1. *Nature* **463**, 906-912, doi:10.1038/nature08765 (2010).

109 Huang, D. T. *et al.* Basis for a ubiquitin-like protein thioester switch toggling E1-E2 affinity. *Nature* **445**, 394-398, doi:10.1038/nature05490 (2007).

110 Haas, A. L., Bright, P. M. & Jackson, V. E. Functional diversity among putative E2 isozymes in the mechanism of ubiquitin-histone ligation. *J Biol Chem* **263**, 13268-13275 (1988).

111 Wang, J. *et al.* The intrinsic affinity between E2 and the Cys domain of E1 in ubiquitin-like modifications. *Mol Cell* **27**, 228-237, doi:10.1016/j.molcel.2007.05.023 (2007).

112 Finley, D., Ciechanover, A. & Varshavsky, A. Thermolability of ubiquitin-activating enzyme from the mammalian cell cycle mutant ts85. *Cell* **37**, 43-55 (1984).

113 Ciechanover, A., Finley, D. & Varshavsky, A. Mammalian cell cycle mutant defective in intracellular protein degradation and ubiquitin-protein conjugation. *Prog Clin Biol Res* **180**, 17-31 (1985).

114 Schlabach, M. R. *et al.* Cancer proliferation gene discovery through functional genomics. *Science* **319**, 620-624, doi:10.1126/science.1149200 (2008).

115 Kulkarni, M. & Smith, H. E. E1 ubiquitin-activating enzyme UBA-1 plays multiple roles throughout C. elegans development. *PLoS Genet* **4**, e1000131, doi:10.1371/journal.pgen.1000131 (2008).

116 Stewart, M. D., Ritterhoff, T., Klevit, R. E. & Brzovic, P. S. E2 enzymes: more than just middle men. *Cell Res* **26**, 423-440, doi:10.1038/cr.2016.35 (2016).

117 Wiener, R. *et al.* E2 ubiquitin-conjugating enzymes regulate the deubiquitinating activity of OTUB1. *Nat Struct Mol Biol* **20**, 1033-1039,

doi:10.1038/nsmb.2655 (2013).

118 Pruneda, J. N. *et al.* E2~Ub conjugates regulate the kinase activity of Shigella effector OspG during pathogenesis. *EMBO J* **33**, 437-449, doi:10.1002/embj.201386386 (2014).

119 Williams, C., van den Berg, M., Sprenger, R. R. & Distel, B. A conserved cysteine is essential for Pex4p-dependent ubiquitination of the peroxisomal import receptor Pex5p. *J Biol Chem* **282**, 22534-22543, doi:10.1074/jbc.M702038200 (2007).

120 Leon, S. & Subramani, S. A conserved cysteine residue of Pichia pastoris Pex20p is essential for its recycling from the peroxisome to the cytosol. *J Biol Chem* **282**, 7424-7430, doi:10.1074/jbc.M611627200 (2007).

121 Wenzel, D. M., Lissounov, A., Brzovic, P. S. & Klevit, R. E. UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids. *Nature* **474**, 105-108, doi:10.1038/nature09966 (2011).

122 Wang, X. *et al.* Ube2j2 ubiquitinates hydroxylated amino acids on ER-associated degradation substrates. *J Cell Biol* **187**, 655-668, doi:10.1083/jcb.200908036 (2009).

123 Zheng, N., Wang, P., Jeffrey, P. D. & Pavletich, N. P. Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. *Cell* **102**, 533-539 (2000).

124 Brzovic, P. S. *et al.* Binding and recognition in the assembly of an active BRCA1/BARD1 ubiquitin-ligase complex. *Proc Natl Acad Sci U S A* **100**, 5646-5651, doi:10.1073/pnas.0836054100 (2003).

125 Plechanovova, A., Jaffray, E. G., Tatham, M. H., Naismith, J. H. & Hay, R. T. Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature* **489**, 115-120, doi:10.1038/nature11376 (2012).

126 Dou, H., Buetow, L., Sibbet, G. J., Cameron, K. & Huang, D. T. BIRC7-E2 ubiquitin conjugate structure reveals the mechanism of ubiquitin transfer by a RING dimer. *Nat Struct Mol Biol* **19**, 876-883, doi:10.1038/nsmb.2379 (2012).

127 Soss, S. E., Klevit, R. E. & Chazin, W. J. Activation of UbcH5c~Ub is the result of a shift in interdomain motions of the conjugate bound to U-box E3 ligase E4B. *Biochemistry* **52**, 2991-2999, doi:10.1021/bi3015949 (2013).

128 Kamadurai, H. B. *et al.* Insights into ubiquitin transfer cascades from a structure of a UbcH5B approximately ubiquitin-HECT(NEDD4L) complex. *Mol Cell* **36**, 1095-1102, doi:10.1016/j.molcel.2009.11.010 (2009).

129 Lechtenberg, B. C. *et al.* Structure of a HOIP/E2~ubiquitin complex reveals

RBR E3 ligase mechanism and regulation. *Nature* **529**, 546-550, doi:10.1038/nature16511 (2016).

130 Buetow, L. *et al.* Activation of a primed RING E3-E2-ubiquitin complex by non-covalent ubiquitin. *Mol Cell* **58**, 297-310, doi:10.1016/j.molcel.2015.02.017 (2015).

131 Ranaweera, R. S. & Yang, X. Auto-ubiquitination of Mdm2 enhances its substrate ubiquitin ligase activity. *J Biol Chem* **288**, 18939-18946, doi:10.1074/jbc.M113.454470 (2013).

132 Sakata, E. *et al.* Crystal structure of UbcH5b~ubiquitin intermediate: insight into the formation of the self-assembled E2~Ub conjugates. *Structure* **18**, 138-147, doi:10.1016/j.str.2009.11.007 (2010).

133 Nguyen, L. *et al.* The ubiquitin-conjugating enzyme, UbcM2, is restricted to monoubiquitylation by a two-fold mechanism that involves backside residues of E2 and Lys48 of ubiquitin. *Biochemistry* **53**, 4004-4014, doi:10.1021/bi500072v (2014).

134 Machida, Y. J. *et al.* UBE2T is the E2 in the Fanconi anemia pathway and undergoes negative autoregulation. *Mol Cell* **23**, 589-596, doi:10.1016/j.molcel.2006.06.024 (2006).

135 Schumacher, F. R., Wilson, G. & Day, C. L. The N-terminal extension of UBE2E ubiquitin-conjugating enzymes limits chain assembly. *J Mol Biol* **425**, 4099-4111, doi:10.1016/j.jmb.2013.06.039 (2013).

136 Freemont, P. S., Hanson, I. M. & Trowsdale, J. A novel cysteine-rich sequence motif. *Cell* **64**, 483-484 (1991).

137 Buetow, L. & Huang, D. T. Structural insights into the catalysis and regulation of E3 ubiquitin ligases. *Nat Rev Mol Cell Biol* **17**, 626-642, doi:10.1038/nrm.2016.91 (2016).

138 Ohi, M. D., Vander Kooi, C. W., Rosenberg, J. A., Chazin, W. J. & Gould, K. L. Structural insights into the U-box, a domain associated with multi-ubiquitination. *Nat Struct Biol* **10**, 250-255, doi:10.1038/nsb906 (2003).

139 Aravind, L. & Koonin, E. V. The U box is a modified RING finger - a common domain in ubiquitination. *Curr Biol* **10**, R132-134 (2000).

140 Li, W. *et al.* Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling. *PLoS One* **3**, e1487, doi:10.1371/journal.pone.0001487 (2008).

141 Dominguez, C. *et al.* Structural model of the UbcH5B/CNOT4 complex revealed by combining NMR, mutagenesis, and docking approaches. *Structure* **12**, 633-644, doi:10.1016/j.str.2004.03.004 (2004).

142 Mace, P. D. *et al.* Structures of the cIAP2 RING domain reveal conformational changes associated with ubiquitin-conjugating enzyme (E2) recruitment. *J Biol Chem* **283**, 31633-31640, doi:10.1074/jbc.M804753200 (2008).

143 Liew, C. W., Sun, H., Hunter, T. & Day, C. L. RING domain dimerization is essential for RNF4 function. *Biochem J* **431**, 23-29, doi:10.1042/BJ20100957 (2010).

144 Yin, Q. *et al.* E2 interaction and dimerization in the crystal structure of TRAF6. *Nat Struct Mol Biol* **16**, 658-666, doi:10.1038/nsmb.1605 (2009).

145 Xu, Z. *et al.* Interactions between the quality control ubiquitin ligase CHIP and ubiquitin conjugating enzymes. *BMC Struct Biol* **8**, 26, doi:10.1186/1472-6807-8-26 (2008).

146 Vander Kooi, C. W. *et al.* The Prp19 U-box crystal structure suggests a common dimeric architecture for a class of oligomeric E3 ubiquitin ligases. *Biochemistry* **45**, 121-130, doi:10.1021/bi051787e (2006).

147 de Moura, T. R. *et al.* Prp19/Pso4 Is an Autoinhibited Ubiquitin Ligase Activated by Stepwise Assembly of Three Splicing Factors. *Mol Cell* **69**, 979-992 e976, doi:10.1016/j.molcel.2018.02.022 (2018).

148 Zhang, M. *et al.* Chaperoned ubiquitylation--crystal structures of the CHIP U box E3 ubiquitin ligase and a CHIP-Ubc13-Uev1a complex. *Mol Cell* **20**, 525-538, doi:10.1016/j.molcel.2005.09.023 (2005).

149 Joukov, V., Chen, J., Fox, E. A., Green, J. B. & Livingston, D. M. Functional communication between endogenous BRCA1 and its partner, BARD1, during Xenopus laevis development. *Proc Natl Acad Sci U S A* **98**, 12078-12083, doi:10.1073/pnas.211427098 (2001).

150 Brzovic, P. S., Rajagopal, P., Hoyt, D. W., King, M. C. & Klevit, R. E. Structure of a BRCA1-BARD1 heterodimeric RING-RING complex. *Nat Struct Biol* **8**, 833-837, doi:10.1038/nsb1001-833 (2001).

151 Linke, K. *et al.* Structure of the MDM2/MDMX RING domain heterodimer reveals dimerization is required for their ubiquitylation in trans. *Cell Death Differ* **15**, 841-848, doi:10.1038/sj.cdd.4402309 (2008).

152 Lydeard, J. R., Schulman, B. A. & Harper, J. W. Building and remodelling Cullin-RING E3 ubiquitin ligases. *EMBO Rep* **14**, 1050-1061, doi:10.1038/embor.2013.173 (2013).

153 Petroski, M. D. & Deshaies, R. J. Function and regulation of cullin-RING ubiquitin ligases. *Nat Rev Mol Cell Biol* **6**, 9-20, doi:10.1038/nrm1547 (2005).

154 Zheng, N. *et al.* Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin

ligase complex. *Nature* **416**, 703-709, doi:10.1038/416703a (2002).

155 Schulman, B. A. *et al.* Insights into SCF ubiquitin ligases from the structure of the Skp1-Skp2 complex. *Nature* **408**, 381-386, doi:10.1038/35042620 (2000).

156 Pruneda, J. N., Stoll, K. E., Bolton, L. J., Brzovic, P. S. & Klevit, R. E. Ubiquitin in motion: structural studies of the ubiquitin-conjugating enzyme approximately ubiquitin conjugate. *Biochemistry* **50**, 1624-1633, doi:10.1021/bi101913m (2011).

157 Saha, A., Lewis, S., Kleiger, G., Kuhlman, B. & Deshaies, R. J. Essential role for ubiquitin-ubiquitin-conjugating enzyme interaction in ubiquitin discharge from Cdc34 to substrate. *Mol Cell* **42**, 75-83, doi:10.1016/j.molcel.2011.03.016 (2011).

158 Plechanovova, A. *et al.* Mechanism of ubiquitylation by dimeric RING ligase RNF4. *Nat Struct Mol Biol* **18**, 1052-1059, doi:10.1038/nsmb.2108 (2011).

159 Pruneda, J. N. *et al.* Structure of an E3:E2~Ub complex reveals an allosteric mechanism shared among RING/U-box ligases. *Mol Cell* **47**, 933-942, doi:10.1016/j.molcel.2012.07.001 (2012).

160 Hamilton, K. S. *et al.* Structure of a conjugating enzyme-ubiquitin thiolester intermediate reveals a novel role for the ubiquitin tail. *Structure* **9**, 897-904 (2001).

161 Wickliffe, K. E., Lorenz, S., Wemmer, D. E., Kuriyan, J. & Rape, M. The mechanism of linkage-specific ubiquitin chain elongation by a single-subunit E2. *Cell* **144**, 769-781, doi:10.1016/j.cell.2011.01.035 (2011).

162 Dou, H. *et al.* Structural basis for autoinhibition and phosphorylation-dependent activation of c-Cbl. *Nat Struct Mol Biol* **19**, 184-192, doi:10.1038/nsmb.2231 (2012).

163 Dou, H., Buetow, L., Sibbet, G. J., Cameron, K. & Huang, D. T. Essentiality of a non-RING element in priming donor ubiquitin for catalysis by a monomeric E3. *Nat Struct Mol Biol* **20**, 982-986, doi:10.1038/nsmb.2621 (2013).

164 Ben-Saadon, R., Zaaroor, D., Ziv, T. & Ciechanover, A. The polycomb protein Ring1B generates self atypical mixed ubiquitin chains required for its in vitro histone H2A ligase activity. *Mol Cell* **24**, 701-711, doi:10.1016/j.molcel.2006.10.022 (2006).

165 Saha, A. & Deshaies, R. J. Multimodal activation of the ubiquitin ligase SCF by Nedd8 conjugation. *Mol Cell* **32**, 21-31, doi:10.1016/j.molcel.2008.08.021 (2008).

166 Deshaies, R. J., Emberley, E. D. & Saha, A. Control of cullin-ring ubiquitin ligase activity by nedd8. *Subcell Biochem* **54**, 41-56, doi:10.1007/978-1-4419-6676-6_4 (2010).

167 Angers, S. *et al.* Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. *Nature* **443**, 590-593, doi:10.1038/nature05175 (2006).

168 Huibregtse, J. M., Scheffner, M., Beaudenon, S. & Howley, P. M. A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. *Proc Natl Acad Sci U S A* **92**, 2563-2567 (1995).

169 Wang, M. & Pickart, C. M. Different HECT domain ubiquitin ligases employ distinct mechanisms of polyubiquitin chain synthesis. *EMBO J* **24**, 4324-4333, doi:10.1038/sj.emboj.7600895 (2005).

170 Lorenz, S. Structural mechanisms of HECT-type ubiquitin ligases. *Biol Chem* **399**, 127-145, doi:10.1515/hsz-2017-0184 (2018).

171 Ingham, R. J., Gish, G. & Pawson, T. The Nedd4 family of E3 ubiquitin ligases: functional diversity within a common modular architecture. *Oncogene* **23**, 1972-1984, doi:10.1038/sj.onc.1207436 (2004).

172 Rotin, D. & Kumar, S. Physiological functions of the HECT family of ubiquitin ligases. *Nat Rev Mol Cell Biol* **10**, 398-409, doi:10.1038/nrm2690 (2009).

173 Coussens, L. *et al.* Multiple, distinct forms of bovine and human protein kinase C suggest diversity in cellular signaling pathways. *Science* **233**, 859-866 (1986).

174 Clark, J. D. *et al.* A novel arachidonic acid-selective cytosolic PLA2 contains a Ca(2+)-dependent translocation domain with homology to PKC and GAP. *Cell* **65**, 1043-1051 (1991).

175 Perin, M. S., Brose, N., Jahn, R. & Sudhof, T. C. Domain structure of synaptotagmin (p65). *J Biol Chem* **266**, 623-629 (1991).

176 Sutton, R. B., Davletov, B. A., Berghuis, A. M., Sudhof, T. C. & Sprang, S. R. Structure of the first C2 domain of synaptotagmin I: a novel Ca2+/phospholipid-binding fold. *Cell* **80**, 929-938 (1995).

177 Essen, L. O., Perisic, O., Cheung, R., Katan, M. & Williams, R. L. Crystal structure of a mammalian phosphoinositide-specific phospholipase C delta. *Nature* **380**, 595-602, doi:10.1038/380595a0 (1996).

178 Plant, P. J., Yeger, H., Staub, O., Howard, P. & Rotin, D. The C2 domain of the ubiquitin protein ligase Nedd4 mediates Ca2+-dependent plasma membrane localization. *J Biol Chem* **272**, 32329-32336 (1997).

179 Zarrinpar, A. & Lim, W. A. Converging on proline: the mechanism of WW

domain peptide recognition. *Nat Struct Biol* **7**, 611-613, doi:10.1038/77891 (2000).

180 Macias, M. J., Wiesner, S. & Sudol, M. WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. *FEBS Lett* **513**, 30-37 (2002).

181 Chen, H. I. & Sudol, M. The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules. *Proc Natl Acad Sci U S A* **92**, 7819-7823 (1995).

182 Bedford, M. T., Chan, D. C. & Leder, P. FBP WW domains and the Abl SH3 domain bind to a specific class of proline-rich ligands. *EMBO J* **16**, 2376-2383, doi:10.1093/emboj/16.9.2376 (1997).

183 Lu, P. J., Zhou, X. Z., Shen, M. & Lu, K. P. Function of WW domains as phosphoserine- or phosphothreonine-binding modules. *Science* **283**, 1325-1328 (1999).

184 Persaud, A. *et al.* Comparison of substrate specificity of the ubiquitin ligases Nedd4 and Nedd4-2 using proteome arrays. *Mol Syst Biol* **5**, 333, doi:10.1038/msb.2009.85 (2009).

185 Fajner, V., Maspero, E. & Polo, S. Targeting HECT-type E3 ligases - insights from catalysis, regulation and inhibitors. *FEBS Lett* **591**, 2636-2647, doi:10.1002/1873-3468.12775 (2017).

186 Yoshida, M. *et al.* Poly(A) binding protein (PABP) homeostasis is mediated by the stability of its inhibitor, Paip2. *EMBO J* **25**, 1934-1944, doi:10.1038/sj.emboj.7601079 (2006).

187 Huang, L. *et al.* Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade. *Science* **286**, 1321-1326 (1999).

188 Salvat, C., Wang, G., Dastur, A., Lyon, N. & Huibregtse, J. M. The -4 phenylalanine is required for substrate ubiquitination catalyzed by HECT ubiquitin ligases. *J Biol Chem* **279**, 18935-18943, doi:10.1074/jbc.M312201200 (2004).

189 Kamadurai, H. B. *et al.* Mechanism of ubiquitin ligation and lysine prioritization by a HECT E3. *Elife* **2**, e00828, doi:10.7554/eLife.00828 (2013).

190 Nuber, U. & Scheffner, M. Identification of determinants in E2 ubiquitin-conjugating enzymes required for hect E3 ubiquitin-protein ligase interaction. *J Biol Chem* **274**, 7576-7582 (1999).

191 Fotia, A. B., Cook, D. I. & Kumar, S. The ubiquitin-protein ligases Nedd4

and Nedd4-2 show similar ubiquitin-conjugating enzyme specificities. *Int J Biochem Cell Biol* **38**, 472-479, doi:10.1016/j.biocel.2005.11.006 (2006).

192 Debonneville, C. & Staub, O. Participation of the ubiquitin-conjugating enzyme UBE2E3 in Nedd4-2-dependent regulation of the epithelial Na+ channel. *Mol Cell Biol* **24**, 2397-2409 (2004).

193 Maspero, E. *et al.* Structure of a ubiquitin-loaded HECT ligase reveals the molecular basis for catalytic priming. *Nat Struct Mol Biol* **20**, 696-701, doi:10.1038/nsmb.2566 (2013).

194 Maspero, E. *et al.* Structure of the HECT:ubiquitin complex and its role in ubiquitin chain elongation. *EMBO Rep* **12**, 342-349, doi:10.1038/embor.2011.21 (2011).

195 Kim, H. C., Steffen, A. M., Oldham, M. L., Chen, J. & Huibregtse, J. M. Structure and function of a HECT domain ubiquitin-binding site. *EMBO Rep* **12**, 334-341, doi:10.1038/embor.2011.23 (2011).

196 Kim, H. C. & Huibregtse, J. M. Polyubiquitination by HECT E3s and the determinants of chain type specificity. *Mol Cell Biol* **29**, 3307-3318, doi:10.1128/MCB.00240-09 (2009).

197 Mund, T. & Pelham, H. R. Control of the activity of WW-HECT domain E3 ubiquitin ligases by NDFIP proteins. *EMBO Rep* **10**, 501-507, doi:10.1038/embor.2009.30 (2009).

198 Riling, C. *et al.* Itch WW Domains Inhibit Its E3 Ubiquitin Ligase Activity by Blocking E2-E3 Ligase Trans-thiolation. *J Biol Chem* **290**, 23875-23887, doi:10.1074/jbc.M115.649269 (2015).

199 Gallagher, E., Gao, M., Liu, Y. C. & Karin, M. Activation of the E3 ubiquitin ligase Itch through a phosphorylation-induced conformational change. *Proc Natl Acad Sci U S A* **103**, 1717-1722, doi:10.1073/pnas.0510664103 (2006).

200 Wiesner, S. *et al.* Autoinhibition of the HECT-type ubiquitin ligase Smurf2 through its C2 domain. *Cell* **130**, 651-662, doi:10.1016/j.cell.2007.06.050 (2007).

201 Kavsak, P. *et al.* Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF beta receptor for degradation. *Mol Cell* **6**, 1365-1375 (2000).

202 Ogunjimi, A. A. *et al.* Regulation of Smurf2 ubiquitin ligase activity by anchoring the E2 to the HECT domain. *Mol Cell* **19**, 297-308, doi:10.1016/j.molcel.2005.06.028 (2005).

203 Aragon, E. *et al.* Structural basis for the versatile interactions of Smad7 with regulator WW domains in TGF-beta Pathways. *Structure* **20**, 1726-

1736, doi:10.1016/j.str.2012.07.014 (2012).

204 Chen, Z. *et al.* A Tunable Brake for HECT Ubiquitin Ligases. *Mol Cell* **66**, 345-357 e346, doi:10.1016/j.molcel.2017.03.020 (2017).

205 Zhu, K. *et al.* Allosteric auto-inhibition and activation of the Nedd4 family E3 ligase Itch. *EMBO Rep* **18**, 1618-1630, doi:10.15252/embr.201744454 (2017).

206 Wang, J. *et al.* Calcium activates Nedd4 E3 ubiquitin ligases by releasing the C2 domain-mediated auto-inhibition. *J Biol Chem* **285**, 12279-12288, doi:10.1074/jbc.M109.086405 (2010).

207 Escobedo, A. *et al.* Structural basis of the activation and degradation mechanisms of the E3 ubiquitin ligase Nedd4L. *Structure* **22**, 1446-1457, doi:10.1016/j.str.2014.08.016 (2014).

208 French, M. E., Kretzmann, B. R. & Hicke, L. Regulation of the RSP5 ubiquitin ligase by an intrinsic ubiquitin-binding site. *J Biol Chem* **284**, 12071-12079, doi:10.1074/jbc.M901106200 (2009).

209 French, M. E. *et al.* Mechanism of ubiquitin chain synthesis employed by a HECT domain ubiquitin ligase. *J Biol Chem* **292**, 10398-10413, doi:10.1074/jbc.M117.789479 (2017).

210 Ogunjimi, A. A. *et al.* The ubiquitin binding region of the Smurf HECT domain facilitates polyubiquitylation and binding of ubiquitylated substrates. *J Biol Chem* **285**, 6308-6315, doi:10.1074/jbc.M109.044537 (2010).

211 Kathman, S. G. *et al.* A Small Molecule That Switches a Ubiquitin Ligase From a Processive to a Distributive Enzymatic Mechanism. *J Am Chem Soc* **137**, 12442-12445, doi:10.1021/jacs.5b06839 (2015).

212 Zhang, W. *et al.* System-Wide Modulation of HECT E3 Ligases with Selective Ubiquitin Variant Probes. *Mol Cell* **62**, 121-136, doi:10.1016/j.molcel.2016.02.005 (2016).

213 Kuhnle, S. *et al.* Physical and functional interaction of the HECT ubiquitin-protein ligases E6AP and HERC2. *J Biol Chem* **286**, 19410-19416, doi:10.1074/jbc.M110.205211 (2011).

214 Huibregtse, J. M., Scheffner, M. & Howley, P. M. A cellular protein mediates association of p53 with the E6 oncoprotein of human papillomavirus types 16 or 18. *EMBO J* **10**, 4129-4135 (1991).

215 Martinez-Zapien, D. *et al.* Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. *Nature* **529**, 541-545, doi:10.1038/nature16481 (2016).

216 Ronchi, V. P., Klein, J. M., Edwards, D. J. & Haas, A. L. The active form of

E6-associated protein (E6AP)/UBE3A ubiquitin ligase is an oligomer. *J Biol Chem* **289**, 1033-1048, doi:10.1074/jbc.M113.517805 (2014).

217 Sander, B., Xu, W., Eilers, M., Popov, N. & Lorenz, S. A conformational switch regulates the ubiquitin ligase HUWE1. *Elife* **6**, doi:10.7554/eLife.21036 (2017).

218 Cao, X. R. *et al.* Nedd4 controls animal growth by regulating IGF-1 signaling. *Sci Signal* **1**, ra5, doi:10.1126/scisignal.1160940 (2008).

219 Kawabe, H. *et al.* Regulation of Rap2A by the ubiquitin ligase Nedd4-1 controls neurite development. *Neuron* **65**, 358-372, doi:10.1016/j.neuron.2010.01.007 (2010).

220 Drinjakovic, J. *et al.* E3 ligase Nedd4 promotes axon branching by downregulating PTEN. *Neuron* **65**, 341-357, doi:10.1016/j.neuron.2010.01.017 (2010).

221 Fouladkou, F. *et al.* The ubiquitin ligase Nedd4-1 is required for heart development and is a suppressor of thrombospondin-1. *J Biol Chem* **285**, 6770-6780, doi:10.1074/jbc.M109.082347 (2010).

222 Boase, N. A. *et al.* Respiratory distress and perinatal lethality in Nedd4-2-deficient mice. *Nat Commun* **2**, 287, doi:10.1038/ncomms1284 (2011).

223 Zhi, X. & Chen, C. WWP1: a versatile ubiquitin E3 ligase in signaling and diseases. *Cell Mol Life Sci* **69**, 1425-1434, doi:10.1007/s00018-011-0871-7 (2012).

224 Zhao, L. *et al.* Tumor necrosis factor inhibits mesenchymal stem cell differentiation into osteoblasts via the ubiquitin E3 ligase Wwp1. *Stem Cells* **29**, 1601-1610, doi:10.1002/stem.703 (2011).

225 Xu, H. *et al.* WWP2 promotes degradation of transcription factor OCT4 in human embryonic stem cells. *Cell Res* **19**, 561-573, doi:10.1038/cr.2009.31 (2009).

226 Yamashita, M. *et al.* Ubiquitin ligase Smurf1 controls osteoblast activity and bone homeostasis by targeting MEKK2 for degradation. *Cell* **121**, 101-113, doi:10.1016/j.cell.2005.01.035 (2005).

227 Bekker-Jensen, S. *et al.* HERC2 coordinates ubiquitin-dependent assembly of DNA repair factors on damaged chromosomes. *Nat Cell Biol* **12**, 80-86; sup pp 81-12, doi:10.1038/ncb2008 (2010).

228 Wu, W. *et al.* HERC2 is an E3 ligase that targets BRCA1 for degradation. *Cancer Res* **70**, 6384-6392, doi:10.1158/0008-5472.CAN-10-1304 (2010).

229 Yi, J. J. *et al.* An Autism-Linked Mutation Disables Phosphorylation Control of UBE3A. *Cell* **162**, 795-807, doi:10.1016/j.cell.2015.06.045 (2015).

230 Yang, Y. *et al.* E3 ubiquitin ligase Mule ubiquitinates Miz1 and is required

for TNFalpha-induced JNK activation. *Proc Natl Acad Sci U S A* **107**, 13444-13449, doi:10.1073/pnas.0913690107 (2010).

231 Zhong, Q., Gao, W., Du, F. & Wang, X. Mule/ARF-BP1, a BH3-only E3 ubiquitin ligase, catalyzes the polyubiquitination of Mcl-1 and regulates apoptosis. *Cell* **121**, 1085-1095, doi:10.1016/j.cell.2005.06.009 (2005).

232 Spratt, D. E., Walden, H. & Shaw, G. S. RBR E3 ubiquitin ligases: new structures, new insights, new questions. *Biochem J* **458**, 421-437, doi:10.1042/BJ20140006 (2014).

233 van der Reijden, B. A., Erpelinck-Verschueren, C. A., Lowenberg, B. & Jansen, J. H. TRIADs: a new class of proteins with a novel cysteine-rich signature. *Protein Sci* **8**, 1557-1561, doi:10.1110/ps.8.7.1557 (1999).

234 Chaugule, V. K. *et al.* Autoregulation of Parkin activity through its ubiquitin-like domain. *EMBO J* **30**, 2853-2867, doi:10.1038/emboj.2011.204 (2011).

235 Kelsall, I. R. *et al.* TRIAD1 and HHARI bind to and are activated by distinct neddylated Cullin-RING ligase complexes. *EMBO J* **32**, 2848-2860, doi:10.1038/emboj.2013.209 (2013).

236 Wauer, T. & Komander, D. Structure of the human Parkin ligase domain in an autoinhibited state. *EMBO J* **32**, 2099-2112, doi:10.1038/emboj.2013.125 (2013).

237 Duda, D. M. *et al.* Structure of HHARI, a RING-IBR-RING ubiquitin ligase: autoinhibition of an Ariadne-family E3 and insights into ligation mechanism. *Structure* **21**, 1030-1041, doi:10.1016/j.str.2013.04.019 (2013).

238 Dove, K. K. & Klevit, R. E. RING-Between-RING E3 Ligases: Emerging Themes amid the Variations. *J Mol Biol* **429**, 3363-3375, doi:10.1016/j.jmb.2017.08.008 (2017).

239 Stieglitz, B. *et al.* Structural basis for ligase-specific conjugation of linear ubiquitin chains by HOIP. *Nature* **503**, 422-426, doi:10.1038/nature12638 (2013).

240 Dove, K. K. *et al.* Structural Studies of HHARI/UbcH7 approximately Ub Reveal Unique E2 approximately Ub Conformational Restriction by RBR RING1. *Structure* **25**, 890-900 e895, doi:10.1016/j.str.2017.04.013 (2017).

241 Smit, J. J. *et al.* The E3 ligase HOIP specifies linear ubiquitin chain assembly through its RING-IBR-RING domain and the unique LDD extension. *EMBO J* **31**, 3833-3844, doi:10.1038/emboj.2012.217 (2012).

242 Stieglitz, B., Morris-Davies, A. C., Koliopoulos, M. G., Christodoulou, E. & Rittinger, K. LUBAC synthesizes linear ubiquitin chains via a thioester

intermediate. *EMBO Rep* **13**, 840-846, doi:10.1038/embor.2012.105 (2012).

243 Dove, K. K., Stieglitz, B., Duncan, E. D., Rittinger, K. & Klevit, R. E. Molecular insights into RBR E3 ligase ubiquitin transfer mechanisms. *EMBO Rep* **17**, 1221-1235, doi:10.15252/embr.201642641 (2016).

244 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, doi:10.1038/s41467-017-00272-6 (2017).

245 Trempe, J. F. *et al.* Structure of parkin reveals mechanisms for ubiquitin ligase activation. *Science* **340**, 1451-1455, doi:10.1126/science.1237908 (2013).

246 Fiesel, F. C. *et al.* Structural and Functional Impact of Parkinson Disease-Associated Mutations in the E3 Ubiquitin Ligase Parkin. *Hum Mutat* **36**, 774-786, doi:10.1002/humu.22808 (2015).

247 Kumar, A. *et al.* Parkin-phosphoubiquitin complex reveals cryptic ubiquitin-binding site required for RBR ligase activity. *Nat Struct Mol Biol* **24**, 475-483, doi:10.1038/nsmb.3400 (2017).

248 Jansen, E. F., Curl, A. L. & Balls, A. K. Reaction of alpha-chymotrypsin with analogues of diisopropyl fluorophosphate. *J Biol Chem* **190**, 557-561 (1951).

249 Kazlauskaite, A. *et al.* Phosphorylation of Parkin at Serine65 is essential for activation: elaboration of a Miro1 substrate-based assay of Parkin E3 ligase activity. *Open Biol* **4**, 130213, doi:10.1098/rsob.130213 (2014).

250 Sauve, V. *et al.* A Ubl/ubiquitin switch in the activation of Parkin. *EMBO J* **34**, 2492-2505, doi:10.15252/embj.201592237 (2015).

251 Clark, I. E. *et al.* Drosophila pink1 is required for mitochondrial function and interacts genetically with parkin. *Nature* **441**, 1162-1166, doi:10.1038/nature04779 (2006).

252 Park, J. *et al.* Mitochondrial dysfunction in Drosophila PINK1 mutants is complemented by parkin. *Nature* **441**, 1157-1161, doi:10.1038/nature04788 (2006).

253 Winklhofer, K. F. & Haass, C. Mitochondrial dysfunction in Parkinson's disease. *Biochim Biophys Acta* **1802**, 29-44, doi:10.1016/j.bbadis.2009.08.013 (2010).

254 Yamano, K. & Youle, R. J. PINK1 is degraded through the N-end rule pathway. *Autophagy* **9**, 1758-1769, doi:10.4161/auto.24633 (2013).

255 Narendra, D., Tanaka, A., Suen, D. F. & Youle, R. J. Parkin is recruited selectively to impaired mitochondria and promotes their autophagy. *J Cell*

Biol **183**, 795-803, doi:10.1083/jcb.200809125 (2008).

256 Narendra, D. P. *et al.* PINK1 is selectively stabilized on impaired mitochondria to activate Parkin. *PLoS Biol* **8**, e1000298, doi:10.1371/journal.pbio.1000298 (2010).

257 Vives-Bauza, C. *et al.* PINK1-dependent recruitment of Parkin to mitochondria in mitophagy. *Proc Natl Acad Sci U S A* **107**, 378-383, doi:10.1073/pnas.0911187107 (2010).

258 Kazlauskaite, A. *et al.* Binding to serine 65-phosphorylated ubiquitin primes Parkin for optimal PINK1-dependent phosphorylation and activation. *EMBO Rep* **16**, 939-954, doi:10.15252/embr.201540352 (2015).

259 Tang, M. Y. *et al.* Structure-guided mutagenesis reveals a hierarchical mechanism of Parkin activation. *Nat Commun* **8**, 14697, doi:10.1038/ncomms14697 (2017).

260 Yagi, H. *et al.* A non-canonical UBA-UBL interaction forms the linear-ubiquitin-chain assembly complex. *EMBO Rep* **13**, 462-468, doi:10.1038/embor.2012.24 (2012).

261 Liu, J. *et al.* Structural Insights into SHARPIN-Mediated Activation of HOIP for the Linear Ubiquitin Chain Assembly. *Cell Rep* **21**, 27-36, doi:10.1016/j.celrep.2017.09.031 (2017).

262 Elliott, P. R. *et al.* Molecular basis and regulation of OTULIN-LUBAC interaction. *Mol Cell* **54**, 335-348, doi:10.1016/j.molcel.2014.03.018 (2014).

263 Elliott, P. R. & Komander, D. Regulation of Met1-linked polyubiquitin signalling by the deubiquitinase OTULIN. *FEBS J* **283**, 39-53, doi:10.1111/febs.13547 (2016).

264 Elliott, P. R. *et al.* SPATA2 Links CYLD to LUBAC, Activates CYLD, and Controls LUBAC Signaling. *Mol Cell* **63**, 990-1005, doi:10.1016/j.molcel.2016.08.001 (2016).

265 Scott, D. C. *et al.* Two Distinct Types of E3 Ligases Work in Unison to Regulate Substrate Ubiquitylation. *Cell* **166**, 1198-1214 e1124, doi:10.1016/j.cell.2016.07.027 (2016).

266 Bernardini, J. P., Lazarou, M. & Dewson, G. Parkin and mitophagy in cancer. *Oncogene* **36**, 1315-1327, doi:10.1038/onc.2016.302 (2017).

267 Harper, J. W., Ordureau, A. & Heo, J. M. Building and decoding ubiquitin chains for mitophagy. *Nat Rev Mol Cell Biol* **19**, 93-108, doi:10.1038/nrm.2017.129 (2018).

268 Imai, Y. *et al.* An unfolded putative transmembrane polypeptide, which can lead to endoplasmic reticulum stress, is a substrate of Parkin. *Cell* **105**,

891-902 (2001).

269 Imai, Y. *et al.* CHIP is associated with Parkin, a gene responsible for familial Parkinson's disease, and enhances its ubiquitin ligase activity. *Mol Cell* **10**, 55-67 (2002).

270 Gupta, A., Anjomani-Virmouni, S., Koundouros, N. & Poulogiannis, G. PARK2 loss promotes cancer progression via redox-mediated inactivation of PTEN. *Mol Cell Oncol* **4**, e1329692, doi:10.1080/23723556.2017.1329692 (2017).

271 Wahabi, K., Perwez, A. & Rizvi, M. A. Parkin in Parkinson's Disease and Cancer: a Double-Edged Sword. *Mol Neurobiol*, doi:10.1007/s12035-018-0879-1 (2018).

272 Tan, N. G. *et al.* Human homologue of ariadne promotes the ubiquitylation of translation initiation factor 4E homologous protein, 4EHP. *FEBS Lett* **554**, 501-504 (2003).

273 Elmehdawi, F. *et al.* Human Homolog of Drosophila Ariadne (HHARI) is a marker of cellular proliferation associated with nuclear bodies. *Exp Cell Res* **319**, 161-172, doi:10.1016/j.yexcr.2012.10.002 (2013).

274 Ho, S. R., Lee, Y. J. & Lin, W. C. Regulation of RNF144A E3 Ubiquitin Ligase Activity by Self-association through Its Transmembrane Domain. *J Biol Chem* **290**, 23026-23038, doi:10.1074/jbc.M115.645499 (2015).

275 Ho, S. R., Mahanic, C. S., Lee, Y. J. & Lin, W. C. RNF144A, an E3 ubiquitin ligase for DNA-PKcs, promotes apoptosis during DNA damage. *Proc Natl Acad Sci U S A* **111**, E2646-2655, doi:10.1073/pnas.1323107111 (2014).

276 Zhang, Y., Liao, X. H., Xie, H. Y., Shao, Z. M. & Li, D. Q. RBR-type E3 ubiquitin ligase RNF144A targets PARP1 for ubiquitin-dependent degradation and regulates PARP inhibitor sensitivity in breast cancer cells. *Oncotarget* **8**, 94505-94518, doi:10.18632/oncotarget.21784 (2017).

277 Zhang, Y. *et al.* Epigenetic silencing of RNF144A expression in breast cancer cells through promoter hypermethylation and MBD4. *Cancer Med*, doi:10.1002/cam4.1324 (2018).

278 Ariffin, J. K. *et al.* The E3 ubiquitin ligase RNF144B is LPS-inducible in human, but not mouse, macrophages and promotes inducible IL-1beta expression. *J Leukoc Biol* **100**, 155-161, doi:10.1189/jlb.2AB0815-339R (2016).

279 Lopez, J. & Tait, S. W. Killing the Killer: PARC/CUL9 promotes cell survival by destroying cytochrome C. *Sci Signal* **7**, pe17, doi:10.1126/scisignal.2005619 (2014).

280 Hishikawa, N. *et al.* Dorfin localizes to the ubiquitylated inclusions in

Parkinson's disease, dementia with Lewy bodies, multiple system atrophy, and amyotrophic lateral sclerosis. *Am J Pathol* **163**, 609-619 (2003).

281 Ito, T. *et al.* Dorfin localizes to Lewy bodies and ubiquitylates synphilin-1. *J Biol Chem* **278**, 29106-29114, doi:10.1074/jbc.M302763200 (2003).

282 Le Quesne, J. P., Spriggs, K. A., Bushell, M. & Willis, A. E. Dysregulation of protein synthesis and disease. *J Pathol* **220**, 140-151, doi:10.1002/path.2627 (2010).

283 Paul, S. Dysfunction of the ubiquitin-proteasome system in multiple disease conditions: therapeutic approaches. *Bioessays* **30**, 1172-1184, doi:10.1002/bies.20852 (2008).

284 Martinon, F., Mayor, A. & Tschopp, J. The inflammasomes: guardians of the body. *Annu Rev Immunol* **27**, 229-265, doi:10.1146/annurev.immunol.021908.132715 (2009).

285 Appella, E. & Anderson, C. W. Post-translational modifications and activation of p53 by genotoxic stresses. *Eur J Biochem* **268**, 2764-2772 (2001).

286 Giangrande, P. H., Kimbrel, E. A., Edwards, D. P. & McDonnell, D. P. The opposing transcriptional activities of the two isoforms of the human progesterone receptor are due to differential cofactor binding. *Mol Cell Biol* **20**, 3102-3115 (2000).

287 Twining, S. S. Regulation of proteolytic activity in tissues. *Critical reviews in biochemistry and molecular biology* **29**, 315-383, doi:10.3109/10409239409083484 (1994).

288 Sanman, L. E. & Bogyo, M. Activity-based profiling of proteases. *Annu Rev Biochem* **83**, 249-273, doi:10.1146/annurev-biochem-060713-035352 (2014).

289 Kisselev, A. F., van der Linden, W. A. & Overkleeft, H. S. Proteasome inhibitors: an expanding army attacking a unique target. *Chem Biol* **19**, 99-115, doi:10.1016/j.chembiol.2012.01.003 (2012).

290 Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* **77**, 383-414, doi:10.1146/annurev.biochem.75.101304.124125 (2008).

291 Wright, M. H. & Sieber, S. A. Chemical proteomics approaches for identifying the cellular targets of natural products. *Nat Prod Rep* **33**, 681-708, doi:10.1039/c6np00001k (2016).

292 Ostrowski, K. & Barnard, E. A. Application of isotopically labelled specific inhibitors as a method in enzyme cytochemistry. *Exp Cell Res* **25**, 465-468

(1961).

293 Liu, Y., Patricelli, M. P. & Cravatt, B. F. Activity-based protein profiling: the serine hydrolases. *Proc Natl Acad Sci U S A* **96**, 14694-14699 (1999).

294 Kobe, B. & Kemp, B. E. Active site-directed protein regulation. *Nature* **402**, 373-376, doi:10.1038/46478 (1999).

295 Tipper, D. J. & Strominger, J. L. Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. *Proc Natl Acad Sci U S A* **54**, 1133-1141 (1965).

296 Suginaka, H., Blumberg, P. M. & Strominger, J. L. Multiple penicillin-binding components in Bacillus subtilis, Bacillus cereus, Staphylococcus aureus, and Escherichia coli. *J Biol Chem* **247**, 5279-5288 (1972).

297 Iwaya, M., Goldman, R., Tipper, D. J., Feingold, B. & Strominger, J. L. Morphology of an Escherichia coli mutant with a temperature-dependent round cell shape. *J Bacteriol* **136**, 1143-1158 (1978).

298 Balls, A. K. & Jansen, E. F. Stoichiometric inhibition of chymotrypsin. *Adv Enzymol Relat Subj Biochem* **13**, 321-343 (1952).

299 Balls, A. K. & Jansen, E. F. Proteolytic enzymes. *Annu Rev Biochem* **21**, 1-28, doi:10.1146/annurev.bi.21.070152.000245 (1952).

300 Powers, J. C., Asgian, J. L., Ekici, O. D. & James, K. E. Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem Rev* **102**, 4639-4750 (2002).

301 Willems, L. I., Overkleeft, H. S. & van Kasteren, S. I. Current developments in activity-based protein profiling. *Bioconjug Chem* **25**, 1181-1191, doi:10.1021/bc500208y (2014).

302 Niphakis, M. J. & Cravatt, B. F. Enzyme inhibitor discovery by activity-based protein profiling. *Annu Rev Biochem* **83**, 341-377, doi:10.1146/annurev-biochem-060713-035708 (2014).

303 Hewings, D. S., Flygare, J. A., Bogyo, M. & Wertz, I. E. Activity-based probes for the ubiquitin conjugation-deconjugation machinery: new chemistries, new tools, and new insights. *FEBS J* **284**, 1555-1576, doi:10.1111/febs.14039 (2017).

304 Saghatelian, A., Jessani, N., Joseph, A., Humphrey, M. & Cravatt, B. F. Activity-based probes for the proteomic profiling of metalloproteases. *Proc Natl Acad Sci U S A* **101**, 10000-10005, doi:10.1073/pnas.0402784101 (2004).

305 Hemelaar, J. *et al.* Chemistry-based functional proteomics: mechanism-based activity-profiling tools for ubiquitin and ubiquitin-like specific proteases. *J Proteome Res* **3**, 268-276 (2004).

306 Borodovsky, A. *et al.* Chemistry-based functional proteomics reveals novel members of the deubiquitinating enzyme family. *Chem Biol* **9**, 1149-1159 (2002).

307 Kidd, D., Liu, Y. & Cravatt, B. F. Profiling serine hydrolase activities in complex proteomes. *Biochemistry* **40**, 4005-4015 (2001).

308 Chan, E. W., Chattopadhaya, S., Panicker, R. C., Huang, X. & Yao, S. Q. Developing photoactive affinity probes for proteomic profiling: hydroxamate-based probes for metalloproteases. *J Am Chem Soc* **126**, 14435-14446, doi:10.1021/ja047044i (2004).

309 Weihofen, A., Binns, K., Lemberg, M. K., Ashman, K. & Martoglio, B. Identification of signal peptide peptidase, a presenilin-type aspartic protease. *Science* **296**, 2215-2218, doi:10.1126/science.1070925 (2002).

310 Yang, P. & Liu, K. Activity-based protein profiling: recent advances in probe development and applications. *Chembiochem* **16**, 712-724, doi:10.1002/cbic.201402582 (2015).

311 Gygi, S. P. *et al.* Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* **17**, 994-999, doi:10.1038/13690 (1999).

312 Weerapana, E. *et al.* Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **468**, 790-795, doi:10.1038/nature09472 (2010).

313 Puri, A. W., Broz, P., Shen, A., Monack, D. M. & Bogyo, M. Caspase-1 activity is required to bypass macrophage apoptosis upon Salmonella infection. *Nat Chem Biol* **8**, 745-747, doi:10.1038/nchembio.1023 (2012).

314 Edgington, L. E., Verdoes, M. & Bogyo, M. Functional imaging of proteases: recent advances in the design and application of substrate-based and activity-based probes. *Curr Opin Chem Biol* **15**, 798-805, doi:10.1016/j.cbpa.2011.10.012 (2011).

315 Fonovic, M. & Bogyo, M. Activity based probes for proteases: applications to biomarker discovery, molecular imaging and drug screening. *Curr Pharm Des* **13**, 253-261 (2007).

316 Vickers, C. J., Gonzalez-Paez, G. E. & Wolan, D. W. Selective detection of caspase-3 versus caspase-7 using activity-based probes with key unnatural amino acids. *ACS Chem Biol* **8**, 1558-1566, doi:10.1021/cb400209w (2013).

317 Heinis, C., Rutherford, T., Freund, S. & Winter, G. Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nat Chem Biol* **5**, 502-507, doi:10.1038/nchembio.184 (2009).

318 Baeriswyl, V. *et al.* Development of a selective peptide macrocycle inhibitor of coagulation factor XII toward the generation of a safe antithrombotic therapy. *J Med Chem* **56**, 3742-3746, doi:10.1021/jm400236j (2013).

319 Fan, F. *et al.* Protein profiling of active cysteine cathepsins in living cells using an activity-based probe containing a cell-penetrating peptide. *J Proteome Res* **11**, 5763-5772, doi:10.1021/pr300575u (2012).

320 Pan, D. *et al.* A general strategy for developing cell-permeable photo-modulatable organic fluorescent probes for live-cell super-resolution imaging. *Nat Commun* **5**, 5573, doi:10.1038/ncomms6573 (2014).

321 Deshpande, P. P., Biswas, S. & Torchilin, V. P. Current trends in the use of liposomes for tumor targeting. *Nanomedicine (Lond)* **8**, 1509-1528, doi:10.2217/nnm.13.118 (2013).

322 Pao, K. C. *et al.* Probes of ubiquitin E3 ligases enable systematic dissection of parkin activation. *Nat Chem Biol* **12**, 324-331, doi:10.1038/nchembio.2045 (2016).

323 Kornahrens, A. F. *et al.* Design of Benzoxathiazin-3-one 1,1-Dioxides as a New Class of Irreversible Serine Hydrolase Inhibitors: Discovery of a Uniquely Selective PNPLA4 Inhibitor. *J Am Chem Soc* **139**, 7052-7061, doi:10.1021/jacs.7b02985 (2017).

324 Martell, J. & Weerapana, E. Applications of copper-catalyzed click chemistry in activity-based protein profiling. *Molecules* **19**, 1378-1393, doi:10.3390/molecules19021378 (2014).

325 Schiapparelli, L. M. *et al.* Direct detection of biotinylated proteins by mass spectrometry. *J Proteome Res* **13**, 3966-3978, doi:10.1021/pr5002862 (2014).

326 Tully, S. E. & Cravatt, B. F. Activity-based probes that target functional subclasses of phospholipases in proteomes. *J Am Chem Soc* **132**, 3264-3265, doi:10.1021/ja1000505 (2010).

327 Lahav, D. *et al.* A Fluorescence Polarization Activity-Based Protein Profiling Assay in the Discovery of Potent, Selective Inhibitors for Human Nonlysosomal Glucosylceramidase. *J Am Chem Soc* **139**, 14192-14197, doi:10.1021/jacs.7b07352 (2017).

328 Kawaguchi, M., Ikegawa, S., Ieda, N. & Nakagawa, H. A Fluorescent Probe for Imaging Sirtuin Activity in Living Cells, Based on One-Step Cleavage of the Dabcyl Quencher. *Chembiochem* **17**, 1961-1967, doi:10.1002/cbic.201600374 (2016).

329 Verdoes, M. *et al.* Improved quenched fluorescent probe for imaging of

cysteine cathepsin activity. *J Am Chem Soc* **135**, 14726-14730, doi:10.1021/ja4056068 (2013).

330 Blum, G., von Degenfeld, G., Merchant, M. J., Blau, H. M. & Bogyo, M. Noninvasive optical imaging of cysteine protease activity using fluorescently quenched activity-based probes. *Nat Chem Biol* **3**, 668-677, doi:10.1038/nchembio.2007.26 (2007).

331 Patricelli, M. P., Giang, D. K., Stamp, L. M. & Burbaum, J. J. Direct visualization of serine hydrolase activities in complex proteomes using fluorescent active site-directed probes. *Proteomics* **1**, 1067-1071, doi:10.1002/1615-9861(200109)1:9<1067::AID-PROT1067>3.0.CO;2-4 (2001).

332 Greenbaum, D. C. *et al.* Small molecule affinity fingerprinting. A tool for enzyme family subclassification, target identification, and inhibitor design. *Chem Biol* **9**, 1085-1094 (2002).

333 Ekkebus, R. *et al.* On terminal alkynes that can react with active-site cysteine nucleophiles in proteases. *J Am Chem Soc* **135**, 2867-2870, doi:10.1021/ja309802n (2013).

334 Stanley, M. *et al.* Orthogonal thiol functionalization at a single atomic center for profiling transthiolation activity of E1 activating enzymes. *ACS Chem Biol* **10**, 1542-1554, doi:10.1021/acschembio.5b00118 (2015).

335 Adam, G. C., Burbaum, J., Kozarich, J. W., Patricelli, M. P. & Cravatt, B. F. Mapping enzyme active sites in complex proteomes. *J Am Chem Soc* **126**, 1363-1368, doi:10.1021/ja038441g (2004).

336 Jessani, N. *et al.* A streamlined platform for high-content functional proteomics of primary human specimens. *Nat Methods* **2**, 691-697, doi:10.1038/nmeth778 (2005).

337 Okerberg, E. S. *et al.* High-resolution functional proteomics by active-site peptide profiling. *Proc Natl Acad Sci U S A* **102**, 4996-5001, doi:10.1073/pnas.0501205102 (2005).

338 Speers, A. E. & Cravatt, B. F. A tandem orthogonal proteolysis strategy for high-content chemical proteomics. *J Am Chem Soc* **127**, 10018-10019, doi:10.1021/ja0532842 (2005).

339 Kramer, H. B., Nicholson, B., Kessler, B. M. & Altun, M. Detection of ubiquitin-proteasome enzymatic activities in cells: application of activity-based probes to inhibitor development. *Biochim Biophys Acta* **1823**, 2029-2037, doi:10.1016/j.bbamcr.2012.05.014 (2012).

340 Galmozzi, A., Dominguez, E., Cravatt, B. F. & Saez, E. Application of activity-based protein profiling to study enzyme function in adipocytes.

Methods Enzymol **538**, 151-169, doi:10.1016/B978-0-12-800280-3.00009-8 (2014).

341 Jiang, X. *et al.* Sensitive and Accurate Quantitation of Phosphopeptides Using TMT Isobaric Labeling Technique. *J Proteome Res* **16**, 4244-4252, doi:10.1021/acs.jproteome.7b00610 (2017).

342 Ting, L., Rad, R., Gygi, S. P. & Haas, W. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nat Methods* **8**, 937-940, doi:10.1038/nmeth.1714 (2011).

343 Berberich, M. J., Paulo, J. A. & Everley, R. A. MS3-IDQ: Utilizing MS3 Spectra beyond Quantification Yields Increased Coverage of the Phosphoproteome in Isobaric Tag Experiments. *J Proteome Res* **17**, 1741-1747, doi:10.1021/acs.jproteome.8b00006 (2018).

344 Kemp, M. Recent Advances in the Discovery of Deubiquitinating Enzyme Inhibitors. *Prog Med Chem* **55**, 149-192, doi:10.1016/bs.pmch.2015.10.002 (2016).

345 Kategaya, L. *et al.* USP7 small-molecule inhibitors interfere with ubiquitin binding. *Nature* **550**, 534-538, doi:10.1038/nature24006 (2017).

346 Ulintz, P. J. *et al.* Comparison of MS(2)-only, MSA, and MS(2)/MS(3) methodologies for phosphopeptide identification. *J Proteome Res* **8**, 887-899, doi:10.1021/pr800535h (2009).

347 Wiedner, S. D. *et al.* Organelle-specific activity-based protein profiling in living cells. *Angew Chem Int Ed Engl* **53**, 2919-2922, doi:10.1002/anie.201309135 (2014).

348 Kato, D. *et al.* Activity-based probes that target diverse cysteine protease families. *Nat Chem Biol* **1**, 33-38, doi:10.1038/nchembio707 (2005).

349 Blankman, J. L., Simon, G. M. & Cravatt, B. F. A comprehensive profile of brain enzymes that hydrolyze the endocannabinoid 2-arachidonoylglycerol. *Chem Biol* **14**, 1347-1356, doi:10.1016/j.chembiol.2007.11.006 (2007).

350 Turk, B., Turk, D. & Turk, V. Protease signalling: the cutting edge. *EMBO J* **31**, 1630-1643, doi:10.1038/emboj.2012.42 (2012).

351 Ortega, C. *et al.* Systematic Survey of Serine Hydrolase Activity in Mycobacterium tuberculosis Defines Changes Associated with Persistence. *Cell Chem Biol* **23**, 290-298, doi:10.1016/j.chembiol.2016.01.003 (2016).

352 Gillet, L. C. *et al.* In-cell selectivity profiling of serine protease inhibitors by activity-based proteomics. *Mol Cell Proteomics* **7**, 1241-1253, doi:10.1074/mcp.M700505-MCP200 (2008).

353 Uhlmann, F., Wernic, D., Poupart, M. A., Koonin, E. V. & Nasmyth, K. Cleavage of cohesin by the CD clan protease separin triggers anaphase in yeast. *Cell* **103**, 375-386 (2000).

354 Borodovsky, A. *et al.* Small-molecule inhibitors and probes for ubiquitin- and ubiquitin-like-specific proteases. *Chembiochem* **6**, 287-291, doi:10.1002/cbic.200400236 (2005).

355 Sieber, S. A., Niessen, S., Hoover, H. S. & Cravatt, B. F. Proteomic profiling of metalloprotease activities with cocktails of active-site probes. *Nat Chem Biol* **2**, 274-281, doi:10.1038/nchembio781 (2006).

356 Patricelli, M. P. *et al.* Functional interrogation of the kinome using nucleotide acyl phosphates. *Biochemistry* **46**, 350-358, doi:10.1021/bi062142x (2007).

357 Yee, M. C., Fas, S. C., Stohlmeyer, M. M., Wandless, T. J. & Cimprich, K. A. A cell-permeable, activity-based probe for protein and lipid kinases. *J Biol Chem* **280**, 29053-29059, doi:10.1074/jbc.M504730200 (2005).

358 Wilke, K. E., Francis, S. & Carlson, E. E. Activity-based probe for histidine kinase signaling. *J Am Chem Soc* **134**, 9150-9153, doi:10.1021/ja3041702 (2012).

359 Garland, M. *et al.* Development of an activity-based probe for acyl-protein thioesterases. *PLoS One* **13**, e0190255, doi:10.1371/journal.pone.0190255 (2018).

360 Heal, W. P., Dang, T. H. & Tate, E. W. Activity-based probes: discovering new biology and new drug targets. *Chem Soc Rev* **40**, 246-257, doi:10.1039/c0cs00004c (2011).

361 Evans, M. J. & Cravatt, B. F. Mechanism-based profiling of enzyme families. *Chem Rev* **106**, 3279-3301, doi:10.1021/cr050288g (2006).

362 Mignatti, P. & Rifkin, D. B. Plasminogen activators and angiogenesis. *Curr Top Microbiol Immunol* **213 (Pt 1)**, 33-50 (1996).

363 Lane, R. M., Potkin, S. G. & Enz, A. Targeting acetylcholinesterase and butyrylcholinesterase in dementia. *Int J Neuropsychopharmacol* **9**, 101-124, doi:10.1017/S1461145705005833 (2006).

364 Yoshida, S. & Shiosaka, S. Plasticity-related serine proteases in the brain (review). *Int J Mol Med* **3**, 405-409 (1999).

365 Nomura, D. K. *et al.* Endocannabinoid hydrolysis generates brain prostaglandins that promote neuroinflammation. *Science* **334**, 809-813, doi:10.1126/science.1209200 (2011).

366 Long, J. Z. & Cravatt, B. F. The metabolic serine hydrolases and their functions in mammalian physiology and disease. *Chem Rev* **111**, 6022-

6063, doi:10.1021/cr200075y (2011).

367 DeClerck, Y. A. *et al.* Proteases and protease inhibitors in tumor progression. *Adv Exp Med Biol* **425**, 89-97 (1997).

368 Nomura, D. K., Dix, M. M. & Cravatt, B. F. Activity-based protein profiling for biochemical pathway discovery in cancer. *Nat Rev Cancer* **10**, 630-638, doi:10.1038/nrc2901 (2010).

369 Simon, G. M. & Cravatt, B. F. Activity-based proteomics of enzyme superfamilies: serine hydrolases as a case study. *J Biol Chem* **285**, 11051-11055, doi:10.1074/jbc.R109.097600 (2010).

370 Pineiro-Sanchez, M. L. *et al.* Identification of the 170-kDa melanoma membrane-bound gelatinase (seprase) as a serine integral membrane protease. *J Biol Chem* **272**, 7595-7601 (1997).

371 Glynn, P., Read, D. J., Guo, R., Wylie, S. & Johnson, M. K. Synthesis and characterization of a biotinylated organophosphorus ester for detection and affinity purification of a brain serine esterase: neuropathy target esterase. *Biochem J* **301 (Pt 2)**, 551-556 (1994).

372 Kam, C. M., Fujikawa, K. & Powers, J. C. Mechanism-based isocoumarin inhibitors for trypsin and blood coagulation serine proteases: new anticoagulants. *Biochemistry* **27**, 2547-2557 (1988).

373 Winkler, U. *et al.* Characterization, application and potential uses of biotin-tagged inhibitors for lymphocyte serine proteases (granzymes). *Mol Immunol* **33**, 615-623 (1996).

374 Bouma, B. N., Miles, L. A., Beretta, G. & Griffin, J. H. Human plasma prekallikrein. Studies of its activation by activated factor XII and of its inactivation by diisopropyl phosphofluoridate. *Biochemistry* **19**, 1151-1160 (1980).

375 Jessani, N. *et al.* Carcinoma and stromal enzyme activity profiles associated with breast tumor growth in vivo. *Proc Natl Acad Sci U S A* **101**, 13756-13761, doi:10.1073/pnas.0404727101 (2004).

376 Milner, J. M. *et al.* Fibroblast activation protein alpha is expressed by chondrocytes following a pro-inflammatory stimulus and is elevated in osteoarthritis. *Arthritis Res Ther* **8**, R23, doi:10.1186/ar1877 (2006).

377 Mahrus, S. & Craik, C. S. Selective chemical functional probes of granzymes A and B reveal granzyme B is a major effector of natural killer cell-mediated lysis of target cells. *Chem Biol* **12**, 567-577, doi:10.1016/j.chembiol.2005.03.006 (2005).

378 Nomura, D. K. *et al.* A brain detoxifying enzyme for organophosphorus nerve poisons. *Proc Natl Acad Sci U S A* **102**, 6195-6200,

doi:10.1073/pnas.0501915102 (2005).

379 Madsen, M. A., Deryugina, E. I., Niessen, S., Cravatt, B. F. & Quigley, J. P. Activity-based protein profiling implicates urokinase activation as a key step in human fibrosarcoma intravasation. *J Biol Chem* **281**, 15997-16005, doi:10.1074/jbc.M601223200 (2006).

380 Jessani, N., Liu, Y., Humphrey, M. & Cravatt, B. F. Enzyme activity profiles of the secreted and membrane proteome that depict cancer cell invasiveness. *Proc Natl Acad Sci U S A* **99**, 10335-10340, doi:10.1073/pnas.162187599 (2002).

381 Chiang, K. P., Niessen, S., Saghatelian, A. & Cravatt, B. F. An enzyme that regulates ether lipid signaling pathways in cancer annotated by multidimensional profiling. *Chem Biol* **13**, 1041-1050, doi:10.1016/j.chembiol.2006.08.008 (2006).

382 Lu, J., Xiao Yj, Y. J., Baudhuin, L. M., Hong, G. & Xu, Y. Role of ether-linked lysophosphatidic acids in ovarian cancer cells. *J Lipid Res* **43**, 463-476 (2002).

383 Nomura, D. K. *et al.* Monoacylglycerol lipase regulates a fatty acid network that promotes cancer pathogenesis. *Cell* **140**, 49-61, doi:10.1016/j.cell.2009.11.027 (2010).

384 Long, J. Z. *et al.* Selective blockade of 2-arachidonoylglycerol hydrolysis produces cannabinoid behavioral effects. *Nat Chem Biol* **5**, 37-44, doi:10.1038/nchembio.129 (2009).

385 Iwata, Y., Mort, J. S., Tateishi, H. & Lee, E. R. Macrophage cathepsin L, a factor in the erosion of subchondral bone in rheumatoid arthritis. *Arthritis Rheum* **40**, 499-509 (1997).

386 Lavrik, I. N., Golks, A. & Krammer, P. H. Caspases: pharmacological manipulation of cell death. *J Clin Invest* **115**, 2665-2672, doi:10.1172/JCI26252 (2005).

387 Crawford, E. D. & Wells, J. A. Caspase substrates and cellular remodeling. *Annu Rev Biochem* **80**, 1055-1087, doi:10.1146/annurev-biochem-061809-121639 (2011).

388 Pop, C. & Salvesen, G. S. Human caspases: activation, specificity, and regulation. *J Biol Chem* **284**, 21777-21781, doi:10.1074/jbc.R800084200 (2009).

389 Joyce, J. A. *et al.* Cathepsin cysteine proteases are effectors of invasive growth and angiogenesis during multistage tumorigenesis. *Cancer Cell* **5**, 443-453 (2004).

390 Yan, S., Sameni, M. & Sloane, B. F. Cathepsin B and human tumor

progression. *Biol Chem* **379**, 113-123 (1998).

391 Goulet, B. *et al.* A cathepsin L isoform that is devoid of a signal peptide localizes to the nucleus in S phase and processes the CDP/Cux transcription factor. *Mol Cell* **14**, 207-219 (2004).

392 Greenbaum, D. C. *et al.* A role for the protease falcipain 1 in host cell invasion by the human malaria parasite. *Science* **298**, 2002-2006, doi:10.1126/science.1077426 (2002).

393 Mason, R. W., Wilcox, D., Wikstrom, P. & Shaw, E. N. The identification of active forms of cysteine proteinases in Kirsten-virus-transformed mouse fibroblasts by use of a specific radiolabelled inhibitor. *Biochem J* **257**, 125-129 (1989).

394 Thornberry, N. A. *et al.* Inactivation of interleukin-1 beta converting enzyme by peptide (acyloxy)methyl ketones. *Biochemistry* **33**, 3934-3940 (1994).

395 Rawlings, N. D. & Barrett, A. J. Families of cysteine peptidases. *Methods Enzymol* **244**, 461-486 (1994).

396 Barrett, A. J. & Rawlings, N. D. Evolutionary lines of cysteine peptidases. *Biol Chem* **382**, 727-733, doi:10.1515/BC.2001.088 (2001).

397 Turk, V., Turk, B. & Turk, D. Lysosomal cysteine proteases: facts and opportunities. *EMBO J* **20**, 4629-4633, doi:10.1093/emboj/20.17.4629 (2001).

398 Yasuda, Y., Kaleta, J. & Bromme, D. The role of cathepsins in osteoporosis and arthritis: rationale for the design of new therapeutics. *Adv Drug Deliv Rev* **57**, 973-993, doi:10.1016/j.addr.2004.12.013 (2005).

399 Razzaque, A., Hanada, Y., Mukai, J. I., Murao, S. & Nishino, T. NAD- and FAD-3'-pyrophosphates--enzymic synthesis and inertness. *FEBS Lett* **93**, 335-338 (1978).

400 Barrett, A. J. *et al.* L-trans-Epoxysuccinyl-leucylamido(4-guanidino)butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. *Biochem J* **201**, 189-198 (1982).

401 Yasothornsrikul, S. *et al.* Cathepsin L in secretory vesicles functions as a prohormone-processing enzyme for production of the enkephalin peptide neurotransmitter. *Proc Natl Acad Sci U S A* **100**, 9590-9595, doi:10.1073/pnas.1531542100 (2003).

402 Greenbaum, D., Medzihradszky, K. F., Burlingame, A. & Bogyo, M. Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. *Chem Biol* **7**, 569-581 (2000).

403 McIlwain, D. R., Berger, T. & Mak, T. W. Caspase functions in cell death and

disease. *Cold Spring Harb Perspect Biol* **5**, a008656, doi:10.1101/cshperspect.a008656 (2013).

404 Thornberry, N. A. *et al.* A novel heterodimeric cysteine protease is required for interleukin-1 beta processing in monocytes. *Nature* **356**, 768-774, doi:10.1038/356768a0 (1992).

405 Blum, G., Weimer, R. M., Edgington, L. E., Adams, W. & Bogyo, M. Comparative assessment of substrates and activity based probes as tools for non-invasive optical imaging of cysteine protease activity. *PLoS One* **4**, e6374, doi:10.1371/journal.pone.0006374 (2009).

406 Green, M. & Loewenstein, P. M. Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein. *Cell* **55**, 1179-1188 (1988).

407 Kirby, J. M. *et al.* Cwp84, a surface-associated cysteine protease, plays a role in the maturation of the surface layer of Clostridium difficile. *J Biol Chem* **284**, 34666-34673, doi:10.1074/jbc.M109.051177 (2009).

408 Borodovsky, A. *et al.* A novel active site-directed probe specific for deubiquitylating enzymes reveals proteasome association of USP14. *EMBO J* **20**, 5187-5196, doi:10.1093/emboj/20.18.5187 (2001).

409 Gopinath, P., Ohayon, S., Nawatha, M. & Brik, A. Chemical and semisynthetic approaches to study and target deubiquitinases. *Chem Soc Rev* **45**, 4171-4198, doi:10.1039/c6cs00083e (2016).

410 Altun, M. *et al.* Activity-based chemical proteomics accelerates inhibitor development for deubiquitylating enzymes. *Chem Biol* **18**, 1401-1412, doi:10.1016/j.chembiol.2011.08.018 (2011).

411 Clague, M. J., Coulson, J. M. & Urbe, S. Cellular functions of the DUBs. *J Cell Sci* **125**, 277-286, doi:10.1242/jcs.090985 (2012).

412 Hicke, L. & Dunn, R. Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins. *Annual review of cell and developmental biology* **19**, 141-172, doi:10.1146/annurev.cellbio.19.110701.154617 (2003).

413 Hassink, G. C. *et al.* The ER-resident ubiquitin-specific protease 19 participates in the UPR and rescues ERAD substrates. *EMBO Rep* **10**, 755-761, doi:10.1038/embor.2009.69 (2009).

414 Li, M. *et al.* Deubiquitination of p53 by HAUSP is an important pathway for p53 stabilization. *Nature* **416**, 648-653, doi:10.1038/nature737 (2002).

415 Sowa, M. E., Bennett, E. J., Gygi, S. P. & Harper, J. W. Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138**, 389-403, doi:10.1016/j.cell.2009.04.042 (2009).

416 Clague, M. J. *et al.* Deubiquitylases from genes to organism. *Physiol Rev* **93**, 1289-1315, doi:10.1152/physrev.00002.2013 (2013).

417 Amerik, A. Y. & Hochstrasser, M. Mechanism and function of deubiquitinating enzymes. *Biochim Biophys Acta* **1695**, 189-207, doi:10.1016/j.bbamcr.2004.10.003 (2004).

418 Mevissen, T. E. T. & Komander, D. Mechanisms of Deubiquitinase Specificity and Regulation. *Annu Rev Biochem* **86**, 159-192, doi:10.1146/annurev-biochem-061516-044916 (2017).

419 Kwasna, D. *et al.* Discovery and Characterization of ZUFSP/ZUP1, a Distinct Deubiquitinase Class Important for Genome Stability. *Mol Cell* **70**, 150-164 e156, doi:10.1016/j.molcel.2018.02.023 (2018).

420 Hewings, D. S. *et al.* Reactive-site-centric chemoproteomics identifies a distinct class of deubiquitinase enzymes. *Nat Commun* **9**, 1162, doi:10.1038/s41467-018-03511-6 (2018).

421 Johnston, S. C., Riddle, S. M., Cohen, R. E. & Hill, C. P. Structural basis for the specificity of ubiquitin C-terminal hydrolases. *EMBO J* **18**, 3877-3887, doi:10.1093/emboj/18.14.3877 (1999).

422 Hemelaar, J., Lelyveld, V. S., Kessler, B. M. & Ploegh, H. L. A single protease, Apg4B, is specific for the autophagy-related ubiquitin-like proteins GATE-16, MAP1-LC3, GABARAP, and Apg8L. *J Biol Chem* **278**, 51841-51850, doi:10.1074/jbc.M308762200 (2003).

423 Hemelaar, J. *et al.* Specific and covalent targeting of conjugating and deconjugating enzymes of ubiquitin-like proteins. *Mol Cell Biol* **24**, 84-95 (2004).

424 Sommer, S., Weikart, N. D., Linne, U. & Mootz, H. D. Covalent inhibition of SUMO and ubiquitin-specific cysteine proteases by an in situ thiol-alkyne addition. *Bioorg Med Chem* **21**, 2511-2517, doi:10.1016/j.bmc.2013.02.039 (2013).

425 Love, K. R., Pandya, R. K., Spooner, E. & Ploegh, H. L. Ubiquitin C-terminal electrophiles are activity-based probes for identification and mechanistic study of ubiquitin conjugating machinery. *ACS Chem Biol* **4**, 275-287, doi:10.1021/cb9000348 (2009).

426 Whedon, S. D. *et al.* Selenocysteine as a latent bioorthogonal electrophilic probe for deubiquitylating enzymes. *J Am Chem Soc*, doi:10.1021/jacs.6b05688 (2016).

427 Naik, E. & Dixit, V. M. Usp9X Is Required for Lymphocyte Activation and Homeostasis through Its Control of ZAP70 Ubiquitination and PKCbeta Kinase Activity. *J Immunol* **196**, 3438-3451,

doi:10.4049/jimmunol.1403165 (2016).

428 Balakirev, M. Y., Tcherniuk, S. O., Jaquinod, M. & Chroboczek, J. Otubains: a new family of cysteine proteases in the ubiquitin pathway. *EMBO Rep* **4**, 517-522, doi:10.1038/sj.embor.embor824 (2003).

429 Pruneda, J. N. *et al.* The Molecular Basis for Ubiquitin and Ubiquitin-like Specificities in Bacterial Effector Proteases. *Mol Cell* **63**, 261-276, doi:10.1016/j.molcel.2016.06.015 (2016).

430 Flierman, D. *et al.* Non-hydrolyzable Diubiquitin Probes Reveal Linkage-Specific Reactivity of Deubiquitylating Enzymes Mediated by S2 Pockets. *Cell Chem Biol* **23**, 472-482, doi:10.1016/j.chembiol.2016.03.009 (2016).

431 Gao, Y. *et al.* Early adipogenesis is regulated through USP7-mediated deubiquitination of the histone acetyltransferase TIP60. *Nat Commun* **4**, 2656, doi:10.1038/ncomms3656 (2013).

432 Harrigan, J. A., Jacq, X., Martin, N. M. & Jackson, S. P. Deubiquitylating enzymes and drug discovery: emerging opportunities. *Nat Rev Drug Discov* **17**, 57-78, doi:10.1038/nrd.2017.152 (2018).

433 Wang, X. *et al.* The proteasome deubiquitinase inhibitor VLX1570 shows selectivity for ubiquitin-specific protease-14 and induces apoptosis of multiple myeloma cells. *Sci Rep* **6**, 26979, doi:10.1038/srep26979 (2016).

434 Turnbull, A. P. *et al.* Molecular basis of USP7 inhibition by selective small-molecule inhibitors. *Nature* **550**, 481-486, doi:10.1038/nature24451 (2017).

435 An, H. & Statsyuk, A. V. Development of activity-based probes for ubiquitin and ubiquitin-like protein signaling pathways. *J Am Chem Soc* **135**, 16948-16962, doi:10.1021/ja4099643 (2013).

436 An, H. & Statsyuk, A. V. Facile synthesis of covalent probes to capture enzymatic intermediates during E1 enzyme catalysis. *Chem Commun (Camb)* **52**, 2477-2480, doi:10.1039/c5cc08592f (2016).

437 Krist, D. T., Park, S., Boneh, G. H., Rice, S. E. & Statsyuk, A. V. UbFluor: A Mechanism-Based Probe for HECT E3 Ligases. *Chem Sci* **7**, 5587-5595, doi:10.1039/C6SC01167E (2016).

438 Park, S., Foote, P. K., Krist, D. T., Rice, S. E. & Statsyuk, A. V. UbMES and UbFluor: Novel probes for ring-between-ring (RBR) E3 ubiquitin ligase PARKIN. *J Biol Chem* **292**, 16539-16553, doi:10.1074/jbc.M116.773200 (2017).

439 Mulder, M. P. *et al.* A cascading activity-based probe sequentially targets E1-E2-E3 ubiquitin enzymes. *Nat Chem Biol* **12**, 523-530, doi:10.1038/nchembio.2084 (2016).

440 Freedy, A. M. *et al.* Chemoselective Installation of Amine Bonds on Proteins through Aza-Michael Ligation. *J Am Chem Soc* **139**, 18365-18375, doi:10.1021/jacs.7b10702 (2017).

441 Dadova, J. *et al.* Vinylsulfonamide and acrylamide modification of DNA for cross-linking with proteins. *Angew Chem Int Ed Engl* **52**, 10515-10518, doi:10.1002/anie.201303577 (2013).

442 Hong, V., Presolski, S. I., Ma, C. & Finn, M. G. Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation. *Angew Chem Int Ed Engl* **48**, 9879-9883, doi:10.1002/anie.200905087 (2009).

443 Lin, D. Y., Diao, J., Zhou, D. & Chen, J. Biochemical and structural studies of a HECT-like ubiquitin ligase from Escherichia coli O157:H7. *J Biol Chem* **286**, 441-449, doi:10.1074/jbc.M110.167643 (2011).

444 Virdee, S., Ye, Y., Nguyen, D. P., Komander, D. & Chin, J. W. Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OTU deubiquitinase. *Nat Chem Biol* **6**, 750-757, doi:10.1038/nchembio.426 (2010).

445 El Oualid, F. *et al.* Chemical synthesis of ubiquitin, ubiquitin-based probes, and diubiquitin. *Angewandte Chemie* **49**, 10149-10153, doi:10.1002/anie.201005995 (2010).

446 Plechanovova, A., Jaffray, E. G., Tatham, M. H., Naismith, J. H. & Hay, R. T. Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature*, doi:10.1038/nature11376 (2012).

447 Emmerich, C. H. *et al.* Activation of the canonical IKK complex by K63/M1-linked hybrid ubiquitin chains. *Proc Natl Acad Sci U S A* **110**, 15247-15252, doi:10.1073/pnas.1314715110 (2013).

448 Kazlauskaite, A. *et al.* Phosphorylation of Parkin at Serine65 is essential for activation: elaboration of a Miro1 substrate-based assay of Parkin E3 ligase activity. *Open biology* **4**, 130213, doi:10.1098/rsob.130213 (2014).

449 Lai, Y. C. *et al.* Phosphoproteomic screening identifies Rab GTPases as novel downstream targets of PINK1. *EMBO J* **34**, 2840-2861, doi:10.15252/embj.201591593 (2015).

450 Wu, P. Y. *et al.* A conserved catalytic residue in the ubiquitin-conjugating enzyme family. *EMBO J* **22**, 5241-5250, doi:10.1093/emboj/cdg501 (2003).

451 Yunus, A. A. & Lima, C. D. Lysine activation and functional analysis of E2-mediated conjugation in the SUMO pathway. *Nat Struct Mol Biol* **13**, 491-499, doi:10.1038/nsmb1104 (2006).

452 Brownell, J. E. *et al.* Substrate-assisted inhibition of ubiquitin-like protein-

activating enzymes: the NEDD8 E1 inhibitor MLN4924 forms a NEDD8-AMP mimetic in situ. *Mol Cell* **37**, 102-111, doi:10.1016/j.molcel.2009.12.024 (2010).

453 Lovell, S. C. *et al.* Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* **50**, 437-450, doi:10.1002/prot.10286 (2003).

454 Dominguez, E. *et al.* Integrated phenotypic and activity-based profiling links Ces3 to obesity and diabetes. *Nat Chem Biol* **10**, 113-121, doi:10.1038/nchembio.1429 (2014).

455 Lu, X. *et al.* Designed semisynthetic protein inhibitors of Ub/Ubl E1 activating enzymes. *Journal of the American Chemical Society* **132**, 1748-1749, doi:10.1021/ja9088549 (2010).

456 Stanley, M. *et al.* Orthogonal Thiol Functionalization at a Single Atomic Center for Profiling Transthiolation Activity of E1 Activating Enzymes. *ACS chemical biology*, doi:10.1021/acschembio.5b00118 (2015).

457 Virdee, S. *et al.* Traceless and site-specific ubiquitination of recombinant proteins. *J Am Chem Soc* **133**, 10708-10711, doi:10.1021/ja202799r (2011).

458 Strickson, S. *et al.* The anti-inflammatory drug BAY 11-7082 suppresses the MyD88-dependent signalling network by targeting the ubiquitin system. *Biochem J* **451**, 427-437, doi:10.1042/BJ20121651 (2013).

459 Hong, V., Presolski, S. I., Ma, C. & Finn, M. G. Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation. *Angewandte Chemie* **48**, 9879-9883, doi:10.1002/anie.200905087 (2009).

460 Cui, C., Zhao, W., Chen, J., Wang, J. & Li, Q. Elimination of in vivo cleavage between target protein and intein in the intein-mediated protein purification systems. *Protein expression and purification* **50**, 74-81, doi:10.1016/j.pep.2006.05.019 (2006).

461 Weissman, A. M. Themes and variations on ubiquitylation. *Nat Rev Mol Cell Biol* **2**, 169-178, doi:10.1038/35056563 (2001).

462 Sarraf, S. A. *et al.* Landscape of the PARKIN-dependent ubiquitylome in response to mitochondrial depolarization. *Nature* **496**, 372-376, doi:10.1038/nature12043 (2013).

463 Dawson, T. M. & Dawson, V. L. The role of parkin in familial and sporadic Parkinson's disease. *Mov Disord* **25 Suppl 1**, S32-39, doi:10.1002/mds.22798 (2010).

464 Pickrell, A. M. & Youle, R. J. The roles of PINK1, parkin, and mitochondrial fidelity in Parkinson's disease. *Neuron* **85**, 257-273, doi:10.1016/j.neuron.2014.12.007 (2015).

465 Martinez, A. *et al.* Quantitative proteomic analysis of Parkin substrates in Drosophila neurons. *Mol Neurodegener* **12**, 29, doi:10.1186/s13024-017-0170-3 (2017).

466 Riley, B. E. *et al.* Structure and function of Parkin E3 ubiquitin ligase reveals aspects of RING and HECT ligases. *Nat Commun* **4**, 1982, doi:10.1038/ncomms2982 (2013).

467 Mackay, C. *et al.* E3 ubiquitin ligase HOIP attenuates apoptotic cell death induced by cisplatin. *Cancer research* **74**, 2246-2257, doi:10.1158/0008-5472.CAN-13-2131 (2014).

468 Schaeffer, V. *et al.* Binding of OTULIN to the PUB domain of HOIP controls NF-kappaB signaling. *Molecular cell* **54**, 349-361, doi:10.1016/j.molcel.2014.03.016 (2014).

469 Heo, J. M., Ordureau, A., Paulo, J. A., Rinehart, J. & Harper, J. W. The PINK1-PARKIN Mitochondrial Ubiquitylation Pathway Drives a Program of OPTN/NDP52 Recruitment and TBK1 Activation to Promote Mitophagy. *Mol Cell* **60**, 7-20, doi:10.1016/j.molcel.2015.08.016 (2015).

470 Perez, F. A. & Palmiter, R. D. Parkin-deficient mice are not a robust model of parkinsonism. *Proc Natl Acad Sci U S A* **102**, 2174-2179, doi:10.1073/pnas.0409598102 (2005).

471 McWilliams, T. G. *et al.* Basal Mitophagy Occurs Independently of PINK1 in Mouse Tissues of High Metabolic Demand. *Cell Metab* **27**, 439-449 e435, doi:10.1016/j.cmet.2017.12.008 (2018).

472 Byrne, R., Mund, T. & Licchesi, J. D. F. Activity-Based Probes for HECT E3 Ubiquitin Ligases. *Chembiochem* **18**, 1415-1427, doi:10.1002/cbic.201700006 (2017).

473 Scheffner, M., Nuber, U. & Huibregtse, J. M. Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade. *Nature* **373**, 81-83, doi:10.1038/373081a0 (1995).

474 Huibregtse, J. M., Scheffner, M., Beaudenon, S. & Howley, P. M. A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 5249 (1995).

475 Speers, A. E. & Cravatt, B. F. Activity-Based Protein Profiling (ABPP) and Click Chemistry (CC)-ABPP by MudPIT Mass Spectrometry. *Curr Protoc Chem Biol* **1**, 29-41, doi:10.1002/9780470559277.ch090138 (2009).

476 Salisbury, C. M. & Cravatt, B. F. Activity-based probes for proteomic profiling of histone deacetylase complexes. *Proc Natl Acad Sci U S A* **104**, 1171-1176, doi:10.1073/pnas.0608659104 (2007).

477 Carroll, R. G., Hollville, E. & Martin, S. J. Parkin sensitizes toward apoptosis induced by mitochondrial depolarization through promoting degradation of Mcl-1. *Cell Rep* **9**, 1538-1553, doi:10.1016/j.celrep.2014.10.046 (2014).

478 Grill, B., Murphey, R. K. & Borgen, M. A. The PHR proteins: intracellular signaling hubs in neuronal development and axon degeneration. *Neural Dev* **11**, 8, doi:10.1186/s13064-016-0063-0 (2016).

479 Guo, Q., Xie, J., Dang, C. V., Liu, E. T. & Bishop, J. M. Identification of a large Myc-binding protein that contains RCC1-like repeats. *Proc Natl Acad Sci U S A* **95**, 9172-9177 (1998).

480 Schaefer, A. M., Hadwiger, G. D. & Nonet, M. L. rpm-1, a conserved neuronal gene that regulates targeting and synaptogenesis in C. elegans. *Neuron* **26**, 345-356 (2000).

481 Wan, H. I. *et al.* Highwire regulates synaptic growth in Drosophila. *Neuron* **26**, 313-329 (2000).

482 Zhen, M., Huang, X., Bamber, B. & Jin, Y. Regulation of presynaptic terminal organization by C. elegans RPM-1, a putative guanine nucleotide exchanger with a RING-H2 finger domain. *Neuron* **26**, 331-343 (2000).

483 D'Souza, J. *et al.* Formation of the retinotectal projection requires Esrom, an ortholog of PAM (protein associated with Myc). *Development* **132**, 247-256, doi:10.1242/dev.01578 (2005).

484 Bloom, A. J., Miller, B. R., Sanes, J. R. & DiAntonio, A. The requirement for Phr1 in CNS axon tract formation reveals the corticostriatal boundary as a choice point for cortical axons. *Genes Dev* **21**, 2593-2606, doi:10.1101/gad.1592107 (2007).

485 Lewcock, J. W., Genoud, N., Lettieri, K. & Pfaff, S. L. The ubiquitin ligase Phr1 regulates axon outgrowth through modulation of microtubule dynamics. *Neuron* **56**, 604-620, doi:10.1016/j.neuron.2007.09.009 (2007).

486 Burgess, R. W. *et al.* Evidence for a conserved function in synapse formation reveals Phr1 as a candidate gene for respiratory failure in newborn mice. *Mol Cell Biol* **24**, 1096-1105 (2004).

487 Grill, B. *et al.* C. elegans RPM-1 regulates axon termination and synaptogenesis through the Rab GEF GLO-4 and the Rab GTPase GLO-1. *Neuron* **55**, 587-601, doi:10.1016/j.neuron.2007.07.009 (2007).

488 Opperman, K. J. & Grill, B. RPM-1 is localized to distinct subcellular compartments and regulates axon length in GABAergic motor neurons. *Neural Dev* **9**, 10, doi:10.1186/1749-8104-9-10 (2014).

489 Abrams, B., Grill, B., Huang, X. & Jin, Y. Cellular and molecular determinants targeting the Caenorhabditis elegans PHR protein RPM-1 to

perisynaptic regions. *Dev Dyn* **237**, 630-639, doi:10.1002/dvdy.21446 (2008).

490 Wu, C., Wairkar, Y. P., Collins, C. A. & DiAntonio, A. Highwire function at the Drosophila neuromuscular junction: spatial, structural, and temporal requirements. *J Neurosci* **25**, 9557-9566, doi:10.1523/JNEUROSCI.2532-05.2005 (2005).

491 Han, S. *et al.* Pam (Protein associated with Myc) functions as an E3 ubiquitin ligase and regulates TSC/mTOR signaling. *Cell Signal* **20**, 1084-1091, doi:10.1016/j.cellsig.2008.01.020 (2008).

492 Saiga, T. *et al.* Fbxo45 forms a novel ubiquitin ligase complex and is required for neuronal development. *Mol Cell Biol* **29**, 3529-3543, doi:10.1128/MCB.00364-09 (2009).

493 Babetto, E., Beirowski, B., Russler, E. V., Milbrandt, J. & DiAntonio, A. The Phr1 ubiquitin ligase promotes injury-induced axon self-destruction. *Cell Rep* **3**, 1422-1429, doi:10.1016/j.celrep.2013.04.013 (2013).

494 Holland, S. *et al.* The ubiquitin ligase MYCBP2 regulates transient receptor potential vanilloid receptor 1 (TRPV1) internalization through inhibition of p38 MAPK signaling. *J Biol Chem* **286**, 3671-3680, doi:10.1074/jbc.M110.154765 (2011).

495 Xiong, X. *et al.* The Highwire ubiquitin ligase promotes axonal degeneration by tuning levels of Nmnat protein. *PLoS Biol* **10**, e1001440, doi:10.1371/journal.pbio.1001440 (2012).

496 Milde, S., Gilley, J. & Coleman, M. P. Axonal trafficking of NMNAT2 and its roles in axon growth and survival in vivo. *Bioarchitecture* **3**, 133-140, doi:10.4161/bioa.27049 (2013).

497 McIntyre, M. C., Frattini, M. G., Grossman, S. R. & Laimins, L. A. Human papillomavirus type 18 E7 protein requires intact Cys-X-X-Cys motifs for zinc binding, dimerization, and transformation but not for Rb binding. *Journal of virology* **67**, 3142-3150 (1993).

498 Wang, X. *et al.* Ubiquitination of serine, threonine, or lysine residues on the cytoplasmic tail can induce ERAD of MHC-I by viral E3 ligase mK3. *J Cell Biol* **177**, 613-624, doi:10.1083/jcb.200611063 (2007).

499 van Wijk, S. J. *et al.* A comprehensive framework of E2-RING E3 interactions of the human ubiquitin-proteasome system. *Mol Syst Biol* **5**, 295, doi:10.1038/msb.2009.55 (2009).

500 Metzger, M. B., Pruneda, J. N., Klevit, R. E. & Weissman, A. M. RING-type E3 ligases: master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. *Biochim Biophys Acta* **1843**, 47-60,

doi:10.1016/j.bbamcr.2013.05.026 (2014).

501 Koliopoulos, M. G., Esposito, D., Christodoulou, E., Taylor, I. A. & Rittinger, K. Functional role of TRIM E3 ligase oligomerization and regulation of catalytic activity. *EMBO J* **35**, 1204-1218, doi:10.15252/embj.201593741 (2016).

502 Wright, J. D., Mace, P. D. & Day, C. L. Secondary ubiquitin-RING docking enhances Arkadia and Ark2C E3 ligase activity. *Nat Struct Mol Biol* **23**, 45-52, doi:10.1038/nsmb.3142 (2016).

503 Allen, G. F., Toth, R., James, J. & Ganley, I. G. Loss of iron triggers PINK1/Parkin-independent mitophagy. *EMBO Rep* **14**, 1127-1135, doi:10.1038/embor.2013.168 (2013).

504 Irwin, D. J., Lee, V. M. & Trojanowski, J. Q. Parkinson's disease dementia: convergence of alpha-synuclein, tau and amyloid-beta pathologies. *Nat Rev Neurosci* **14**, 626-636, doi:10.1038/nrn3549 (2013).

505 Stefanis, L. alpha-Synuclein in Parkinson's disease. *Cold Spring Harb Perspect Med* **2**, a009399, doi:10.1101/cshperspect.a009399 (2012).

506 Arora, A. & Scholar, E. M. Role of tyrosine kinase inhibitors in cancer therapy. *J Pharmacol Exp Ther* **315**, 971-979, doi:10.1124/jpet.105.084145 (2005).

507 Gross, S., Rahal, R., Stransky, N., Lengauer, C. & Hoeflich, K. P. Targeting cancer with kinase inhibitors. *J Clin Invest* **125**, 1780-1789, doi:10.1172/JCI76094 (2015).

508 Cohen, P. & Tcherpakov, M. Will the ubiquitin system furnish as many drug targets as protein kinases? *Cell* **143**, 686-693, doi:10.1016/j.cell.2010.11.016 (2010).

509 Huang, X. & Dixit, V. M. Drugging the undruggables: exploring the ubiquitin system for drug development. *Cell Res* **26**, 484-498, doi:10.1038/cr.2016.31 (2016).

510 Zhao, R. *et al.* DDB2 modulates TGF-beta signal transduction in human ovarian cancer cells by downregulating NEDD4L. *Nucleic Acids Res* **43**, 7838-7849, doi:10.1093/nar/gkv667 (2015).

511 Qu, M. H. *et al.* miR-93 promotes TGF-beta-induced epithelial-to-mesenchymal transition through downregulation of NEDD4L in lung cancer cells. *Tumour Biol* **37**, 5645-5651, doi:10.1007/s13277-015-4328-8 (2016).

512 Kito, Y. *et al.* Pathobiological properties of the ubiquitin ligase Nedd4L in melanoma. *Int J Exp Pathol* **95**, 24-28, doi:10.1111/iep.12051 (2014).

513 Veeriah, S. *et al.* Somatic mutations of the Parkinson's disease-associated

gene PARK2 in glioblastoma and other human malignancies. *Nat Genet* **42**, 77-82, doi:10.1038/ng.491 (2010).

514 Poulogiannis, G. *et al.* PARK2 deletions occur frequently in sporadic colorectal cancer and accelerate adenoma development in Apc mutant mice. *Proc Natl Acad Sci U S A* **107**, 15145-15150, doi:10.1073/pnas.1009941107 (2010).

515 Cesari, R. *et al.* Parkin, a gene implicated in autosomal recessive juvenile parkinsonism, is a candidate tumor suppressor gene on chromosome 6q25-q27. *Proc Natl Acad Sci U S A* **100**, 5956-5961, doi:10.1073/pnas.0931262100 (2003).

516 Cappadocia, L. & Lima, C. D. Ubiquitin-like Protein Conjugation: Structures, Chemistry, and Mechanism. *Chem Rev* **118**, 889-918, doi:10.1021/acs.chemrev.6b00737 (2018).

517 Guidotti, G., Brambilla, L. & Rossi, D. Cell-Penetrating Peptides: From Basic Research to Clinics. *Trends Pharmacol Sci* **38**, 406-424, doi:10.1016/j.tips.2017.01.003 (2017).

518 Kristensen, M., Birch, D. & Morck Nielsen, H. Applications and Challenges for Use of Cell-Penetrating Peptides as Delivery Vectors for Peptide and Protein Cargos. *Int J Mol Sci* **17**, doi:10.3390/ijms17020185 (2016).

519 Ellis, H. R. & Poole, L. B. Roles for the two cysteine residues of AhpC in catalysis of peroxide reduction by alkyl hydroperoxide reductase from Salmonella typhimurium. *Biochemistry* **36**, 13349-13356, doi:10.1021/bi9713658 (1997).

520 Alpi, A. F., Chaugule, V. & Walden, H. Mechanism and disease association of E2-conjugating enzymes: lessons from UBE2T and UBE2L3. *Biochem J* **473**, 3401-3419, doi:10.1042/BCJ20160028 (2016).

521 McDowell, G. S. & Philpott, A. Non-canonical ubiquitylation: mechanisms and consequences. *Int J Biochem Cell Biol* **45**, 1833-1842, doi:10.1016/j.biocel.2013.05.026 (2013).

522 Wang, X., Herr, R. A. & Hansen, T. H. Ubiquitination of substrates by esterification. *Traffic* **13**, 19-24, doi:10.1111/j.1600-0854.2011.01269.x (2012).

523 Coscoy, L. & Ganem, D. PHD domains and E3 ubiquitin ligases: viruses make the connection. *Trends Cell Biol* **13**, 7-12 (2003).

524 Hewitt, E. W. *et al.* Ubiquitylation of MHC class I by the K3 viral protein signals internalization and TSG101-dependent degradation. *EMBO J* **21**, 2418-2429, doi:10.1093/emboj/21.10.2418 (2002).

525 Cadwell, K. & Coscoy, L. Ubiquitination on nonlysine residues by a viral E3

ubiquitin ligase. *Science* **309**, 127-130, doi:10.1126/science.1110340 (2005).

526 Williams, C., van den Berg, M., Geers, E. & Distel, B. Pex10p functions as an E3 ligase for the Ubc4p-dependent ubiquitination of Pex5p. *Biochem Biophys Res Commun* **374**, 620-624, doi:10.1016/j.bbrc.2008.07.054 (2008).

527 Platta, H. W. *et al.* Pex2 and pex12 function as protein-ubiquitin ligases in peroxisomal protein import. *Mol Cell Biol* **29**, 5505-5516, doi:10.1128/MCB.00388-09 (2009).

528 Vosper, J. M. *et al.* Ubiquitylation on canonical and non-canonical sites targets the transcription factor neurogenin for ubiquitin-mediated proteolysis. *J Biol Chem* **284**, 15458-15468, doi:10.1074/jbc.M809366200 (2009).

529 Wang, Y. & Tjandra, N. Structural insights of tBid, the caspase-8-activated Bid, and its BH3 domain. *J Biol Chem* **288**, 35840-35851, doi:10.1074/jbc.M113.503680 (2013).

530 Tait, S. W. *et al.* Apoptosis induction by Bid requires unconventional ubiquitination and degradation of its N-terminal fragment. *J Cell Biol* **179**, 1453-1466, doi:10.1083/jcb.200707063 (2007).

531 Ishikura, S., Weissman, A. M. & Bonifacino, J. S. Serine residues in the cytosolic tail of the T-cell antigen receptor alpha-chain mediate ubiquitination and endoplasmic reticulum-associated degradation of the unassembled protein. *J Biol Chem* **285**, 23916-23924, doi:10.1074/jbc.M110.127936 (2010).

532 Bays, N. W., Gardner, R. G., Seelig, L. P., Joazeiro, C. A. & Hampton, R. Y. Hrd1p/Der3p is a membrane-anchored ubiquitin ligase required for ER-associated degradation. *Nat Cell Biol* **3**, 24-29, doi:10.1038/35050524 (2001).

533 Shimizu, Y., Okuda-Shimizu, Y. & Hendershot, L. M. Ubiquitylation of an ERAD substrate occurs on multiple types of amino acids. *Mol Cell* **40**, 917-926, doi:10.1016/j.molcel.2010.11.033 (2010).

534 Herr, R. A., Harris, J., Fang, S., Wang, X. & Hansen, T. H. Role of the RING-CH domain of viral ligase mK3 in ubiquitination of non-lysine and lysine MHC I residues. *Traffic* **10**, 1301-1317, doi:10.1111/j.1600-0854.2009.00946.x (2009).

535 Song, J. *et al.* Stability of thioester intermediates in ubiquitin-like modifications. *Protein Sci* **18**, 2492-2499, doi:10.1002/pro.254 (2009).

536 Holland, S. & Scholich, K. Regulation of neuronal functions by the E3-

ubiquitinligase protein associated with MYC (MYCBP2). *Commun Integr Biol* **4**, 513-515, doi:10.4161/cib.4.5.15967 (2011).

537 Ji, R. R., Samad, T. A., Jin, S. X., Schmoll, R. & Woolf, C. J. p38 MAPK activation by NGF in primary sensory neurons after inflammation increases TRPV1 levels and maintains heat hyperalgesia. *Neuron* **36**, 57-68 (2002).

538 Nakata, K. *et al.* Regulation of a DLK-1 and p38 MAP kinase pathway by the ubiquitin ligase RPM-1 is required for presynaptic development. *Cell* **120**, 407-420, doi:10.1016/j.cell.2004.12.017 (2005).

539 Sakuma, H. *et al.* Molecular cloning and functional expression of a cDNA encoding a new member of mixed lineage protein kinase from human brain. *J Biol Chem* **272**, 28622-28629 (1997).

540 Inoki, K., Li, Y., Zhu, T., Wu, J. & Guan, K. L. TSC2 is phosphorylated and inhibited by Akt and suppresses mTOR signalling. *Nat Cell Biol* **4**, 648-657, doi:10.1038/ncb839 (2002).

541 Swaney, D. L. *et al.* Global analysis of phosphorylation and ubiquitylation cross-talk in protein degradation. *Nat Methods* **10**, 676-682, doi:10.1038/nmeth.2519 (2013).

542 Ordureau, A. *et al.* Quantitative proteomics reveal a feedforward mechanism for mitochondrial PARKIN translocation and ubiquitin chain synthesis. *Mol Cell* **56**, 360-375, doi:10.1016/j.molcel.2014.09.007 (2014).

543 Heap, R. E., Gant, M. S., Lamoliatte, F., Peltier, J. & Trost, M. Mass spectrometry techniques for studying the ubiquitin system. *Biochem Soc Trans* **45**, 1137-1148, doi:10.1042/BST20170091 (2017).

544 Sylvestersen, K. B., Young, C. & Nielsen, M. L. Advances in characterizing ubiquitylation sites by mass spectrometry. *Curr Opin Chem Biol* **17**, 49-58, doi:10.1016/j.cbpa.2012.12.009 (2013).

545 Ohsaki, Y., Cheng, J., Fujita, A., Tokumoto, T. & Fujimoto, T. Cytoplasmic lipid droplets are sites of convergence of proteasomal and autophagic degradation of apolipoprotein B. *Mol Biol Cell* **17**, 2674-2683, doi:10.1091/mbc.E05-07-0659 (2006).

546 Yoshida, Y. *et al.* E3 ubiquitin ligase that recognizes sugar chains. *Nature* **418**, 438-442, doi:10.1038/nature00890 (2002).

547 Pao, K. C. *et al.* Activity-based E3 ligase profiling uncovers an E3 ligase with esterification activity. *Nature* **556**, 381-385, doi:10.1038/s41586-018-0026-1 (2018).

548 Udeshi, N. D. *et al.* Refined preparation and use of anti-diglycine remnant (K-epsilon-GG) antibody enables routine quantification of 10,000s of

ubiquitination sites in single proteomics experiments. *Mol Cell Proteomics* **12**, 825-831, doi:10.1074/mcp.O112.027094 (2013).