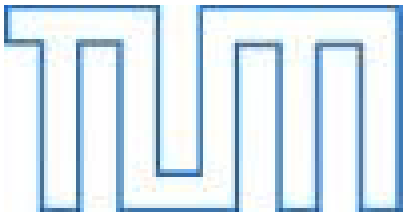


Technische Universität München

Fakultät Chemie

Mechanisms of polyubiquitin chain formation

Joanna Michalina Liwocha

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades einer

Doktorin der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Priv.-Doz. Dr. Gerd Gemmecker
Prüfer*innen der Dissertation:

1. Hon. Prof. Dr. Brenda Schulman
2. Prof. Dr. Danny Nedialkova
3. Prof. Dr. Gary Kleiger

Die Dissertation wurde am 16.05.22 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie
am 01.07.22 angenommen.

TABLE OF CONTENT

SUMMARY

Ubiquitylation regulates the vast majority of cellular processes through the modification of protein substrates. In addition to substrate lysine residues, internal lysines within ubiquitin undergo ubiquitylation, creating structurally distinct ubiquitin polymers which signal for specific biological outcomes. For example, K48 and K11 chains drive proteasomal degradation whereas K63 chains plays essential roles in DNA repair, protein trafficking and signaling. Numerous diseases including cancer and neurodegenerative disorders are caused by a breakdown in the Ubiquitin Proteasome System (UPS). Additionally, there is emerging interest in reprograming UPS to target disease-causing proteins for degradation, using molecular glues or Proteolysis Targeting Chimeras (PROTACs). Thus, understanding the basic principles of ubiquitin chain formation on biochemical and structural level has broad and emerging implications for human health and disease.

Sequential activity of the E1-E2-E3 enzymes cascade controls this process. It is known that E2 or E3 enzymes are responsible for the assembly of specific ubiquitin chain types. However, the mechanisms and key determinants of UB chain specificity remain elusive. The first part of this thesis moves the field one step closer towards understanding these basic principles of ubiquitin modifications. The project employed a multidisciplinary, cooperative effort to elucidate striking, atomic-level specificity of E2s and E3s. Using chemically synthetized UB analogs, we showed that shortening or extending the lysine acceptor by one hydrocarbon almost entirely abolishes ubiquitylation. Thus, revealing that the UB's acceptor lysine is an unexpected determinant of UB specificity. The phenomenon was extensively characterized using biochemistry, enzyme kinetics, NMR, MS and molecular dynamics on multiple different components of the ubiquitin system. Together, this multidisciplinary approach provided insights into the importance of proper geometry in UB chain formation and revealed pleiotropic effects of changing the acceptor lysine length, where catalysis or binding can be influenced. UB^{A}

It seems possible that in the cellular environment many natural components might impact presentation of the acceptor lysine. Binding to the protein partners or linkage within a chain might result in E2 and E3 enzymes specificity. Additionally, strong preference for the native lysine geometry may play a role in robust and sufficient

ubiquitylation for proteasomal degradation, even when the favored targeted lysines are inaccessible. Such features, may also have an effect on the success or failure of molecular glues and PROTACs.

CRLs, being responsible for more than 20% of all degradation events in the eukaryotic system remain not fully understood. The mechanism of K48-linked UB chain formation, catalyzed by UBE2R, in complex with CRL, remains elusive as well. Moreover, finding favored CRL-E2 pairs awaits further investigation. Recent studies, along with data from this study, have shown exquisite preference of UBE2R for CRL2, not only in the presence of native substrates but also in the context of neo-substrates. Thus, the second part of this thesis focuses on the structural studies of UBE2R-mediated K48-linked UB chain formation in complex with $CRL2^{VHL}$. Importantly, proceeding structural studies of $CRL2^{VHL}$ complex with Hif1α as a substrate is to the great interest of many fields. The complex is associated with von Hippel-Lindau syndrome (VHL), a disease associated especially with kidney and pancreatic cancers. Additionally, Hif1α was identified as an important transcription factor responsible for oxygen homeostasis in the cellular environment. All in all, the mechanism of Hif1α polyubiquitylation to this date is not known, despite the importance of the complex and its role in tumorigenesis. Even though the UB code writers function on millisecond time-scale, it is possible to visualize the active conformations with the help of chemical biology. As such, the electrophilic traps were designed to mimic native intermediate of K48-linked diUB formation. As a result, 4.2 Å map was solved giving an insight into the mechnism of UBE2R1-mediated polyubiquitylation. The obtained density allowed to fit crystal structures, nonetheless due to the lack of high resolution data in crucial parts of the complex, especially catalytic core, the structure could not be build and extensively analyzed on side chain level. However, many interesting features arose from the cryo-EM data. First, the CUL2 CR3, 4HB and C/R domains with RBX1 undergo major rearrangements in comparison to apo crystal structure (PDB 5N4W) to place UBE2R1-UB^D in proximity of UB^A and acceptor lysine for efficient catalysis. Second, the position of UB^A bound to UBE2R1 helix α3 could be determined. Third, essential for chain formation, acidic loop could be visualized, wraping around the active site. Fourth, required for UBE2R1's catalysis and binding, acidic tail was visible in CUL2's CTD. Lastly, the collected data suggested the allosteric role of UBE2R1's NEDD8 activation. As far as the captured model gives more insight into the UBE2R1's mode of action many questions remain unanswered. Precise position of UB^A's K48, catalytic cysteine

and UB^D's C-terminus await to be resolved. Obtained map does not provide a sufficient information to conclude catalytic site conformation. Additional structural efforts, either cryo-EM or crystal studies, need to be proceeded. High resolution data confirming many of the conclusions and visualizing not visible parts will be the next goal of this project. Furthermore, extensive kinetics studies need to be performed to verify the role of many mutants identified in this study.

ZUSAMMENFASSUNG

Die Ubiquitinierung reguliert die überwiegende Mehrheit der zellulären Prozesse durch die Modifikation von Proteinsubstraten. Zusätzlich zu den Lysinresten des Substrates durchlaufen auch interne Lysine innerhalb von Ubiquitin eine Ubiquitinierung, wodurch strukturell unterschiedliche Ubiquitinpolymere entstehen, die Signale für spezifische biologische Resultate sind. Beispielsweise treiben K48- und K11-Ketten den proteasomalen Abbau voran, während K63-Ketten eine wesentliche Rolle bei der DNA-Reparatur, dem Proteintransport und der Signalübertragung spielen. Zahlreiche Krankheiten, darunter Krebs und neurodegenerative Erkrankungen, werden durch einen Zusammenbruch des Ubiquitin-Proteasom-Systems (UPS) verursacht. Darüber hinaus gibt es ein wachsendes Interesse an der Umprogrammierung von UPS, um krankheitsverursachende Proteine für den Abbau unter Verwendung von molekularen Klebstoffen oder Proteolyse-hervorrufenden Chimeras (PROTACs) zu markieren. Daher hat das Verständnis der Grundprinzipien der Ubiquitin-Kettenbildung auf biochemischer und struktureller Ebene weitreichende und sich abzeichnende Auswirkungen auf die menschliche Gesundheit und Krankheit. Die sequentielle Aktivität der E1-E2-E3-Enzymkaskade steuert diesen Prozess. Es ist bekannt, dass E2- oder E3-Enzyme für den Aufbau spezifischer Ubiquitin-Kettentypen verantwortlich sind. Die Mechanismen und Schlüsseldeterminanten der UB-Kettenspezifität bleiben jedoch schwer fassbar. Der erste Teil dieser Arbeit bringt das Feld einen Schritt näher an das Verständnis der Grundprinzipien von Ubiquitin-Modifikationen heran. Das Projekt setzte eine multidisziplinäre, kooperative Anstrengung ein, um die bemerkenswerte Spezifität von E2s und E3s auf atomarer Ebene aufzuklären. Unter Verwendung chemisch synthetisierter UB-Analoga zeigten wir, dass die Verkürzung oder Verlängerung des Lysin-Akzeptors um einen Kohlenwasserstoff die Ubiquitinierung fast vollständig aufhebt. Dies zeigt, dass das Akzeptor-Lysin von UB eine unerwartete Determinante der UB-Spezifität ist. Das Phänomen wurde mithilfe von Biochemie, Enzymkinetik, NMR, MS und Molekulardynamik an mehreren verschiedenen Komponenten des Ubiquitin-Systems umfassend charakterisiert. Zusammengenommen lieferte dieser multidisziplinäre Ansatz Einblicke in die Bedeutung der richtigen Geometrie bei der UB-Kettenbildung und enthüllte pleiotrope Effekte der Änderung der Lysinlänge des Akzeptors, durch die die Katalyse oder Bindung beeinflusst wurden. Es scheint möglich, dass viele natürliche Komponenten

in der zellulären Umgebung die Präsentation des Akzeptor-Lysins beeinflussen. Die Bindung an die Proteinpartner oder die Verknüpfung innerhalb einer Kette könnte zur Spezifität der E2- und E3-Enzyme führen. Darüber hinaus kann eine starke Präferenz für die native Lysingeometrie eine Rolle bei der robusten und ausreichenden Ubiquitinierung für den proteasomalen Abbau spielen, selbst wenn die bevorzugten Ziellysine nicht zugänglich sind. Solche Merkmale könnten sich auch auf den Erfolg oder Misserfolg von molekularen Klebstoffen und PROTACs auswirken. CRLs, die für mehr als 20 % aller Abbauereignisse im eukaryotischen System verantwortlich sind, sind noch nicht vollständig verstanden. Auch der Mechanismus der K48-verknüpften UB-Kettenbildung, katalysiert durch UBE2R im Komplex mit CRL, ist weiterhin unklar. Darüber hinaus besteht weiterhin Forschungsbedarf im Auffinden bevorzugter CRL-E2-Paare. Jüngste Studien haben, zusammen mit Daten aus dieser Studie, eine herausragende Präferenz von UBE2R für CRL2 gezeigt, nicht nur in Gegenwart nativer Substrate, sondern auch im Zusammenhang mit Neosubstraten. Daher konzentriert sich der zweite Teil dieser Arbeit auf die Strukturstudien der UBE2R-vermittelten K48-verknüpften UB-Kettenbildung im Komplex mit $CRL2^{VHL}$. Maßgeblich ist, dass die Fortführung von Strukturstudien des $CRL2^{VHL}$-Komplexes mit Hif1α als Substrat von großem Interesse für viele Bereiche ist. Der Komplex ist mit dem von Hippel-Lindau-Syndrom (VHL) assoziiert, einer Erkrankung, die insbesondere mit Nieren- und Bauchspeicheldrüsenkrebs einhergeht. Darüber hinaus wurde Hif1α als wichtiger Transkriptionsfaktor identifiziert, der für die Sauerstoffhomöostase in der zellulären Umgebung verantwortlich ist. Insgesamt ist der Mechanismus der Hif1α-Polyubiquitinierung trotz der Bedeutung des Komplexes und seiner Rolle bei der Tumorentstehung bis heute nicht bekannt. Obwohl die Schreiber des Ubiquitincodes im Millisekunden-Maßstab arbeiten, ist es möglich, die aktiven Konformationen mithilfe der chemischen Biologie zu visualisieren. Als solche wurden die elektrophilen Fallen so gestaltet, dass sie das native Intermediat der K48-verknüpften diUB-Bildung nachahmen. Als Ergebnis wurde eine 4,2-Å-Karte gelöst, die einen Einblick in den Mechanismus der UBE2R1-vermittelten Polyubiquitinierung gibt. Die erhaltene Dichte ermöglichte das Einsetzen von Kristallstrukturen, jedoch konnte die Struktur aufgrund des Mangels an hochauflösenden Daten in entscheidenden Teilen des Komplexes, insbesondere im katalytischen Kern, nicht aufgebaut und auf Seitenkettenebene umfassend analysiert werden. Aus den CryoEM-Daten ergaben sich jedoch viele interessante Merkmale. Erstens werden die CUL2-CR3-, 4HB- und C/R-Domänen mit

RBX1 im Vergleich zur Apo-Kristallstruktur (PDB 5N4W) erheblichen Umlagerungen unterzogen, um UBE2R1-UB^D für eine effiziente Katalyse in die Nähe von UB^A und Akzeptor-Lysin zu bringen. Zweitens konnte die Position von an die Helix α3 von UBE2R1 gebundenem UB^A bestimmt werden. Drittens konnte eine für die Kettenbildung wesentliche Säureschleife sichtbar gemacht werden, die sich um das aktive Zentrum wickelt. Viertens war der für die Katalyse und Bindung von UBE2R1‘s erforderliche saure Schwanz in der CTD von CUL2 sichtbar. Schließlich deuteten die gesammelten Daten auf die allosterische Rolle der NEDD8-Aktivierung von UBE2R1 hin. Obwohl das erfasste Modell mehr Einblick in die Wirkungsweise von UBE2R1 gibt, bleiben viele Fragen unbeantwortet. Die genaue Position von UB^A‘s K48, katalytischem Cystein und UB^D‘s C-Terminus muss noch geklärt werden. Die erhaltene Karte liefert keine ausreichenden Informationen, um die Konformation des katalytischen Zentrums zu schlussfolgern. Zusätzliche strukturelle Bemühungen, entweder durch CryoEM oder Kristallistationsstudien, müssen fortgesetzt werden. Hochauflösende Daten, die viele der Schlussfolgerungen bestätigen und nicht sichtbare Teile visualisieren, werden das nächste Ziel dieses Projekts sein. Darüber hinaus müssen umfangreiche Kinetikstudien durchgeführt werden, um die Rolle vieler in dieser Studie identifizierter Mutanten zu verifizieren.

ACKNOWLEDGEMENTS

It is difficult to describe my gratitude to Prof. Brenda Schulman, my project's principal investigator and mentor, who gave me this amazing opportunity to work in a top world-class laboratory. I want to especially thank Brenda for making me a better scientist, changing my way of thinking about experiments and complicated enzymatic systems. I know, that the possibility of doing my PhD in Brenda's laboratory not only massively transformed my career but also transformed me as a person. For that Brenda, I will be always grateful. I wouldn't be where I am without your help along the way.

The success of these projects wouldn't be possible without support and mentoring from Prof. Gary Kleiger, who was extremely helpful throughout my whole PhD. Thank you, Gary, for being in my Thesis Advisory Committee and your availability to discuss experiments, trouble-shoot and brainstorm, as well as all of your experimental inputs.

I would like to express my gratitude to Prof. Elena Conti, third member of my Thesis Advisory Committee, for all the advice on the project.

I would like to thank all my collaborators on the lysine ruler project - the great, multidisciplinary insight into the chain formation mechanisms wouldn't be possible without all the people involved in the project. Our study strictly relied on chemical biology tools, thus I want to express my great appreciation to Prof. Huib Ovaa and Dr. Gerbrand J.van der Heden van Noort. Additionally, I want to thank: Dr. David Krist, Fynn Hansen, Vinh Truong, Dr. Ozge Karayel, Nicholas Purser, Daniel Houston, Nicole Burton, Dr. Mark Bostock, Prof. Michael Sattler, Prof. Matthias Mann and Prof. Joseph Harrison for making our paper possible.

I greatly appreciate all the help from Schulman Laboratory members, especially Dr. Jesuraj Prabu, Susanne von Gronau and the CRL team members for scientific and technical support.

I am thankful Dr. Daniel Bollschweiler and Dr. Tillman Schaefer from the Cryo-EM facility of the Max Planck Institute for all the help with the microscopes.

I am happy to be able to have an assistance of many facilities: Bioorganic Chemistry & Biophysics Core Facility, Crystallography Facility and Mass Spectrometry Facility.

Many thanks to IMPRS coordinators: Dr. Hans-Joerg Schaeffer, Dr. Marta Cipinska, Dr. Ingrid Wolf, and Maximilian Reif for all the help.

I can not imagine to have better friends by my side. Kuba, Gosia and Blazej, you are my polish family in Munich. Laura and Ishi – it was a pleasure to have you by my side. Thank you all for your amazing support scientifically and personally.

Last, but not least, I am touched beyond words by the care and help of my boyfriend, Renato and my family. They were always there for me.

ABBREVIATIONS

4-HB - 4-Helix Bundle
ABP - Activity-based Probe
APC/C - Anaphase-Promoting Complex
BARD1 - BRCA1-Associated RING domain 1
BRCA1- Breast Cancer 1
c-Cbl - casitas b-lineage lymphoma
CR - CUL-Repeat
CRL - Cullin-RING Ligase
CTD - C-Terminal Domain
CUL - Cullin
CV - Column Volume
DUB - Deubiquitylating enzyme
E1 - UB-activating enzyme
E2 - UB-conjugating enzyme
E3 - Ubiquitin ligase
E6AP - E6-Associated Protein
EGFR - Epidermal Growth factor receptor
ELOB/C - Elongin B and Elongin C
FEM-1 - Feminization-1
FT - Flow-Through
GST - Glutathione S-Transferase
HECT - Homologous to E6AP C terminus
Hif1 - Hypoxia-inducible factor 1
HOIP - HOIL-1L Interacting Protein
hPRC1L - Polycomb Repressive Complex 1-Like
HRMS - High Resolution Mass Spectrometry
HTH - Helix-Turn-Helix
IBR - In-Between-RING
IEX - Ion-Exchange
LC/MS - Liquid Chromatography- Mass Spectrometry
LRR-1 - Leucine-Rich Repeat protein 1
LUBAC - Linear UB Chain Assembly

MD - Molecular Dynamics

MS - Mass Spectrometry

NTD - N-Terminal Domain

O/N - overnight

PRAME - Preferentially Expressed Antigene in Melanome

PROTACs - Proteolysis Targeting Chimeras

RACK1 - Receptor for the Activated C Kinase 1

RBR - RING-between-RING

RING - Really Interesting New Gene

RMSD - Root Mean Square Deviation

RTK - Receptor Tyrosine Kinase

SCF - Skp, Cullin, F-box

SEC - Size-Exclusion Chromatography

SIL - Stable Isotope-Labeled

SIM - Selected Ion Monitoring

SPPS - Solid Peptide Synthesis

SUMO1 - Small UB-related Modifier 1

TEM - Transmission Electron Microscope

TEV - Tobacco Etch Virus Protease

UB - Ubiquitin

UB^A - Acceptor Ubiquitin

UB^A - UB-Associated

UBC - UB-Conjugating

UB^D - donor Ubiquitin

UB^D - Ubiquitin-Binding domain

UBL - UB-Like

UEV - UB E2 Variant

UPS - Ubiquitin Proteasome System

USP - Ubiquitin-Specific Protease

VHL - von Hippel-Lindau

WHB - Winged-Helix B

WT - Wild Type

ZF - Zinc-Finger

1. INTRODUCTION

1.1 The UB system

Ubiquitylation is a post-translational modification, which controls the vast majority of eukaryotic, cellular processes, such as protein trafficking, quality control, signal transduction, differentiation and cell division[1-3]. The outcome of ubiquitylation depends on the chain topology, substrate and enzyme localization and interactions with other effectors[4,5]. The most known function of ubiquitin is targeting proteins to the proteasome and their subsequent degradation. As such, specific K48 polyubiquitin chains need to be assembled on the protein substrate[6]. Additionally, the failure in ubiquitin system was found to be involved in a variety of cancers, inflammatory diseases and neurodegenerative disorders[7].

1.1.1 UB

Ubiquitin (UB) is a small, 76-amino acid, highly stable and conserved protein[8]. The UB core structure adopts a β-grasp fold and it contains a six-residue long C-terminal, flexible extension[9] (Figure 1.1a). UB's tail conjugates with catalytic cysteines of UB cascade enzymes and other UBs to form chains[10-12] (see section 1.1.2, Ubiquitylation cascade). This small protein contains a couple of hydrophobic recognition surfaces (so called F4, I36, I44 patches and TEK-box), which play an important role in UB's interactome (Figure 1.1b)[13]. The F4 patch (Q2, F4 and T12) seems to function in trafficking[14] and is essential for yeast's cell division[14] and interaction with Ubiquitin-Specific Protease (USP) domain belonging to Deubiquitylating enzymes (DUBs)[15]. Containing L71, L73 and UB's tail, the I36 patch moderates interactions between UB moieties in chains and it is being recognized by DUBs[15], Ubiquitin-Binding domains (UB^Ds)[16] and Homologous to E6AP C terminus (HECTs) E3s[17]. Residues L8, I44, H68 and V70 belong to the I44 patch. This hydrophobic surface binds to most UB^Ds and the proteasome, thus being crucial for cell division[13,14,18]. Lastly, the TEK-box of UB carry K6, K11, T12, T14 and E34 and it is involved in mitotic degradation[19].The most essential parts of UB are its seven lysines, distributed throughout the whole protein and N-terminus, which are modified to form UB chains (Figure 1.1c)[8]. Details regarding UB chain formation are described in 1.1.3.

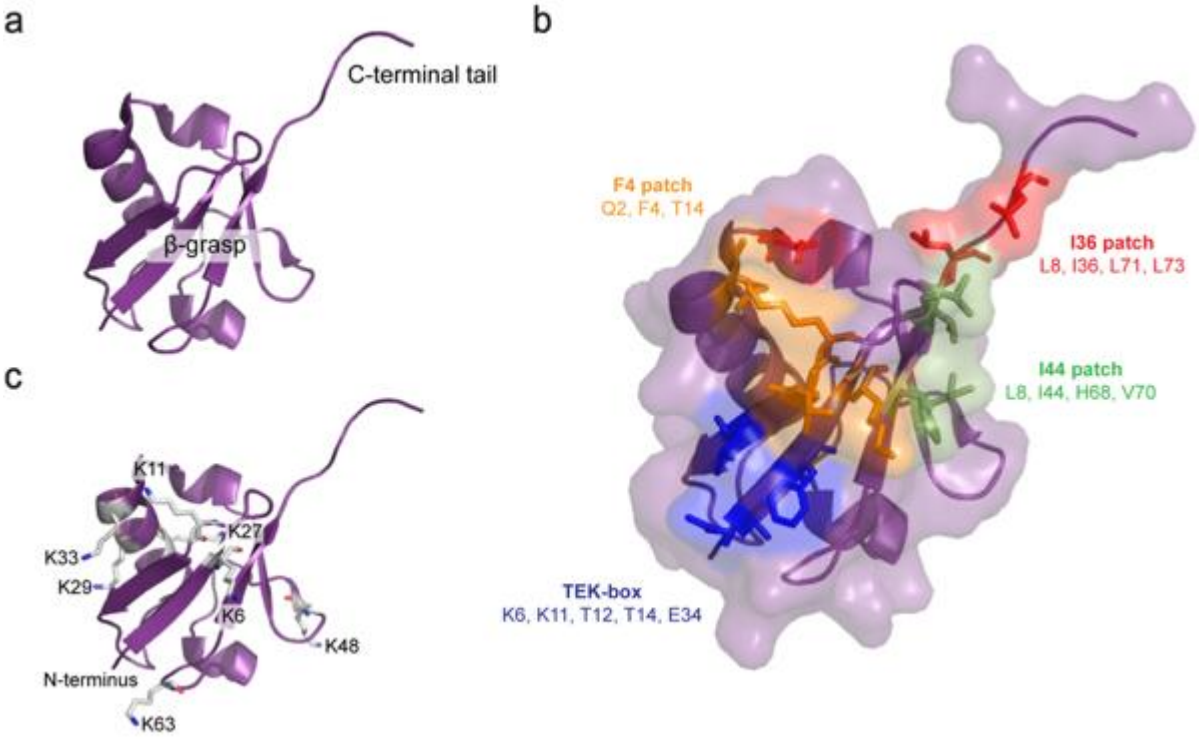


Figure 1.1 Structure of UB with its essential features (PDB 3CMM). (a) Structure of UB depicting β-grasp and C-terminal tail. (b) Structure of UB with marked hydrophobic patches. (c) Structure of UB with all seven lysines and the N-terminus depicted.

1.1.2 Ubiquitylation cascade

The ubiquitylation cascade starts with UB activation. First, catalyzed by UB-activating enzyme (E1), UB's C-terminus undergoes acyl-adenylation in an ATP- and Mg-dependent manner (Figure 1.2, step 1-2)[11]. Subsequently, E1's catalytic cysteine attacks UB-adenylate to form thioester-bound UB (E1~UB, where "~" refers to a thioester bond) (Figure 1.2, step 1-2)[11]. Secondly, activated UB (here called donor UB (UB^D)) is transferred to the catalytic cysteine of a UB-conjugating enzyme (E2) in a transthiolation reaction (Figure 1.2, step 3)[20]. Further, E2 can either interact with Really Interesting New Gene (RING) type of UB ligase (E3) to directly transfer UB to the substrate lysine and form an isopeptide bond between the acceptor lysine and C-terminus of UB or pass UB in another transthiolation reaction to the catalytic cysteine of an E3 (Figure 1.2, step 4)[21]. These types of E3s are divided into two classes: HECTs and RING-between-RINGs (RBRs) (Figure 1.2, step 4)[21]. Finally, HECT and RBR E3s catalyze isopeptide bond formation on a target lysine either on substrate protein or another UB (here called acceptor UB (UB^A))((Figure 1.2, step 5)[21]. More about specific

E2s and E3s is discussed in section 1.2 and 1.3. Taken together, the UB system is orchestrated by sequential activity of different classes of enzymes and it is uniquely complex with 2 E1s, around 40 E2s and around 600 of E3s[10,11]. The ubiquitylation system also contains around 100 DUBs, removing this post-translation modification from protein substrates[22].

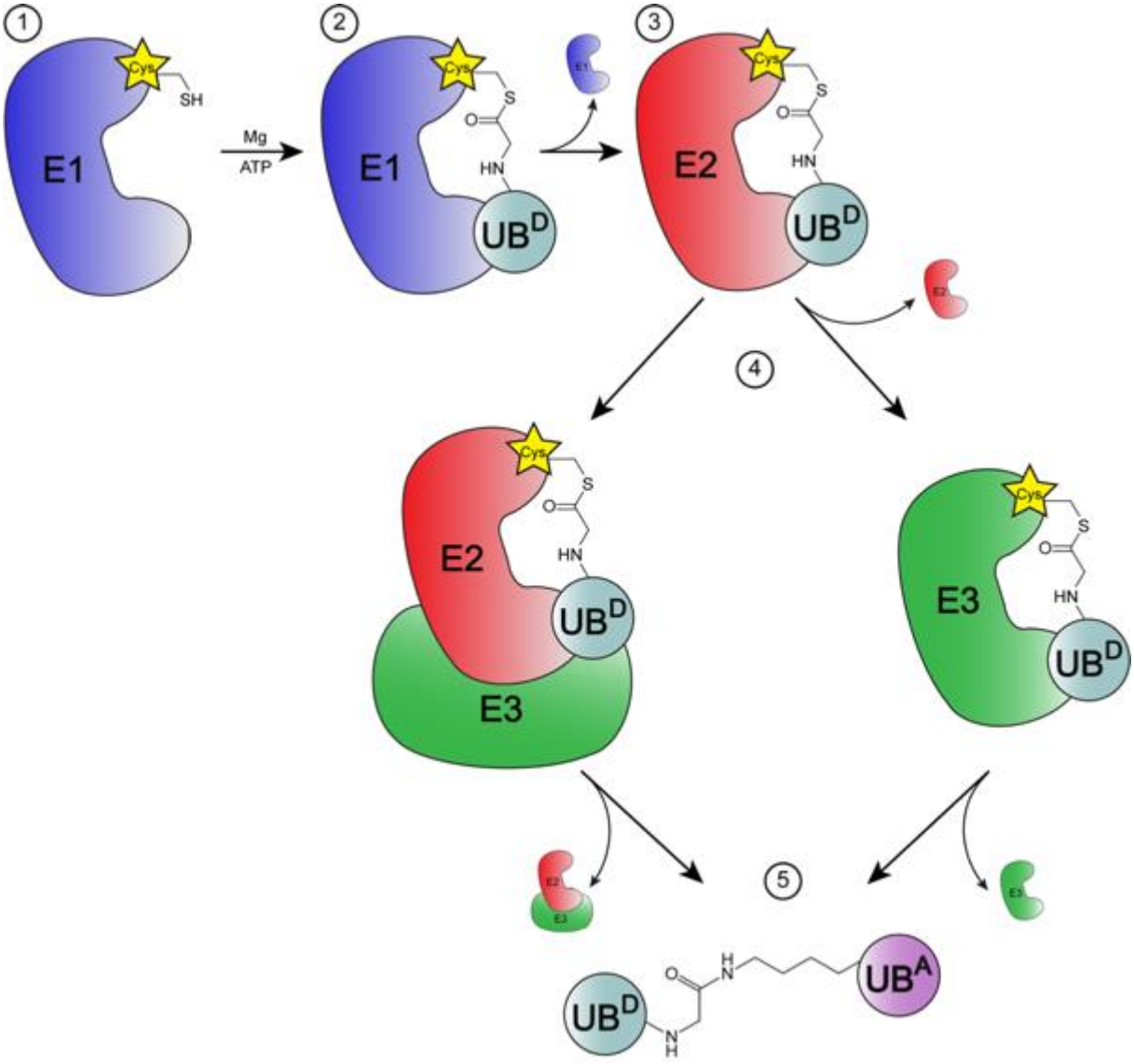


Figure 1.2 Ubiquitylation cascade. Step 1-2: UB gets activated in an ATP and Mg-dependent manner and gets conjugated to E1's catalytic cysteine to form a thioester bond. Step 3: E1 catalyzed transthiolation reaction to transfer UB C-terminus to E2's catalytic cysteine to form a thioester bond. Step 4: E2 can either be recruited by RING E3 ligases to stimulate UB transfer to a targeted lysine (left site) or pass UB to E3's catalytic cysteine (HECT and RBR type of E3 ligases) in a transthiolation reaction (right site). Step 5: UB's C-terminus gets conjugated to acceptor lysine on another UB or substrate to form an isopeptide bond.

1.1.3 UB chains

Ubiquitylation cascade may result only in the monoubiquitylation of a target protein, where just a single UB moiety is attached to the substrate's lysine (Figure 1.3a)[8]. That is the case for modification of histone H2A by human Polycomb Repressive Complex 1-Like (hPRC1L)[23]. The E3 ligase, Bmi1/Ring1b, which is a part of hPRC1L, collaborates with the E2 UBE2D to transfer UB onto H2A's K119[24]. Histone monoubiquitylation plays a role in gene activation and it is one of the major histone modification in eukaryotic cells[25].

Some substrates, like the member of Receptor Tyrosine Kinases (RTK) family: Epidermal Growth Factor Receptor (EGFR) undergo multi-monoubiquitylation (Figure 1.3)[26]. EGFR is associated with many important cellular functions such as cell proliferation, differentiation and survival[27]. The ligand-dependent ubiquitylation is involved in the receptor's endocytosis and degradation by the lysosome[26,28].

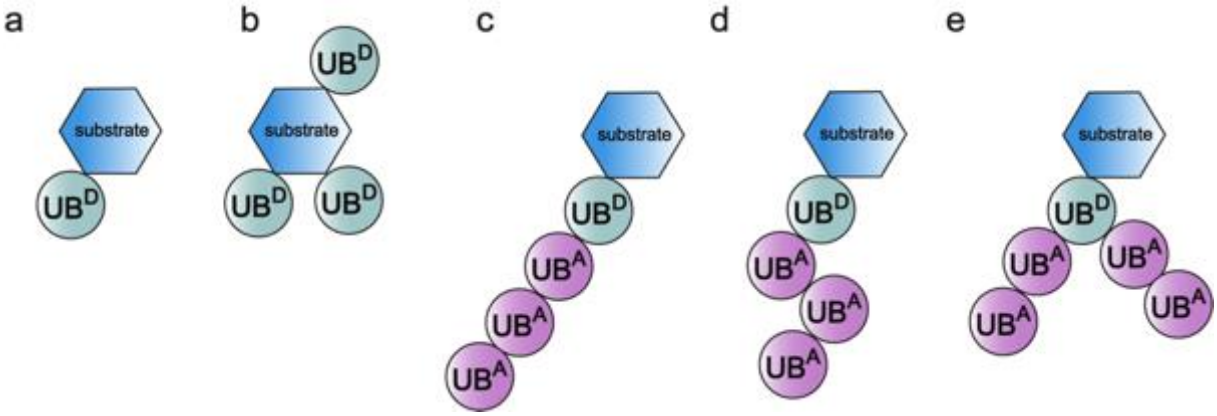


Figure 1.3 The complexity of UB code. (a) Monoubiquitylation of the substrate, targeting single lysine. (b) Multi-monoubiquitylation of the substrate, targeting many lysines by single UB. (c) Single linkage chain formation on the substrate, targeting one type of UB's lysine. (d,e) Branched chain formation, targeting multiple lysines of UB.

In most of the cases, sequential activity of E1-E2-E3 cascade leads to protein polyubiquitylation (Figure 1.3 c-e)[8]. As mentioned above, UB chains can be formed via one of the seven lysines and the N-terminus of UB (Section 1.1.1 Figure 1.1c)[8]. Two different types of chains can be determined: homotypic, where UBs are connected by a single linkage type and heterotypic, where UBs are bridged by multiple linkages[29]. The most abundant homotypic UB connections in cells are K48-linked chains, targeting

substrate proteins for degradation by the 26S proteasome[30]. There are multiple E2-E3 pairs assembling K48 linkages. SCF E3 ligases are one of the best characterized K48 chain building complexes responsible for a cell cycle regulation and it will be further discussed in section 1.6[31]. Another well described UB chain topology is a K63-linked one, with its many non-proteasomal functions[29]. In contrast, K63-linked chains modulate DNA repair[32], NF-κB transcription factor activation[33] and innate immune response[34]. The UBE2N/UBE2V1 E2 pair is a model K63-linked chain former[35]. K11 linkages play a role in cell cycle regulation, via proteasomal degradation and were detected to be preferentially produced in early G1 phase as well as mitosis[36]. A well-known K11-linked chain former is the Anaphase-Promoting Complex (APC/C) with its designated E2, UBE2S[37]. N-terminally (M1)-linked linear chains were shown to be quickly assembled upon inflammatory signaling cascades activation[38-40], by a RBR E3 ligase LUBAC. Other UB linkages remain poorly characterized. K6-linked chains were associated with removal of damaged mitochondria[41], K27-linked chains were associated with proteasomal degradation[42] and K33 linkages are thought to regulate trafficking via trans-Golgi network[43]. Moreover, heterotypic chain types are further divided into mixed chains, where UB is modified with only one additional UB moiety but different lysines and branched chains, where a single UB can be modified with multiple UBs (Figure 1.3 d,e)[29]. Formation of K11/K48 branched chains was shown to enhance proteolytic signal[44]. A compelling example of branched chain former is the HECT E3 ligase, TRIP12, which was recently discovered to build K29-linked UBs on a preassembled K48 chain in the context of small-molecule-degrader and its neo-substrate, BRD4[45]. Many functions of branched UB chains still remain unclear.

1.2 UB-conjugating enzymes

1.2.1 E2 structure

The E2 family commonly contains a highly homologous UB-Conjugating (UBC) core domain with the active site cysteine and E1 interaction site (Figure 1.4)[12]. Contrary, UB E2 Variant (UEV) possesses a UBC domain without the catalytic cysteine[35]. The UBC domain consists of a classic α/β-fold with four α-helixes and anti-parallel four-stranded β-sheet (Figure 1.4). The catalytic cysteine is located in a groove in the central part of the protein (Figure 1.4)[46-48]. The backside of the UBC domain can non-

covalently bind UB or UB-Like (UBL) proteins to further stimulate the activity[49,50]. Apart from the canonical fold, a couple of E2s accommodate additional insertions[20]. As such, UBE2R (which is discussed in details in section 1.5) and UBE2G contain an extra loop nearby the active site[51]. Moreover, extensions of the UBC domain are a common feature of many E2s[20]. For example, UBE2R consists of a C-terminal acidic tail (section 1.5.5) and a UBE2K UB-Associated domain (UB^A)[52,53]. These auxiliary insertions and extensions encompass for additional features and functions of different E2s[20].

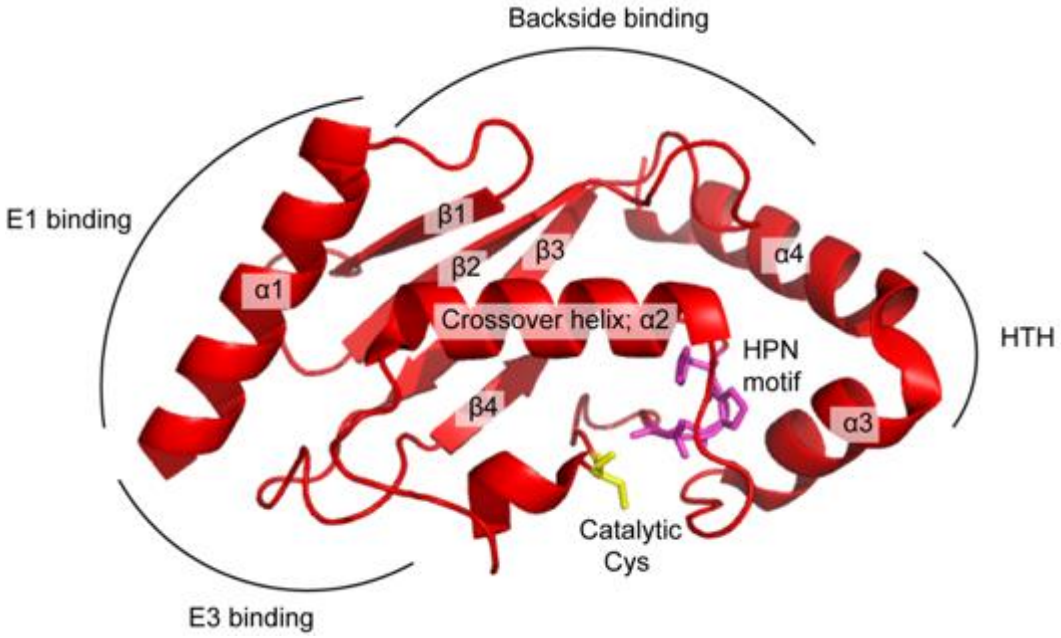


Figure 1.4 Conserved structure of E2's catalytic core (UBC domain) (PDB 2ESK). The catalytic cysteine is depicted with conserved HPN motif. E1, E3, and backside binding sites and HTH motif are shown. Helices and β-sheets are marked.

1.2.2 E2~UB thioester

E2~UB conjugate is highly flexible with multiple conformations occurring between the two proteins (Figure 1.5 a)[54]. Thus, this open state is not beneficial for further transfer of UB^D. One E2 can associate with multiple E3s and upon binding, the E2~UB closed conformation is favored (Figure 1.5 a,b). Multiple crystal structures have shown closed conformation between E2~UB and the RING domain of an E3 ligase. In a closed state, UB^D's hydrophobic I44 patch is interacting with E2's crossover helix (Figure 1.5 b). This is further stabilized by an arginine from the RING domain (so called "Linchpin

Arginine"), bridging E2 and UB. Additionally, with UB folded back to the E2, the thioester bond is positioned for the nucleophilic attack of the acceptor's lysine and primed for catalysis (more about the mechanism of E2 catalysis is described in section 1.4.1)[55-57].

Interestingly, recent studies showed that UBE2S's UBC's domain HTH (Helix-Turn-Helix) motif (Figure 1.4) functions to recruit UB^D, thus promoting the open state[58]. UBE2S~UB binding to APC/C interrupts HTH-UB^D interactions and supports the closed conformation[58].

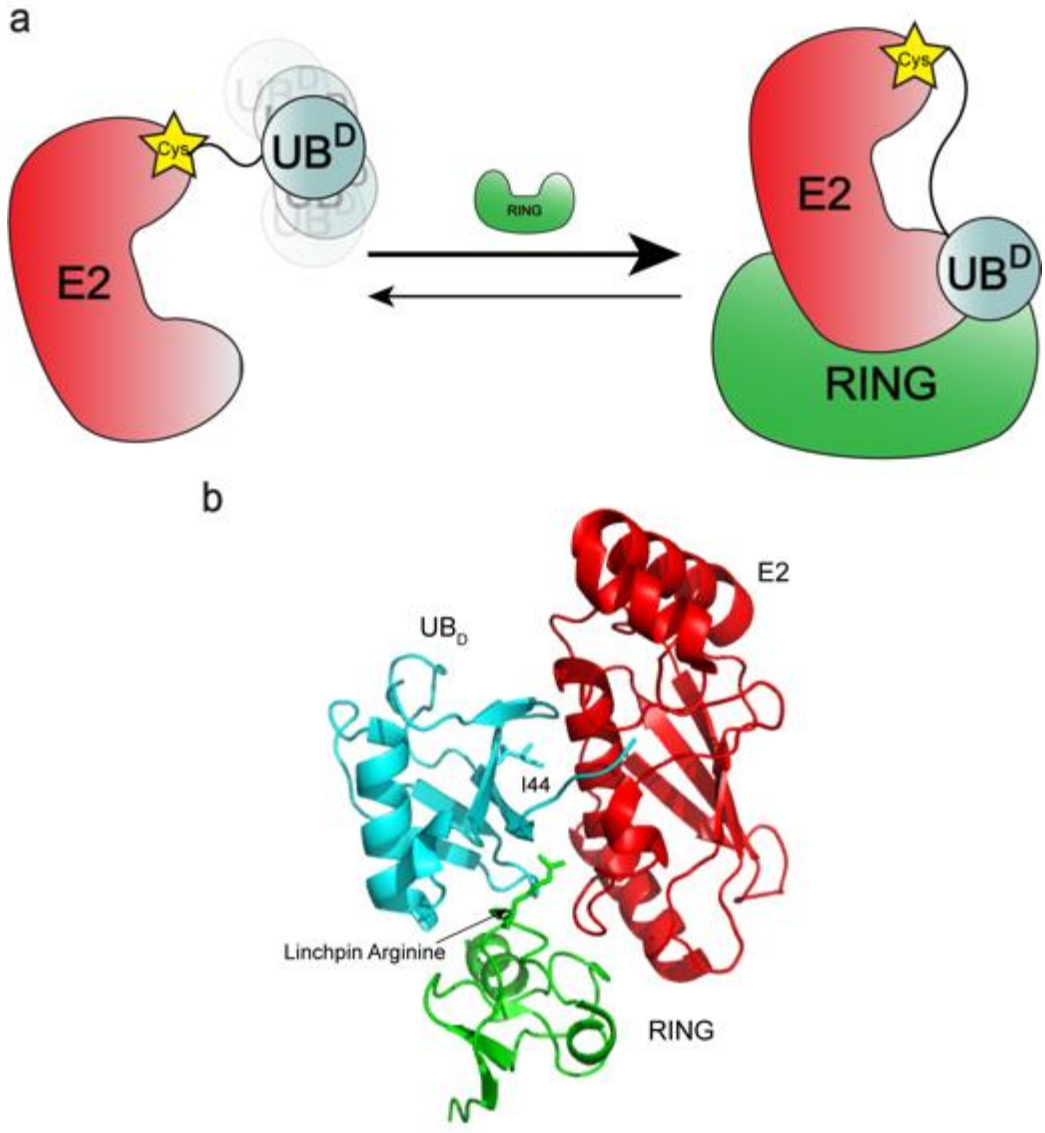


Figure 1.5 RING domain stimulates closed conformation. (a) Cartoon representation of RING-stimulated stabilization of flexible E2~UB. (b) Structural representation of E2~UB bound to the RING domain. Linchipin arginine and UB^D I44 are depicted (PDB 4AUQ).

1.2.3 Intrinsic activity of E2 enzymes

Even though most of the E2s share high sequence and structure homology, their activity varies substantially[20]. As demonstrated in section 3.1.1, the easiest assay to assess such intrinsic reactivity is E2~UB discharge towards small-molecules[20]. As such, most of the E2s catalyze release of UB^D to free nucleophile – lysine, which is consistent with their ability to work with RING E3 ligases[20]. In contrast, UBE2L3 was shown to discharge only to free cysteine, thus confirming their ability to work with HECT and RBR types of E3s[59]. Interestingly, UBE2W is known to react in the presence of lysine-less peptides, transferring UB^D to an α-amino group[60]. Indeed, studies have shown UBE2W involvement in substrate's disordered N-terminal recognition and modification[60].

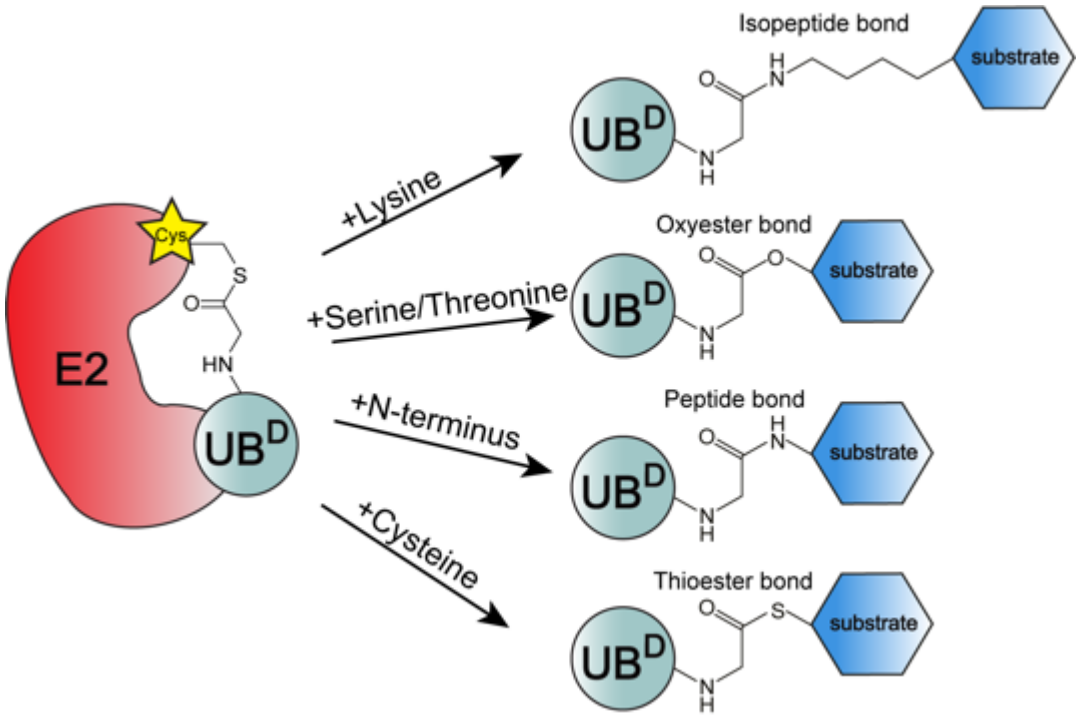


Figure 1.6 Intrinsic activity of the E2 enzymes. Thioester bound UB^D can be transferred to the substrate's: (1) lysine to form isopeptide bond, (2) serine or threonine to form oxyester bond, (3) N-termius amine to form peptide bond, (4) cysteine to form thioester bond.

Additionally, another E2, UBE2J2 functions to transfer UB^D to serines and threonines of the histocompatibility complex together with the viral murine K3 RING E3 ligase[61,62]. Taken together E2s can catalyze the formation of isopeptide, oxyester, peptide and

thioester bonds as shown in Figure 1.6[20]. At the top of different E2's activities towards various amino acids and groups, most of the E2s exhibit exceptional specificity towards specific lysines on the substrate and/or acceptor UB. More about the E2 catalysis mechanism as well as the lysine specificity is described in section 1.4.

1.3 UB ligases

At the end of the ubiquitylation cascade E3 UB ligases come to play[63]. With around 600 enzymes and three major classes (RING, HECT and RBR), E3s are most diverse amongst UB enzymes[21].

RING E3 ligases are the largest family characterized by the presence of a RING or U-box domain[10]. They act as a scaffold, bringing substrate and E2~UB together for catalysis. In some cases, the catalytic and substrate-recruiting part are a single protein as for Casitas b-lineage lymphoma (c-Cbl) ligase[64]. In other cases, these modules appear as multiprotein complexes like for: Breast Cancer 1 (BRCA1) and BRCA1-Associated RING domain 1 (BARD1), APC and Cullin-RING Ligases (CRLs)[65-68]. CRLs are described in more detail in section 1.6. The mechanism of E2~UB activation by RING E3 is characterized above (section 1.2.2).

With around 30 members, all HECT E3 ligases contain homologous, C-terminal, catalytic HECT domain, with its most characterized members: E6-Associated Protein (E6AP) and NEDD4 family[69,70]. The reactive core consists of two lobes: N- and C-lobe with a short linker between them. This highly flexible domain accounts for: (1) binding E2~UB^D to the N-lobe, (2) superposition of N-lobe-bound E2~UB^D and C-lobe's catalytic cysteine, (3) further movement to modify substrate (Figure 1.7 a-c)[17,71-75]. In more detail, first, E2~UB^D adopts an open conformation upon association to the HECT domain (Figure 1.7 a)[17]. Secondly, the C-lobe positions UB^D by contacting its C-terminus for a transthiolation reaction. After the transfer, interaction between UB^D and C-lobe is maintained, however, the tail conformation extends. This resembles RING-E2~UB^D pre-thioester aminolysis state (Figure 1.7 b)[74]. Lastly, C-lobe rotates to position HECT~UB^D in vicinity of substrate's lysine (Figure 1.7 c)[75].

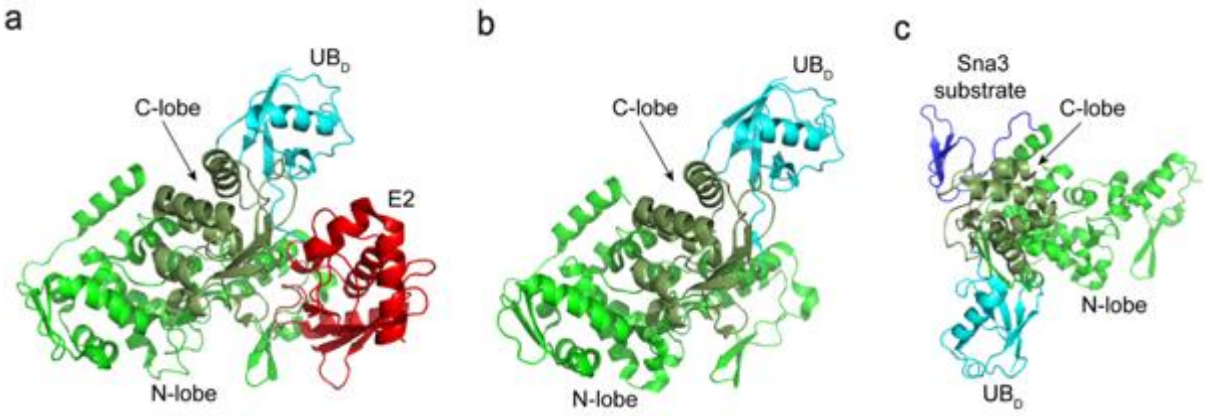


Figure 1.7 Structural representation of HECT E3 ligases mode of action. (a) N-lobe HECT domain stimulates E2-UB^D closed conformation and positions UB^D in vicinity of C-lobe (PDB 3JW0). (b) C-lobe-bound UB^D structure (PDB 4BBN). (c) C-lobe-UB^D rotation for substrate ubiquitylation (PDB 4LCD).

RBRs were discovered to be a hybrid between RING and HECT E3 ligases[59,76,77]. With two RING domains (RING1 and RING2) and the central In-Between-RINGs (IBR) domain, the catalytic architecture differs from other E3 ligases. RING1 recruits E2~UB and RING2, containing the catalytic cysteine, receives UB^D, which is then further transferred to the substrate's lysine[21]. More than ten RBRs were found in the human genome[78]. Most known and well-studied RBRs include: (1) PARKIN (protein mutated in Parkinson disease), (2) HOIL-1L Interacting Protein (HOIP), part of the LUBAC (Linear UB Chain Assembly) complex playing a role in NF-κB signaling, (3) ARIH1 and TRIAD, proteins associated with CRLs[79-82]. The first crystal structure of PARKIN RBR revealed a compact conformation, with buried RING2's catalytic cysteine, suggesting an autoinhibited state (Figure 1.8 a)[76]. Different topology was found in a co-crystal structure of activated HOIP and UBCH5B-UB^D. Surprisingly, additional UB moiety was observed, packing against RING1-IBR and allosteric activation/release from autoinhibition mechanism was proposed (Figure 1.8 b)[77]. Another example of RBR activation by UBL is shown and described in more detail in section 1.6.4. In summary, RBRs exhibit different topology, mechanism of action and activation[21].

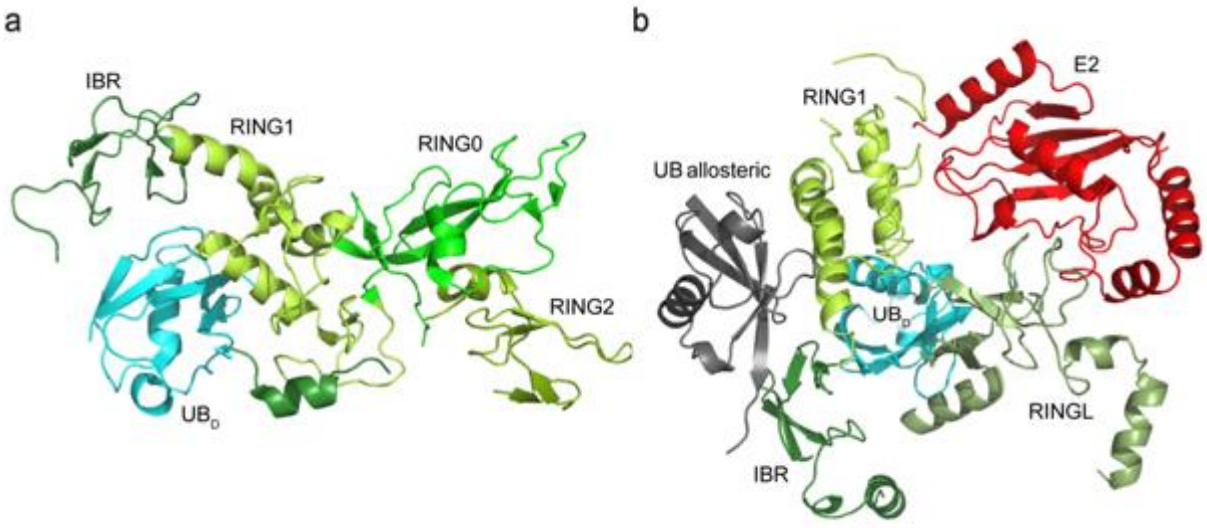


Figure 1.8 Structural representation of RBR E3 ligase's mode of action. (a) Autoinhibited state of PARKIN RBR (PDB 4K95). (b) Co-crystal structure of UBCH5B-UB and HOIP RBR in activated state with allosterically bound UB (PDB 5EDV).

1.4 Acceptor lysine

1.4.1 Mechanism of acceptor lysine activation

Final step of UB cascade involves transfer of UB^D's C-terminus from E2/E3's catalytic cysteine most often to the ε-amino group of substrate's acceptor lysine and isopeptide bond formation[83]. Pioneering studies to decipher the mechanism of this reaction catalyzed by E2/E3 have focused on Small UB-related Modifier 1 (SUMO1) – UBL and its cognate E2 Ubc9[83]. The co-crystal structure of Ubc9 with SUMO-RanGAP1 substrate revealed the acceptor lysine in the active site (Figure 1.9). Close inspection of the active site as well as careful biochemical characterization reported the role of E2 active site residues in the lysine activation. First, it requires nucleophile formation, through deprotonation of the attacking lysine's ε-amino group (and/or protonation of the leaving group of E2's catalytic cysteine). This deprotonation occurs thanks to the unique microenvironment of E2 active site, which allows to suppress lysine pK_a[83]. Second, lysine is guided and positioned towards the thioester bond. Several conserved residues (N85, Y87 and D127) from E2 active site play a role in these two steps (Figure 1.9).

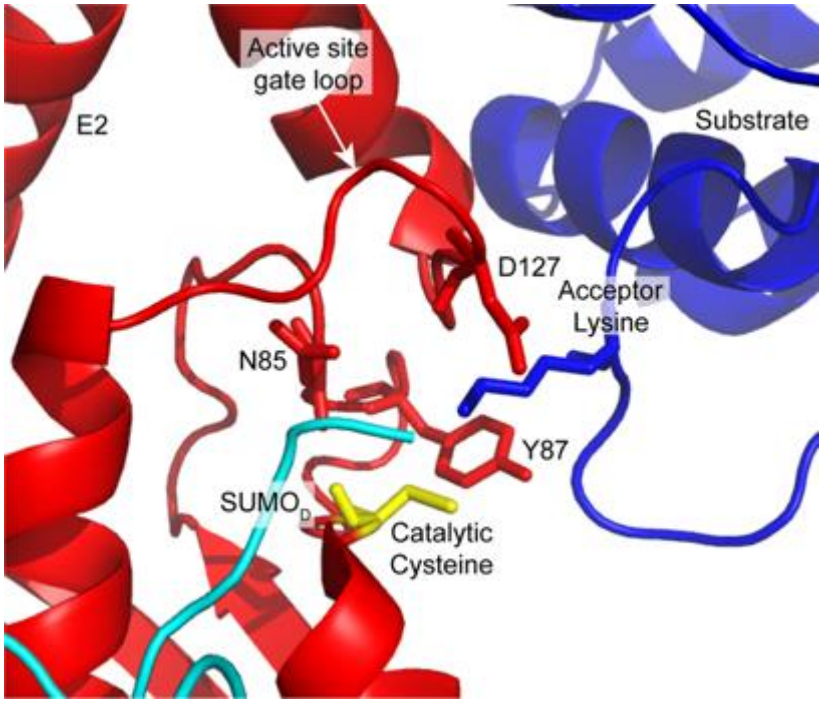


Figure 1.9 Acceptor lysine is activated by multiple residues in E2 active site. Amino acids playing a role in isopeptide bond formation catalysis are depicted. Substrate's acceptor lysine, catalytic cysteine and active site gate loop are shown (PDB 1Z5S).

Upon mutation of N85Q, Y87A and D127A, the isopeptide bond formation between SUMO and substrate was impaired, but not binding (Figure 1.10 a). Additional assays revealed, that previously mentioned mutations failed to effectively suppress pK_a of the acceptor lysine (Figure 1.10 b), thus accounting for slower catalysis. Computational data suggested pK_a suppression via desolvation. In detail, the optimal water surrounding lysine and hydrogen bonds with solvent get replaced by suboptimal contacts with E2's active site side chains and solvent, which promotes deprotonation[83]. Additionally, D127 as well as Y87 (as a hydrophobic platform) play a role in proper lysine coordination (Figure 1.10 c). Next, the nucleophile attacks the thioester bond and an oxyanion intermediate is formed[83]. The highly conserved N85 was proposed to serve as an oxyanion hole and functions to stabilize the negative tetrahedral intermediate (Figure 1.10 c)[84,85]. Furthermore, N85 was shown to position the active site loop containing D127, which functions as a gate to the active site[83,85,86]. Taken together, many E2's conserved residues participate in lysine activation to achieve reactive configuration and further to catalyze isopeptide bond formation[83-85].

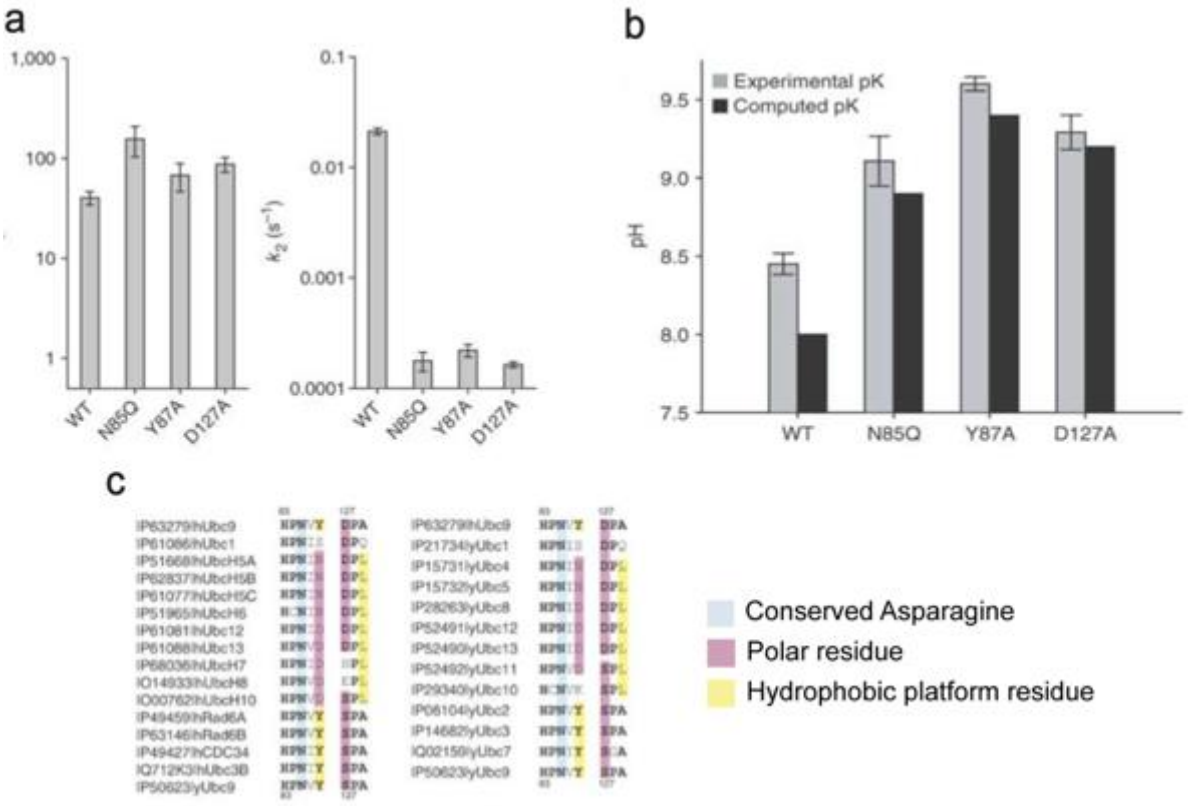


Figure 1.10 The role of E2 active site residues in acceptor lysine activation. (a) Mutations of N85, Y87 and D127 in Ubc9 active site impairs catalysis. (b) N85, Y87 and D127 are playing a role in acceptor lysine *pK*$_a$ suppression. (c) E2 multiple sequence alignment depicting conserved asparagine, polar residues and hydrophobic platform residues important for catalysis. Adapted from Yunus et al 2006[83].

1.4.2 Lysine specificity

The UB code relies on accuracy and specificity. First, in many cases, it is crucial which targeted protein's lysine gets ubiquitinylated along with its recruitment and recognition by E3. Second, as the fate of the substrate often depends on the topology of the built UB chain, it is essential that proper lysine on the acceptor UB is modified. In general, the last enzyme transferring UB^D to the acceptor lysine dictates specificity: E2s in the presence of RING E3s, HECTs and RBRs[87].

1.4.2.1 Target protein recognition and substrate lysine specificity

In some cases, target selection limits to its recruitment by UB machinery, usually via a domain outside of the catalytic part of E3[87]. Degrons, sequences of substrate proteins

that gets recognized by E3s, mediate this interaction[88,89]. Interestingly, in many instances these degrons can be post-translationally modified (for example hydroxylated and phosphorylated), which adds additional layer of regulation[90-92]. Three different lysine selectivity mechanisms were described: (1) E3 ligases with low lysine specificity and large ubiquitylation zone, like in case of Skp, Cullin, F-box (SCF) ARIH1 and p27[93-97]; (2) lysine-specific E3 ligases having limited area of action via structural constraints like for $SCF^{\beta\text{-}TrCP}$ and $I\kappa\beta\alpha$[90]; (3) E3 ligases achieving specificity, when only one lysine is available for ubiquitylation like for the HECT E3 Rsp5p and Sna3p[98].

1.4.2.2 Lysine specificity in chain formation

Polyubiquitin chain forming E2s and E3s catalyze a reaction, where UB itself is a substrate. Frequently, only one specific lysine, out of seven available, gets targeted. To achieve specificity, E2s are known to bind and orient UB^A in a way that precisely one lysine is placed in vicinity of the active site. As mentioned before, UBE2K, a K48-linked chain former, possesses an additional UB^A domain[99]. A recent crystal structure of trapped complex revealed the mode of UB^A recruitment by UB^A to orient K48 for UBE2K~UB^D attack (Figure 1.11 a)[53]. Moreover, UBE2N cooperates with UBE2V, a catalytically inactive E2 variant, to modify UB^A's K63[35]. The backside of UBE2V's UBC domain was found to non-covalently bind UB^A and direct K63 for modification as shown in Figure 1.11 b[48,100]. On the other hand, UBE2S that is forming K11-linked chains, uses a different mechanism to ensure specificity[101]. Upon UB^D recruitment, UBE2S binds UB^A via electrostatic interactions to bring K11 into the active site. Further, E34 on UB^A suppresses acceptor's lysine pK_a. This substrate-assisted catalysis supports effective isopeptide bond formation between UBs (Figure 1.12)[101]. For HECT E3 ligases,the C-terminal part of the C-lobe was found to participate in chain specificity[74,102]. Crystal structure of HOIP, a member of RBR family showed RING2 and LDD domain to position UB^A for linear UB chain formation (Figure 1.11 c)[103,104]. Thus, many different mechanisms are adapted by E2s and E3s to catalyze specific chain formation.

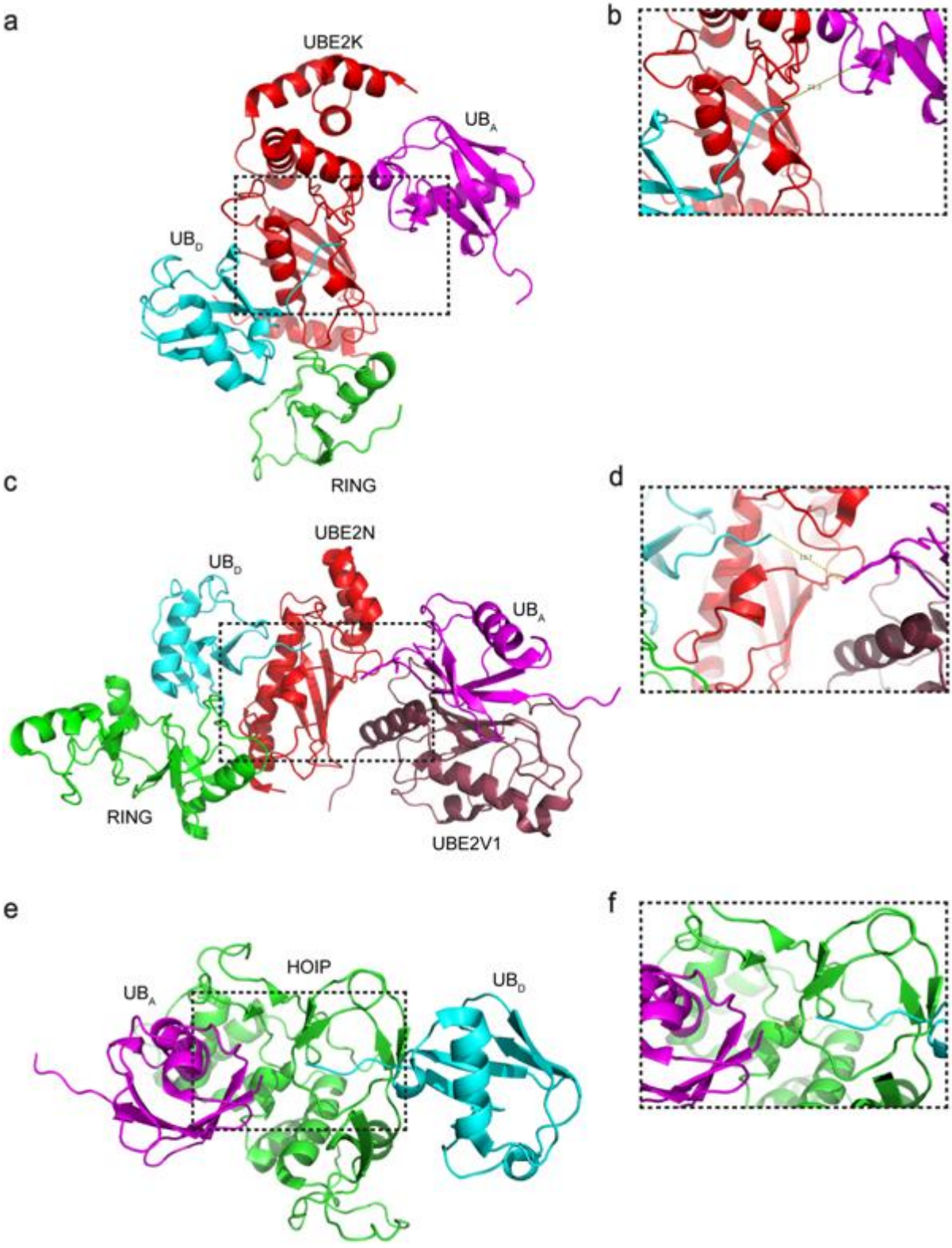


Figure 1.11 Structural studies of UB chain formation. (a) Crystal structure of RING-bound UBE2K with UB^D and UB^A (PDB 7OJX). (b) Close-out view at the UBE2K's active site with UB^D's C-terminus and UB^A's acceptor lysine (21 Å apart). (c) Crystal structure of UBE2N/UBE2V1-RNF4 RING dimer with UB^D and UB^A (PDB 5AIT). (d) Close-out view at the UBE2N's active site with UB^D's C-terminus and UB^A's acceptor lysine (15 Å apart). (e) Crystal

structure of HOIP RBR with UB^D and UB^A (PDB 4LJO). (f) Close-out view at the HOIP's active site with UB^D's C-terminus and UB^A's N-terminus (6.8 Å apart).

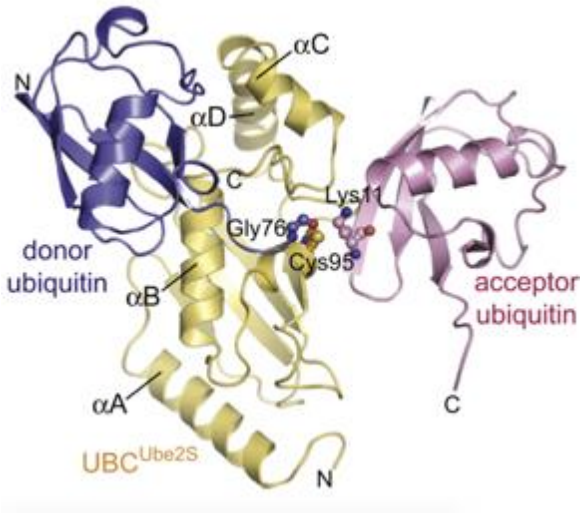


Figure 1.12 UBE2S UB^D and UB^A model. Adapted from Wickliffe et al. 2011[101].

1.5 UBE2R

UBE2R, also known as CDC34, is one of the best characterized, K48-linked chain forming enzymes. Initial studies in yeast *Saccharomyces cerevisiae* showed that mutants lacking the *CDC34* gene were defective in cell cycle G_1 to the S phase transition and DNA replication initiation[105]. This discovery in 1988 led to the identification of a new UB-conjugating enzyme. Further research identified CDC34p to be essential for degradation of many substrates including the cyclin-dependent kinase inhibitor (Sic1p). Sic1p has to be removed by the proteasome for the cell to enter S phase[106-108]. CDC34p was early associated to work with SCF complexes, part of the CRL family[109,110]. Unique architecture of UBE2R, which will be the topic of this chapter, allows its fast processivity to assemble K48-linked UB chains on CRL-bound substrates[31].

1.5.1 UBE2R catalysis

UBE2R was shown to catalyze rapid K48-linked chain formation on a substrate primed already with one UB[31]. Targeted protein priming was found to be rate limiting. Later

studies on Iκβα ubiquitylation by SCF$^{\beta TRCP2}$ complex has revealed handoff mechanism of SCF substrate polyubiquitylation [31,111]. UBE2D3 was assigned to efficiently prime the target protein with one UB, followed by a rapid chain elongation catalyzed by UBE2R[111]. Moreover, in the context of the whole assembled SCF complex, UBE2R exhibits a significantly lower K_M in the presence of UB-modified substrate compared to substrate alone[112].

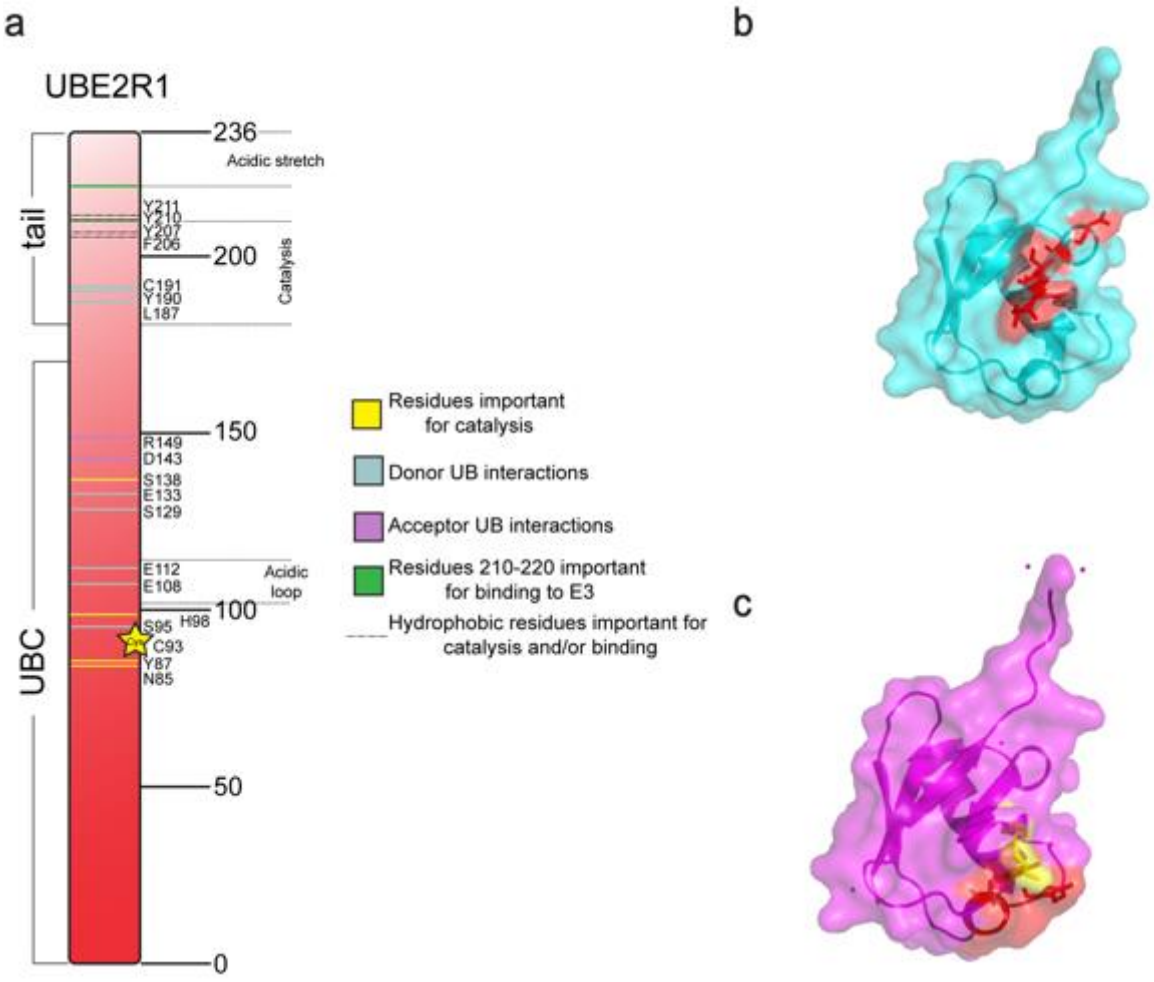


Figure 1.13 Previously identified residues important for UBE2R1 catalysis. (a) Scheme depicting the architecture and residues important for UBE2R1 catalysis, UB^D and UB^A interactions and E3 binding. (b) UB^D residues, essential for UBE2R1 interactions, colored in red. (c) UB^A residues essential for UBE2R1 interactions, colored in red. Acceptor K48 colored in yellow.

Along with what is described in section 1.4.1 UBE2R's N85, Y87 as well as S138 (D127 residues in model presented in Figure 1.9) function to catalyze efficient isopeptide

bond formation[113]. Multiple residues and UBC extensions afford for UBE2R activity (Figure 1.13), which will be discussed in the following sections.

1.5.2 UBE2R – UB^D interactions

Recruitment of UBE2R~UB by CRL accelerates the rate of UB^D discharge ~40-fold, by RBX1's RING-mediated closed conformation stabilization[31]. UBE2R's S95, E108 and E112 were suggested to participate in UB^D's C-terminal tail orientation and stabilization[113]. Moreover, another two important interaction points with UB^D were identified. First, I44A donor UB mutation was found to abolish substrate monoubiqutylation and chain extension[114]. I44 packs against S129 on UBE2R's surface and S129L mutant partially compensates the I44A mutation (Figure 1.14). Secondly, UB^D's R42E mutant exhibited defects in UBE2R~UB discharge, which was fully rescued by elegant UBE2R E133R charge swap mutation (Figure 1.14). Taken together, proper interactions between UBE2R and UB^D are crucial for efficient thioester discharge and proper organization of the active site for nucleophilic attack[114]. A recent crystal structure with UB^D-bound UBE2R2 explained how the proximal part of the tail stabilizes E2~UB closed conformation[115]. Specifically, Y190, C191 and L187 contact UB^D's A46 and G47 (Figure 1.14)[115].

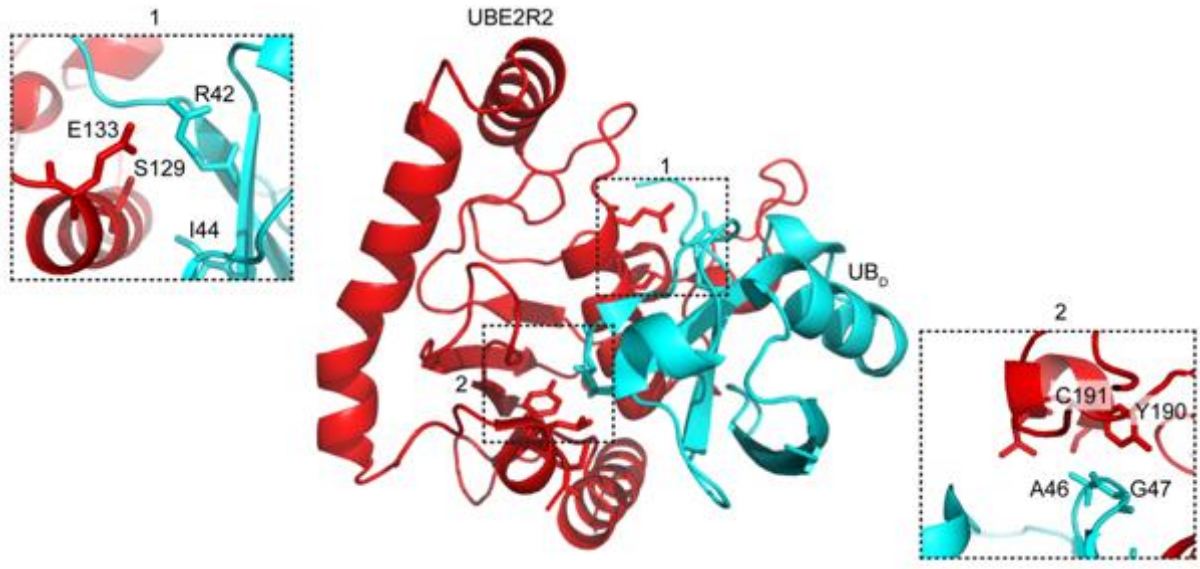


Figure 1.14 Crystal structure of UBE2R2 in complex with UB^D (PDB 6NYO). Important residues for UBE2R2-UB^D interactions are depicted (panel 1 and 2).

1.5.3 UBE2R – UB^A interactions

Robust chain formation by UBE2R depends on its interactions with UB^A. By screening UB residues essential for yeast vegetative growth, early studies have identified important residues on acceptor UB to play a role in chain extension[14,31]. I44 hydrophobic patch residues I44 and G47, as well as T12 and the C-terminal part (70-76) seemed to contribute to efficient diUB synthesis by UBE2R (Figure 1.15 a)[31].

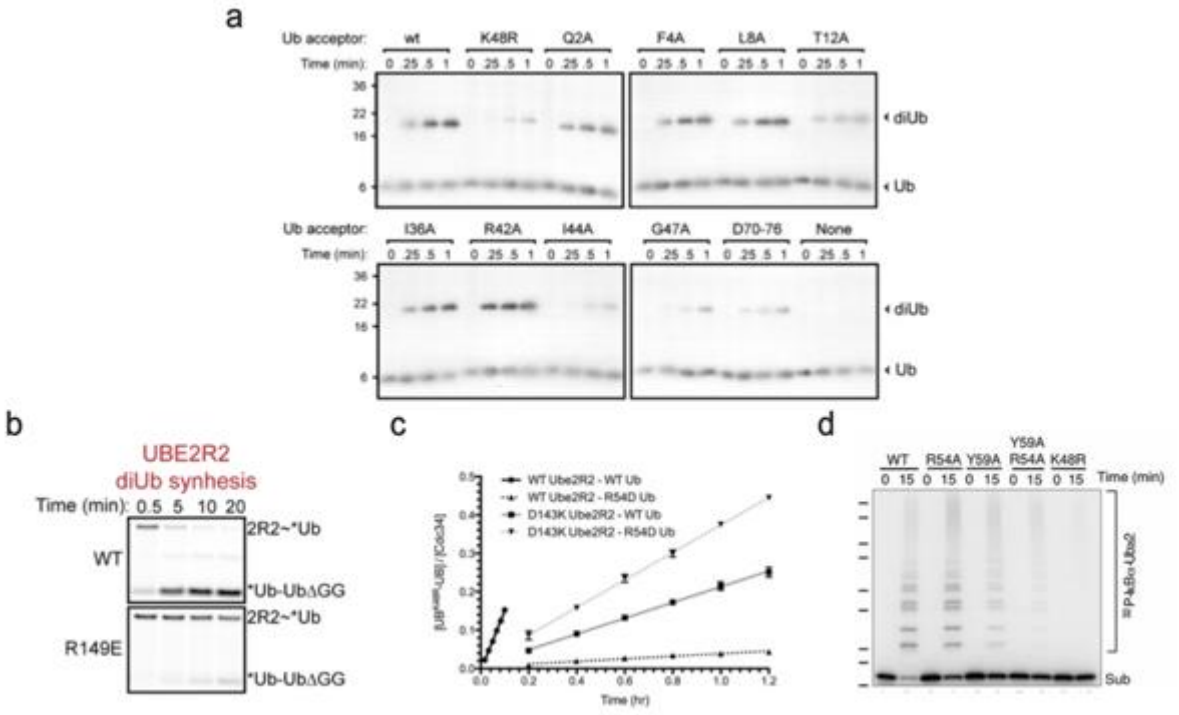


Figure 1.15 Previously identified important residues for UBE2R-UB^A interactions. (a) UB^A's I44 hydrophobic patch residues, T12, and the C-terminal residues 70-76 contribute to efficient diUB synthesis by UBE2R. Adapted from Petroski et al. 2005[31]. (b) UBE2R's R149 plays a role in UB^A binding. Adapted from Welsh et al. 2022[58]. (c) UBE2R's D143 and UB^A's R54 pair as an important interaction for UB^A recruitment. Adapted from Hill et al. 2016[116]. (d) R54A/Y59A UB^A's mutants disrupt UBE2R-mediated chain formation. Adapted from Chong et al. 2014[117].

Furthermore, a charged, conserved stretch on UBE2R (residues 143-153) was found to play a role in chain formation[113]. Recent experiments carefully examined HTH residues and proved that R149 is important for UBE2R-UB^A interaction (Figure 1.15 b)[58]. Additionally, another investigation proved importance of D143 in acceptor UB

recruitment and orientation via its contact with R54 of UB^A (Figure 1.15 c)[116]. R54 is a part of E51-Y59, which was identified to play an important role in K48-linked chain formation (Figure 1.15d). Introduction of R54A/Y59A loop disruption mutants into mammalian cells caused a decrease in K48-linked chain production[117].

1.5.4 Acidic loop

The acidic loop, a unique UBC extension found only in UBE2R and UBE2G E2s, has an essential role in processivity and specificity of UBE2R[31].

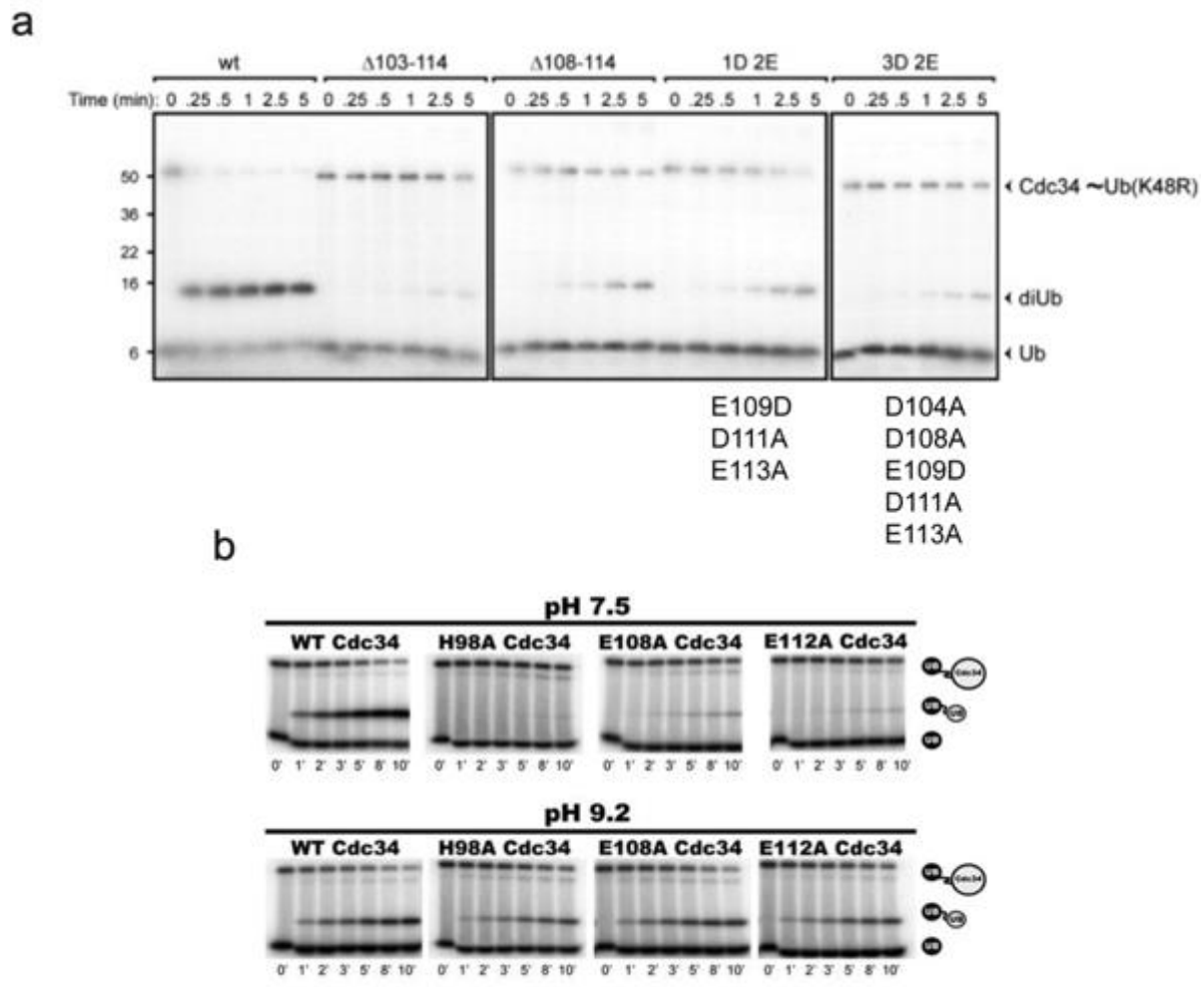


Figure 1.16 The importance of UBE2R's acidic loop in K48-linked UB chain formation. (a) Deletion or acidic stretch mutation of the acidic loop abolishes diUB formation. Adapted from Petroski et al. 2005[31]. (b) H98, E108 and E112 play a role in UBE2R's catalysis, by lowering pK_a of the acceptor lysine. Adapted from Ziemba et al. 2013[118].

It was hypothesized that it might support chemical steps of nucleophilic attack and further isopeptide bond formation[31]. Additionally, it might position K48 in proximity of the active site, thus determining specificity[31]. Deletion of the loop (residues 102-113) along with its mutation to alanines abolished diUB formation (Figure 1.16 a). Interestingly, the acidic loop was not required for efficient substrate monoubiquitylation[31,119]. A careful biochemical study revealed the importance of H98, E108 and E112 in acceptor lysine pK_a suppression, thus their role in catalysis (Figure 1.16 b)[118]. The highly acidic loop was also proposed to interact with Rbx1, thus mediating binding to CRL[116,120]. Conversely, other work has suggested acidic loop collaboration with UB^A's E51-Y59 loop[117].

1.5.5 Acidic tail

Following catalytic core, UBE2R contains a 40 amino acid long C-terminal tail. Deletion of the whole UBE2R's extension was shown to be lethal in yeast[121,122]. The acidic tail was proved to be essential for UBE2R binding to SCF and its catalytic function[52,109,123]. A kinetic study has identified ~100-fold increase in K_M for Δtail UBE2R mutant, proving the importance of the tail binding to SCF for effective substrate polyubiquitylation (Figure 1.17a)[52]. Residues between amino acids 210 and 220 (CDC34p) were described, as playing a role in direct interaction with SCF (Figure 1.17 a)[52]. Additionally, the tail interaction with E3 stabilizes the catalytic domain recruitment and it is responsible for fast association with the complex[124]. Moreover, the rate of substrate ubiquitylation was significantly reduced (~33-fold) for UBE2R Δtail construct (Figure 1.17 b) and proximal region of the CDC34p tail (residues 191-210) was found to affect catalysis[52]. Furthermore, the acidic part of UBE2R's tail was found to engage with SCF's basic canyon via electrostatic interactions[52,124]. A crosslinking experiment proved acidic tail binding in multiple conformations to SCF[125]. Other investigations argued the importance of four hydrophobic residues (F206, Y207, Y210, Y211) in efficient chain formation[126]. The previously mentioned crystal structure (PDB 6NYO) revealed the proximal part of the human tail (V181, P184 and T185) packing against UBE2R2's backside (Figure 1.14)[115]. To date, structural basis of UBE2R tail recruitment by E3 remains elusive.

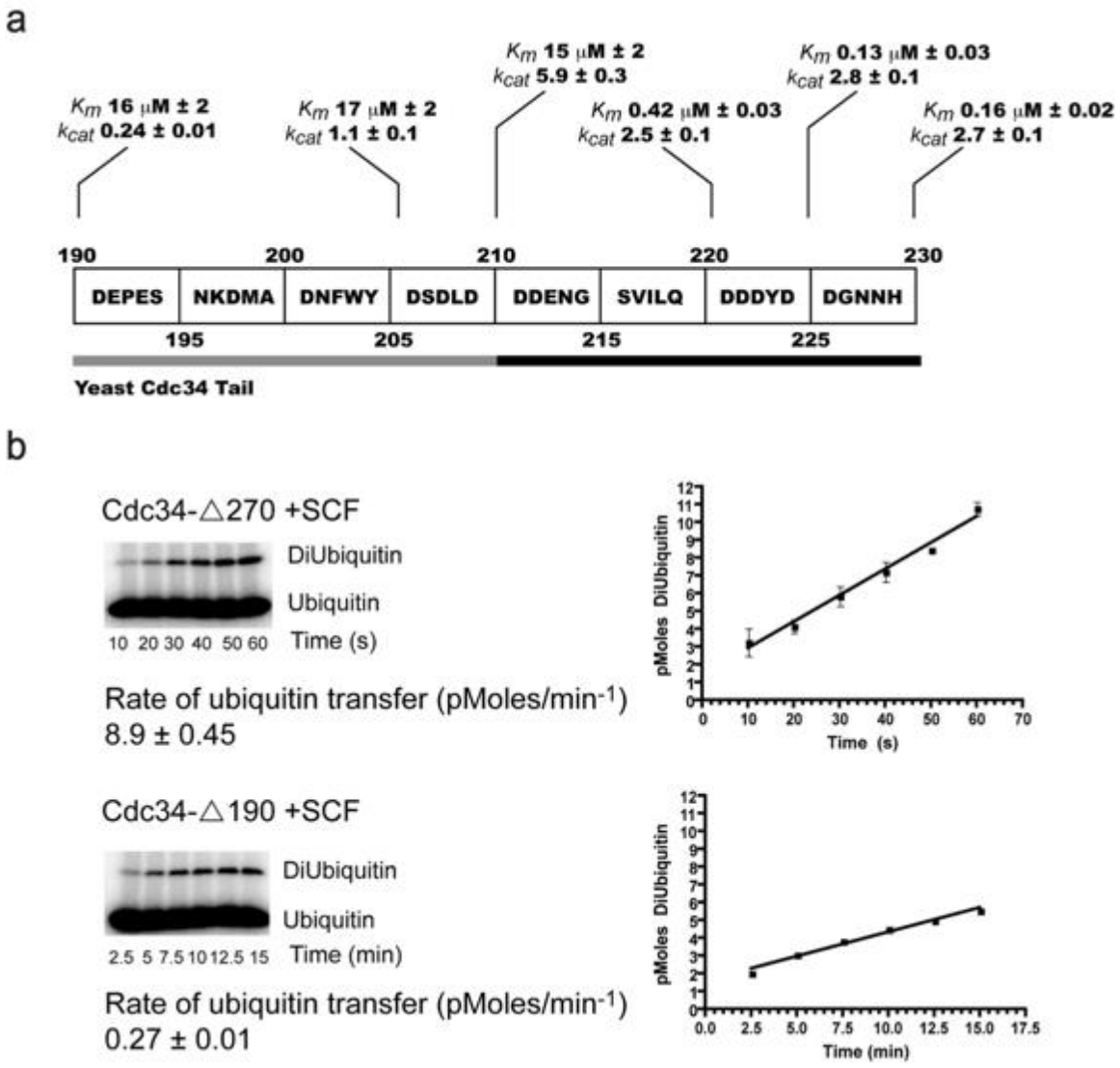


Figure 1.17 Acidic tail importance in UBE2R's E3 binding and catalysis. (a) K_M and k_{cat} determination of yeast CDC34's tail truncations. Adapted from Kleiger et al. 2009[52]. (b) CDC34 ∆tail construct significantly reduces the rate of ubiquitylation in the presence of SCF. Adapted from Kleiger et al. 2009[52].

1.5.6 Allosteric UBE2R inhibitor

The UB system is often perturbed in many neurological disorders and cancers. As such, it is of great interest to develop drugs, selectively modulating activity of UB cascade enzymes[127,128]. CC0651 was found to inhibit the UBE2R1 activity *in vitro* and *in vivo*[129]. Biological relevance of the drug was confirmed by inhibiting proliferation and accumulation of one of the SCF substrates p27 in human cancer cell lines[129]. The crystal structure revealed binding of the inhibitor away from the catalytic site and

allosterically stabilizing E2-UB non-covalent complex (Figure 1.18). Thus, CC0651 works by slowing down UB^D's discharge from the catalytic cysteine[129]. Recent studies have identified a molecule with ~1000-fold better efficiency[129,130]. Interestingly, the discovered inhibitor selectively works with UBE2R1 but not with its close homolog - UBE2R2. Apart from their biological relevance, inhibitors became an important tool for structural biology to stabilize UBE2R-UB intermediates[115].

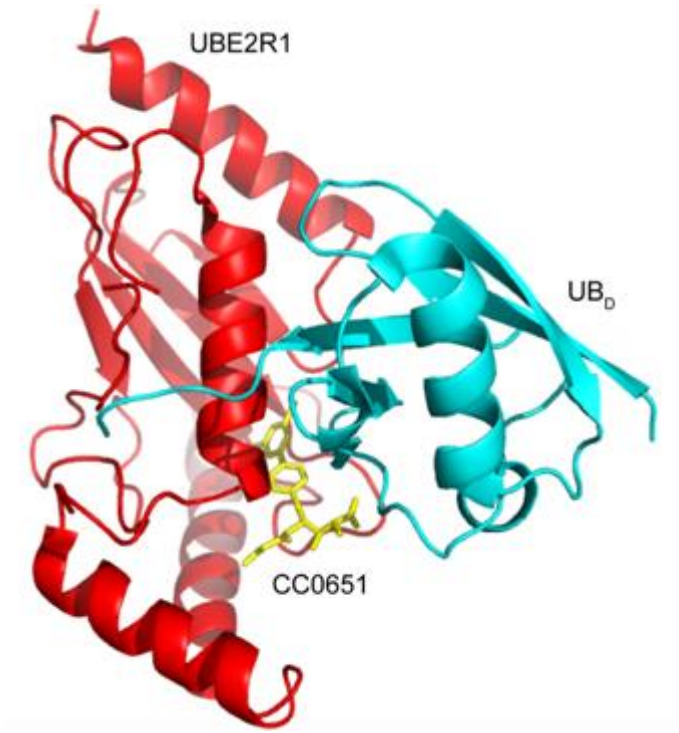


Figure 1.18 Crystal structure of UBE2R1-UB^D with CC0651 bound (PDB 3RZ3). CC0651 stabilized UBE2R1-UB^D closed conformation.

1.6 Cullin RING ligases (CRLs)

1.6.1 Introduction to CRLs

CRLs control the vast majority of important cellular functions including development, cell division, metabolism, signal transduction, hormone perception, transcription differentiation, and circadian rhythm[131]. Moreover, they are responsible for more than 20% of all degradation events by the 26S proteasome in cells[132]. CRL superfamily is divided to five canonical families named based on the Cullin (CUL) protein present in the complex. As such: CUL1 (CRL1), CUL2 (CRL2), CUL3 (CRL3), CUL4A/B (CRL4)

and CUL5 (CRL5) (Figure 1.19)[131]. Cullin, a 100kDa protein, acts as a scaffold recruiting a substrate receptor module on the N-terminus and a RBX RING protein on the C-terminus. Each CRL family associates with specific substrate adaptors and/or substrate receptors[133]. CRL1 partners with SKP1 as substrate adaptor, which is known to bind F-BOX substrate receptor[134].

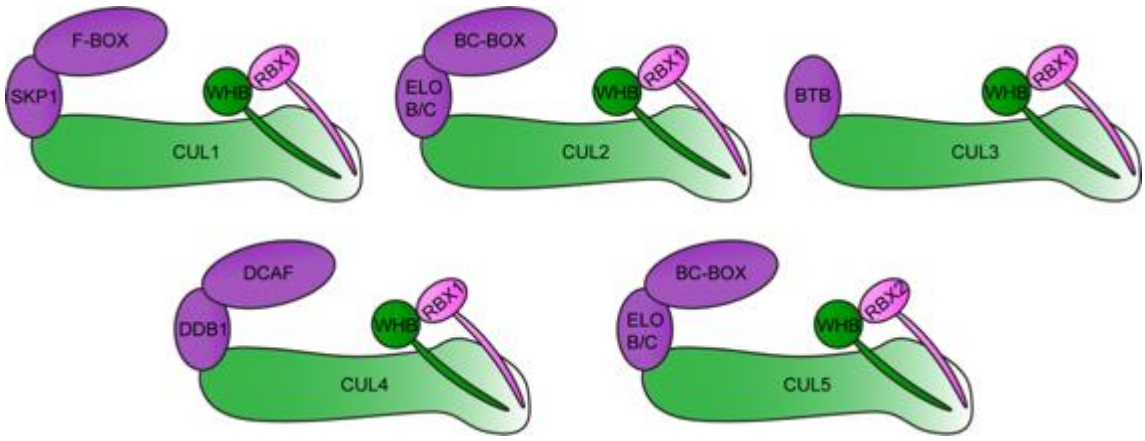


Figure 1.19 Schematic representation of canonicals CRLs. CUL1-CUL5 (in green) associate with different substrate adaptors and substrate receptors (in purple) as well as various RING proteins: RBX1 and RBX2 (in pink).

Elongin B and Elongin C (ELOB/C) is a substrate adaptor designated to CRL2 and CRL5, recruiting the substrate binding BC-BOX[135]. CRL3 consists of the BTB-3-BOX substrate binding module, a single protein homologous to SKP1 and F-BOX, known to dimerize[136-138]. Finally, CRL4 comprises DDB1 adaptor with the DCAF substrate receptor[139,140]. Additionally, CRLs collaborate with their designated RING domain proteins: RBX1 for CRL1, CRL2, CRL3 and CRL4 and RBX2 for CRL5 (Figure 1.19)[131]. Interestingly, RBX1/2 can either recruit E2~UB^D or RBR~UB^D ARIH1 (RBX1) and ARIH2 (RBX2)[81,82]. Apart from ARIH1/2 working specifically to monoubiquitinylate CRL-bound substrates, a couple of E2 enzymes were assigned to work with CRLs: UBE2D to prime a substrate with UB, UBE2R1/2 and UBE2G1 for processive K48-linked chain formation[50,112].

1.6.2 CUL architecture

All canonical CUL proteins (CUL1-5) show a similar architecture. First, the N-Terminal Domain (NTD) consists of three tandem CUL-repeat domains (CR1, CR2 and CR3) followed by a 4-Helix Bundle (4HB) (Figure 1.20)[141]. CR1 recruits a substrate adaptor and/or substrate receptor as described in section 1.6.1 and the NTD engages with CAND1, a substrate exchange protein[141,142]. Second, the C-Terminal Domain (CTD) is composed of a C/R domain and a Winged-Helix B (WHB) domain (Figure 1.20)[141]. The C/R domain binds RBX1's N-terminus and it was identified to engage with the CSN deneddylation machinery[141,143,144]. Following C/R domain and helix 29, WHB accommodates the conserved lysine, which is known to be modified by NEDD8[145]. WHB as well as RBX1's RING domain were shown to be the most flexible parts of CUL assembly[50,146]. WHB and RING domains pack against each other, remaining in an autoinhibited state in the absence of NEDD8[146,147]. More about CUL regulation is described in section 1.6.3.

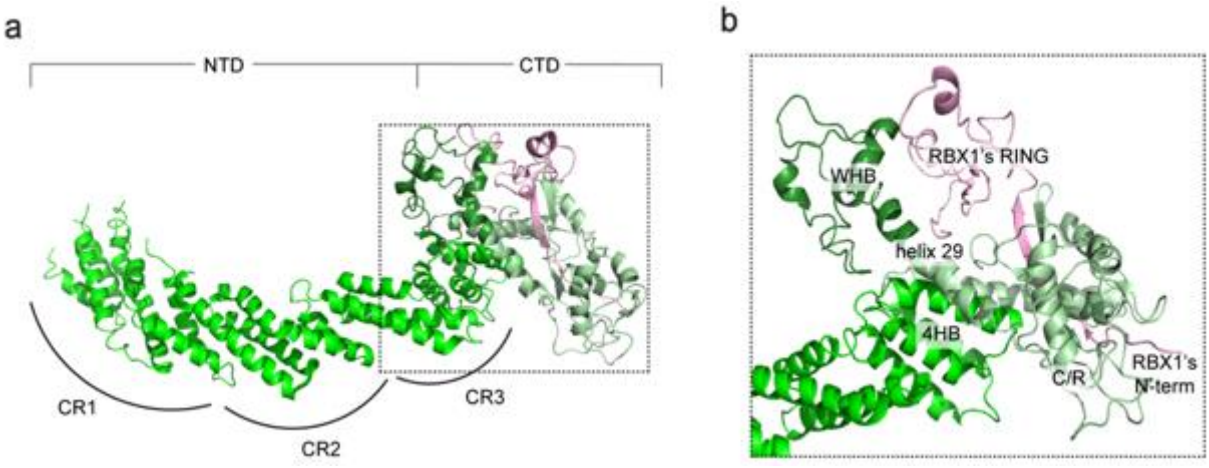


Figure 1.20 Crystal structure of CUL1 (PDB 1LDK) with main structural components depicted. The architecture of NTD and CTD is shown.

1.6.3 CUL regulation

Multiple additional proteins and complex machineries assemble with CRLs to control its activation and subsequently substrate degradation[131]. First, the NEDD8-specific E1-E2-E3 cascade acts to conjugate NEDD8's C-terminus to WHB's lysine (Figure 1.21)[148-150]. Neddylation is triggered by a substrate being bound to the substrate

receptor. Further, this leads to conformational changes, promoting E2/E3~UB recruitment as well as UB transfer catalysis stimulation[50,146,151]. Second, COP9 signalosome (CSN) cleaves the isopeptide bond between WHB and NEDD8 (Figure 1.21)[143,144]. This occurs only when no substrate is bound to the substrate receptor[143,144,152,153]. Lastly, CAND1 binds to the unneddylated CUL and promotes substrate module exchange (Figure 1.21)[142,154].

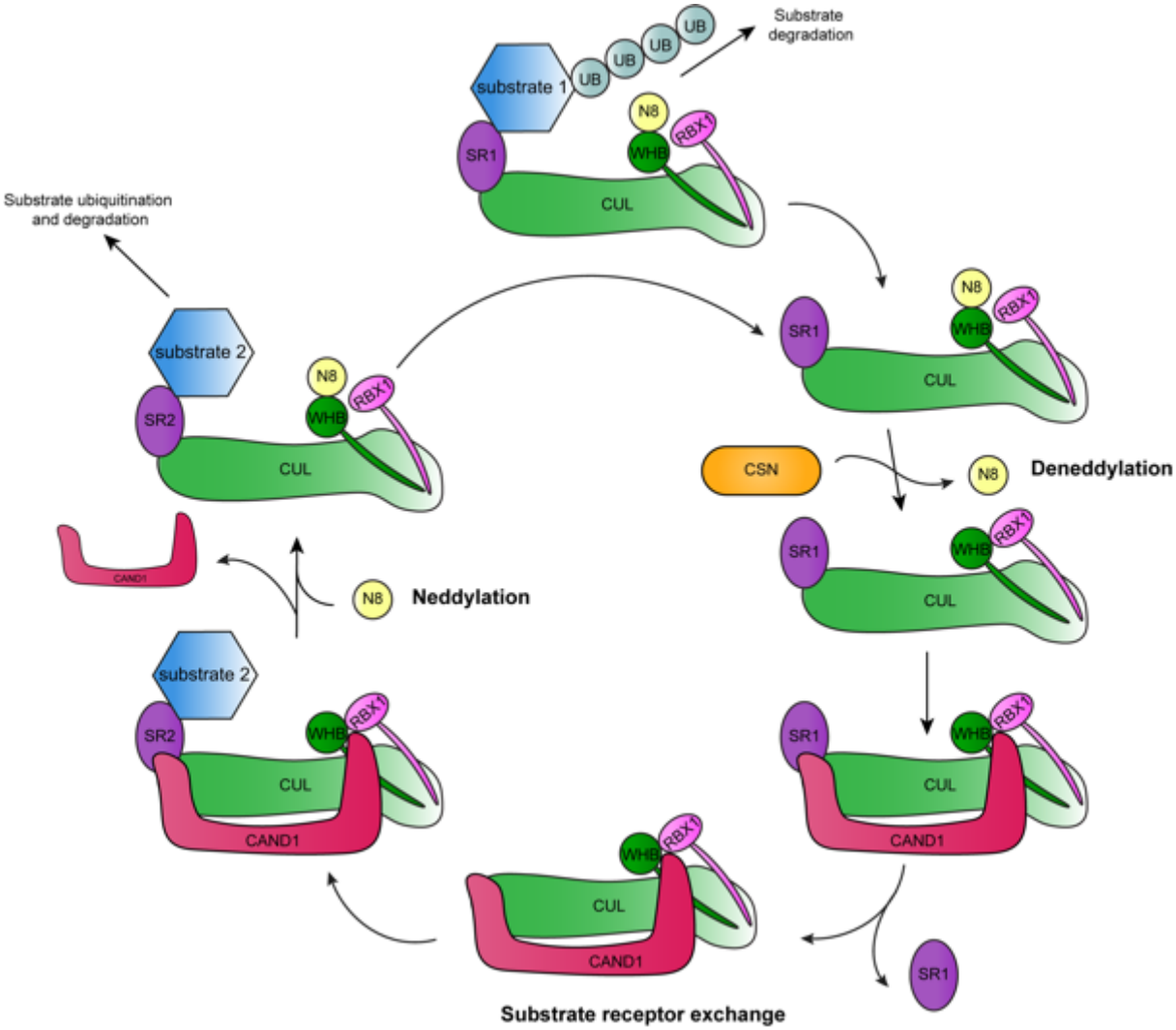


Figure 1.21 Schematic representation of CUL activation and deactivation cycle. Substrate receptor exchange cycle is depicted.

1.6.4 Structures of active CUL complexes

Recent Cryo-EM development allowed to visualize CRLs in action. Specifically, CRL-E2 and CRL-RBRs structures were solved demonstrating a snapshot of substrate

monoubiquitylation and/or E2/RBR activation by CRL-NEDD8[50,82,97]. A chemical disulfide probe allowed to visualize a ubiquitylation intermediate of the NEDD8-CRL1$^{\beta TRCP1}$-UBE2D3-UB^{D} complex. The structure revealed the basis of NEDD8 activation for CRL1$^{\beta TRCP1}$-UBE2D3. NEDD8 interacted with βTRCP1, CUL1's WHB and the backside of UBE2D3 to position and activate E2-UB^{D} for IκBα ubiquitylation (Figure 1.22 a)[50]. A similar approach was taken to visualize ARIH1-mediated substrate priming. A UBE2L3-UB electrophilic probe was used to capture the first transition state of UB^{D} transfer from UBE2L3 to ARIH1's catalytic cysteine.

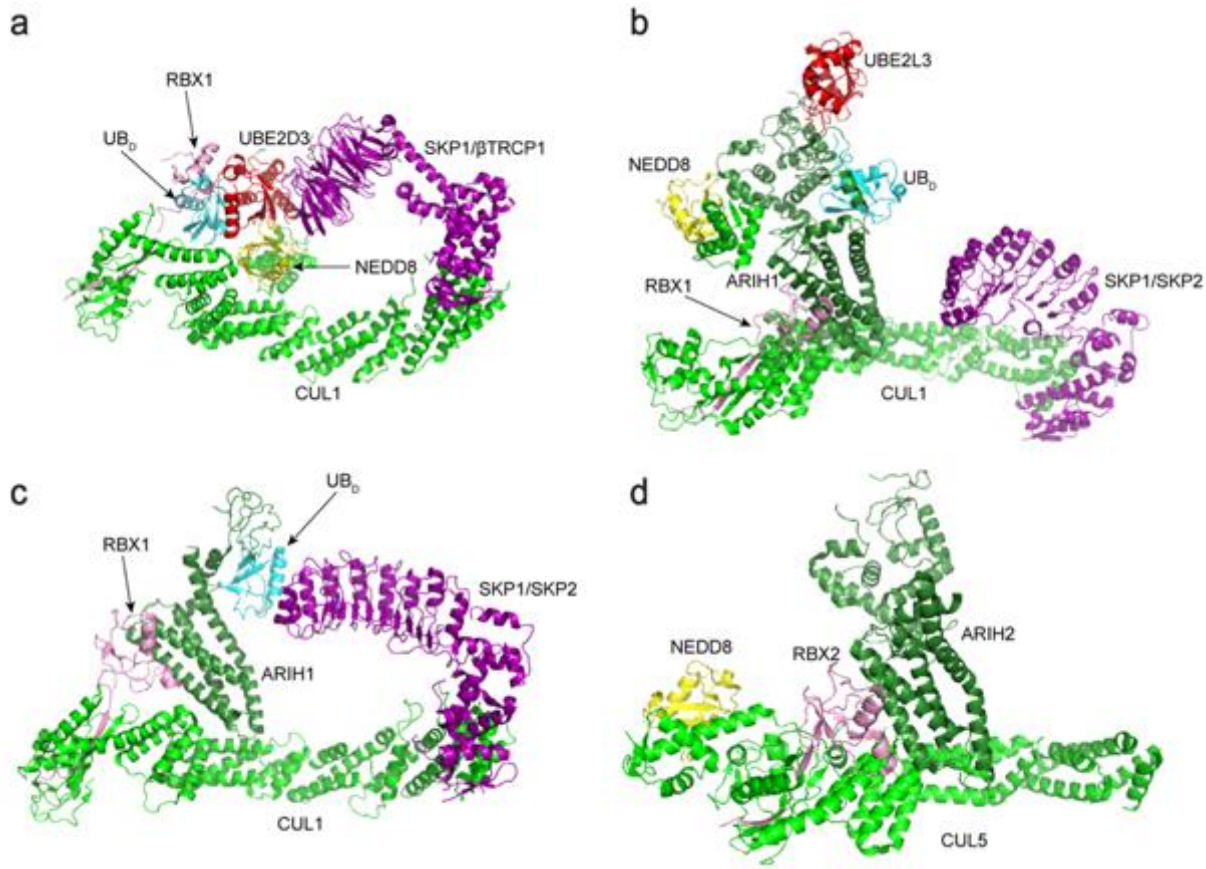


Figure 1.22 Structures of active CUL complexes. (a) CryoEM structure of trapped NEDD8-CRL1$^{\beta TRCP1}$-UBE2D3-UB^{D} complex (PDB 6TTU). (b) CryoEM structure of the first transition state of UB^{D} transfer from UBE2L3 to ARIH1's catalytic cysteine in the presence of NEDD8-CRL1^{SKP2} (PDB 7B5L). (c) CryoEM structure of the second transition state of UB^{D} transfer from ARH1's catalytic cysteine to substrate peptide in the presence of NEDD8-CRL1^{SKP2} (PDB 7B5M). (d) CryoEM structure of NEDD8-CUL5-ARIH2 complex (PDB 7ONI).

Cryo-EM maps explained how ARIH1 is recruited to NEDD8-CRL1 and activated in a NEDD8-dependent manner, by contacting and releasing ARIH1 from its autoinhibited state (Figure 1.22 b). The second transition state, visualized using DHA substrate probes, showed how E3-E3 assembly places UB-bound ARIH1's active site in vicinity of the substrate's lysine (Figure 1.22 c)[97]. Additionally, recent Cryo-EM investigation explained how RBR ARIH2 associates with NEDD8-CRL5. In contrast to ARIH1, ARIH2 was allosterically activated by NEDD8. NEDD8 conjugation to CUL5's WHB stimulated major conformational changes, which allowed ARIH2 tail binding to CUL5 (Figure 1.22 d)[82]. Thus, these approaches proved how activity-based probes can be used to capture transient ubiquitylation intermediates. However, structural basis of CRL-based substrate polyubiquitylation remains elusive.

1.6.5 CRL2

CRL2 ligase is mostly known to associate with the von Hippel-Lindau (VHL) substrate receptor[155]. VHL was named after its crucial role in VHL syndrome development[156]. Identified at the beginning of the last century, VHL disease characterizes in development of vascular tumors and mutation of both VHL gene alleles[157]. Most of the VHL tumor suppressor gene products were not able to bind to the substrate adaptor: ELOB/C[158,159]. Further studies identified Hypoxia-Inducible Factor 1 (Hif1α), as a VHL substrate, being recruited to CRL2 for proteasomal degradation[160,161]. Hif1α is a part of the Hif-1 transcription factor, which activity is tightly dependent on oxygen concentration. When the oxygen level is normal (normoxia), two of Hif1α's proline residues are hydroxylated by Hif prolyl hydroxylases (PHD1-3)[162,163]. This triggers hydroxylated Hif1α binding to VHL's β-domain and subsequent substrate polyubiquitylation[161]. Structural basis of Hif1α recognition by VHL is discussed further in this section. Conversely, when the cell is deprived of oxygen (hypoxia), PHD's activity towards Hif1α is inhibited, which leads to Hif1α accumulation. Hif1α dimerization with Hif1β activates transcription of many genes is response to hypoxia[164]. It is estimated, that Hif-1 controls the expression of hundreds of genes[165]. As such, in the VHL disease, no matter the oxygen concentration, Hif1α gets accumulated, which is thought to be a major oncogenic force. Many cancers as clear cell renal carcinoma, pheochromocytoma or hemangioblastoma were related to VHL syndrome[166]. Other substrates like Sprouty2 (Spry2), subunits of RNA polymerase II Rbp1 and hsRBP7 or

EGFR were also identified to associate with VHL and get targeted for degradation by CRL2[167-170]. However, the most known and studied function of CRL^{VHL} is Hif1α degradation. Apart from VHL, CUL2 binds to numerous other substrate receptors via ELOB/C such as Leucine-Rich Repeat protein 1 (LRR-1), Feminization-1 (FEM-1), Preferentially Expressed Antigene in Melanome (PRAME), ZYG-11, BAF250 and Receptor for the Activated C Kinase 1 (RACK1)[171-176].

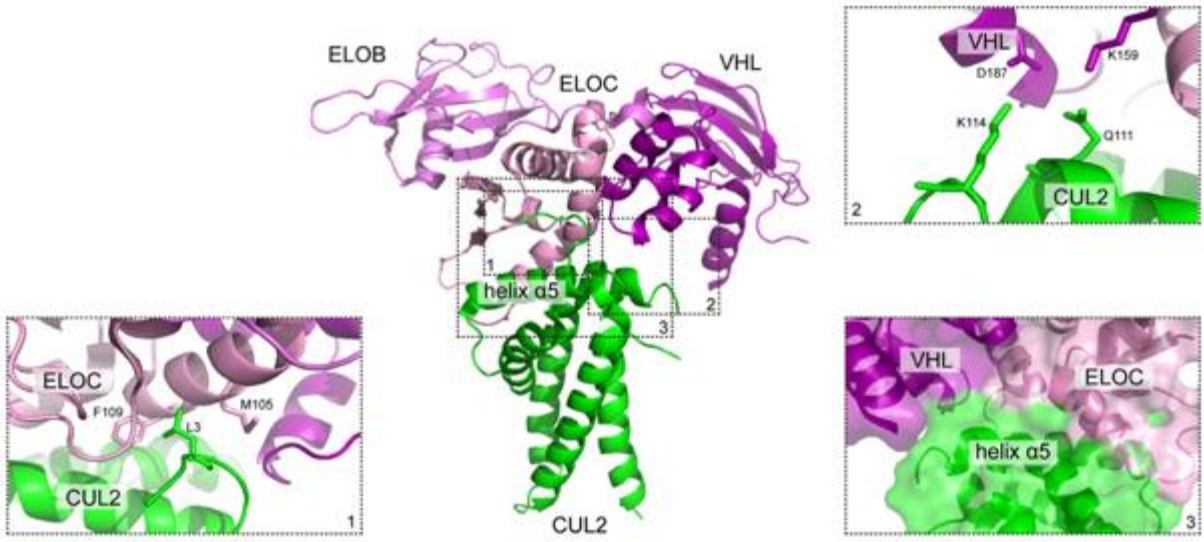


Figure 1.23 Crystal structure of CUL2 NTD with ELOB/C-VHL bound (PDB 4WQO). Key interactions are depicted in panels 1-3.

Crystallography studies revealed the basis of ELOB/C-VHL recognition by CUL2, CUL2 scaffold as well as Hif1α recruitment to VHL (Figure 1.23)[155,177]. ELOC was found to bind to CUL2 via hydrophobic interactions. CUL2's L3 inserts into the hydrophobic pocket of ELOC (M105, F109) (Figure 1.23, panel 1)[177]. Additionally, P5 of ELOC interacts with VHL's V181 and ELOC's M105 (Figure 1.23, panel 1)[177]. The complex formation is further mediated by CUL2's helix α5 contacts with VHL (CUL2's Q111 and K114 with VHL's K159 and D187) and this interface is thought to be crucial for ELOB/C-VHL recruitment by CUL2 over CUL5 (Figure 1.23, panel 2)[177]. The same helix α5 also participates in ELOC binding (Figure 1.23, panel 3)[177].

Full length CUL2 crystal structure revealed inter-domain flexibility between CUL2's NTD and CTD, which was previously observed for other CULs (Figure 1.24)[142,155,178].

The hinge points between CUL repeats as well as NTD and CTD allow CUL to adapt different conformations in response to binding distinct substrate receptor modules, substrates and for processive chain formation (Figure 1.24)[144,155,178,179].

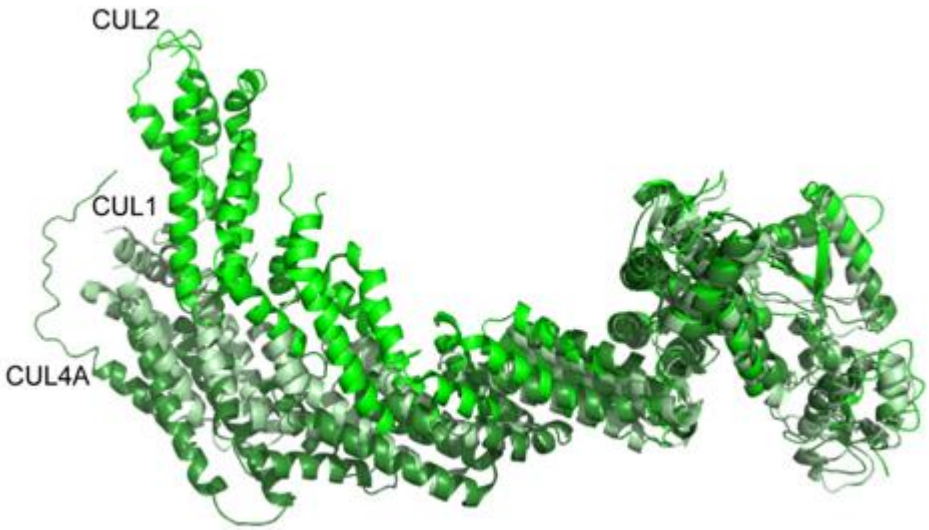


Figure 1.24 CUL1, CUL2 and CUL4A crystal structure comparison showing inter-domain flexibility (PDB 4A0K, 5N4W, 1U6G).

VHL- Hif1α structure showed how post-translationally modified Hif1α's proline is in the center of substrate recognition (Figure 1.25 a)[92]

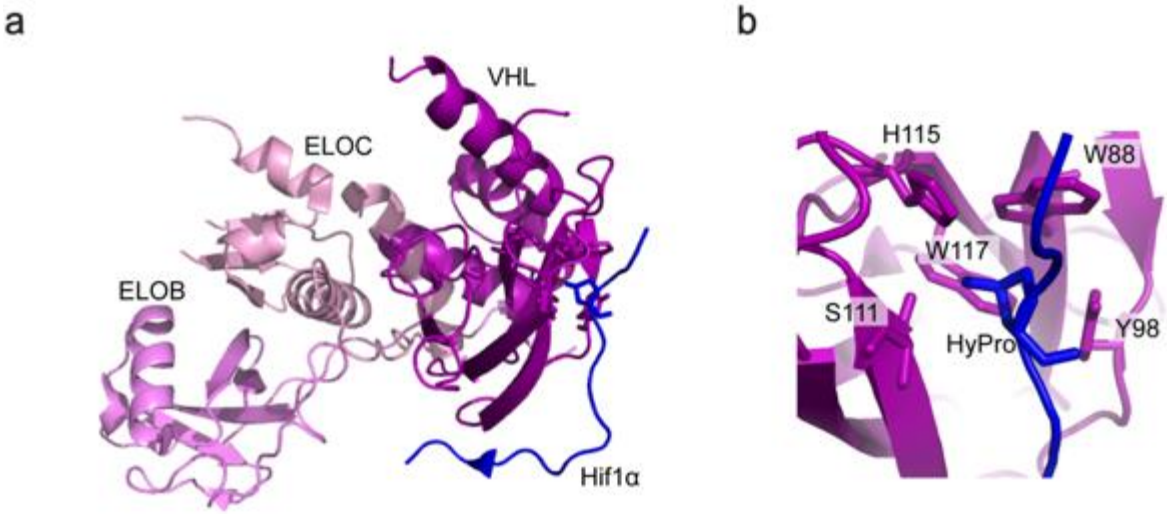


Figure 1.25 Crystal structure of ELOB/C-VHL with Hif1α substrate peptide bound (PDB 1LM8). (a) Hif1α's hydroxylated proline is in the center of substrate recognition. (b) Multiple contacts between Hif1α's hydroxylated proline and VHL drive its specific recognition.

First, the pyrrolidine ring of hydroxylated proline makes multiple van der Waals contacts with VHL's hydrophobic core containing W88, Y98 and W117 (Figure 1.25 b). Second, hydrogen bonds between His115, Ser111 and the hydroxyl group of proline Lastly, the Hif1α backbone in proximity of hydroxylated proline is stabilized by multiple hydrogen bonds with VHL[92]. Despite many partial crystal structures of CRL2^{VHL} complex, the structural basis of the Hif1α ubiquitylation with the whole assembled complex remains unknown.

1.6.6 Targeted protein degradation

In the last two decades, Proteolysis-targeting chimeras (PROTACs) and molecular glues emerged as a great tool to drug undruggable proteins[180]. PROTACs are heterobifunctional molecules on one site binding the substrate receptor of an E3 ligase and on the other site the target protein. Both sites of the molecule (ligands) are connected by a linker[181-184]. Whereas in molecular glues one ligand interacts both with substrate and substrate receptor[180]. The function of the drug is to recruit a non-endogenous substrate (neo-substrate) for degradation. Thus, these molecules emerged as a great opportunity to degrade disease-causing proteins[131]. Importantly, they are working in substoichiometric range, as after one target polyubiquitylation, the drug can associate with another substrate[185]. Their activity to degrade was proved *in vivo* and in 2019 first PROTAC entered clinical trials[186]. PROTACs, as well as molecular glues, were initially designed to associate with CRL4^{CRBN} and CRL2^{VHL}[180]. Multiple crystal structures elucidate the molecular basis of drug recognition by the E3 ligase along with substrate recruitment. CRBN binds molecular glues: thalidomide and its analogs, lenalidomide and pomalidomide. These FDA-approved immunomodulatory drugs were used for multiple myeloma treatment, before their association with CRBN[187]. Further investigations showed their function in Ikaros-family transcription factors degradation[188,189]. The crystal structure revealed how CRBN-bound pomalidomide interacts with Zinc-Finger 2 (ZF2) of IKZF1 (Figure 1.26 a)[187]. Furthermore, VHL-based PROTACS, like MZ1, were shown to efficiently degrade BET family members, such as BRD4[182,190]. The crystal structure reported MZ1 sandwiched in between substrate receptor and targeted protein, explaining ternary complex formation (Figure 1.26 b)[190]. Along with new PROTACs development, understanding the molecular mechanism of drug-dependent degradation is of great interest. UBE2G1

was proved to build K48-linked UB chains on CRBN-bound neo-substrates[191]. Recent genome-wide CRISPR/Cas9 screens have identified UBE2R2 as an important E2 for VHL-based targeted protein degradation[192].

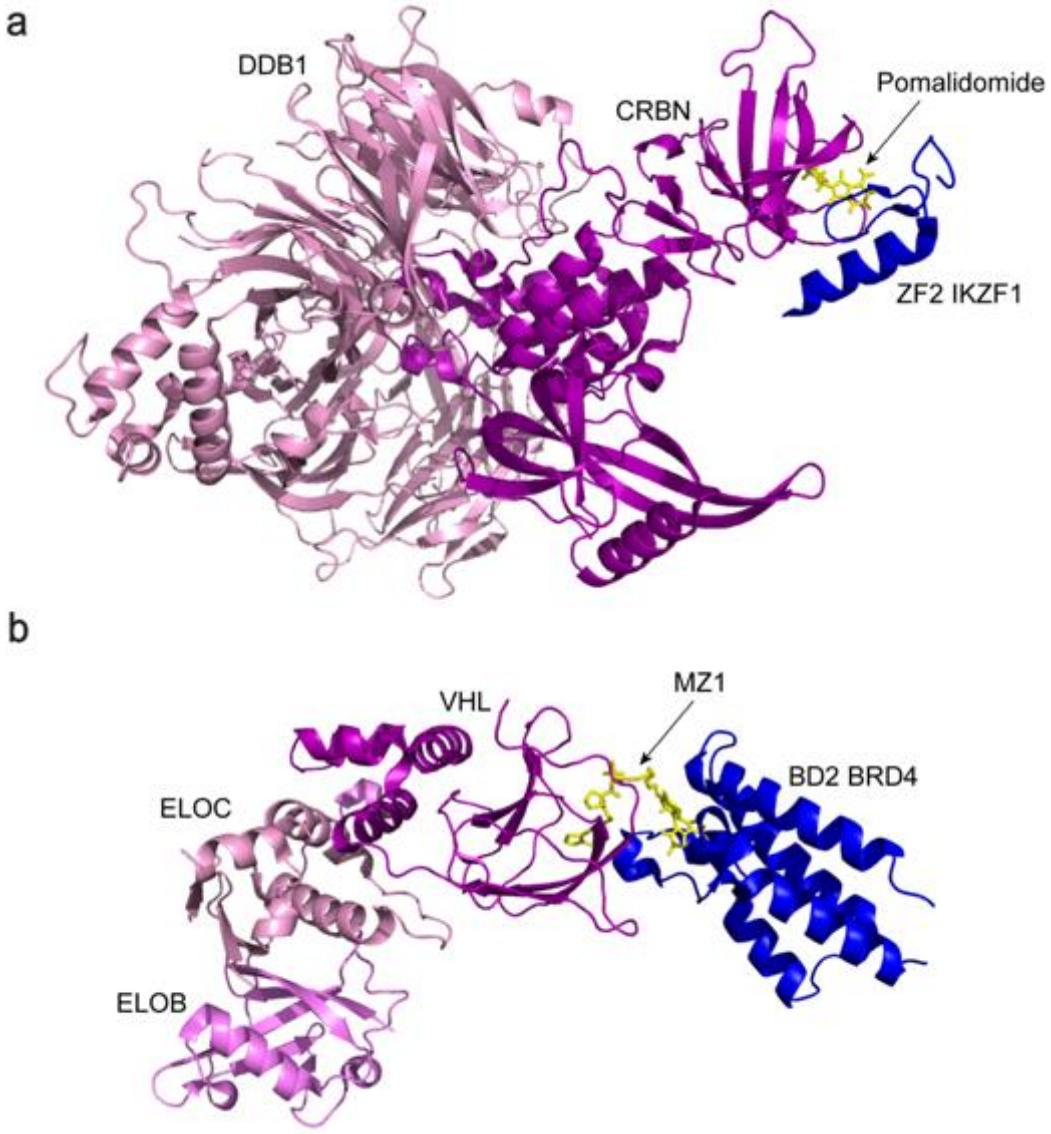


Figure 1.26 Crystal structures showing neo-substrate recruitment to the substrate receptor. (a) Crystal structure of DDB1-CRBN bound to the molecular glue, Pomalidomide and IKZF1's ZF2 (PDB 6H0G). (b) Crystal structure of ELOB/C-VHL bound to PROTAC, MZ1 and BRD4's BD2 (PDB 5T35).

2. MATERIALS AND METHODS

2.1 Plasmid and bacmid generation

2.1.1 Plasmid generation

All constructs listed in Table 2.1 were generated either using Gibson assembly[193] or QuikChange Site-Directed Mutagenesis Kit (Agilent). After Gibson assembly or QuikChange PCR reaction DNA was transformed into DH5α or XL1-Blue. Briefly, 10μl reaction was added to 100μl of the competent cell and incubated on ice for 5min. Next, for heat-shock, cells were put into 42°C for 45sec. Afterwards 2min incubation on ice was performed and 900μl of SOC medium added. After 1h recovery in 37°C, 800rpm all cells were plated on the agarose plates, with respective antibiotics and incubated overnight (O/N) in 37°C. Next day, 5ml of LB medium was inoculated with a single colony, respective antibiotics were added (50μg/ml Kanamycin; 100μg/ml Ampicillin; 34μg/ml Chloramphenicol always used) and the cultures were grown O/N, 37°C, 180rpm. DNA was isolated using QIAprep Spin Miniprep Kit (Qiagen) and sequenced.

2.1.2 Bacmid generation

Plasmid was added to the 50μl of emBacY competent cells and incubated on ice for 30min. Next, for heat-shock, cells were put into 42°C for 1min and afterwards chilled on ice for 2min. 900μl of SOC medium was added and the cells were grown for 6h in in 37°C, 800rpm. Further, 100μl was plated on the agarose plates containing 200μg/ml Ampicillin, 50μg/ml Kanamycin, 7μg/ml Gentamycin, 10μg/ml Tetracycline, 100μg/ml X-gal and 100μM IPTG. The plates were incubated 24h in 37°C or until the color developed. The white colonies were used to inoculate 5ml LB with 200μg/ml Ampicillin, 50μg/ml Kanamycin, 7μg/ml Gentamycin, 10μg/ml Tetracycline and grown O/N, 37°C, 180rpm. Next day standard bacmid isolation protocol was used to isolate DNA.

2.2 Recombinant protein generation

2.2.1 Bacterial protein purification

Proteins depicted in Table 2.1 were expressed in bacterial cells. Table 2.1 contains an information about the tag used, cleavage site as well as an origin of the protein. All protein productions were performed in either BL21-Gold(DE3) or Rossetta (DE3) cells, working under T7 RNA polymerase promoter. First 1μl of plasmid was transformed into the competent cells. After the recovery, 30μl of the cells was plated on the agarose plates with respective antibiotics. Next day, 100ml of LB medium (pre-culture), containing respective antibiotic (50μg/ml Kanamycin; 100μg/ml Ampicillin; 34μg/ml Chloramphenicol always used) was inoculated with a single colony and grown at 37°C, 180rpm. After O/N incubation, 1L TB medium cultures were started with 10ml of preculture and respective antibiotics were added (50μg/ml Kanamycin; 100μg/ml Ampicillin; 34μg/ml Chloramphenicol always used). The cells were grown at 37°C 180rpm to OD_{600} ~0.8 and the temperature was lowered to 18°C or 22°C for UB constructs. The 0.5mM IPTG was added to induce protein expression. The next day, cells were spinned down for 15min 7240rpm 4°C and the pellet was resuspended in Buffer A (Table 2.2). Bacterial cells were either flash frozen or directly used for purification. Protease inhibitor, 2.5mM PMSF, was added before sonication, which was typically performed in 2 cycles, 30% amplitude, 1 sec on 2 sec off. The length of the cell disruption varied depending on the sample volume. Next, lysate was centrifuged at 20000rpm for 30min, 4°C. The soluble fraction was collected and incubated with previously equilibrated beads. The purification scheme for tagless UB is described in section 2.2.2. Typically, 2ml of HIS-Select Nickel Affinity Gel (Sigma) and Glutathione Sepharose 4B (Cytiva) per 1l of culture was used. For proteins overexpressed in large quantities (like E2 and UB constructs) the volume of the beads used was scaled up. After 1h incubation at 4°C on the tube roller, the Flow-Through (FT) was collected and beads washed on the column with 5 Column Volume (CV) either Buffer B (Table 2.2) for His-tagged proteins, or Buffer A (Table 2.2) for Glutathione S-Transferase (GST)-tagged proteins. All E2s were cleaved O/N on beads with respective proteases. All the other proteins were eluted with either Buffer C for His-tagged or with Buffer D for GST-tagged proteins and cleaved O/N. On the next day, cleaved proteins were subject to

Ion-Exchange chromatography (IEX). In general, either Q or S IEX was run, depending on protein's pI. The pH of the buffer was also adjusted accordingly (Buffer E, Table 2.2). Proteins were eluted with 1M NaCl (Buffer F, Table 2.2), concentrated and subjected to Size-Exclusion Chromatography (SEC). SEC Buffer G is listed in Table 2.2. Each purification step was controlled by running SDS-PAGE. Proteins were concentrated, flash frozen and stored at -80°C.

Protein name	Tag and cleavage site	Origin
UBE2R1	GST_TEV	Human
UBE2R1 1-187	GST_TEV	Human
UBE2R1 1-191	GST_TEV	Human
UBE2R1 1-196	GST_TEV	Human
UBE2R1 1-201	GST_TEV	Human
UBE2R1 1-206	GST_TEV	Human
UBE2R1 1-207	GST_TEV	Human
UBE2R1 1-208	GST_TEV	Human
UBE2R1 1-209	GST_TEV	Human
UBE2R1 1-210	GST_TEV	Human
UBE2R1 1-211	GST_TEV	Human
UBE2R1 1-212	GST_TEV	Human
UBE2R1 1-213	GST_TEV	Human
UBE2R1 1-214	GST_TEV	Human
UBE2R1 1-215	GST_TEV	Human
UBE2R1 1-216	GST_TEV	Human
UBE2R1 1-221	GST_TEV	Human
UBE2R1 1-226	GST_TEV	Human
UBE2R1 1-231	GST_TEV	Human
UBE2R1 C223A	GST_TEV	Human
UBE2R1 C191G	GST_TEV	Human
UBE2R1 C191A	GST_TEV	Human
UBE2R1 C191S	GST_TEV	Human
UBE2R1 C191V	GST_TEV	Human
UBE2R1 C191L	GST_TEV	Human

UBE2R1 C191I	GST_TEV	Human
UBE2R1 C191T	GST_TEV	Human
UBE2R1 C191M	GST_TEV	Human
UBE2R1 C191V C223A	GST_TEV	Human
UBE2R1 C191I C223A	GST_TEV	Human
UBE2R1 C191A C223A	GST_TEV	Human
UBE2R1 C93A	GST_TEV	Human
UBE2R1 F206G	GST_TEV	Human
UBE2R1 Y207G	GST_TEV	Human
UBE2R1 Y210G	GST_TEV	Human
UBE2R1 Y211G	GST_TEV	Human
UBE2R1 F206G Y207G	GST_TEV	Human
UBE2R1 Y210G Y211G	GST_TEV	Human
UBE2R1 F206G Y207G Y210G Y211G	GST_TEV	Human
UBE2R1 R149E		
UBE2R2	GST_TEV	Human
UBE2N	GST_TEV	Human
UBE2N K92R	GST_TEV	Human
UBE2D2	GST_TEV	Human
UBE2D3	GST_TEV	Human
UBE2V1	His_GST_Ps3C	Human
UBE2G1	GST_TEV	Human
UBE2S	GST_TEV	Human
UBE2S 1-196 IsoT	GST_TEV	Human
NEDD4	GST_TEV	Human
Rsp5p 383-C	GST_TEV	Yeast
SKP1-FBW7	GST_TEV	Human
NEDD8	GST_Ps3C	Human
APPBP1-**UB^{A}**3	His_Thrombin	Human
UBE2M	GST_Thrombin	Human
UB	Tagless	Human
UB A46V	Tagless	Human

UB A46M	Tagless	Human
UB A46L	Tagless	Human
UB A46I	Tagless	Human
UB A46S	Tagless	Human
UB D58R		
UB G76S GSGSLPETGG	Tagless	Human
UB G76S GSGSLPETGG K48R	Tagless	Human
UB G76S GSGSLPETGG K11R	Tagless	Human
UB G76S GSGSLPETGG K63R	Tagless	Human
UB G76S GSGSLPETGG K48C	Tagless	Human
C_UB	GST_Ps3C	Human
C_UB K11R	GST_Ps3C	Human
C_UB K48R	GST_Ps3C	Human
C_UB K63R	GST_Ps3C	Human
UB_intein ΔG76 A46V	His	Human
RRASVG_UB K48R	GST_TEV	Human
RRASVG_UB K63R	GST_TEV	Human
RING-RING RNF4	GST_TEV	Human
VHL*-ELOB/ELOC	Coexpressed; GST_thrombin	Human

Table 2.1 List of bacterially expressed, recombinant proteins used in this study. The asterix depicts tagged protein, which was used to purify coexpressed complex.

Buffer name	Buffer components
Buffer A	• 50mM TRIS pH 7.5 • 200mM NaCl • 5mM DTT
Buffer B	• 50mM TRIS pH 7.5

	• 200mM NaCl • 1mM β-mercaptoethanol
Buffer C	• 50mM TRIS pH 7.5 • 200mM NaCl • 1mM β-mercaptoethanol • 300mM imidazole
Buffer D	• 50mM TRIS pH 7.5 • 200mM NaCl • 5mM DTT • 10mM Glutathione
Buffer E	• 25mM TRIS/HEPES/MES pH depends on protein's pI • 1mM DTT
Buffer F	• 25mM TRIS/HEPES/MES pH depends on protein's pI • 1M NaCl • 1mM DTT
Buffer G	• 25mM HEPES pH 7.5 • 150mM NaCl • +/- 1mM DTT
Buffer H	• 25mM NaoAc pH 4.5
Buffer I	• 25mM NaoAc pH 4.5 • 250mM NaCl
Buffer J	• 50mM TRIS pH 7.5 • 200mM NaCl • 5mM DTT • 100μM Leupeptin • 2μg/ml Aprotinin • Protease Inhibitor Cocktail Tablets (1 Tablet per 100ml of the buffer)
Buffer K	• 50mM TRIS pH 7.5 • 200mM NaCl • 5mM DTT • 2.5mM Desthiobiotin

Table 2.2 List of buffers used for purification of recombinant proteins derived from *E.coli* and insect cells.

2.2.2 Tagless UB purification

All the step till lysate centrifugation at 20000rpm for 30min, 4°C were followed as in section 2.2.1. Afterwards, the acetic acid was added to the soluble fraction which was continuously stirred, to precipitate all impurities. When pH 4.5 was reached, the solution was still incubated and stirred at RT for 20min. Next, to get rid of precipitated proteins, solution was centrifuged at 20000rpm for 30min, 4°C. Supernatant was collected and dialyzed O/N into 25mM sodium acetate pH 4.5 at 4°C. Next day, the dialyzed solution was centrifuged at 3700rpm, for 10 min, 4°C to remove precipitant. Further the protein was subjected to S IEX and eluted with Buffer I (Table 2.2). Next, SEC was performed for final purification step into the Buffer G (Table 2.2). Each purification step was controlled by running SDS-PAGE. Proteins were concentrated, flash frozen and stored at -80°C.

2.2.3 Insect cell protein purification

Proteins depicted in Table 2.3 were expressed or coexpressed in insect cells. The Table 2.3 contains an information about the tag used, cleavage site as well as an origin of the protein. After bacmid generation (section 2.1.2), baculoviruses were made and isolated from Sf9 cells. Proteins were expressed by infection or coinfection of Hi5 cells. The cells were harvested (15min, 720rpm, 4°C) and resuspended in Buffer J (Table 2.2). Insect cells were either flash frozen or directly used for purification. Protease inhibitor, 2.5mM PMSF, was added before sonication, which was typically performed in 2 cycles, 30% amplitude, 1 sec on 2 sec off. The length of the cell disruption varied depending of the sample volume. In general, the length of sonication was significantly shorter than for bacterial cells. Next, lysate was centrifuged at 20000rpm for 30min, 4°C. The soluble fraction was collected and incubated with previously equilibrated beads. Typically, 2ml Glutathione Sepharose 4B (Cytiva) and 0.5ml StrepTactin Sepharose High Performance (Cytiva) per 1l of culture was used. After 1h incubation at 4°C on the tube roller, the beads were collected by spinning them down for 10min, 1800rpm, 4°C. The supernatant (FT) was discarded and beads were washed on the

column with 5 Column Volume (CV) of Buffer A (Table 2.2). Proteins were eluted with either Buffer D for GST-tagged proteins or Buffer K (Table 2.2) for Strep-tagged proteins and cleaved O/N. On the next day cleaved proteins were subject to ion-exchange chromatography (IEX). In general, either Q or S IEX was run, depending on protein pI. The pH on the buffer was also adjusted accordingly (Buffer E, Table 2.2). Proteins were eluted with 1M NaCl (Buffer F, Table 2.2), concentrated and subjected to Size-Exclusion Chromatography (SEC). SEC Buffer G is listed in Table 2.2 Each purification step was controlled by running SDS-PAGE. Proteins were concentrated, flash frozen and stored at -80°C.

Protein name	Tag and cleavage site	Origin
CRBN-DDB1*	Coexpressed; His_TEV	Human
CUL1-RBX1 5-C*	Coexpressed; GST_TEV	Human
GGGG IKZF1 141-243 Δ197-238 $ZF23^{187}$	GST_TEV	Human
CUL2*-RBX1 5-C	Coexpressed; CUL2: 2xStrep_DAC_TEV RBX1: His_MBP_TEV	Human
CUL2 D660K E664K D675K*-RBX1 5-C	Coexpressed; CUL2: 2xStrep_DAC_TEV RBX1: His_MBP_TEV	Human
CUL2 1-660*-RBX1 5-C	Coexpressed; CUL2: 2xStrep_DAC_TEV RBX1: His_MBP_TEV	Human
CUL4A 38-C – RBX1 5-C*	Coexpressed; GST_TEV	Human
UB^A1	GST_TEV	Human

Table 2.3 List of recombinant proteins expressed in insect cells, used in this study. The asterix depicts tagged protein, which was used to purify coexpressed complex.

2.2.4 UB MESNa generation

UB_intein plasmid was transformed into Rossetta (DE3) cells. Typically, 6x1L TB medium cultures containing 100μg/ml Ampicillin and 34μg/ml of Chloramphenicol were started with 10ml of preculture. The cells were grown at 37°C 180rpm to OD_{600} ~0.8 and the temperature was lowered to 18°C. The 2.5mM IPTG was added to induce protein expression. The next day, cells were spinned down for 15min 7240rpm 4°C and the pellet was resuspended in 20mM HEPES pH 6.8, 50mM NaOAc, 100mM NaCl buffer. Bacterial cells were either flash frozen or directly used for purification. Protease inhibitor, 2.5mM PMSF, was added before sonication, which was typically performed in 2 cycles, 30% amplitude, 1sec on 2sec off for 10min. Next, lysate was centrifuged at 20000rpm for 30min, 4°C. The soluble fraction was collected and incubated with previously equilibrated His beads (20ml). After 1h incubation at 4°C on the tube roller, the Flow-through (FT) was collected and beads washed on the column with 5 Column Volume (CV) of in 20mM HEPES pH 6.8, 50mM NaOAc, 100mM NaCl buffer and eluted with 20mM HEPES pH 6.8, 50mM NaOAc, 100mM NaCl, 300mM Imidazole buffer. Elutions were collected, concentrated to 50ml and diluted to 500ml with cleavage buffer (20mM HEPES pH 6.8, 50mM NaOAc, 100mM NaCl, 100mM MESNa). After O/N MESNa cleavage at RT, on the roller, sample was concentrated and subjected to SEC in 50mM HEPES pH 6.8, 100mM NaCl buffer. Final product was analyzed by intact mass. Each purification step was controlled by running SDS-PAGE.

2.2.5 Fluorescent labelling of UB

UB, UB K11R, UB K48R and UB K63R with a N-terminal Cysteine, for fluorophore modification, were purified from *E.coli*. After SEC, proteins were concentrated to mM range and incubated 15 min on ice with 1mM DTT to reduce the cysteines. To remove reducing reagent, UB's were desalted two times using PD10 (GE Healthcare) desalting columns and incubated for 2h at RT with 5 molar excess of Fluorescin-5-maleimide (AnaSpec Inc.). To get rid of an excess of fluorophore, proteins were doubly desalted, concentrated and submitted for SEC (Buffer G, Table 2.2). To make sure the full removal of the dye, SEC was repeated. Each purification step was controlled by running SDS-PAGE. Proteins were concentrated, flash frozen and stored at -80°C.

2.2.6 Radiolabelling of UB

UB K48R and UB K63R with N-terminal protein kinase A consensus sequence (RRASVG), were purified from *E.coli*. After SEC proteins were concentrated to 50 or 100μM. Radiolabeling was performed using [γ32P]-adenosine triphosphate (ATP) and 5 kU of cyclic adenosine monophosphate-dependent protein kinase (New England Biolabs) for 1h at 30°C. Proteins were concentrated, flash frozen and stored at -80°C.

2.2.7 CUL neddylation

Purified CUL1-RBX1, CUL2-RBX1 and CUL4A-RBX1 were covalently modified by UBL NEDD8 (neddylated) in the presence of APPBP1-UB^A3 (NEDD8 E1) and UBE2M (NEDD8 E2) and excess of NEDD8. The reaction buffer contained 25mM HEPES pH 7.5, 100mM NaCl, 2.5mM $MgCl_2$ and 1mM ATP. First, the efficiency of the reaction was checked. As such, the small scale reaction was run, where 12μM of CUL-RBX1 was incubated with 0.2μM of APPBP1-UB^A3, 2μM of UBE2M and 25μM NEDD8. The samples were taken and quenched with reducing SDS-PAGE sample buffer after 5,8,10 and 20min. The products were analyzed by SDS-PAGE and the time of the reaction was adjusted accordingly to obtain fully neddylated CULs and avoid overneddylation. The large scale reaction was run using the same reagents concentrations and quenched with 10mM DTT. Neddylated CULs were subjected to SEC (Buffer G, Table 2.2) for a final purification step, purity was analyzed by SDS-PAGE. Proteins were concentrated, flash frozen and stored at -80°C.

2.3 Peptides

Table 2.4 contains the list of the peptides used for transpeptidation reactions, cryo-EM and assays. All peptides were synthetized in Max Planck Institute of Biochemistry Core Facility and had >95% purity by HPLC.

N°	Name of the peptide	Sequence
1	cyclin E	H-GGGGLPSGLL(pT)PPQ(pS)GKKQSSDYKDDDDK-OH
2	cyclin E	Ac-KAMLSEQNRASP LPSGLL(pT)PPQ(pS)GKKQSSDYKDDDDK-OH
3	Sna4p	H-GGGGQSLVESPPPYVPENLYFQGDYKDDDDK-OH
4	Hif1α	H-GGGGEMLA(hP)YIPMDDDLQLRSFDDYKDDDDK-OH
5	Hif1α	Ac-KNPFSTQDTDLDLEMLA(hP)YIPMDDDLQLR-OH
6	Hif1α*	Ac- KLRREPDALTLLA(hP)AAGDTIISLDFGSNK*-OH

Table 2.4 List of peptides used in this study. pT- phosphorylated Threonine, pS- phosphorylated Serine, hP- hydroxylated Proline. Asterix depicts fluorescent label.

2.4 Transpeptidation reaction

Sortase-mediated transpeptidation reaction was performed to link C-terminus of acceptor UBs to the N-terminus of various peptides: cycline E phosphopeptide, Sna4p, Hif1α and GGGG IKZF1141-243 Δ197-238 ZF23. The exact sequences of the peptides are listed in Table 2.4. UBs were either synthetized or recombinantly expressed in *E.coli* (Section 2.2.2). All UB versions contained C-terminal G76S mutation followed by a GSGS linker and LPETGG sortase recognition sequence. 50μM UB was added to 100μM peptides or IKZF1 ZF23. Reaction was started with 10μM His-sortase and incubated on ice for 1h in a 50mM TRIS pH 8.0, 150mM NaCl, 10mM $CaCl_2$ buffer. Next, the mixture was submitted to His-passback to remove His-sortase and final products were separated on SEC in 25mM HEPES pH 7.5, 200mM NaCl, 1mM DTT buffer. Additionally, UB-Sna4p fusions were incubated O/N with TEV protease for FLAG tag removal, followed by SEC.

2.5 Activity-Based Probes (ABPs) generation

2.5.1 UB BmDPA generation

UB BmDPA was a basis for all ABPs generated in this study. First, previously made UB MESNa was incubated O/N, 30°C, 300rpm with Molecule 1 ((E)-3-[2-

(bromomethyl)-1,3-dioxolan-2-yl]prop-2-en-1-amine (BmDPA) (>95% purity, ChiroBlock), Figure 2.1). Reaction solution contained 90mg of Molecule 1 per 1ml, 5mg/ml UB MESNa (final concertation, non-hydrolyzed), 50mM HEPES pH 6.7, 1mM N-hydroxysuccinimide and 10% DMSO. Full conversion of UB MESNa to Molecule 2 (Figure 2.1) was checked by mass spectrometry and Molecule 2 was dialyzed O/N to 50mM HEPES pH 6.7, 100mM NaCl buffer. Next deprotection step was performed. As such, Molecule 2 was mixed with 0.04M p-TsOH dissolved in 54% TFA (v/v) with the ratio 1:4 and incubated 1h at RT. Next, to get rid of TFA from the mixture, UB was precipitated and the protein flakes were washed with cold ether (45-fold reaction volume). Afterwards, precipitant was air-dried and respuspnded in 100mM Phosphate buffer pH 6.0, 500mM NaCl and 8M urea. Subsequently, UB was refolded into the 100mM Phosphate buffer pH 6.0, 100mM NaCl by O/N dialysis. Generated UB BMDPa (Molecule 3, Figure 2.1) was controlled by mass spectrometry.

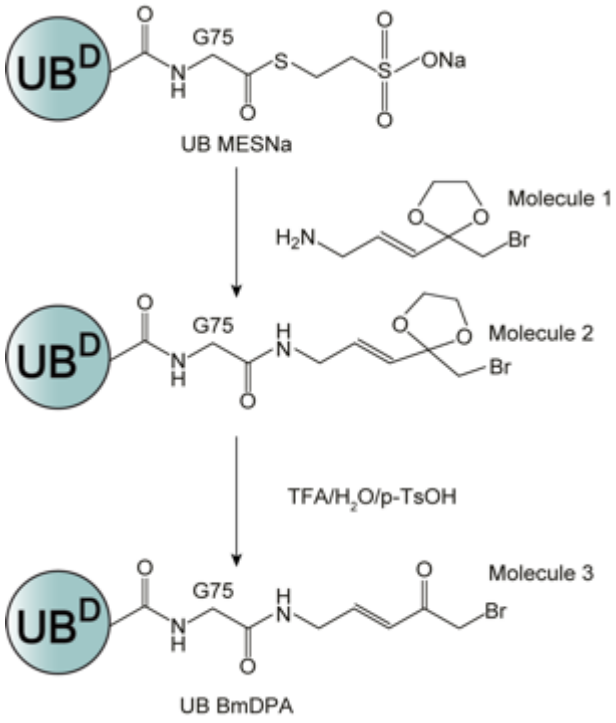


Figure 2.1 Schematic representation of chemical reactions performed to generate UB BmDPA.

2.5.2 diUB probe generation

To synthetize diUB probes, following reagents were firstly prepared: (1) N-terminus of Hif1α peptide 4 (Table 2.4) was linked in a sortase-mediated transpeptidation reaction

into the C-terminus of UB K48C (details of the reaction are described in section 2.4); (2) UB BMDPa.

Hif1α-UB K48C was incubated with 1mM TCEP on ice for 20min for cysteine reduction. Afterwards, Hif1α-UB K48C was desalted using Zeba desalting columns into 25mM HEPES pH 7.5, 150mM NaCl buffer and immediately added to UB BMDPa. Both reagents were mixed in 1:3 ratio and incubated at 30°C for 1h. Final product was purified by SEC in 25mM HEPES pH 7.5, 150mM NaCl buffer and analyzed by SDS-PAGE.

2.6 Biochemical assays

2.6.1 Discharge assay to free amino acids (pulse-chase)

20μM UBE2N was charged with 20μM fluorescently-labeled donor UB K63R (UB*) using 0.3μM UB^A1 (E1) in 50mM HEPES pH 7.5, 100mM NaCl, 2.5mM $MgCl_2$, 1.5mM ATP and 0.05mg/ml BSA buffer for 30min. The pulse reaction was stopped by EDTA addition to a final concentration of 30mM. Next, UBE2N~UB* (final concertation 0.5μM) was added to initiate chase reaction in a mixture containing 0.5μM RNF4 RING dimer, 0.5μM UBE2V1 and 35mM free amino acids acceptors (L-Serine, L-Dap, Nα-acetyl-L-ornithine, L-lysine, D-lysine, Nα-acetyl-L-lysine, Nε-acetyl-L-lysine and L-homolysine) in 25mM HEPES pH 7.5, 150mM NaCl buffer. The reducing and non-reducing SDS-PAGE sample buffer was used to quench the reactions after 0, 5, 10, 20, 30, 45, 60, 120 or 180min. Products and substrates were separated on SDS-PAGE and the Amersham Typhoon system (GE Healthcare) was used to scan the gels. Fluorescent bands intensities were quantified using ImageQuantTL (GE Healthcare). Further, UBE2N~UB* band intensities were divided by each lane's total fluorescence. The data was normalized to the 0 time point, plotted in GraphPad Prism 8 (GraphPad Software) and fitted, using nonlinear regression, to an exponential decay function. Assays were performed in duplicate.

2.6.2 Discharge assay to UB analogs (pulse-chase)

20μM E2 was charged with 20μM fluorescently-labeled donor UB (UB*) using 0.3μM UB^A1 (E1) in 50mM HEPES pH 7.5, 100mM NaCl, 2.5mM $MgCl_2$, 1.5mM ATP and 0.05mg/ml BSA buffer for 30min. The pulse reaction was stopped by EDTA addition to a final concentration of 30mM. Next, E2~UB* (final concertation 0.5μM) was added to initiate chase reaction in a mixture containing UB acceptor analogs and in some cases E3s in 25mM HEPES pH 7.5, 150mM NaCl buffer. Tables 2.5-2.7 describe exact concentrations of UB analogs and E3s used in specific assays. All reactions where quenched with non-reducing SDS-PAGE sample buffer after indicated time. Products and substrates were separated on SDS-PAGE and the Amersham Typhoon system (GE Healthcare) was used to scan the gels. Fluorescent bands intensities were quantified using ImageQuantTL (GE Healthcare). Further, diUB* band intensities were divided by each lane's UB* total fluorescence intensity and then multiplied by the amount of UB* used in the reactions. Assays were performed in duplicate.

For UBE2S assays, to minimize E2 autoubiquitylation, E1~UB* pulse reaction was performed. 10μM UB^A1 was charged with 10μM fluorescently-labeled donor UB (UB*) for 10min in 50mM HEPES pH 7.5, 100mM NaCl, 2.5mM $MgCl_2$, 1.5mM ATP and 0.05mg/ml BSA buffer. The pulse reaction was stopped by EDTA addition to a final concentration of 30mM and desalted twice using Zeba desalting columns. Next, ~5μM E1-UB* was added to initiate chase reaction in a mixture containing UB acceptor analog, 0.2μM UBE2S and in some assays APC/C with its coactivator CDH1. All proteins concentrations used in these assays can be found in Tables 2.5 and 2.8 UBE2S reactions were analyzed as described above.

For substrate-dependent reactions, either phosphorylated cyclin E (CRL1) or IKZF1 ZF23 (CRL4) fused to UB analogs were used with its respective substrate receptors: SKP1-FBW7 and CRBN. To facilitate complex formation between IKZF1 ZF23 and CRBN 2μM small molecule pomalidomide was added to the reactions. Similarly, 0.5μM E2~UB* was added to the chase reactions. Table 2.8 describes exact concentrations of substrate-UB analogs and E3s used in specific assays. Substrate-dependent reactions were analyzed as described above.

E2 enzyme	UB [µM]
UBE2N/UBE2V1	300
UBE2R2	750
UBE2G1	550
UBE2S_IsoT	50

Table 2.5 Final acceptor UB concentrations, used in chase reactions, in the absence of an E3.

E2 enzyme	UB [µM]	RNF4 [µM]
UBE2N/UBE2V1	30	0.5
UBE2D3	100	1.0

Table 2.6 Final acceptor UB and E3 concentrations, used in chase reactions, in the presence of RNF4 E3.

Assay type	UB [µM]	HECT E3 [µM]	UB-substrate [µM]
UBE2D2/NEDD4	30	1.0	-
UBE2D2/Rsp5p	10	0.3	-
UBE2D2/Rsp5p	-	0.3	0.5

Table 2.7 Final acceptor UB and E3 concentrations, used in chase reactions, in the presence of HECT E3s NEDD4 and Rsp5p.

E2 enzyme	UB [μM]	UB-substrate [μM]	NEDD8-CRL1 [μM]	SKP1/FBW7 [μM]	NEDD8-CRL4 [μM]	CRBN [μM]	APC/C+CDH1 [μM]
UBE2R2	40	-	0.4	-	-	-	-
UBE2R2	-	0.4	0.2	0.2	-	-	-
UBE2G1	100	-	-	-	1.0	-	-
UBE2G1	-	2.0	-	-	1.0	1.0	-
UBE2S	20	-	-	-	-	-	0.1

Table 2.8 Final acceptor UB and E3 concentrations used in chase reactions, in the presence of $CRL1^{FBW7}$, $CRL4^{CRBN}$ and APC with its coactivator CDH1.

2.6.3 Multiturnover assays

Multiturnover assays were performed in 50mM HEPES pH 7.5, 100mM NaCl, 2.5mM $MgCl_2$, 1.5mM ATP and 0.05mg/ml BSA buffer. Typically, 0.3μM UB^A1, 1μM UBE2R1, 20μM UB* K48R, 0.5μM E3 (either N8C2R1+ELOB/C-VHL or N8C1R1 + SKP1/FBW7) and 2μM substrate peptide, were used. All reactions were quenched with reducing SDS-PAGE sample buffer after indicated time. Products and substrates were separated on SDS-PAGE and the Amersham Typhoon system (GE Healthcare) was used to scan the gels.

2.6.4 CUL2 pulse-chase assays

CUL2 pulse-chase assays were performed as described in section 2.6.2. Several following changes were made: the chase reactions were run in 25mM HEPES pH.7.5, 20mM NaCl buffer. For diUB formation assays, 100μM acceptor UB and 0.5μM NEDD8-CUL2 were used. In substrate-dependent experiments, 2μM Hif1α-UB and 0.5μM NEDD8-$CRL2^{VHL}$ (or different CUL2 mutats) were added. Reactions were analyzed as described above. For the assays, where non-labeled UB^D was used, final 10μM E2~UB* was added to the chase reactions.

2.6.5 Steady-state assays

Steady-state assays were performed in 30mM TRIS pH 7.5, 100mM NaCl, 5mM $MgCl_2$, 2mM ATP and 1mM DTT buffer. 1μM E1, 1μM E2 and 20μM UB were mixed and incubated in RT for 15 min (tube 1). Parallelly, 0.5μM NEDD8-$CRL2^{VHL}$ and 0.2μM Hif1α* were pre-mixed (tube 2). Ubiquitylation reactions were started with mixing 10μl of the reaction components from tube 1 and tube 2, resulting in 2x dilution. All reactions were quenched with reducing SDS-PAGE sample buffer after indicated time. Products and substrates were separated on SDS-PAGE and the Amersham Typhoon system (GE Healthcare) was used to scan the gels.

2.7 Cryo-EM

2.7.1 Sample preparation

Previously generated diUB ABPs were used to trap and form the complex mimicking catalyzed by UBE2R1 transfer of UB^D to substrate-linked UB^A. As such, first UBE2R1 was incubated with 1mM TCEP on ice for 20 min. Afterwrads, UBE2R1 was desalted using Zeba desalting columns into 25mM HEPES pH 7.5, 150mM NaCl buffer and immediately added to the rest of complex components. Reaction mixture contained 7.5μM UBE2R1, 7.5μM N8C2R1, 7.5μM ELOB/C-VHL and 30μM diUB ABP. Proteins were incubated together for 30min in 30°C and subsequently complex was purified by SEC into 25mM HEPES pH 7.5, 100mM NaCl, 1mM TCEP buffer. The purity and reaction efficiency was assessed by SDS-PAGE and appropriate fractions pooled and concentrated to around 0.4 mg/ml. To further stabilize the complex, CC0651 to the final concentration of 200μM was added[129] and complex was crosslinked with 0.05% Glutaraldehyde for 15min on ice. Crosslinking was quenched with 50mM Tris pH 7.5. 3.5μl of the sample was applied onto R1.2/1.3 holey carbon grids (Quantifoil), blotted for 3.5s with blot force 4 (100% humidity, 4°C) and plunge-frozen in liquid ethane. Vitrobot Mark IV was used for all plunging experiments.

2.7.2 Data collection

All screening datasets were collected on a 200kV Glacios Transmission Electron Microscope (TEM), containing Gatan K2 direct detection camera with counting mode. On average, per dataset, 2000 micrographs were collected with a 1.885Å pixel size. Data collection scheme included defocus range from -1.5μm to -2.7μm and total exposure between 60 to 70 e^{-} $Å^{-2}$ (45 frames). High resolution dataset was collected on a 300kV Titan/Krios TEM, containing a post-GIF Gatan K3 Summit direct electron camera with counting mode. Around 20000 micrographs were collected with a 0.885 Å pixel size. Data collection scheme included defocus range from -0.7μm to -2.2μm and total exposure 70 e^{-} $Å^{-2}$.

2.7.3 Data processing

All screening datasets were processed in RELION 3.0[194] and Titan/Krios dataset in RELION 4.0[195]. First, raw movie frames were align and dose-weighted. Second, contrast-transfer-function was used via Gctf to correct micrographs[196]. Particles were picked using Gautomatch (K.Zhang). All further operations including 2D classification, 3D classification, local and global 3D refinement, post-processing were done using RELION 3.0 or RELION 4.0 software.
Ab-initio reconstruction for initial model generartion was performed in cryoSPARC[197]. Additionally, DeepEMhancer was used to generate post-processed final maps[198].

2.8 Liquid chromatography-tandem mass spectrometry

The Liquid Chromatography-tandem Mass Spectrometry (LC/MS) experiments were done in collaboration with Fynn Hansen, Ozge Karayel and Matthias Mann from Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Martinsried, Germany.

2.8.1 In-gel digestion

UBE2D3-madiated pulse-chase reactions were done as described in section 2.6.2. First, UBE2D3 was charged with either GST-UB or fluorescent UB (UB*). The GST-

tagged UB^D was used to differentiate between $diUB^D$, which is a side product of pulse reaction and diUB between UB^D and UB^A. Second, 0.5μM UBE2D3~UB^{D}* was added to the reactions containing 100μM UB_{C4}, $^{K11}UB_{C5}$, $^{K48}UB_{C5,}$ $^{K63}UB_{C5}$ and 1μM RNF4 RING dimer. Assays were run 15min for UB_{C4} and 1h for UB_{C5}s and stopped with SDS-PAGE sample buffer. Products and substrates were separated on SDS-PAGE and scanned on an Amersham Typhoon system (GE Healthcare) for reactions consisting UB*. For assays run in the presence of GST-UB^D, gels were stained with Coommassie brilliant blue to visualize GST-UB^D-UB^A product band. Subsequently, the gels were destained in 40% methanol, 10% acetic acid and 60% de-ionized water for several hours. The destaining solution was changed at least two times to achieve a clear background. Desired GST-UB^D-UB^A gel band was cut from the gel and chopped into smaller pieces (~1x1mm). Next, gel pieces were washed twice in 50% EtOH/50% 50 mM ammonium bicarbonate, pH 8.0 (ABC) followed by their dehydration in absolute EtOH. Further, gel pieces were dried in Speed-vac system (Eppendorf, Concentrator plus), rehydrated by incubation at 37°C O/N in 200 µl of 1% (wt/vol) Sodium deoxycholate (SDC) buffer (40 mM 2-chloroacetamide (CAA), 10 mM Tris-(2-carboxyethyl)phosphine (TCEP), 0.5 µg LysC, 0.5 µg trypsin, in 100 mM Tris-HCl pH 8.5). To extract peptides from gel pieces, isopropanol buffer (1% 2,2,2-trifluoroacetic acid (TFA) in isopropanol) was added in two consecutive rounds and liquid phase was collected. Next, Stable Isotope-Labeled (SIL) analogs of native, chain-specific di-Gly were added to the samples. These peptide standards provided chromatographic orientation for light, endogenous counterparts detection. In UB_{C4} samples, for di-Gly peptides absolute quantification, SIL analogs amounts were modified to yield peptide quantification ratios between 0.1 and 10 (2fmol and 20fmol for K27_GG, K29_GG, K33_GG and K11_GG, K48_GG, K63_GG, respectively).

2.8.2 LC/MS sample preparation

In-house prepared syringe device (described previously in[199,200]) was used to insert three layers of a Styrenedivinylbenzene-reverse phase sulfonate (SDB-RPS) matrix (Empore) into a pipette tip for StageTips preparation. Next, previously mixed with isopropanol buffer peptides, were loaded onto the StageTips followed by tips wash with (1) isopropanol buffer and (2) 0.2% TFA/2% ACN. 1.25% NH_4OH/80% ACN was used for sample elution and collected fractions were dried in a Speed-vac centrifuge.

Before LC/MS analysis, peptides were resuspended in 0.2% TFA/2% CAN and shortly sonicated (Branson Ultrasonics).

2.8.3 LC/MS sample measurements

Prepared peptides were loaded onto 50-cm, 75 µm inner diameter reversed-phase column (in-house packed with eproSil-Pur C18-AQ 1.9-µm resin (Dr Maisch)), which temperature was kept at 60°C by a homemade column oven. The mass spectrometer (Exactive HF-X, Thermo Fisher Scientific) was connected online into an EASY-nLC 1200 system (Thermo Fisher Scientific) via a nano-electrospray source. Binary buffer system of buffer A (0.1% formic acid (FA)) and buffer B (0.1% FA, 80% acetonitrile) was used for peptides elution at 300 nl min^{-1} flow rate. Elution was done as follows: first, 7% buffer B gradient was used (0.1% (vol/vol) FA, 80% (vol/vol) ACN), second the gradient was stepwise increased to 14% in 4min, then to 26% in 22min. Lastly, the gradient was developed to 95% in 2min and maintained for 6min.

The programmed targeted scan mode was used by mass spectrometer. Every full acquisition scan (resolution of 60,000 at 200 m/z (3 × 106 ions accumulated with a maximum injection time of 20 ms)), 20 multiplexed Selected Ion Monitoring (SIM) scans were done (multiplexing degree of four). Heavy and light (endogenous) counterparts were simultaneously recorded. In every SIM scan m/z of 150–2,000 range was covered with resolutions of 120,000 (5 × 104 ions accumulated with a maximum injection time of 65 ms, 1.4-m/z isolation window and 0.4-m/z isolation offset). Detected tagerted peptide with their respective values are listed in Table 2.9.

Peptide	Sequence	Mass [m/z]	Charge State [z]
NleQIFVK	NleQIFVK	374.24181	2
K06_GG (light)	NleQIFVK^{GG}TLTGK	454.60989	3
K11_GG (light)	TLTGK^{GG}TITLEVEPSDTIENVK	801.42688	3
K11_GG (heavy)	TLTGK^{GG}TI***T***LEVEPSDTIENVK	803.09703	3
K27_GG (light)	TITLEVEPSDTIENVK^{GG}AK	701.03895	3
K27_GG (heavy)	TITLEVEPSDTIEN***V***K^{GG}AK	703.04356	3
K29_GG (light)	AK^{GG}IQDK	408.73233	2
K29_GG (heavy)	AK^{GG}***I***QDK	412.24092	2
K33_GG (light)	IQDK^{GG}EGIPPDQQR	546.61291	3
K33_GG (heavy)	IQDK^{GG}EGIP***P***DQQR	548.61751	3
K48_GG (light)	LIFAGK^{GG}QLEDGR	487.60005	3
K48_GG (heavy)	LIFAGK^{GG}QLEDG***R***	490.93614	3
K63_GG (light)	TLSDYNIQK^{GG}ESTLHLVLR	561.80503	4
K63_GG (heavy)	TLSDYNIQK^{GG}ESTLHLV***L***R	563.55932	4

Table 2.9 List of targeted peptides. Superscripted “GG” mark Gly-Gly modified lysines and bold/italic characters point stable isotope labeled amino acids.

2.8.4 LC/MS data analysis

Collected raw MS data was analyzed in an open-source software project, Skyline[199,200]. Proper peak picking of MS1 filtered light (endogenous) peptides based on their co-elution with heavy (SIL) peptides, was manually inspected on graphical representation of chromatographic traces. Quantifications were done only on the most abundant peaks of the isotope cluster, based on area and on the precursor level.

Subsequent bioinformatics analyses were implemented in Microsoft Excel and plotted in GraphPad Prism 8 (GraphPad Software). From the signals, of the samples, run in the presence of UB^{A}s, background signal, detected in sample where no UB^{A} was added was subtracted. Further, the first tryptic peptide “NleQIFVK” of N-terminally modified UB^{A} (Met1 mutated to Norleucine (Nle)) was used to normalize each chain peptide. As

the normalization peptide contained K6, this chain type was not measured in the protocol. Relative to WT fold change of each chain peptide was calculated using equations specified below.

Equation 1:

Background correction:

$$\boldsymbol{AUC_{rep\ i}^{UB^{Acceptor,BGcorr}} = AUC_{rep\ i}^{UB^{Acceptor}} - \overline{AUC}^{UB^{no\ Acceptor}}\ \{i|1 \leq i \leq 3\}}$$

Equation 2:

Normalization to “NleQIFVK” peptide:

$$\boldsymbol{_{GG-Pep.}AUC_{rep\ i}^{UB^{Acceptor,Norm}} = \frac{_{GG-Pep.}AUC_{rep\ i}^{UB^{Acceptor,BGcorr}}}{_{NleQIFVK}AUC_{rep\ i}^{UB^{Acceptor,BGcorr}}}\ \{i|1 \leq i \leq 3\}}$$

Equation 3:

Fold change calculation relative to WT:

$$\boldsymbol{_{GG-Pep.}FC_{rep\ i}^{UB^{Acceptor,Norm}} = \frac{_{GG-Pep.}AUC_{rep\ i}^{UB_{C5}^{Acceptor,Norm}}}{_{GG-Pep.}\widetilde{AUC}^{UB_{C4}^{Acceptor,Norm}}}\ \{i|1 \leq i \leq 3\}}$$

2.9 Kinetics

Kinetics experiments were done in collaboration with Nicholas Purser, Daniel Houston, Nicole Burton and Gary Kleiger from Department of Chemistry and Biochemistry, University of Nevada, Las Vegas, Nevada, USA.

2.9.1 pK_a^{app} estimation

Steady-state kinetics assays were used to determine pK_a^{app} for synthetic UB_{C4} and $^{K63}UB_{C5}$ in the presence of UBE2N/UBE2V1 complex. The assay monitors isopeptide bond formation between radiolabeled UB^D and unlabeled UB^A. Initially, BIS-TRIS propane buffer titration series was created with pH values 5.7, 6.1, 6.5, 6.9, 7.3, 7.7,

8.1, 8.5, 8.9, 9.3, 9.7 and 10.1. The reaction mixtures were prepared with BIS-TRIS stock buffer (final concentration 50mM), 5µM radiolabeled UB^D K63R, 2µM UBE2N (WT or K92R)/UBE2V1 and 0.25µM UB^A1 in 50mM NaCl, 2mM ATP, 5mM $MgCl_2$,1mM DTT. All the components were incubated for 1min before addition of synthetic UB^As (either UB_{C4} or $^{K63}UB_{C5,}$ final concertation 100µM). Reactions run in a presence of UB_{C4} were quenched after 2min 45sec and with $^{K63}UB_{C5}$ after 15min with either reducing or non-reducing SDS-PAGE sample buffer. Products and substrates were separated on SDS-PAGE, followed by audioradiography. Gels were scanned on a Typhoon 9410 Imager (GE Healthcare) and bands intensities were quantified using ImageQuantTL (GE Healthcare). Next, diUB band intensities were divided by each lane's total signal intensity and then multiplied by the amount of radiolabeled UB^D used in the reactions. Finally, obtained value was divided by the time of incubation and UBE2N/UBE2V1 complex concentration. GraphPad Prism software was used to plot velocities as a function of the reaction's pH and fit to a sigmoidal four-parameter logistic curve with the Hill slope constrained to 1. Importantly, this model is based on an assumption, that the reaction velocity depends on a single ionizing species being activated under higher pH. A similar approach was used to determine pK_a^{app} for synthetic UB_{C4} and $^{K48}UB_{C5}$ in the presence of UBE2R. Following changes were made: (1) 15µM radiolabeled UB^D K48R, 10µM UBE2R2 and 0.5µM UB^A1was used, (2) UB_{C4} reactions were stopped after 5 min and $^{K48}UB_{C5}$ ones after 60min. The incubation times were adjusted to make sure all velocities were within the linear range. Assays were performed in duplicate.

2.9.2 K_M estimation

Steady-state kinetics assays were used to determine K_M for synthetic UB_{C4} and $^{K63}UB_{C5}$ in the presence of UBE2N/UBE2V1 complex with or without its E3, RNF4 RING dimer. The assay monitored isopeptide bond formation between radiolabeled UB^D and unlabeled UB^A. First, UB^As were dialyzed into 30 mM BIS-TRIS propane, pH 7.3. Next, the two fold UB^A's dilution series was prepared, starting from 1.4mM for UB_{C4} made in *E.coli*, 1.6mM for synthetic UB_{C4} and 1.3mM for $^{K63}UB_{C5}$. UB^As were further diluted by two fold in the reaction mixture, which was composed of 0.25µM UB^A1, 5µM radiolabeled UB^D K63R, 2µM UBE2N /UBE2V1 for UB_{C4}s and 15µM radiolabeled UB^D K63R, 10µM UBE2N /UBE2V1 for $^{K63}UB_{C5}$ in 50mM BIS-TRIS pH 7.3, 50mM NaCl, 2mM ATP, 5mM $MgCl_2$, 1mM DTT buffer. After 2 min incubation of a

reaction mixture, ubiquitylation was initiated by adding acceptor UBs. Reactions were quenched with either non-reducing or reducing SDS-PAGE sample buffer after either 15 or 30min for UB_{C4} or $^{K63}UB_{C5}$, respectively. Further processing was done as described in section 2.9.1. Obtained velocities values were fitted to the Michaelis-Menten equation for K_M estimation (GraphPad Prism). A similar approach was used to determine K_M in the presence of RNF4 RING dimer E3. Following changes were made: (1) reaction mixture contained 0.5µM UB^A1, 15µM radiolabeled UB^D K63R, 10µM UBE2N /UBE2V1 and 1µM RNF4, (2) dilution series was started from 1.3mM UB_{C4} and 1.25mM $^{K63}UB_{C5}$, (3) incubation was changed to 0.5 or 2.5min for UB_{C4} or $^{K63}UB_{C5}$, respectively. The incubation times were adjusted to make sure all velocities were within the linear range. Assays were performed in duplicate.

Steady-state kinetics assays to determine K_M for synthetic UB_{C4} and $^{K48}UB_{C5}$ in the presence of UBE2R2 were measured similarly to the experiments described above with the following modifications. Acceptor UB dilution series was started from 8.5mM for synthetic UB_{C4} and 12.5mM for $^{K48}UB_{C5}$. Reaction mixture was composed of 0.5µM UB^A1, 15µM radiolabeled UB^D K48R and 10µM UBE2R2. The reactions were run for either 1 or 2.5 min for UB_{C4}'s replicates and either 15 or 16 min for $^{K48}UB_{C5}$'s. Steady-state kinetics assays to determine K_M for synthetic UB_{C4} and $^{K63}UB_{C5}$ in the presence of HECT E3 Rsp5p, was followed in a similar pattern. The starting concentrations for dilution series were 1.6mM, for both, synthetic UB_{C4} and $^{K63}UB_{C5}$. Reaction mixture contained 0.5µM UB^A1, 7µM radiolabeled UB^D K63R 5µM UBE2D2 and 5µM Rsp5p. Reactions were quenched after 5 or 30 min in the presence of UB_{C4} or $^{K63}UB_{C5}$, respectively. The incubation times were adjusted to make sure all velocities were within the linear range. Assays were performed in duplicate.

2.10 UB analogs synthesis

UB analogs synthesis was performed by Gerbrand J. van der Heden van Noort and Huib Ovaa from Oncode Institute and Department of Cell and Chemical Biology, Chemical Immunology, Leiden University Medical Centre, Leiden, the Netherlands.

2.10.1 Solid peptide synthesis (SPPS)

Syro II MultiSyntech Automated Peptide synthesizer was used for SPPS. 25µM scale syntheses were performed with standard 9-fluorenylmethoxycarbonyl (Fmoc) based solid phase peptide chemistry, where four-fold amino acids excess was used over the pre-loaded Fmoc amino acid trityl resin (0.2 mM/g, Rapp Polymere GmbH). Following reagents were purchased from Iris Biotech GmbH: Fmoc-Dap(Boc)-OH (C1), Fmoc-Dab(Boc)-OH (C2) and Fmoc-hLys(Boc)-OH (C5). Fmoc-Orn(Boc)-OH (C3) was bought from ChemImpex Int'l Inc. UB analogs were synthetized on resin according to the previously described protocol followed by 2.5h incubation in TFA/TIS/H_2O/Phenol[201]. Next, UBs were precipitated from Et2O/Pentane and purified by RP/HPLC. The fractions were analyzed by LC/MS, pooled, lyophilized and stored at -20°C.

2.10.2 RP/HPLC

RP-HPLC Water preparative system was used with a C18-Xbridge 5 µm OBD column (30 x 150 mm). Purification was performed using 3 mobile phases: Buffer A: miliQ, Buffer B CH_3CN and Buffer C: 1% TFA in miliQ. Separation was done at a flow rate of 37.5ml/min with the gradient 20-45% Buffer B, 5% Buffer C.

2.10.3 LC/MS

LC/MS measurements were performed on Waters Acquity UPLC (Waters Acquity BEH C18-column (2.1x50 mm, 1.7 µm)) connected with LCTTM ESI-Mass Spectrometer. Samples were analyzed at a flow rate of 0.6ml/min (run time: 6min; column temperature: 60°C). Two mobile phases were used: Buffer A: 1% CH_3CN, 0.1% formic acid in miliQ and Buffer B: 1% miliQ and 0.1% formic acid in CH_3CN, with the gradient 0-95% Buffer B. Data was processed on Waters MassLynx Mass Spectrometry Software 4.1 (deconvolution with MaxEnt1 function).

2.10.4 High resolution Mass Spectrometry (HRMS)

Waters XEVO-G2 XS Q-TOF mass spectrometer was used for UB analogs HRMS. Spectra were recorded with an electrospray ion source in positive mode (temperature 250°C, desolvation gas flow 900L/h, source voltage 3.0kV) with resolution (R) 22000 (mass range m/z 50-2000) and 200 pg/uL Leu-Enk (m/z = 556.2771, “lock mass”).

2.11 1D and 2D proton NMR

NMR was done in collaboration with Mark J. Bostock and Michael Sattler from Biomolecular NMR and Center for Integrated Protein Science Munich at Department Chemie, Technical University of Munich, Garching, Germany and Institute of Structural Biology, Helmholtz Zentrum München, Neuherberg, Germany.

Bruker Avance AVIII 600 MHz spectrometer (^{1}H frequency) was used to record NMR experiments at 298 K or 310 K with a 5 mm QCI cryoprobe. First, ~200 µM UB_{C4}, $UB_{C4\ Bio}$ and $^{K48}UB_{C5}$ samples were dialyzed into 50mM sodium phosphate pH 6.2, 50mM NaCl buffer and supplemented with 10% D_2O. The watergate W5 double echo sequence with gradients was used to acquire 64 scans for 1D spectra[202]. To record 2D [^{1}H,^{1}H] spectra, 2D NOESY pulse sequence with water flip-back and watergate water-suppression was performed (32 scnas, mixing time 120ms)[203,204]. Acquisition for NOESY spectra were done with a ^{1}H spectral width of 8417.5 Hz and 1024 x 256 complex points. Data was processed with the Azara package (W. Boucher, unpublished) and analyzed in CcpNmr Analysis[205].

2.12 Molecular modelling

Molecular dynamics simulations were performed by Vinh H. Truong and Joseph S. Harrison from Department of Chemistry, University of the Pacific, Stockton, CA, USA.

2.12.1 C5 incorporation for Rossetta and Molecular Dynamics (MD)

Neither the molecular dynamics software nor the Rosetta[206] molecular modelling program consist a noncanonical homolysine (C5) amino acid. To incorporate C5 into UB, Rosetta protocol was used to incorporate non-natural amino acids[207]. PubChem

chemical database was a source of iSDF file with dealized atom coordinates for 2,7-Diaminoheptanoic Acid (C5). Subsequently, SDF file was converted to a params file by the molfile2params script in Rosetta. Additionally, Spartan was used to create a deprotonated version of 2,7-Diaminoheptanoic acid with a neutral terminal amine. Finally, C5 was installed into the UB with the help of fixbb application in Rosetta.

In order to performed MD simulations all amino acid residues within the protein must be characterized by its topology file, containing the type, charge and mass of every atom within the simulated protein. Spartan software was applied for C5 partial charges calculation. Full ab-initio calculations were done in Hartree-Fock 6-31* G basis set (both for neutral C5 residues and individual charges) and inserted into the topology file.

2.12.2 UBE2N~UB^{D}/UBE2V1/UB^{A} complex building

To create a thioester bond between UB^{D} and UBE2N, Roesetta's UBQ_E2_Thioester protocol was used. The protocol not only sampled E2 enzyme in thioester-linked state but also positioned a nucleophile approaching the active site[116]. Importantly, existing version of this code positions the amine in plane with the carbonyl bond, which does not resemble tetrahedral intermediate geometry. Thus, 3-D tetrahedron geometry calculation were applied to estimate acceptor position with respect to the active site. In details, UB^{A}'s K63 amine nitrogen and UBE2V1 were placed at specific coordinated, which enabled tetrahedral conformation with UB^{D}'s carbon atom of G76 carbonyl carbon.

With the modified protocol, 10000 models were created containing thioester bond and transition-state-like geometry. First, model was selected by looking at lowest-RMDS comparing UB^{D} position to the closed conformation found in 4AUQ structure. Second, RMSD for the UBE2V1 position in 5AIT UBE2V1/UB^{A} structure was used. Next, the lowest RMSDs for UBE2N/UB^{D} and UBE2V1/UB^{A} were combined. Finally, UBE2V1 protein from the new complex was subjected to the Rosetta Relax protocol, where all the other chains were fixed to get rid of any side chain clashes after structure merge.

Generated complex was further used as a template to make a version of the model with incorporated C5, while keeping acceptor amine position. As such, $^{K63}UB_{C5}$ was created by Rosetta's fixbb application. All atomic coordinates were kept in the same position except K63's side chain. To specify $^{K63}UB_{C5}$ side chain torsion angles, set of

them were taken from free $^{K63}UB_{C5}$ simulations. Best side chain rotamer was selected based on an overlap in nitrogen atom position in the side chain, in comparison to the native UB^{A} C4. Additionally, side chain placement was manually assessed to reduce clashing with surrounding residues. Developed model was used to replace native UB^{A} structure with $^{K63}UB_{C5}$ in the complex.

To perform molecular dynamics simulations, more values must be specified regarding thioester bond. Parameters concerning bond, angle and dihedral entries were accommodated from Amaro *et al.* thioester bond values and inserted into the parameter files[208]. Thioester bond partial charges were adapted from Jones *et al*[86]. and used for PSF generation by formatting them into a patch residue topology entry.

2.12.3 MD simulations

MD simulations were run in the CHARMM36 force field and NAMD software[209,210]. For simulations performed for free UB, the protein was solvated in a rectilinear water box consisting TIP3P water molecules, r=5.0 Å. To generate PSF and solvate the structure, autopsf and solvate plug-ins were used in VMD molecular visualization software. At the time of simulation, system was exposed to optimization stabilization, heated to 298K and subsequently equilibrated (50ns, 2fs time steps). The Particle Mesh Ewald methods, Langevin dynamics and constant pressure control was used to execute the runs. Simulations performed for UBE2N~UB^{D}/UBE2V1/UB^{A} complex were executed similarly. Complex was solvated in a rectilinear water box consisting TIP3P water molecules, r=12.5 Å. System optimization involved heating to 298 K and equilibration (25ns, 2fs time steps). All the other settings remained the same.

2.12.4 MD simulations analysis

MDAnalysis from the Python library was used to assess completed MD simulations[211]. Additionally, Python script was developed in MDAnalysis module for RMSD alignment and calculations to determine structure preservation[212]. Moreover, separate script was written to calculate the gate loop region RMSD in UBE2N~UB^{D}/UBE2V1/UB^{A} complex. To examine free UB residues flexibility, Chi torsion angles were analyzed at the position 11, 48 and 63. Dihedral module, from MDAnalysis package, was used to retrieve the torsion angles from each frame of trajectory files, generated in MD runs.

Other Python scripts were established to assess each torsion angle frequency and extent of its change. To inspect the degrees of freedom, conformational space was divided into bins, as it was done before for libraries containing side chain rotamers[213]. Briefly, chi angles were divided into 120 degrees bins (3 bins in total) and phi and psi angles were divided into 10 degrees bins (36 bins in total). The frequency of each torsion angle movement between bins was measured for free UB and complex trajectories. Same framework was used to catalogue each residues's different rotameric states. C4 versus C5 was always measured from the Cα atom to the side chain nitrogen atom. For the UBE2N~UB^D/UBE2V1/UB^A complex simulations, the distance between $UB^{A'}$s acceptor side chain nitrogen atom and UB^D's G76 carbonyl atom was plotted as a function of simulation time.

Additional Python scripts were developed to calculate the time of different geometries, in the complex simulations, created between $UB^{A'}$s C4/C5 at the position 63 and thioester bond. Relative to the active site acceptor's side chain angle was measured by the torsion angle between UBE2N's Cβ and S and $UB^{A'}$s side chain amine Cα. The angle was recorded at every simulation's frame using a MDAnalysis-utilizing script together with distance between the cysteine and amine. Angles/distances pairs were filtered with 2 types of binning. First combinations were classified by 3Å distance bins (ranged from 2-20Å). Second, 30 degrees bins (total 12) were also used to categorize the combinations.

3. RESULTS

3.1 Role of acceptor lysine

3.1.1 E2~UB discharges to nucleophiles of various free amino acids

In order to examine geometric constraints of the polyubiquitin chain formation reaction and the role of acceptor lysine, first the assays for the simplest activity of an E2 were performed. A discharge assay monitors the transfer of UB^D from E2~UB^D to a free nucleophilic amino acid. For some UB-conjugating enzymes, a strong correlation between preferred residue-type modified in the presence of protein substrate and the free amino acid acceptor like lysine, serine or cysteine can be observed[59,214]. One of the best characterized K63-linked chain forming enzyme is the heterodimeric UBE2N/UBE2V1 E2 complex, where UBE2N partners with catalytically inactive, E2-like UBE2V1[85,100,215].

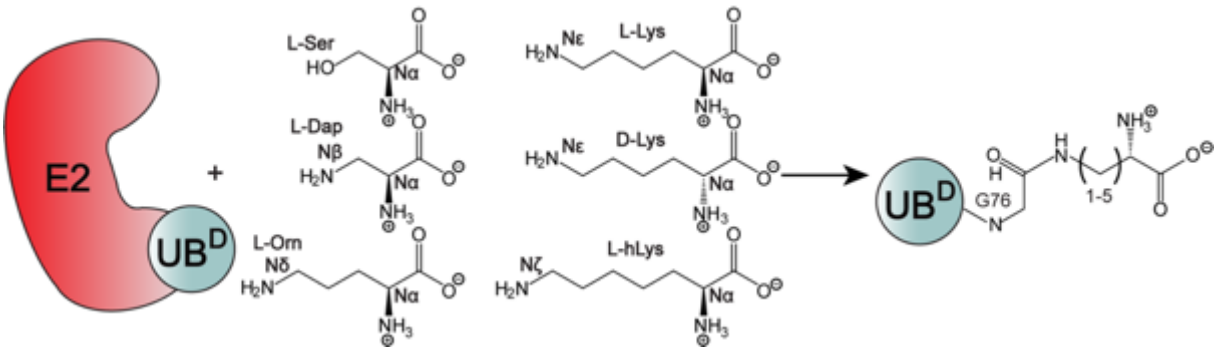


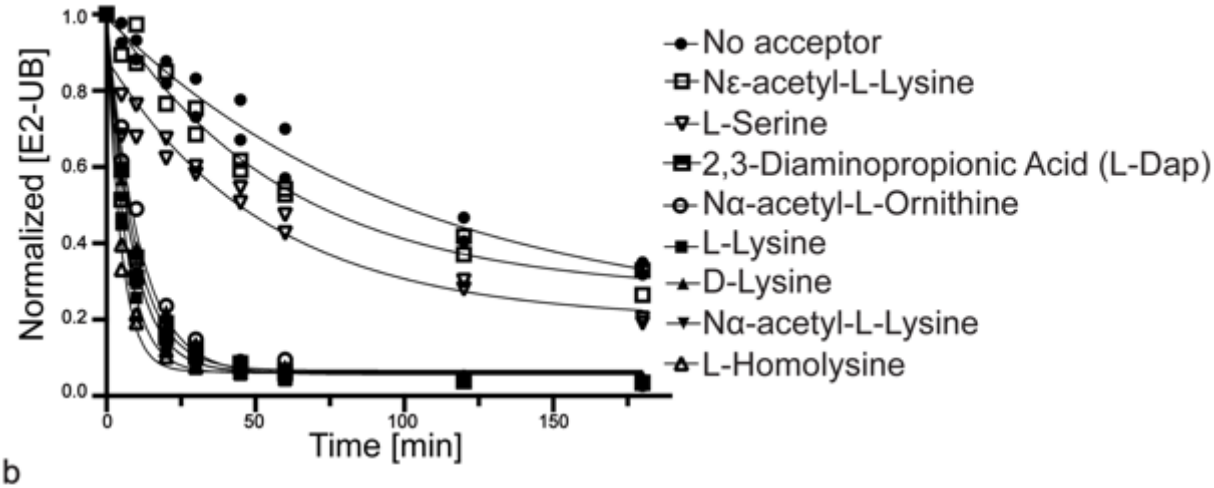
Figure 3.1 Discharge assay reaction scheme. E2~UB reactivity is monitored to various amino acids.

The role of UBE2V1 is to guide acceptor UB's K63 towards the thioester linkage between C-terminus of the donor UB and catalytic cystine of UBE2N[48,215]. The reaction is further catalyzed by the RING domain of RNF4 (here used RNF4 RING dimer) via active conformation stabilization of UBE2N~UB^{D100}. The ability of the UBE2N/UBE2V1-RNF4 complex to forge isopeptide bonds between donor UB and various free amino acid acceptors was examined using a pulse-chase assay format. In the first step of the assay, UBE2N was charged with fluorescently labeled donor UB

K63R using E1. After quenching E1 activity with EDTA, the discharge of UBE2N~UB was observed in the presence of UBE2V1, RNF4 RING dimer and acceptor amino acids. The reaction scheme is depicted in Figure 3.1.

Initially, reactions were carried out with L-lysine (native acceptor, with four methylene groups in the side chain, here called C4), Nα-acetyl-L-lysine with blocked backbone amino group, Nε-acetyl-L-lysine with blocked side chain amino group and L-serine. As expected, discharge occurred only to the native lysine and Nα-acetyl-L-lysine with available side chain amino group (Figure 3.2).

a



b

Acceptor	Rate [min^{-1}]	95% Confidence limit
No acceptor	0.010	0.006-0.014
Nε-acetyl-L-Lysine	0.017	0.013-0.021
L-Serine	0.018	0.010-0.030
L-Dap	0.120	0.106-0.135
Nα-acetyl-L-Ornithine	0.092	0.079-0.107
L-Lysine	0.154	0.138-0.172
D-Lysine	0.106	0.095-0.118
Nα-acetyl-L-Lysine	0.120	0.107-0.134
L-Homolysine	0.210	0.188-0.236

Figure 3.2 UBE2N~UB^{D} discharges to various nucleophilic free amino acids. (a) Time course of UBE2N~UB^{D}* discharge to indicated amino acids. Reactions were run in the presence of UBE2V1 and RNF4 RING. Signal was normalized to the starting UBE2N~UB^{D}* amount. (b) Calculated rates of UBE2N~UB^{D}* discharge.

No UB transfer was observed for Nε-acetyl-L-lysine and L-serine, as these two amino acids do not harbor a nucleophile in the side chain (Figure 3.2). Further, other amino acids were assayed: L-2,3-diaminopropionic acid, Nα-acetyl-L-ornithine and L-homolysine with one (C1), three (C3) and five (C5) methylene groups in the side chain, respectively. All mentioned free amino acids showed native-like reactivity, demonstrating lack of aliphatic chain length specificity (Figure 3.2).

3.1.2 A K63-specific E2 enzyme displays exquisite sensitivity on acceptor lysine architecture

Since removal or addition of methylene groups to the acceptor's side chain had only a very modest effect on the amount of donor UB discharged from UBE2N/UBE2V1-RNF4, we next wondered more generally how L-lysine length within the folded acceptor UB would influence UBE2N/UBE2V1-RNF4 activity.

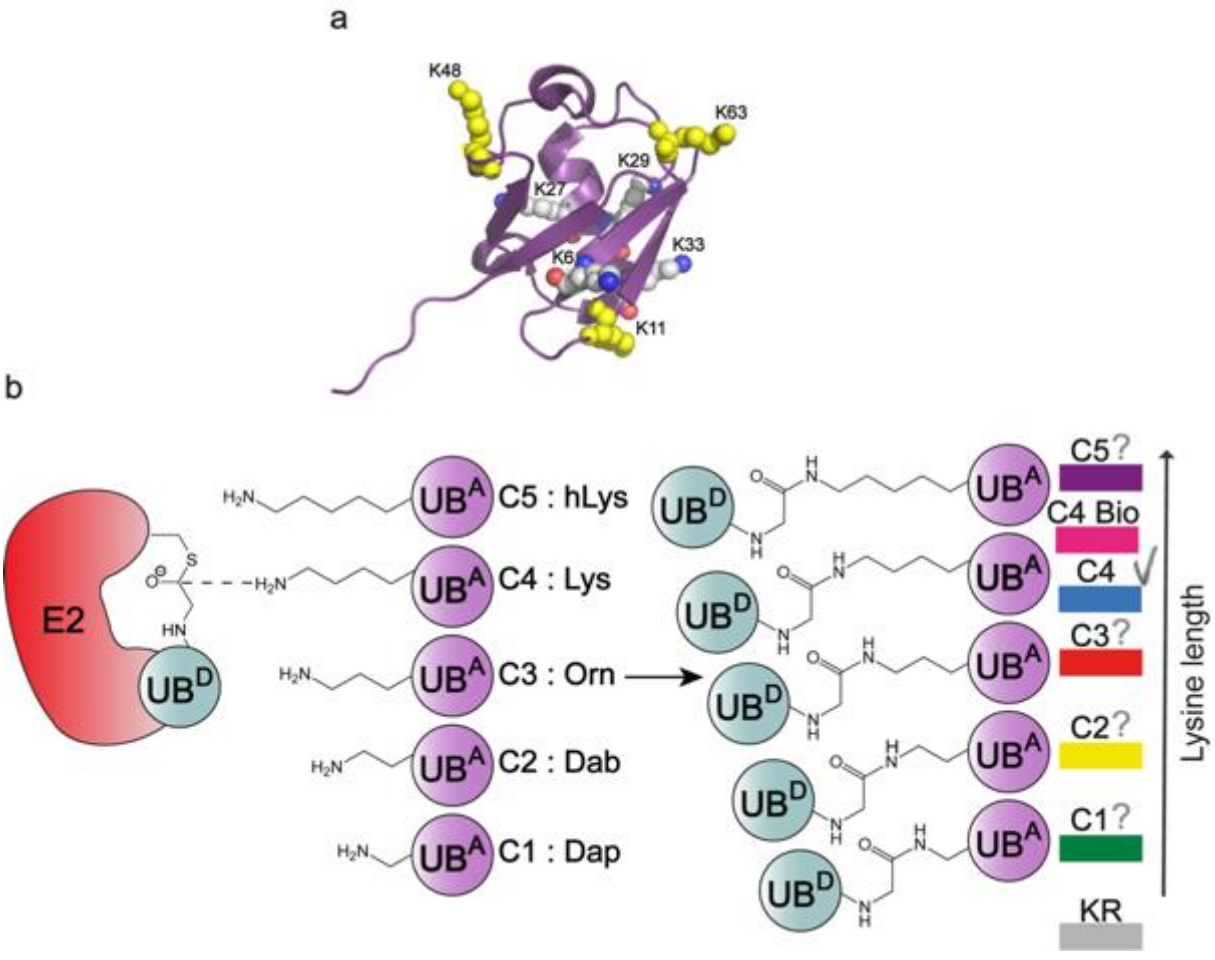


Figure 3.3 Experimental design for UB analogs test. (a) Structure of UB (PDB 3CMM) with all seven lysines depicted. Lysine mutants used in this study are colored in yellow. (b) Experimental scheme of UB analogs test, monitoring E2~UB^D discharges to different acceptors. Color code is used to specific UB mutants (UB^A C1-C5). KR- lysine mutation to arginine.

Solid-phase peptide synthesis was employed to generate a series of UBs with different K63 lysine side chain lengths. UB analogs differed by the number of methylene groups between side chain amino group and alpha carbon: L-2,3-diaminopropionic acid (Dap, here called $^{K63}UB_{C1}$, with one methylene group in the side chain replacing K63), L-2,4-diamonobutyric acid (Dab, here called $^{K63}UB_{C2}$, with two methylene groups in the side chain replacing K63), L-ornithine (Orn, here called $^{K63}UB_{C3}$ for three methylene groups in the side chain replacing K63), L-lysine (Lys, here called $^{K63}UB_{C4}$ for four methylene groups in the side chain, native acceptor), L-homolysine (hLys, here called $^{K63}UB_{C5}$ for five methylene groups in the side chain replacing K63) (Figure 3.3).

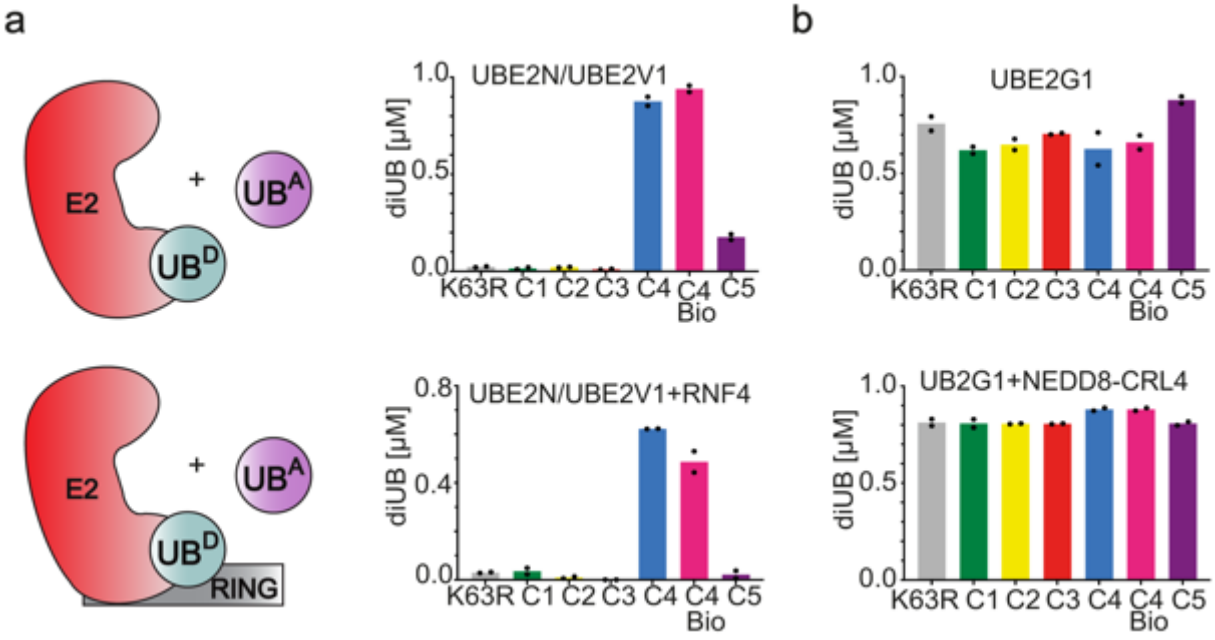


Figure 3.4 A K63-specific UBE2N/UBE2V1 displays striking sensitivity to acceptor lysine architecture. (a) The K63 UB chain forming UBE2N/UBE2V1 strongly prefers the $^{K63}UB_{C4}$ acceptor in the absence (upper panel) and in the presence (lower panel) of the E3 RNF4 RING domain (b) The $^{K63}UB_{C1-5}$ series demonstrates equal reactivity with the K48 UB chain forming E2 UBE2G1 in the absence (upper panel) and in the presence (lower panel) of NEDD8-CRL4 catalytic platform.

The activity of UBE2N/UBE2V1 was again assayed in a pulse-chase format. This time UBE2N~UB^D discharged to different K63 UB analogs and formation of diUB chain was observed. Remarkably, in contrast to the discharge to free amino acids, side chain length change by only one methylene group essentially abolished free UB chain formation. DiUB formation was only observed in the presence of native K63 ($^{K63}UB_{C4}$ or $^{K63}UB_{C4Bio}$ derived from *E.coli*) (Figure 3.4 a). Moreover, while the RING domain of the E3 enzyme RNF4 accelerated the reaction, the striking preference for the lysine chain persisted (Figure 3.4 a). On the other hand, the levels of free diUB formation between the $^{K63}UB_{C1-5}$ analogs and a donor UB in complex with the K48-specific E2 UBE2G1 were all similar, demonstrating the proper folding for all of the synthetic UBs harboring K63 substitutions (Figure 3.4 b). In summary, UBE2N/UBE2V1 showed exquisite specificity for the attacking lysine architecture, but only when in the context of an acceptor UB.

3.1.3 A K48-specific E2 enzymes display exquisite sensitivity on acceptor lysine architecture

Recruitment of acceptor UB by UBE2N is unique, in that it collaborates with its partner subunit UBE2V1 to bind and orient the acceptor[100]. As UBE2V1 specifically binds acceptor UB to position K63 within UBE2N's active site, the system might be especially sensitive to the acceptor length change. Thus, the next set of experiments was carried out using E2s which orient the acceptor UB with their own surface. Hence two K48 linkage-specific E2s, UBE2G1 and UBE2R2, were assayed in the presence of the $^{K48}UB_{C1-5}$ suite[31,51,116]. Interestingly, once again diUB formation was only observed in the presence of native $^{K48}UB_{C4}$, for both intrinsic activity as well as UBE2R2 and UBE2G1 together with their cognate cullin-RING ligase E3s CRL1 and CRL4 respectively[31,51,112,116,191]. (Figure 3.5 a,b)

CRLs are multiprotein complexes, wherein a specific receptor protein recruits a substrate degron motif. A well-studied CRL1 receptor is FBW7, a tumor suppressor protein that recruits the phosphopeptide from various substrates including phosphorylated cell cycle regulator, Cyclin E[216]. Another well-studied CRL4 receptor is CRBN, for which chemotherapeutic thalidomide analogs such as pomalidomide act

as molecular glues to neo substrates including zinc finger domain transcription factors in the Ikaros family[188,189].

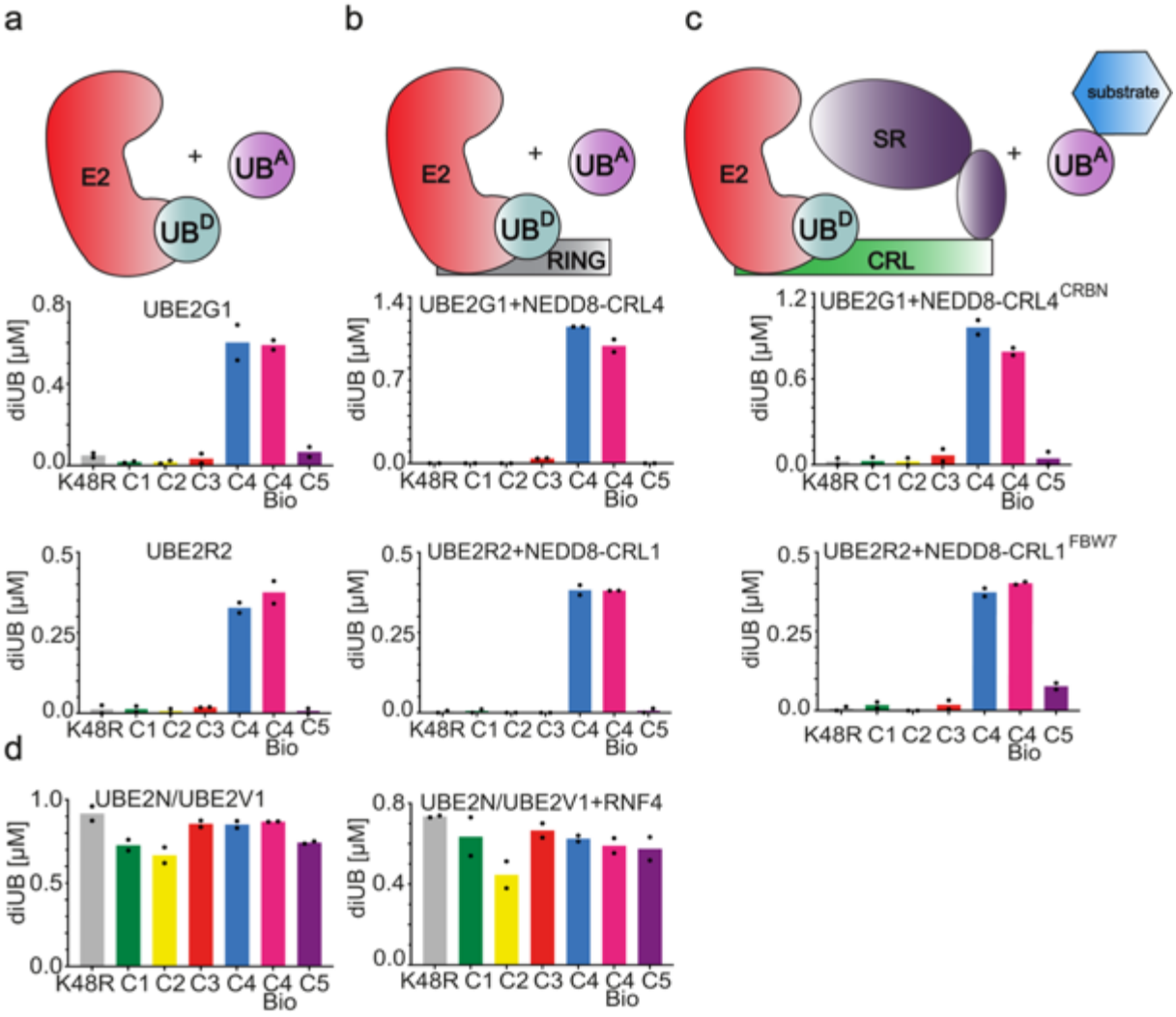


Figure 3.5 K48-specific UBE2G1 and UBE2R2 displays striking sensitivity to acceptor lysine architecture. (a) The K48 UB chain forming UBE2G1 and UBE2R2 enzymes strongly prefer the $^{K48}UB_{C4}$ acceptor in the absence of the E3 (upper panel). The $^{K48}UB_{C1\text{-}5}$ series demonstrates equal reactivity with the K63 UB chain forming E2 UBE2N/UBE2V1 in the absence of the E3 (lower panel) (b) The K48 UB chain forming UBE2G1 and UBE2R2 enzymes strongly prefer the $^{K48}UB_{C4}$ acceptor in the presence of NEDD8-CRL4 and NEDD8-CRL1 respectively (upper panel). The $^{K48}UB_{C1\text{-}5}$ series demonstrates equal reactivity with the K63 UB chain forming E2 UBE2N/UBE2V1 in the presence of the RNF4 E3 (lower panel). (c) The K48 UB chain forming UBE2G1 and UBE2R2 enzymes strongly prefer the native acceptor, in the presence of $^{K48}UB_{C1\text{-}5}$ analogs were sortased to the substrate peptides: IKZF1 ZF23 and phospho-cyclineE respectively.

One drawback to the diUB synthesis assay is that the reactivity of the acceptor UB may differ substantially when attached to an E3-bound protein substrate[188,189]. To address this, sortase-mediated transpeptidation was employed to append both CRL1- and CRL4-based substrates (a cyclin E phosphopeptide and IKZF1 zinc fingers 23, respectively) to a UB from the $^{K48}UB_{C1-5}$ suite and reacted in the presence of E3s and UBE2R2 or UBE2G1, respectively[112,191]. Even in this context, only in the presence of a native lysine considerable UB-chain elongation was observed (Figure 3.5 c).

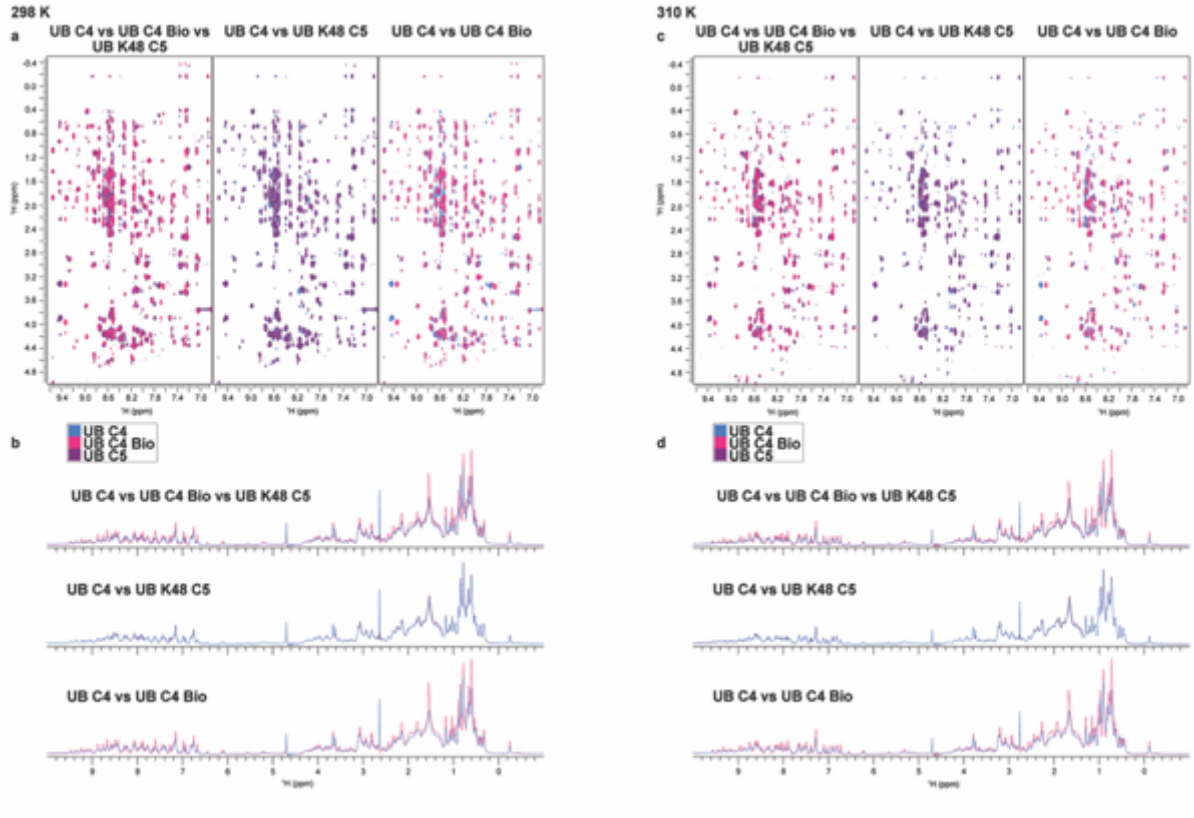


Figure 3.6 UB_{C4}'s recombinant UB ($UB_{C4\ Bio}$) and $^{K48}UB_{C5}$ 1D and 2D proton NMR spectra are greatly superimposable. (a) 2D Nuclear Overhauser effect spectroscopy (NOESY) spectra recorded for UB_{C4} (blue), $UB_{C4\ Bio}$ (pink) and $^{K48}UB_{C5}$ (purple) at 298K. (b) 1D spectra recorded for UB_{C4} (blue), $UB_{C4\ Bio}$ (pink) and $^{K48}UB_{C5}$ (purple) at 298K. (c) 2D Nuclear Overhauser effect spectroscopy (NOESY) spectra recorded for UB_{C4} (blue), $UB_{C4\ Bio}$ (pink) and $^{K48}UB_{C5}$ (purple) at 310K. (d) 1D spectra recorded for UB_{C4} (blue), $UB_{C4\ Bio}$ (pink) and $^{K48}UB_{C5}$ (purple) at 310K. The 1D spectra show signals from amide protons (6 – 10 ppm), Cα-protons (3.5 – 6 ppm), and methyl protons (-0.5–1.0 ppm). The 2D NOESY spectra show NOE interactions between amino acid side chain protons (y-axis) and amide protons (x-axis).

Moreover, as a control, $^{K48}UB_{C1-5}$ suite was assayed with K63 specific UBE2N/UBE2V1 complex, producing nearly equivalent amounts of free diUB chains with UB mutants, with or without RNF4 E3 (Figure 3.5 a,b). As an additional control for proper synthetic UBs folding, proton nuclear magnetic resonance spectroscopy (NMR) was carried out. The measurement was run for $UB_{C4\ Bio}$ (recombinant UB derived from *E.coli*), synthetic UB_{C4} and $^{K48}UB_{C5}$. Overall, the spectra were superimposable and showed good dispersion. Few resonances variability, presumably were caused by sequence differences between three UBs (Met1 in $UB_{C4\ Bio}$ was replaced by NorLeu in synthetic UBs, longer K48 side chain in $^{K48}UB_{C5}$) (Figure 3.6). In summary, the synthetic UBs show wild type-like folding and behavior in the assays. K63- and K48-specific E2 enzymes employing different modes of acceptor UB recruitment are nevertheless highly specific for lysine architecture.

3.1.4 A HECT E3 enzymes display exquisite sensitivity on acceptor lysine architecture

Next, we examined if this remarkable preference for the native lysine also occurs in the context of HECT E3 ligase family. These E3 ligases are known to transfer UB from E2 to HECT domain's catalytic cysteine and then to the substrate lysine[69]. The NEDD4-1 ligase, the eponymous member of the NEDD4 family of HECT ligases, and its yeast ortholog Rsp5p[75] were assayed. Both E3s are well-known to form K63-linked chains[74,102]. Nevertheless, efficient diUB formation was only found for $^{K63}UB_{C4}$ acceptor, while no detectable amount of diUB was formed with $^{K63}UB_{C1-3}$ or $^{K63}UB_{C5}$ (Figure 3.7 a). Additionally, Rsp5p's substrate degron peptide of Sna4p was linked to $^{K63}UB_{C5}$ in a sortase-mediated transpeptidation reaction. Substrate targeting was only observed in the presence of native lysine (Figure 3.7 c). Similar to the other reactions assayed here, discharge is equally maintained when NEDD4-1and Rsp5p were assayed in the presence of the $^{K48}UB_{C1-5}$ analogs (Figure 3.7 b). In summary, K48- and K63 specific E2s and E3s showed striking specificity for lysine architecture within an acceptor UB. The phenomenon was discovered, not only for different modes of acceptor UB recruitment but for distinct active site architectures as well (E2~UB^D versus HECT E3~UB^D).

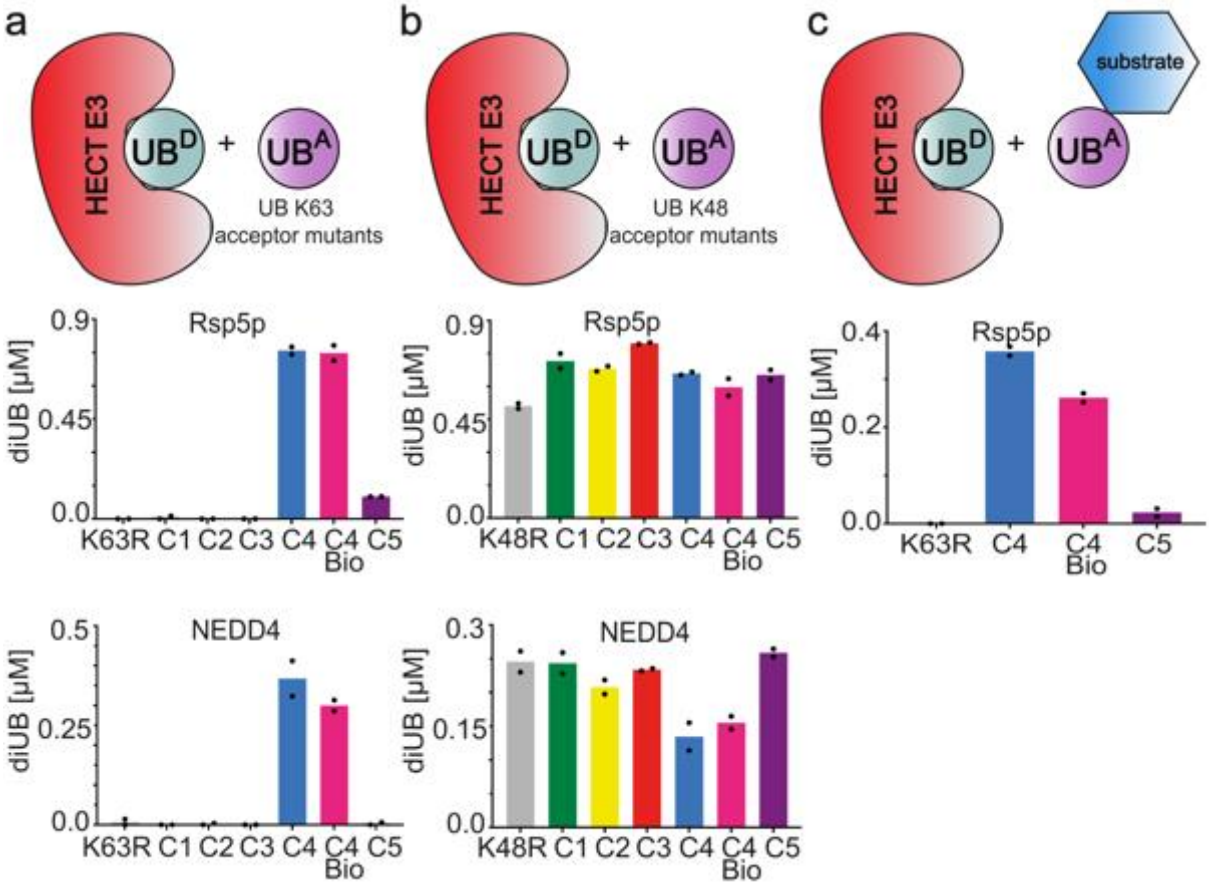


Figure 3.7 K63-specific Rsp5p and NEDD4 HECT E3 ligases display striking sensitivity to acceptor lysine architecture. (a) The K63 UB chain forming Rsp5p and NEDD4 HECT E3 ligases strongly prefer the $^{K63}UB_{C4}$ acceptor. (b) The $^{K48}UB_{C1\text{-}5}$ series demonstrates equal reactivity with the K63 UB chain forming Rsp5p and NEDD4 HECT E3 ligases (c) The K63 UB chain forming Rsp5p strongly prefers native acceptor geometry, in the presence of $^{K63}UB_{C14\text{-}5}$ analogs sortased to Sna4p substrate peptide.

3.1.5 K11-specific enzyme UBE2S displays lack of preference for acceptor lysine architecture

Given the results so far, we wondered, whether different linkage-specific enzymes might also rely on native lysine architecture. The E2 UBE2S, which forms K11 diUB chains, seemed like an interesting candidate, as it is thought that UBE2S harbors an incomplete active site which is complemented by several acceptor UB side chains[101]. UB-assisted catalysis of UBE2S was not easy to assay, as this E2 poorly synthetizes diUB chains on its own. This was overcome by fusion of UBE2S's catalytic domain to

the UB-binding domain: IsoT (UBE2S_IsoT)[101]. Intriguingly, UBE2S_IsoT did not show any

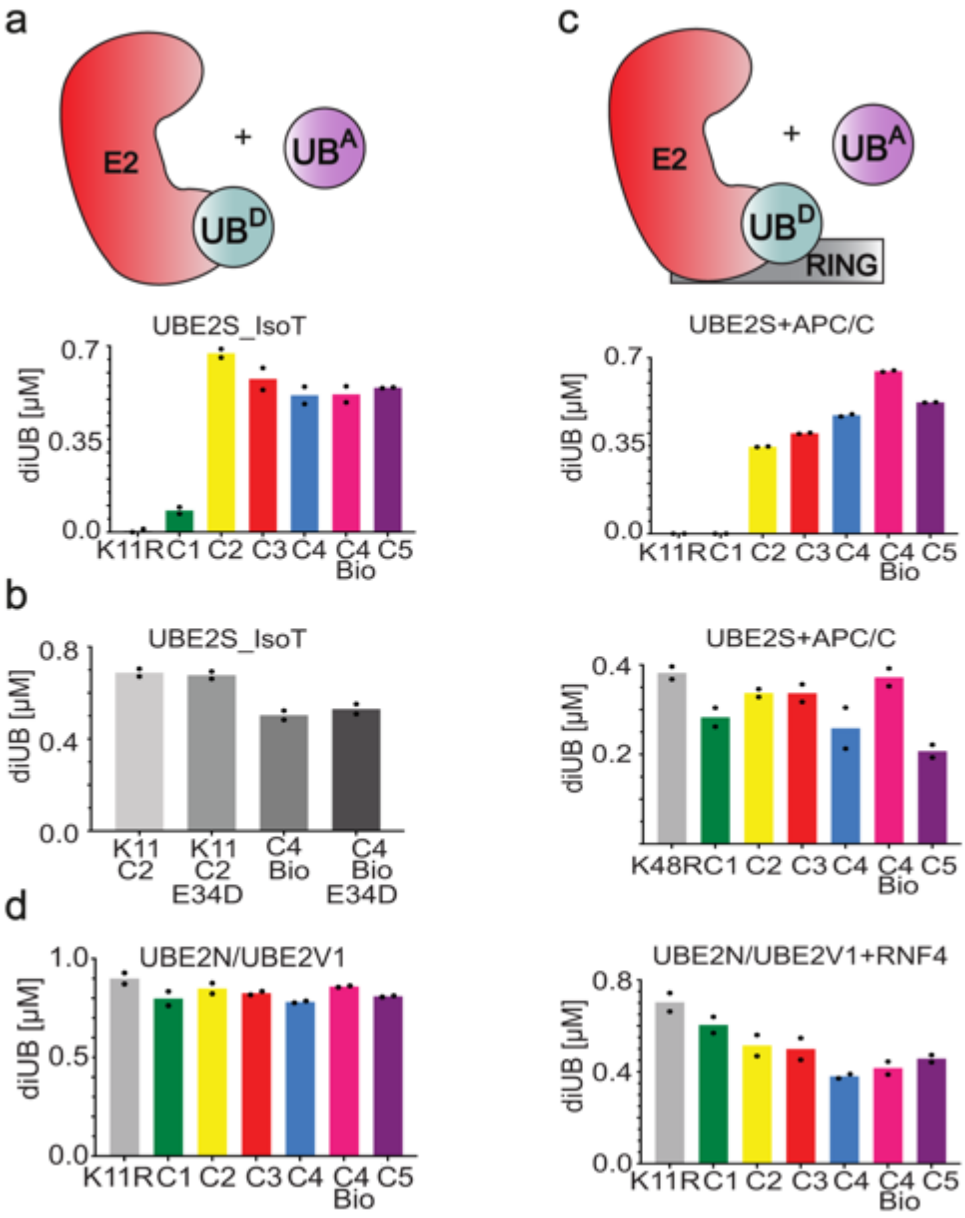


Figure 3.8 K11-specific UBE2S displays lack of preference for acceptor lysine geometry. (a) The K11 UB chain forming UBE2S_IsoT displays equal reactivity towards $^{K11}UB_{C2-C5}$ acceptors. (b) E34D $^{K11}UB_{C2}$ mutation does not perturb diUB formation catalyzed by UBE2S_IsoT. (c) The K11 UB chain forming UBE2S displays equal reactivity towards $^{K11}UB_{C2-C5}$ acceptors in the presence of APC (upper panel). The $^{K48}UB_{C1-5}$ series demonstrates equal reactivity with the K11 UB chain forming UBE2S in the presence of APC (lower panel). (d) The $^{K11}UB_{C1-5}$ series demonstrates equal reactivity with the K63 UB chain forming UBE2N/UBE2V1 in the absence (left panel) and presence (right panel) of RNF4 RING E3 ligase.

specificity for acceptor side chain length, when tested with $^{K11}UB_{C2-5}$ analogs (Figure 3.8 a). It was previously shown that E34 of the acceptor UB plays a role in UBE2S's diUB catalysis, via K11 orientation and pK_a suppression[101]. Thus, the E34D mutation was designed to test, whether shortening of the "guiding" amino acid would influence $^{K11}UB_{C2}$'s ability to act as an acceptor. All tested mutants did not show any significant defects, suggesting that UBE2S mode of action strongly relies on more than one amino acid and can not be simply examined (Figure 3.8 b). Further, $^{K11}UB_{C1-5}$ suite was tested in the presence of the APC/C E3, which stimulates UBE2S activity via RING domain recruitment of acceptor UB[37,217]. Similarly, no strong preference was observed for any acceptor lysine side chain length (Figure 3.8 c). Detected activities were not due to inappropriate UBs folding, as $^{K11}UB_{C1-5}$ analogs were fully active as acceptors in K63-dependent UBE2N/UBE2V1 reactions (Figure 3.8 d).

3.1.6 Promiscuous UBE2D3 activity is impacted by K48 side chain

Motivated by the observation that UBE2S is less sensitive to lysine architecture, we next sought to characterize the ubiquitin chain forming enzyme UBE2D3 (aka UBCH5C). UBE2D3 activity is particularly perplexing, as it collaborates with a large swath of E3s, can transfer UB to numerous protein substrates as well as to active site cysteines in HECT- and RBR-family E3s, and generates a range of branched UB chain linkage types involving UB residues such as K11, K48 and K63[20,111,218-220]. First, absolute quantitation by Mass Spectrometry (MS) was used to determine diUB products formed by UBE2D3 in a pulse-chase assay format. Under the tested conditions, K63 and K11 linkages were made preferentially (Figure 3.9 a). Next, the same assay was performed in the presence of $^{K11}UB_{C5}$, $^{K48}UB_{C5}$, or $^{K63}UB_{C5}$ acceptors. We were intrigued by the observation that when a $^{K48}UB_{C5}$ acceptor is added to UBE2D3~UB^{D}, the mobility of the diUB products differed on an SDS-PAGE gel, in comparison to the wild type UB. Because ubiquitin's high stability, coupled with varying conformations and compactness of different diUB chains could cause variable migrations on SDS-PAGE, the result hinted at the formation of several different linkage types amongst the diUB products (Figure 3.9 b).

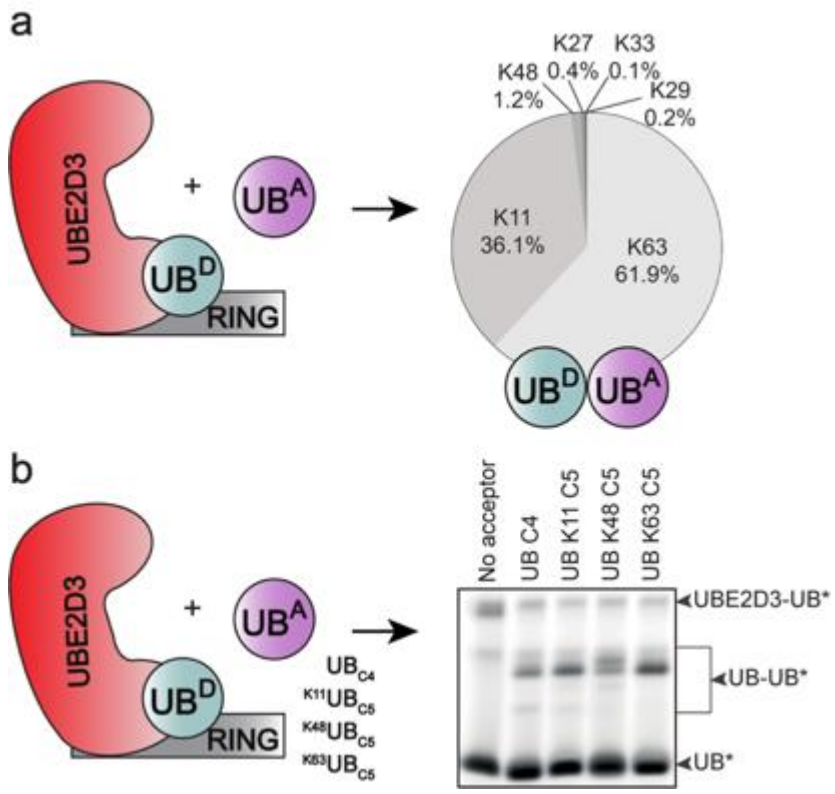


Figure 3.9 UBE2D3 produces different diUB linkages in the presence of $^{K48}UB_{C5}$ analog. (a) Distribution of diUB linkages produced by UBE2D3 with wild type UB^{A}. (b) Fluorescent gel scan after the UBE2D3 reaction with various UB_{C5} analogs. Different diUB gel mobility can be noticed in the presence of $^{K48}UB_{C5}$, indicating chain redistribution

To compare these reaction products, a targeted mass spectrometry strategy was developed to quantify the distribution of UB chain linkage types in the presence of $^{K11}UB_{C5}$, $^{K48}UB_{C5}$, or $^{K63}UB_{C5}$ compared to the distribution of UB chain linkage types produced in the presence of wild type acceptor UB. Addition of $^{K11}UB_{C5}$ and $^{K63}UB_{C5}$ did not change diUB's linkage-type in comparison to UB_{C4} (Figure 3.10 a-c). Remarkably, the increase of K48 by one methylene group drastiacally altered, which residue preferentially accepts UB in the course of chain building. First, the redistribution of K63 and K11 linkages was observed. Second, there was an increase in formation of non-preferred linkage types: K27, K29 and K33 (Figure 3.10 c,d). Thus, the promiscuous ubiquitylating enzyme depends on the location of the UB_{C4} acceptor lysines.

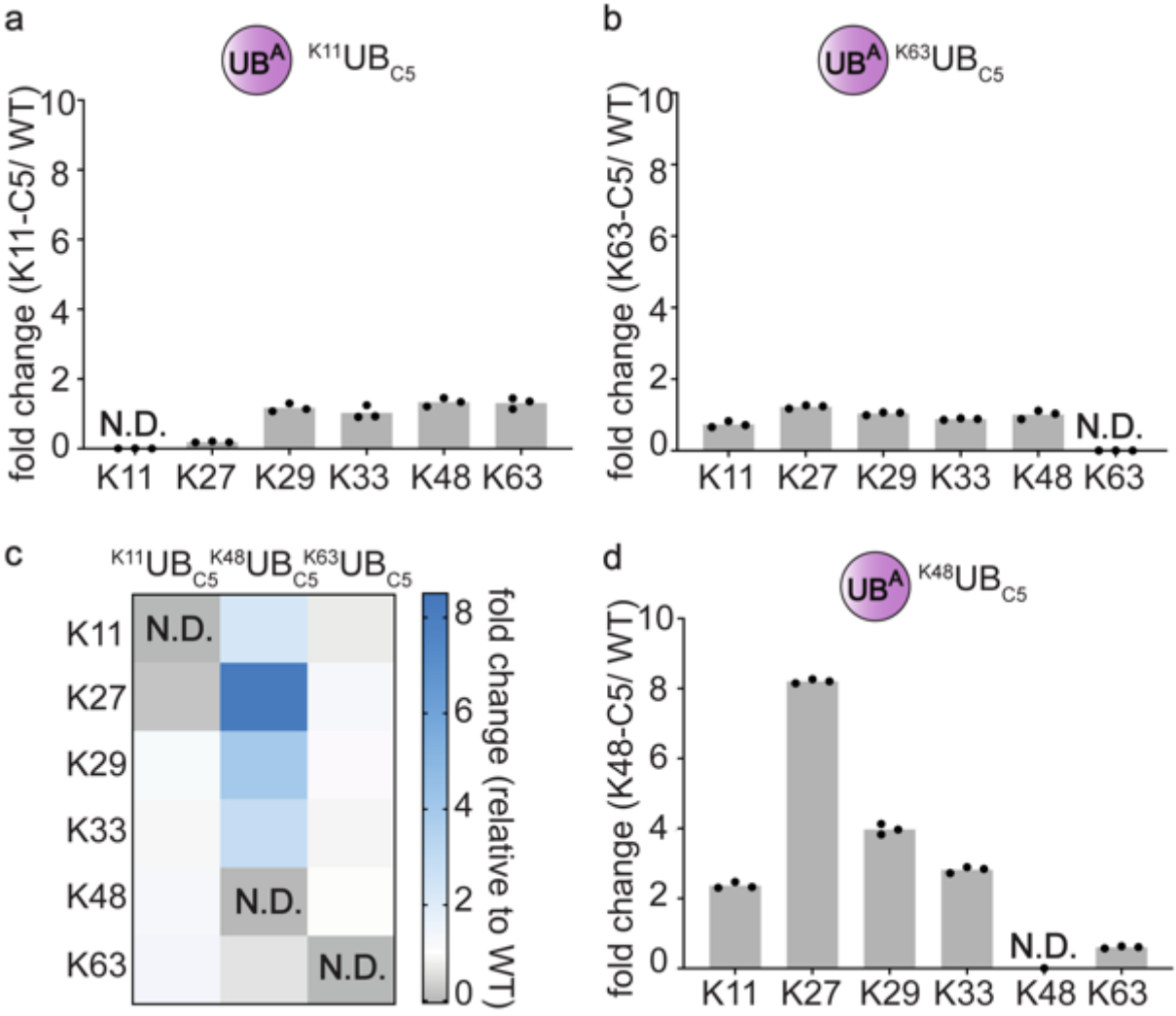


Figure 3.10 The location of the UB lysine mutation determines the distribution of diUB linkage types produced by UBE2D3. (a) The bar graph comparing chain types produced with $^{K11}UB_{C5}$ to wild type UB as an acceptor. (b) The bar graph comparing chain types produced with $^{K63}UB_{C5}$ to wild type UB as an acceptor. (c) Heatmap showing relative fold-change of diUB linkage type produced (rows) by UBE2D3/RNF4 when assayed with C5 acceptors (columns) compared to reactions with a UB_{C4}. (d) The bar graph comparing chain types produced with $^{K48}UB_{C5}$ to wild type UB as an acceptor. N.D- not detected.

3.1.7 Molecular dynamics reveal impact of side chain architecture

Molecular dynamics (MD) simulations were run to reveal how addition of a hydrocarbon to the acceptor side chain at the position 11, 48, and 63 may influence UB structure. Two separate, 50-ns long runs showed that the overall UB globular fold (amino acids 1-70) was maintained with 1.618, 1.271, 1.209 and 1.494 Å average Cα Root Mean Square Deviation (RMSD) for native UB_{C4}, $^{K11}UB_{C5}$, $^{K48}UB_{C5}$ and $^{K63}UB_{C5}$, respectively.

However, other relative differences for all UB_{C5}s were observed. Firstly, the range of potential distances between side chain amine for C_5 and α carbon was

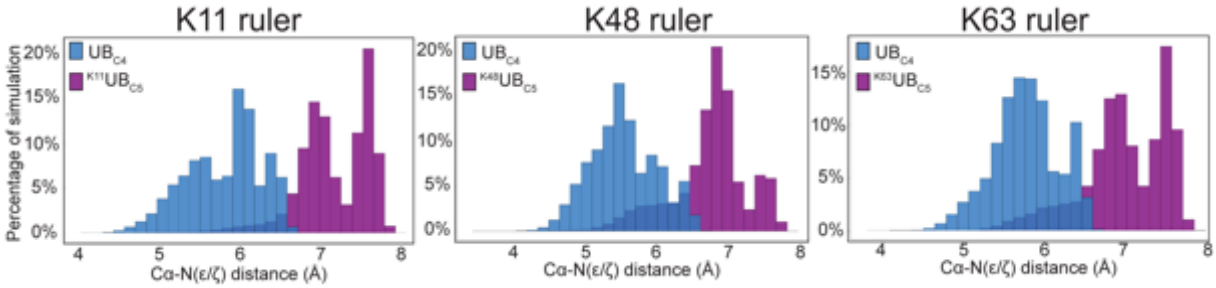


Figure 3.11 C5 acceptor side chain displays increased range of distances from α carbon.

increased (Figure 3.11). Thus, this would mean the greater radius to the C_5 side chain when the backbone would be considered the axis of rotation. Secondly, the potential number of rotamers was expanded from 81 (C_4) to 273 for C_5 side chain (C_4: C_5 rotamer ratios were 56:82, 65:117 and 43:96 at the positions 11, 48 and 63, respectively). Lastly, the dynamics of χ angles changed, especially χ4, which fluctuated more frequently, for C_5, between the three rotamer bins.

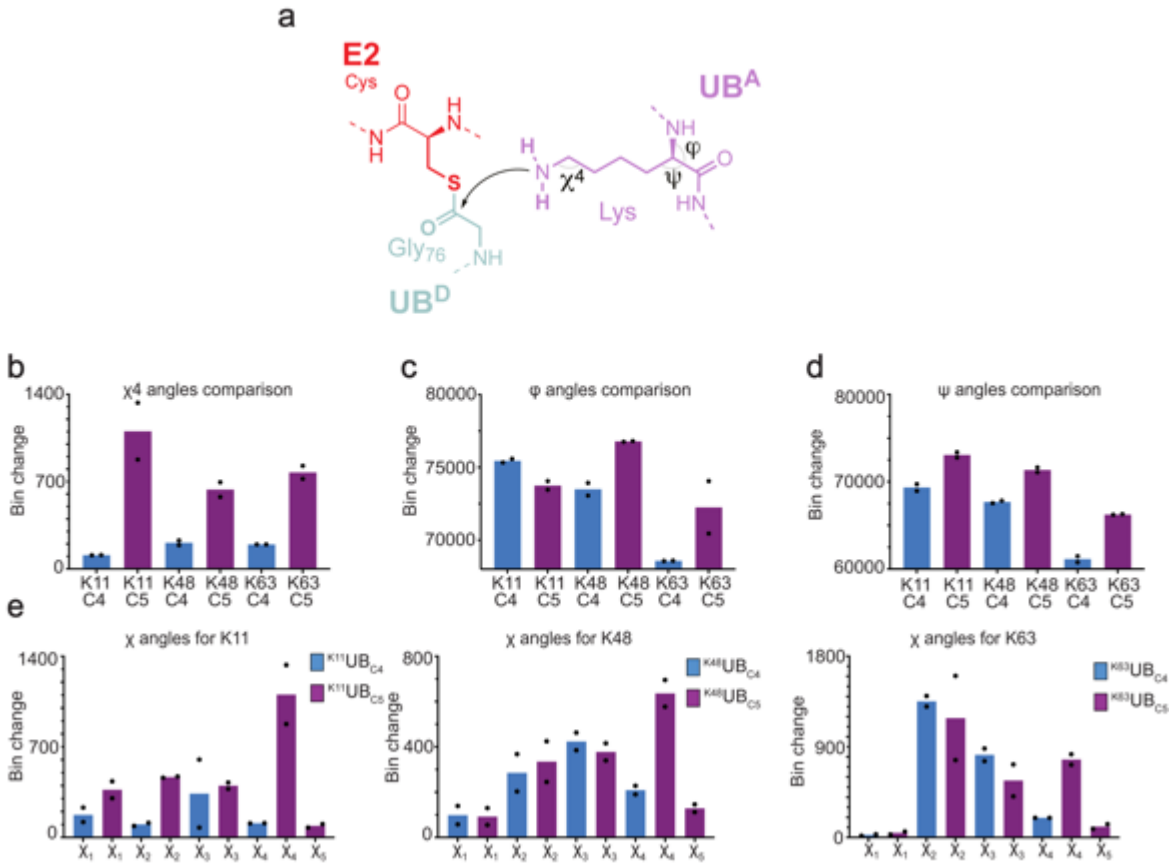


Figure 3.12 Molecular dynamics reveal various structural perturbations for UB_{C5} analogs. (a) Schematic representation of nucleophilic attack of acceptor lysine on E2~UB^D with χ4, φ and ψ angles depicted. (b) Bar graph showing χ4 angles dynamics for UB_{C5} analogs in comparison to UB_{C4}. (c) Bar graph showing φ angles dynamics for UB_{C5} analogs in comparison to UB_{C4}. (d) Bar graph showing ψ angles dynamics for UB_{C5} analogs in comparison to UB_{C4}. (e) Bar graphs showing rotamer interconversions for various side chains in UB_{C5} analogs.

Hence, this would result in more rapid fluctuation of side chain amine (Figure 3.12 b). Additionally, the φ and ψ angles fluctuated more for $^{K48}UB_{C5}$ and $^{K63}UB_{C5}$ (Figure 3.12 c,d), increasing number of allowable φ/ψ combinations (C_4: C_5 ratios were 185:175, 138:169 and 73:90 for residues 11, 48 and 63, respectively). Taken together, more states were accessible for C_5 (backbone and rotamer combinations), with over 1000 additional states for $^{K48}UB_{C5}$ and $^{K63}UB_{C5}$ (C_4: C_5 ratios were 2,942:3,016, 2,942:4,261 and 1,188:2,561 at the residues 11, 48 and 63, respectively) (Figure 3.12 e).

Further, additional simulations were performed to unveil a potential effect of longer lysine side chain in diUB formation. UBE2N~UB^D/UBE2V1/UB^A complex is one of the structurally characterized complexes, where acceptor lysine, in this case K63, points towards the thioester-linked donor UB, but it is 12.5 Å away[53,100]. Nonetheless, the acceptor K63 intermediate was adapted, based on enzymology constraints, previous crystal structures and modeling[221,222]. Similarly, to the simulations with UB alone, three independent, 25 ns long MD simulations demonstrated that C_5 acceptor favorably adopts extended conformations. Furthermore, rotamers were fluctuating more frequently (Figure 3.13 b,c). Even though side chain's amines of C_4 and C_5 maintained a comparable distance to the active site of UBE2N~UB^D (Figure 3.13 a), two important differences could be noticed between the simulations. Firstly, for the longer time of simulation, the native lysine employed more favorable trajectory to approach the active site. In contrast, the rotamers of C_5 were changing more often and the active site was accessed from different angles (Figure 3.13 d,e). Secondly, the active site gate loop of UBE2N (residues 115-120) deviated greatly in the presence of longer acceptor. Intriguingly, the gate loop stabilizes interactions between UBE2N and donor UB's C-terminus, placing the acceptor lysine in the proper position for nucleophilic attack and ensuring proper configuration of catalytic residues[221,223]. Disruption of the active site gate loop conformation, as found for C_5, possibly could favor non-catalytic configurations (Figure 3.13 f,g).

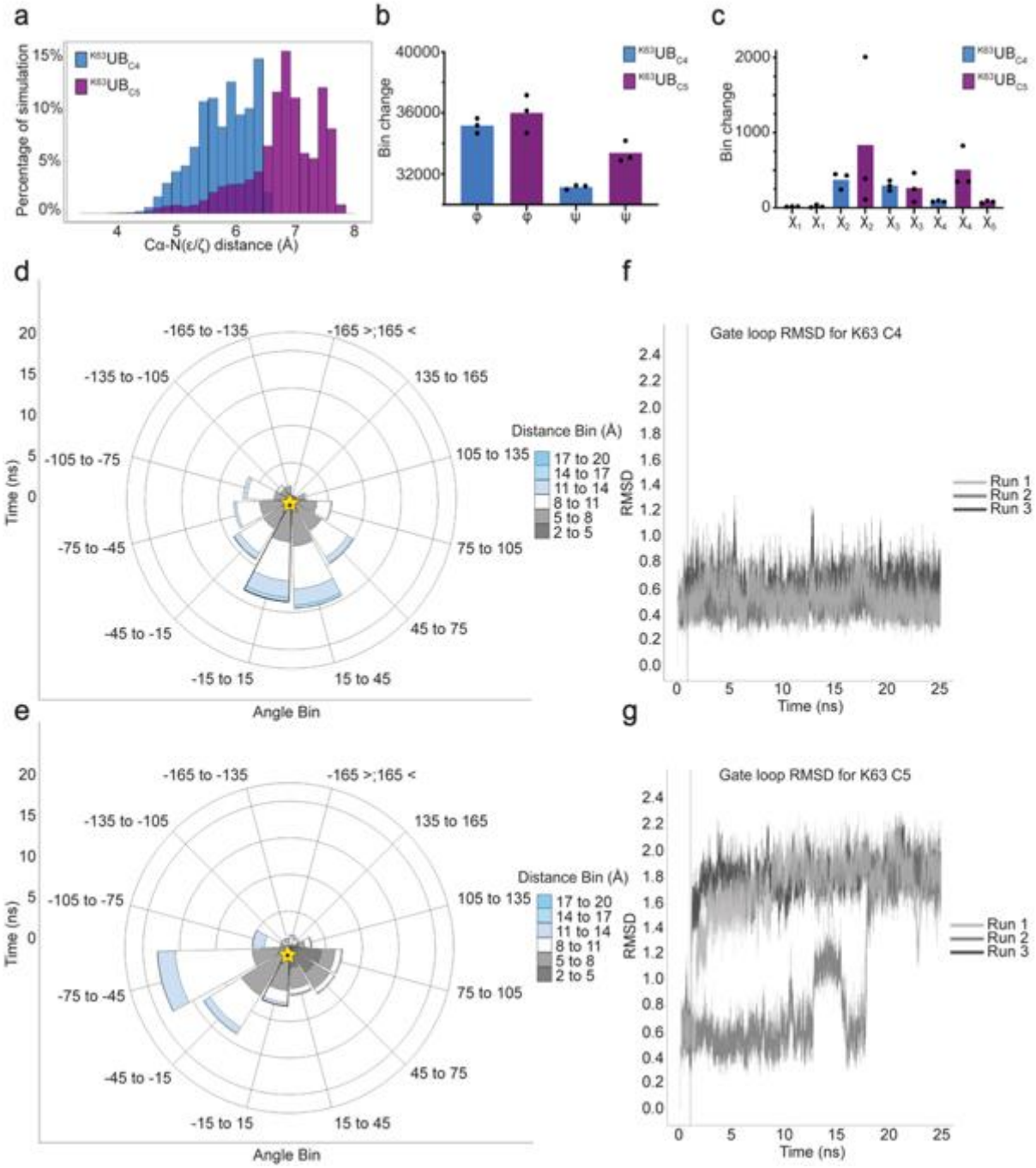


Figure 3.13 Molecular dynamics reveal various structural perturbations for UB_{C5} analogs in the presence of the UBE2N~UB^{D}/UBE2V1/UB^{A} complex. (a) Distance distribution for $^{K63}UB_{C5}$ and UB_{C4} from acceptor lysine amine to α carbon. (b) Bar graph showing φ and ψ angles dynamics for $^{K63}UB_{C5}$ in comparison to UB_{C4} within UBE2N~UB^{D}/UBE2V1/UB^{A} complex (c) Bar graph showing rotamer interconversions for $^{K63}UB_{C5}$ in comparison to UB_{C4} within UBE2N~UB^{D}/UBE2V1/UB^{A} complex. (d) Rose plot showing angle and distance of UB_{C4}'s acceptor lysine side chain amine relative to UBE2N's active site. (e) Rose plot showing angle and distance of $^{K63}UB_{C5}$'s acceptor lysine side chain amine relative to the UBE2N's active site.

(f) RMSD of UBE2N's gate loop, when run with UB_{C4}. (g) RMSD of UBE2N's gate loop, when run with $^{K63}UB_{C5}$.

3.1.8 Kinetic parameters are impacted by UB acceptor lysine

To provide mechanistic insight into a preference for UB_{C4}, quantitative kinetic experiments were carried out. These required conditions, in which it was possible to obtain quantifiable amounts of product for UB_{C5}. Considerable increase of reaction times as well as protein concentrations allowed to detect diUB formation by the E2s UBE2R2, UBE2N/UBE2V1 (with or without E3 RNF4) and by the HECT E3 Rsp5p.

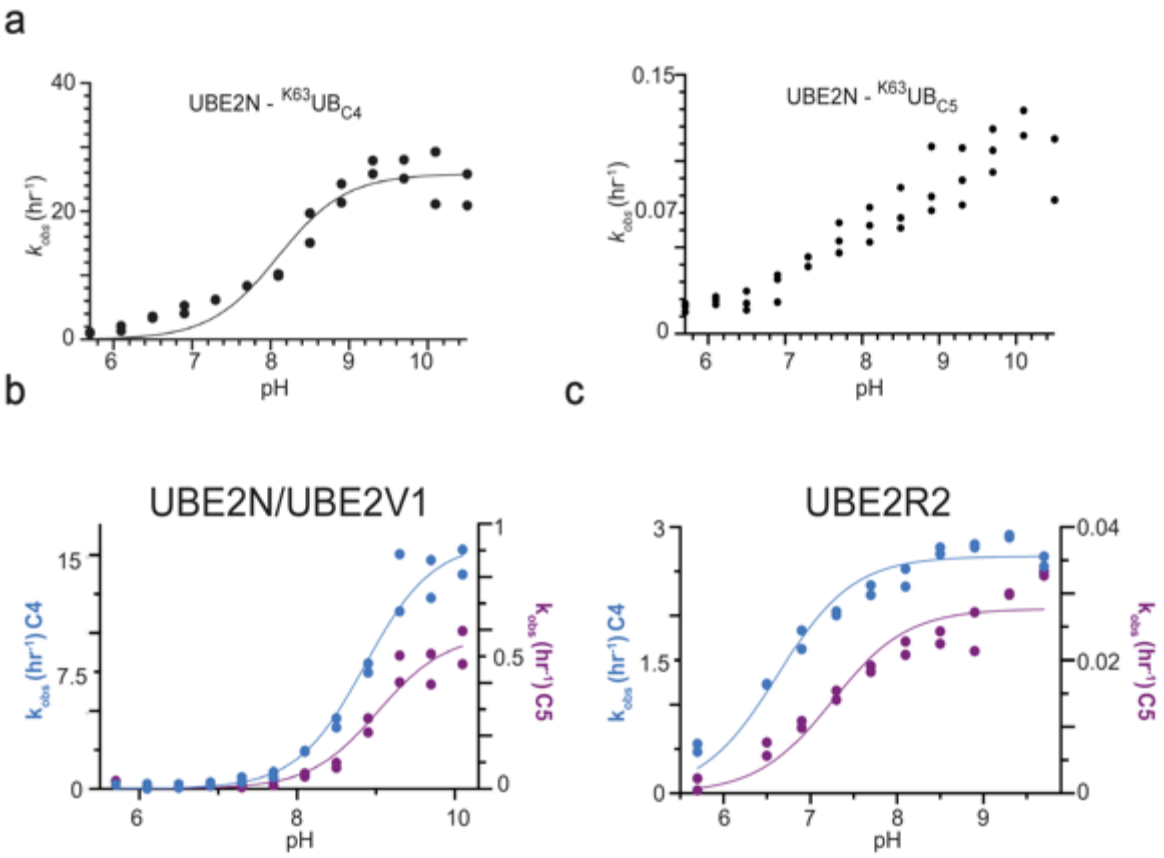


Figure 3.14 UB_{C5} analogs impact the rate of UBE2N/UBE2V1's and UBE2R2's catalysis but not the *pK*$_a$. (a) Reaction velocities, performed in the presence of wild type UBE2N/UBE2V1, plotted as a function of pH. Either UB_{C4} (left panel) or $^{K63}UB_{C5}$ (right panel) were used as acceptor. (b) Reaction velocities for UB_{C4} and $^{K63}UB_{C5}$, performed in the presence of UBE2N K92R/UBE2V1, plotted as a function of pH. (c) Reaction velocities for UB_{C4} and $^{K48}UB_{C5}$, performed in the presence of UBE2R2, plotted as a function of pH.

The experiments with both E2s demonstrated similar profiles and the results were consistent with the striking effects from the pulse-chase assays (Figure 3.14 b,c). Specifically, the values of k_{cat} for UB_{C5}s were lower by 16-fold and 14-fold for UBE2N/UBE2V1 and UBE2R2, respectively (Table 3.1; Figure 3.14 b,c). While a variety of perturbations can affect enzyme activity, inability to activate attacking acceptor lysine amine or weakening of the binding affinity of the acceptor UB for the E2 are quite frequent[48,83,101]. A cutting-edge study on UBL SUMO proposed that pK_a suppression of the acceptor lysine is, at least in part, achieved through the active site complementation[83]. More specifically, several E2 active site residues considerably suppress the substrate lysine pK_a through stabilizing interactions that compensate for the energetic penalty of desolvation[83,85]. In spite of being unable to measure apparent pK_a for E3-dependent reactions, as a result of loss of enzymatic activity at high pH, apparent pK_a values were determined for E2-catalyzed reactions.

E2/E3	UB	Lys	pK_a^{app}	k_{obs} (hr^{-1}) top pH	K_M (10^{-6} M)	k_{cat} (hr^{-1})
UBE2N/V1	$C4_{bio}$	K63			190	6.1
UBE2N/V1	C4	K63	8.9	15.8	398	3.4
UBE2N/V1	C5	K63	9.0	0.58	284	0.21
UBE2N/V1 + RNF4	C4	K63			23	39.1
UBE2N/V1 + RNF4	C5	K63			58	9.3
UBE2R2	C4	K48	6.6	2.67	528	15.8
UBE2R2	C5	K48	7.3	0.028	1940	1.1
Rsp5p	C4	K63			21	1.11
Rsp5p	C5	K63			335	0.44

Table 3.1 Estimates of k_{cat}, K_M, and pK_a^{app} for various E2s and E3s, assayed with UB_{C4}, $^{K48}UB_{C4}$ and $^{K63}UB_{C4}$.

As such, UBE2N/UBE2V1 activity (where UBE2N harbors a K93R mutation to reduce auto-ubiquitylation at higher pH)[100], was assayed with either UB_{C4} or $^{K63}UB_{C5}$ under different pH conditions (Table 3.1; Figure 3.14 b). As both, k_{cat} and K_M values may

exhibit pH-dependent effects alone, the apparent pK_a values should be interpreted with caution. However, the model where a single ionizing residue is responsible for the k_{cat} dependency on pH, fits the best to the data (Figure 3.14 b).

Yet surprisingly, both UB_{C4} and $^{K63}UB_{C5}$ share similar apparent pK_a values (8.9 and 9.0, respectively; Table 3.1; Figure 3.14 b). Alongside experiments with UBE2R2, again demonstrated similar apparent pK_a values (6.6 and 7.3 for UB_{C4} and $^{K48}UB_{C5}$ respectively Table 3.1; Figure 3.14 c). For both, UBE2N/UBE2V1 and UBE2R2, apparent pK_a values differences were not sufficient to explain the change of diUB formation rate while using UB_{C5} in comparison to UB_{C4} at elevated pH (almost 100-fold at pH 9.7 for UBE2R2, Table 3.1). Further, the K_M values were estimated for UB_{C4} and $^{K63}UB_{C5}$ in the presence of UBE2N/UBE2V1 and UB_{C4} and $^{K48}UB_{C5}$ for UBE2R2. The observed two-fold and four-fold binding defects for $UB_{C5,}$ respectively, suggested similar affinities for acceptor UB bearing either native lysine or C_5 (Table 3.1; Figure 3.15 a,c). Taken together, the observed defects in catalysis emerge from other effects of the longer side chain in the acceptor UB.

An E3 stimulating E2 activity, may influence the mechanism responsible for acceptor UB lysine specificity. RNF4 RING greatly impacted diUB formation catalyzed by UBE2N/UBE2V1(K_M for acceptor UB was lowered and k_{cat} was increased by 17-fold and 11-fold for $^{K63}UB_{C4}$ and $^{K63}UB_{C5}$ respectively (Table 3.1, Figure 3.15 b). However, only mild effects were found for k_{cat} (~4 fold) and K_M (~2.5 fold) while carrying out the assay with $^{K63}UB_{C5}$ (Figure 3.15 b). Thus, the detected effects were not greater than those without E3.

Conversely, a kinetic measurement that was done for the HECT E3 Rsp5p highlighted a significantly lower K_M (16-fold) for the UB_{C4} in comparison to $^{K63}UB_{C5}$, while k_{cat} was mildly affected (~2.5-fold) (Table 3.1, Figure 3.15 d). Collectively, a diverse spectrum of outcomes of the acceptor lysine side chain on different E2s and E3s was observed in the kinetic results.

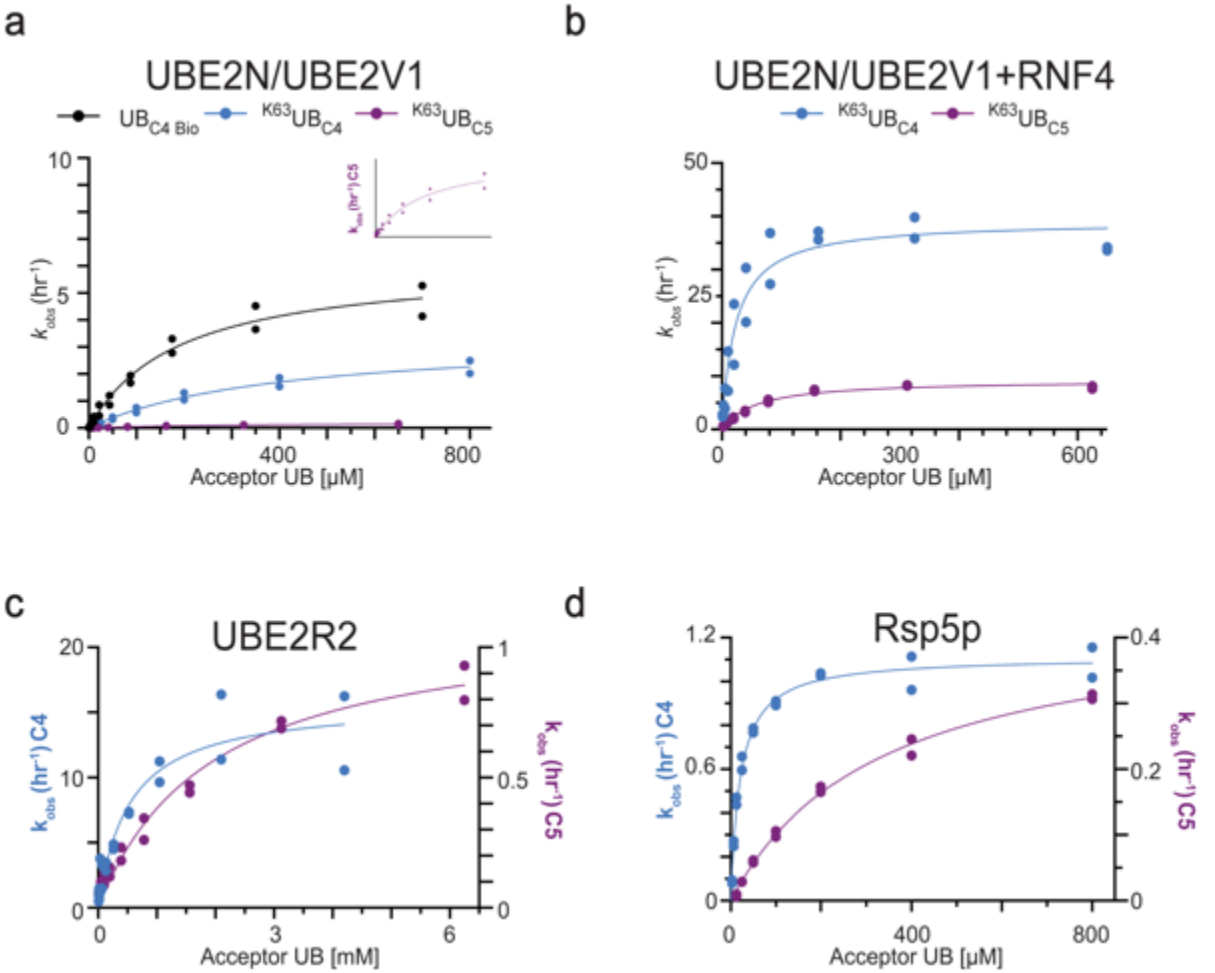


Figure 3.15 UB_{C5} analogs differently impact K_M for UBE2N/UBE2V1, UBE2N/UBE2V1+RNF4, UBE2R2 and Rsp5p (a) Reaction velocities, performed in the presence of UBE2N/UBE2V1, plotted as a function of UB^A concentration. (b) Reaction velocities, performed in the presence of UBE2N/UBE2V1+RNF4, plotted as a function of UB^A concentration. (c) Reaction velocities, performed in the presence of UBE2R2, plotted as a function of UB^A concentration. (d) Reaction velocities, performed in the presence of Rsp5p, plotted as a function of UB^A concentration.

3.2 Structural studies of K48-linked UB chain forming complex

3.2.1 UBE2R preference for CRL2

K48-linked UB chains control proteasomal degradation of plenty of substrates, thus playing an important role in maintaining cellular homeostasis. One of the major classes of E3 ligases, CRLs, are responsible for more than 20% of all degradation events[132].

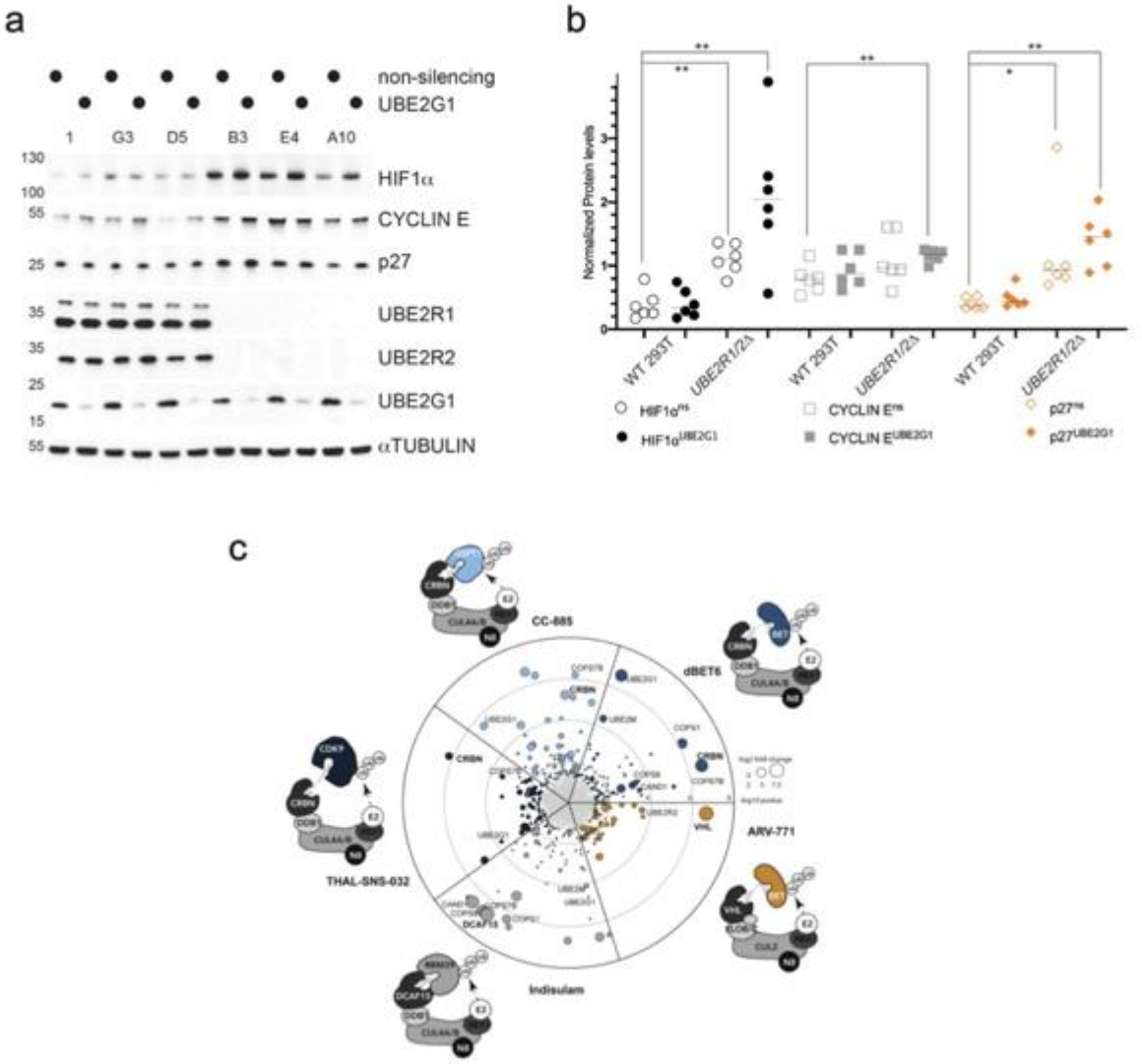


Figure 3.16 UBE2R preference for CRL2. (a) CRL2's Hif1α substrate gets stabilized in UBE2R1/2 double knockout cell lines. Adapted from Hill et al. 2019[112]. (b) Quantitative comparison of CRL2's and CRL1's substrates in HEK 293T and UBE2R1/2 double knockout cell lines. Adapted from Hill et al. 2019[112]. (c) Main genetic determinants of targeted protein degradation identified by CRISPR/Cas9. Adapted from Mayor-Ruiz et al. 2019[192].

CRLs were first discovered, over a quarter century ago, and studied in budding yeast *Saccharomyces cerevisiae*[106,110,134,224]. Initial observations included the importance of different F-box proteins for proteasomal degradation and an essential role of Cdc53 (yeast CUL1) and Cdc34 (yeast ortholog of UBE2R) in this process[134]. To this date most of the studies on mechanisms of CRL-mediated K48-linked chain formation focus on UBE2R in the context of SCF[31,52,81,111-113,116,124,125,225]. Thus, initial structure determination efforts, for this project, were targeting SCF system. In more details, trapped complexes were made, consisting either $CUL1^{FBW7}$ with a CyclinE substrate peptide or $CUL1^{\beta TRCP1}$ with a Iκβα substrate peptide and UBE2R-diUB intermediate. Similar stabilization strategies were used, as described in sections 3.2.2 and 3.2.3. The lack of rigidity of the examined complexes limited the resolution and did not lead to a < 7Å reconstructions.

Interestingly, recent studies have shown that no model SCF substrate (CyclinE, p27, Iκβα and β-Catenin) was stabilized in UBE2R1/2 double knockout HEK 293T cell lines[112]. This surprising result suggested that SCF's substrate degradation must be complemented by another K48-linked chain forming enzyme. UBE2G1, known to work with CRL4, was found to ubiquitinylate SCF substrates in the absence of UBE2R1/2 and buffer their loss in the cells[112,131]. Strikingly, Hif1α, a CRL2 substrate was stabilized in UBE2R1/2 double knockout cell lines (Figure 3.15 a,b)[112].

Another study, investigating genetic determinants of targeted protein degradation, performed a CRISPR/Cas9 knockout screen in the presence of different PROTACs[192]. Consistent with well-established knowledge, CRL4 and CRL2 substrate receptors were identified as well as UBE2G1, which was previously characterized to target CRL4 neomorphic substrates[191,192,226]. Notably, CRL2-dependent neosubstrate degradation was dependent on UBE2R2 (Figure 3.16 c)[192].

Intrigued by these results, suggesting specific UBE2R collaboration with CRL2, site-by-site polyubiquylation assays were performed. Hif1α and the SCF substrate CyclinE were tested in a multiturnover assay in the presence of UBE2R1 and UBE2R2 (Figure 3.17 a). The experiment has shown significantly faster polyubiquitylation of the CRL2 substrate for both versions of UBE2R (Figure 3.17 a). Moreover, a pulse-chase assay was run, where CUL-dependent UBE2R~UB^{D} discharge and formation of diUB* was

observed. Once more, CUL2-stimulated UBE2R chain synthesis was more efficient (Figure 3.17 b). Additionally, kinetics studies in collaboration with Gary Kleiger from the Department of Chemistry and Biochemistry, University of Nevada, Las Vegas, NV, USA identified a substantial 28-fold decrease in UBE2R2 K_M to CRL2 in comparison to CRL1 (data not shown). These experiments, as well as previously published data, constituted a rational to switch to a CRL2-VHL complex to structurally study mechanisms of K48-linked chain formation by UBE2R.

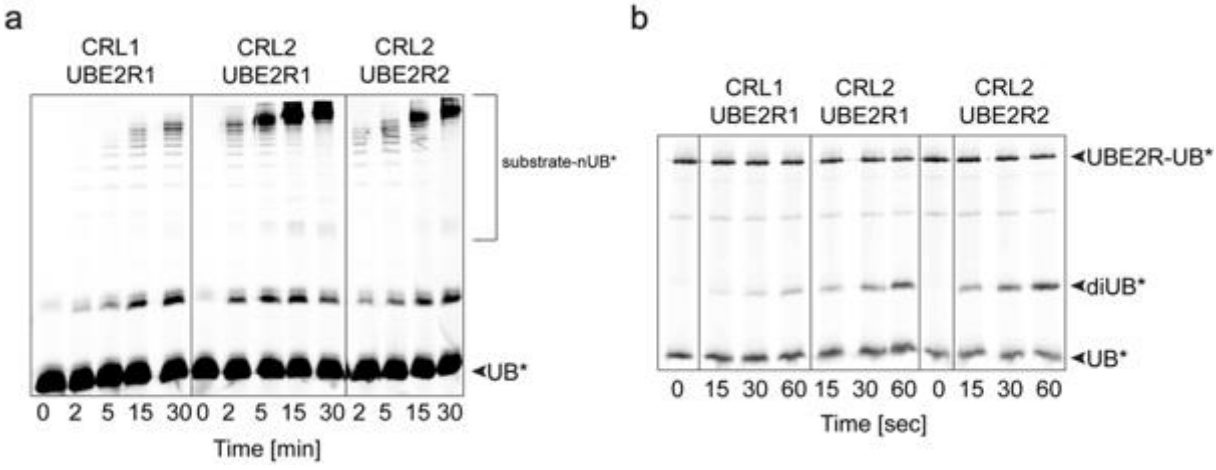


Figure 3.17 UBE2R catalysis is faster in the presence of CRL2. (a) Multiturnover assay comparing UBE2R-catalyzed CyclinE and Hif1α substrates polyubiqutination in the presence of CRL1 and CRL2, respectively. (b) Pulse-chase assay comparing CRL-stimulated diUB formation catalyzed by UBE2R.

3.2.2 Electrophilic probe enables mimicking native tetrahedral intermediate

In order to visualize fast enzymatic reactions, protein complexes have to be trapped in a certain, active conformation. Recent advances in cryo-EM, as well as chemical biology tools allowed to capture E2s and E3s in action. Cysteine chemistry was used to form a disulfide bond between activated substrate-UB^D fusion and the catalytic cysteine of UBE2D3, which allowed to solve the first structure of the whole monoubiquitylating $CRL1^{\beta TRCP1}$ complex[50]. Moreover, $CRL1^{SKP2}$-RBR ARIH1 complex structure was determined explaining the mechanism of UB^D transfer from UBE2L3 to ARIH1 in CRL1-dependent manner and ARIH1-mediated substrate monoubiquitylation[97]. Capturing these transient intermediates was enabled by the use of a UBE2L3-UB^D activity-based probe to trap ARIH1's catalytic cysteine and

dehydroalanine chemistry to mimic the substrate-UB^D-ARIH1 transition state[97]. Stability of the intermediates was ensured by three-way crosslinks between enzymes, UB^D and substrate. A similar approach was taken to peruse structural studies of polyubiquitin chain formation catalyzed by UBE2R1.

To mimic a tetrahedral intermediate between UBE2R1~UB^D and the attacking acceptor K48 of substrate-linked UB^A's (Figure 3.18 a) the activity-based diUB probe strategy was employed, previously described and designed for DUBs[227]. As discovered and shown in the previous result section (3.1) the acceptor lysine geometry is a major determinant of UB code[228], thus the probe had to fully resemble native intermediate (Figure 3.18 b).

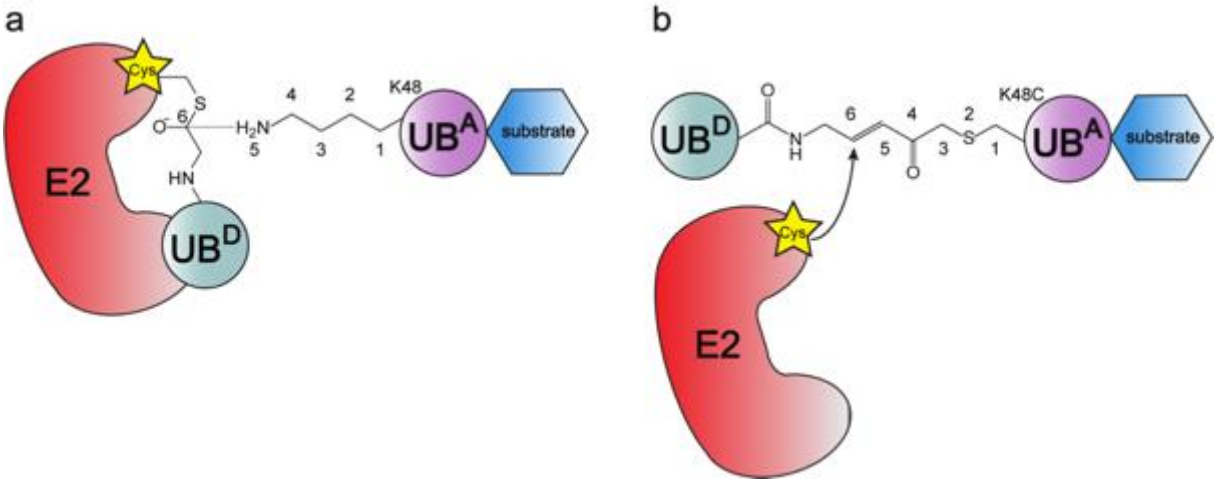


Figure 3.18 Electrophilic probe enables mimicking native tetrahedral intermediate. (a) Cartoon representation of the native tetrahedral intermediate. (b) Cartoon representation of the electrophilic probe used in this study.

3.2.2.1 UBE2R1 single cysteine version generation

The trapping strategy used in this study was based on cysteine chemistry. As such, a UBE2R1 single cysteine version had to be generated. UBE2R1 contains 2 additional, tail-based cysteines: proximal C191 and distal C223, apart from the catalytic one. First, C223A and C191 mutations to smaller hydrophobic residues were tested in a multiturnover assay, to test for obvious catalytic defects (Figure 3.19).

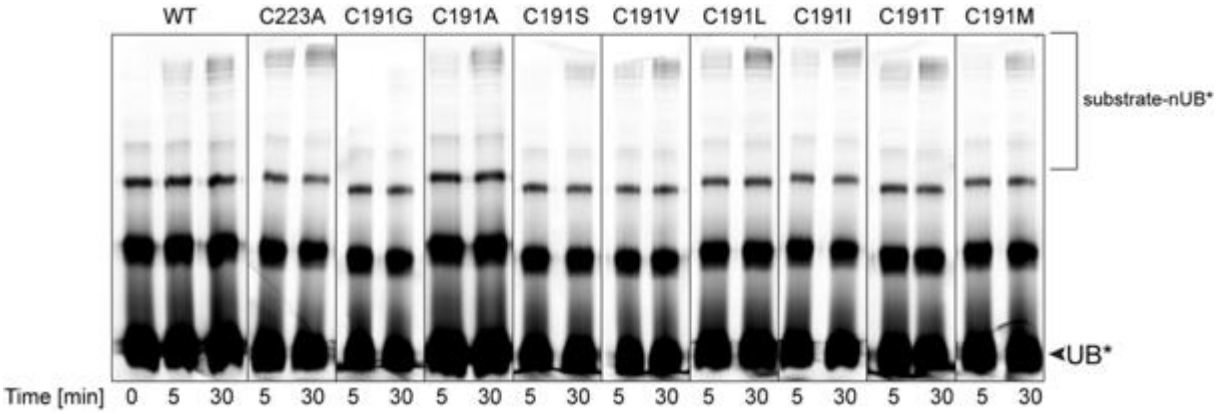


Figure 3.19 Multiturnover assay screen for single cysteine mutants of UBE2R1.

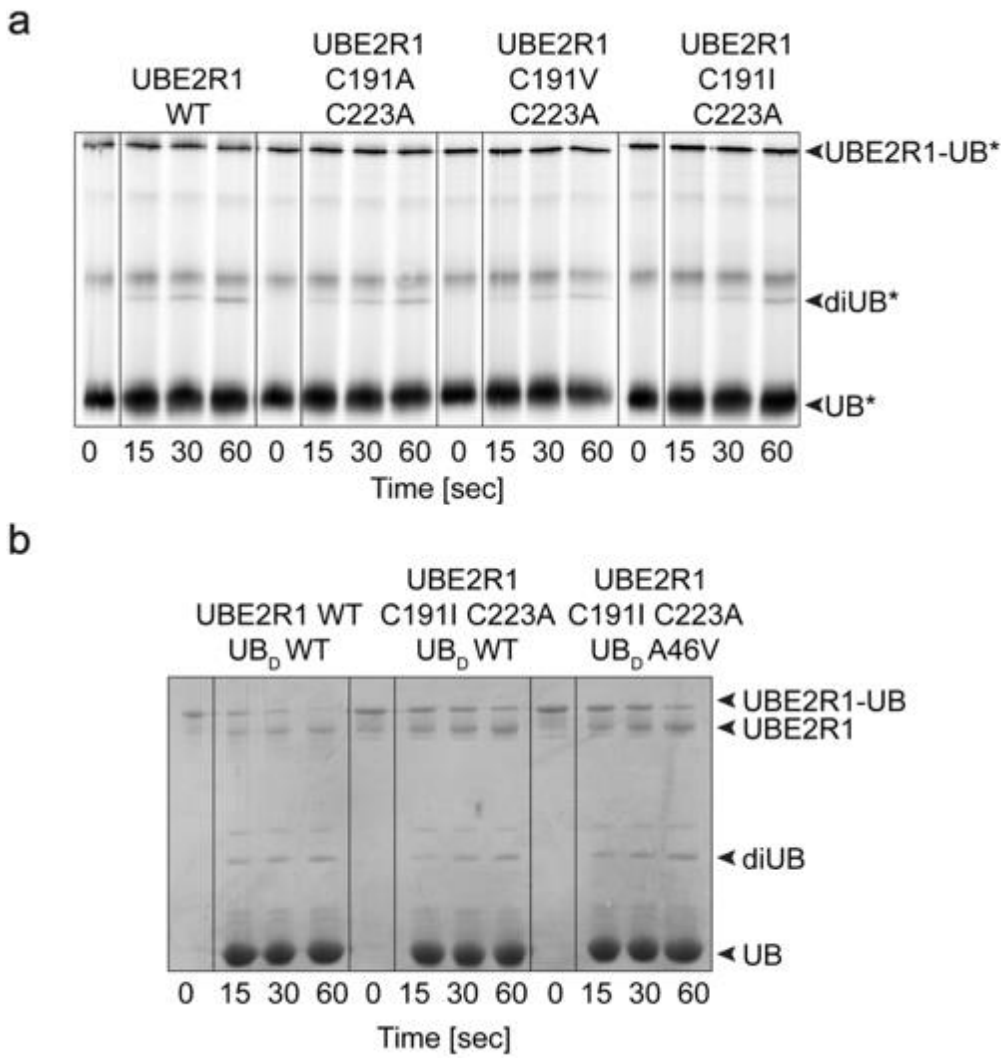


Figure 3.20 (a) Pulse-chase assay screen for double cysteine mutants of UBE2R1. (b) UB^{D} A46V mutation rescues activity of UBE2R1 C191I C223A

An initial screen revealed a couple of cysteine mutant pairs, which were further tested in the more sensitive pulse-chase assay format (C191A C223A; C191V C223A; C191I C223A) (Figure 3.20 a). As the UBE2R1 double cysteine mutant's activity was not fully rescued, additional compensatory mutations were introduced on the UB^D site. A recent crystal structure pointed out the importance of C191 in UB^D A46 binding[115]. As such: A46M, A46V, A46S, A46L, A46I UB^D were made, tested for E1 loading and further assayed with UBE2R1 C191I C223A (data not shown). In summary, the WT-like activity was fully rescued by UB^D's A46V mutation (Figure 3.20 b) and these versions were further used for structural studies.

3.2.2.2 Electrophilic probe validation

First, UB^A K48C was linked in a sortase-mediated transpeptidation reaction to the model substrate peptide, the N-terminus of Hif1α [229]. Second, a diUB-Hif1α probe was made, installing a reactive electrophile between UBs (see Method section for details, Figure 2.1).

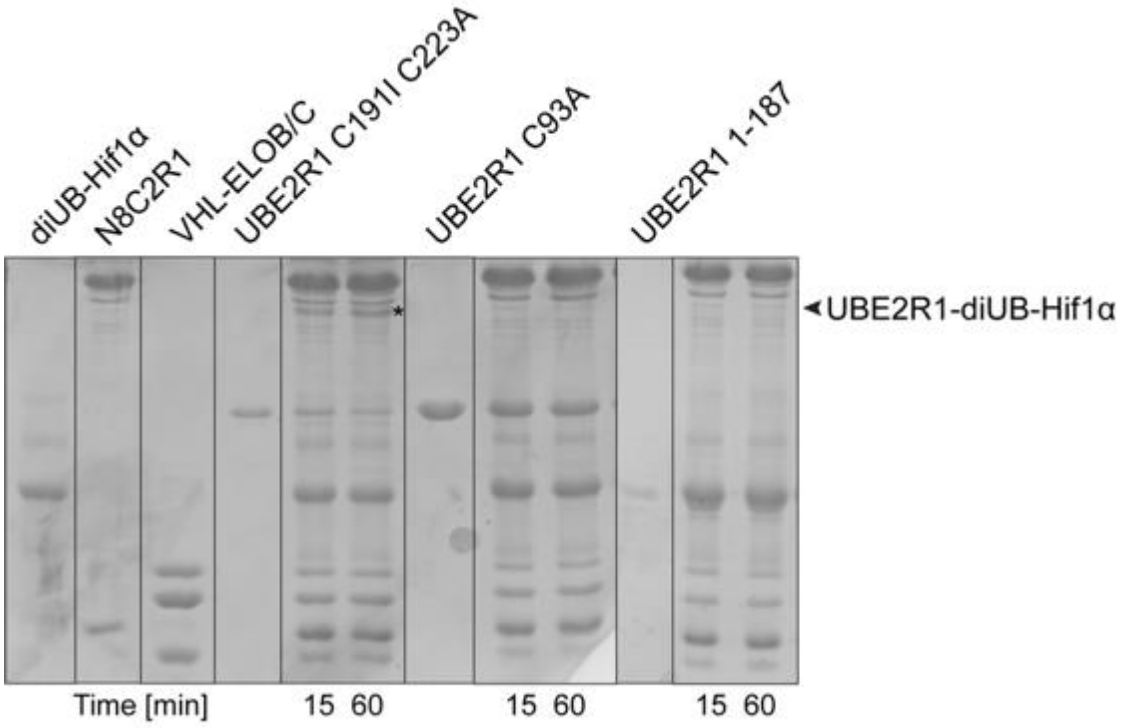


Figure 3.21 The electrophilic probe specifically reacts with catalytic cysteine of UBE2R1. UBE2R1 trapping depends on $CRL2^{VHL}$-UBE2R1 complex assembly.

To validate the probe reactivity, a control experiment was performed. Importantly, the trapping reaction depended on the catalytic cysteine, as UBE2R1 C93A did not show any reactivity (Figure 3.21). Additionally, trapping relied on the whole $CRL2^{VHL}$ complex assembly, as UBE2R1 1-187 (Δtail, deficient in binding to CUL2) failed to react with the probe (Figure 3.21). All in all, validated probes as well as the UBE2R1 C191I C223A UB^{D} A46V mutants were taken further to peruse cryo-EM of UBE2R1-mediated UB chain formation in the presence of $CRL2^{VHL}$.

3.2.3 Structural studies of polyubiquitylation by cryo-EM

In order to ensure maximum complex stabilization for transient K48-linked UB chain formation, apart from the chemical probes, additional strategies were used. First, CC0651 inhibitor[129] was added to stabilize UBE2R1-UB^{D} closed conformation. Second, the complex was crosslinked with glutaraldehyde. Figure 3.22 demonstrates representative gels from independent complex preps before crosslinking and SEC. Relatively higher probe reactivity (Figure 3.22, lane 4), in comparison to other samples, resulted in more stable complex formation, underlining the importance of highest possible trap quality.

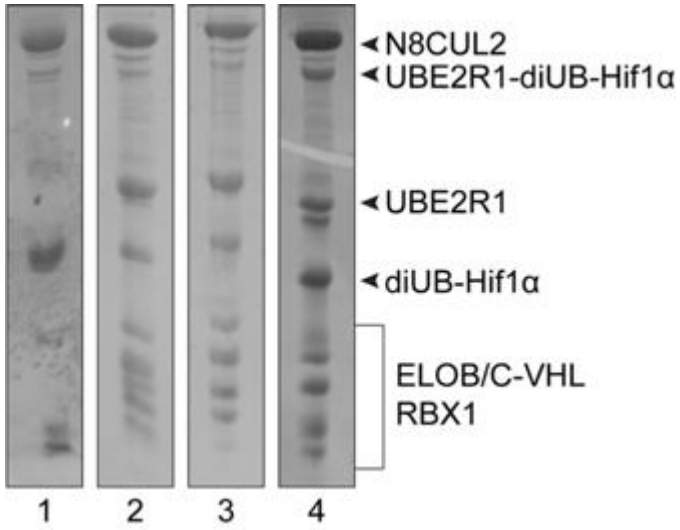


Figure 3.22 Representative gels depicting complex formation reactions used for cryo-EM. Samples were submitted to SEC before plunging.

Obtained Krios dataset, from the sample shown in Figure 3.22, lane 4, led to a map of 4.2 Å resolution, giving an insight into the mechanims of UB chain formation catalyzed by UBE2R1. Briefly, four maps were generated, one global one (map 1) and three focused ones concentrating on substrate receptor VHL-ELOB/C (map 2), CUL2 (map3) and catalytic core with CUL2 CTD (map 4). The processing scheme is depicted in Figure 3.23 and parameters of all maps are included in Figure 3.24. All complex components are illustrated in Figure 3.25a. Following sections discuss important features of the generated model.

3.2.3.1 Cryo-EM data analysis

Following crystal structures were fitted to the obtained density: 1LM8 (ELOB/C-VHL-Hif1α), 4MDK (UBE2R1-UB^D), 5N4W (CUL2-RBX1) and 3CMM (UB^A)[92,129,155,230]. Due to the lack of high resolution data in crucial parts of the complex, especially in the catalytic core, the structure couldn't be build and extensively analyzed on side chain position level. As such, the Namdinator software was used to place crystal structures into the density[231]. A global map with colour coded, fitted proteins is depicted in Figure 3.25 b. Overall, almost all complex components are visible expect CUL2's WHB domain and NEDD8. Crystal structure of substrate receptor VHL and substrate adaptor ELOB/C fits very well into the map and the Hif1α peptide density, bound to VHL, can be observed (Figure 3.26 b)[92]. There is no obvious conformational change. Additionally, CUL2 NTD-ELOC-VHL binding mode seems to be preserved as well[177]. Catalytic core components: UBE2R1, UB^D, UB^A and RBX1 were docked separately, as there is no structure available visualizing UBE2R1-diUB active conformation. The map contains clear density for both UBs, however the crucial loops and side chains, especially for UB^A are not present (Figure 3.26 c). UBE2R1's secondary structures, especially helix1, crossover helix and helix-turn-helix motif can be easily spotted and placed into the density (Figure 3.26 c).

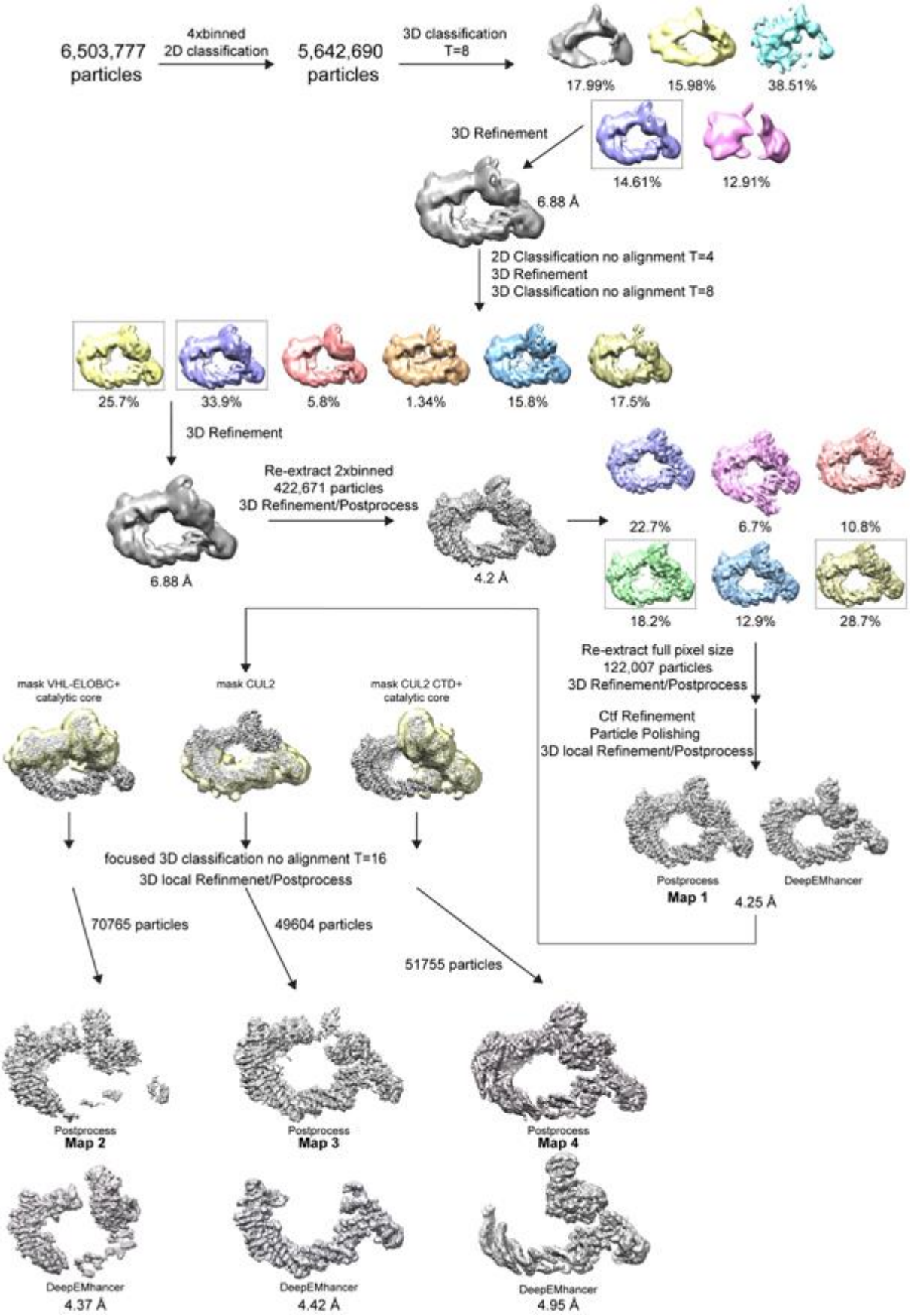


Figure 3.23 Processing workflow of Titan dataset, done in RELION 4.0[195].

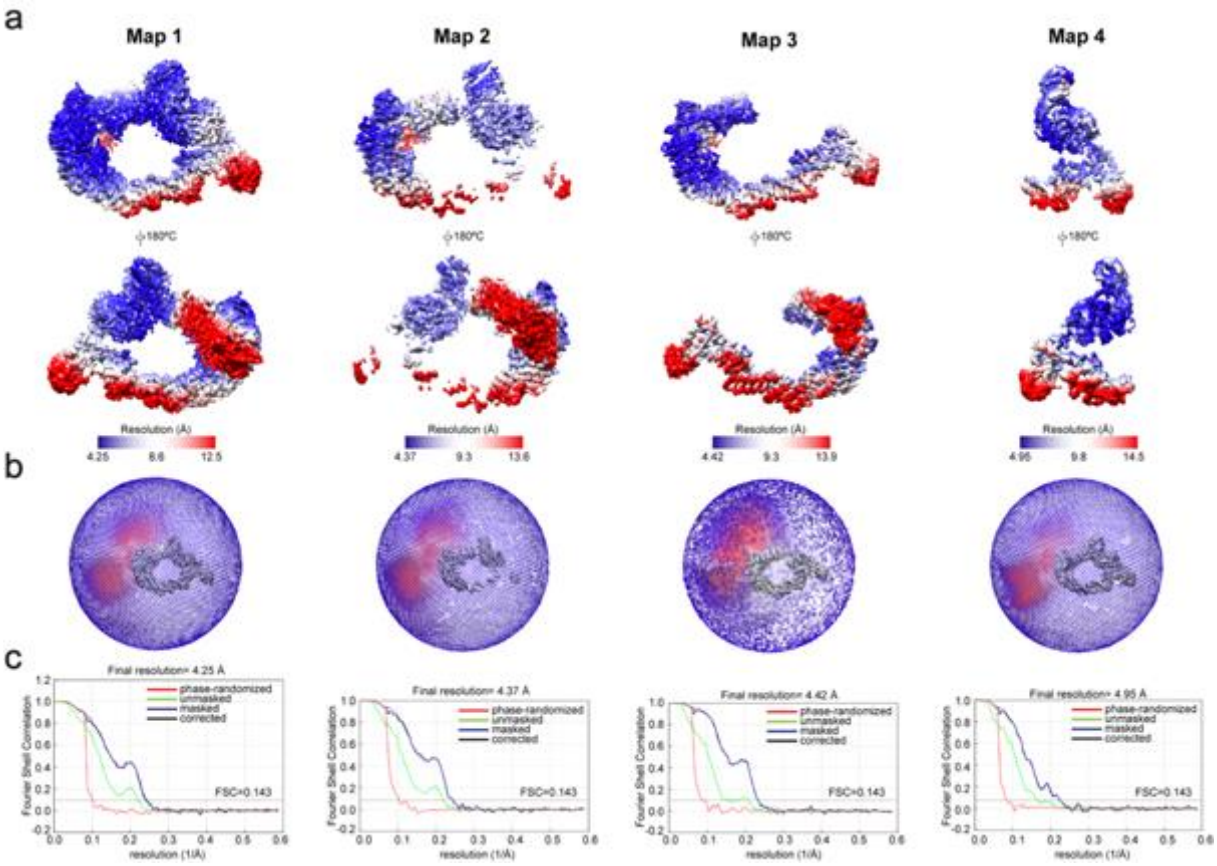


Figure 3.24 Obtained maps parameters. (a) Local resolution plot. (b) Angular distribution of particles. (c) FSC curves.

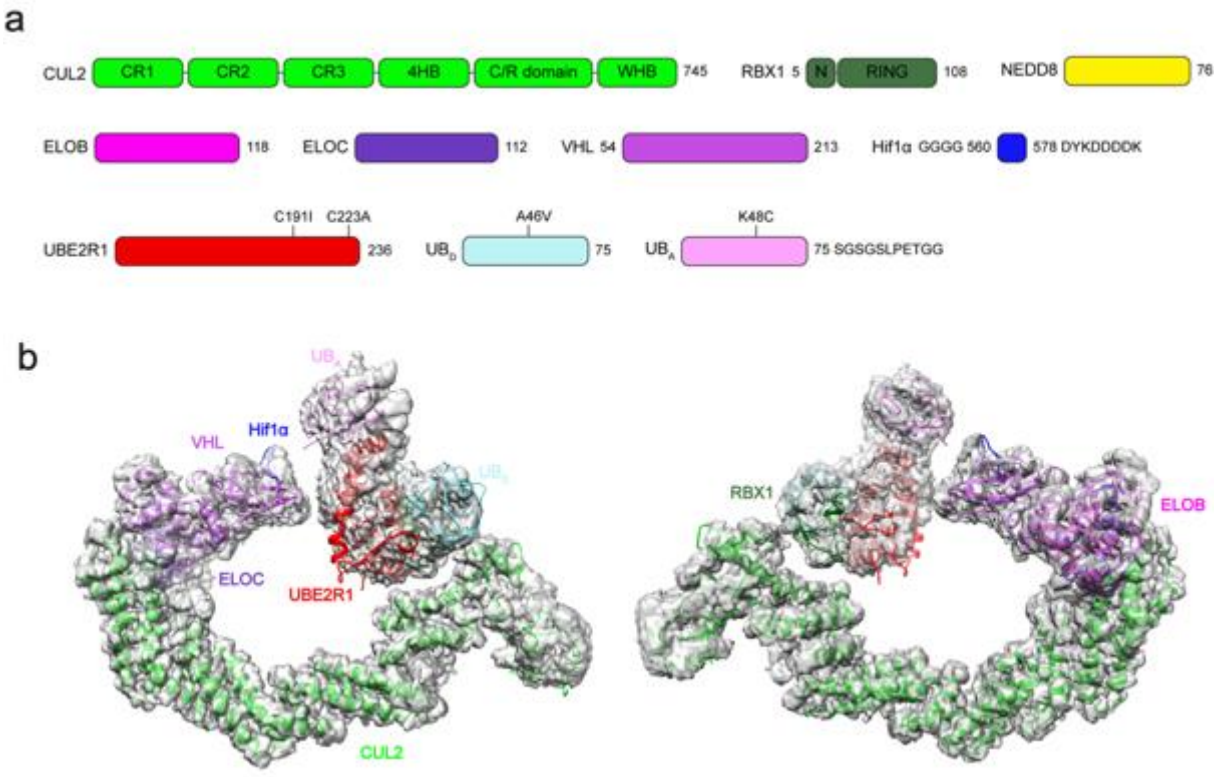


Figure 3.25 (a) Schematic representation of all the complex components. (b) Cryo-EM map with fitted crystal structures: 1LM8 (ELOB/C-VHL-Hif1α), 4MDK (UBE2R1-UB^D), 5N4W (CUL2-RBX1) and 3CMM (UB^A).

Nonetheless, the loops and some parts of the β-sheet modelling awaits higher resolution structure. Finally, RBX1's RING domain was easily docked, binding UBE2R1, UB^D and CUL2 CTD. Furthermore, the 5N4W CUL2-RBX1 crystal structure was placed into the map (Figure 3.27 b)[155]. The apo E3 conformation did not resemble observed CUL-Repeat 3 (CR3) and CTD architecture. The CR3, 4HB and C/R domains active arrangement substantially differs from the previous structure (Figure 3.27 c)[155]. Additional loops, not present in the crystal structure can be observed. Moreover, significant RBX1's RING domain displacement can be noticed (Figure 3.27 c). Both, CUL2's CTD and RBX1's rearrangements seem to happen to bring E2-UB^D into the proximity of substrate-UB^A-bound VHL for efficient catalysis.

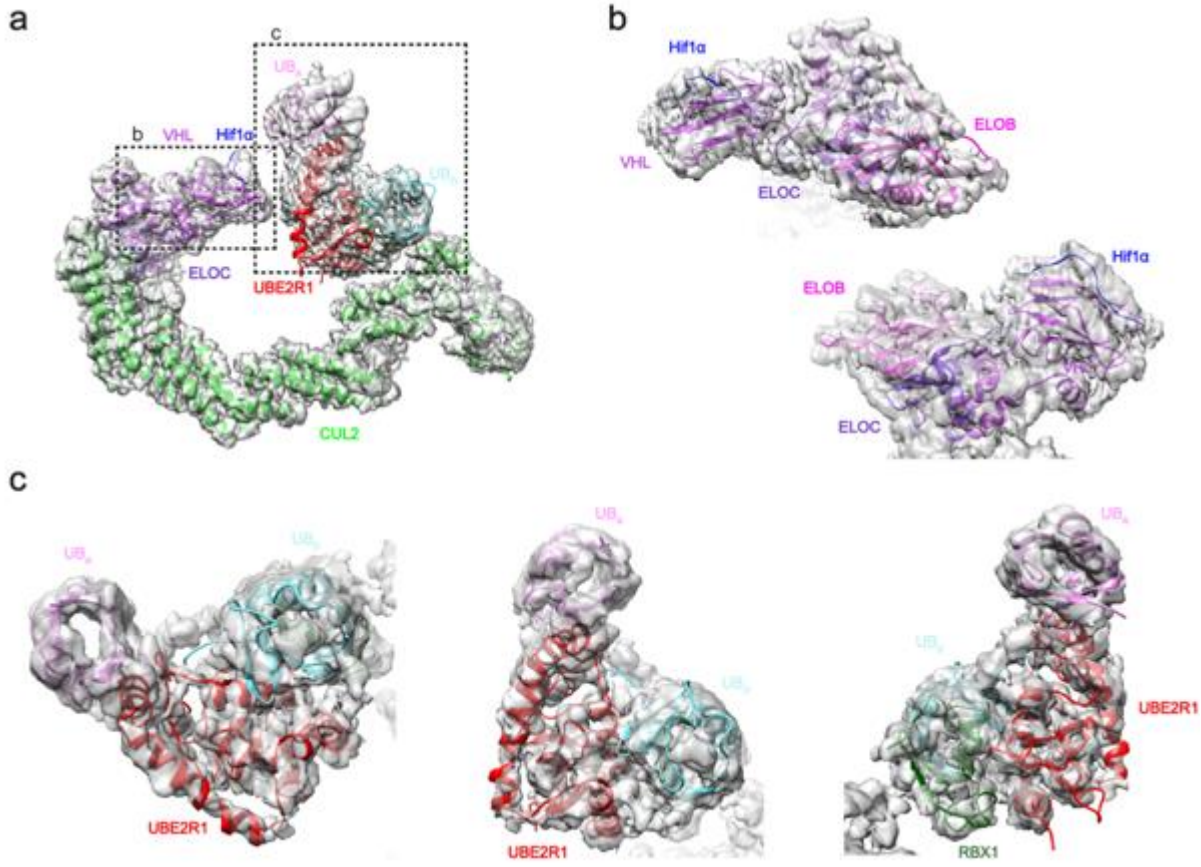


Figure 3.26 (a) Cryo-EM map with fitted crystal structures and marked area of interests shown in this figure. (b) ELOB/C-VHL and Hif1α density with fitted crystal structure (PDB 1LM8). (c)

UBE2R1-UB^D-UB^A-RBX1 density with fitted crystal structures (4MDK (UBE2R1-UB^D), 5N4W (RBX1) and 3CMM (UB^A)).

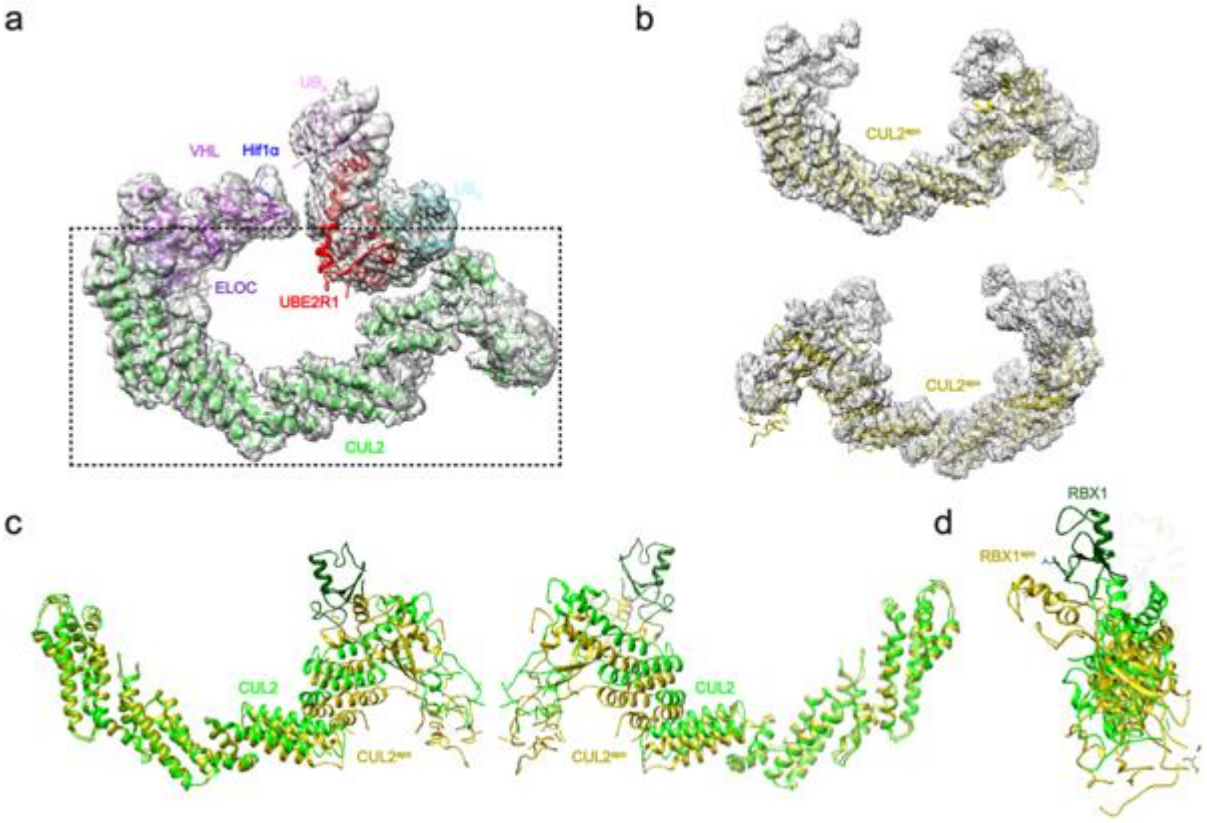


Figure 3.27 (a) Cryo-EM map with fitted crystal structures and marked area of interests shown in this figure. (b) Cryo-EM map CUL2 density with fitted apo crystal structure of CUL2-RBX1 (PDB 5N4W). (c) Comparison of apo crystal structure of CUL2 and Namdinator-modelled CUL2, based on obtained Cryo-EM map. (d) Comparison of apo crystal structure of RBX1 and Namdinator-modelled RBX1, based on obtained Cryo-EM map.

3.2.3.2 UBE2R1's catalytic site, UB^D and UB^A interactions

As noted above, the catalytic core resolution did not allow side chain modelling. However, many important features arise from docked UBE2R1-UB^D-UB^A-RBX1. First, the UB^D position fairly resembles previously observed E2-UB^D conformations[56,115,129]. The I44 and R54 point towards S129 and E133, previously identified as important for UB^D positioning (Figure 3.28 a). Additionally, the RBX1's RING domain situates UB^D in well known, closed conformation with linchpin R46 sandwiched between UBE2R1 and UB^D (Figure 3.28 b). The position and orientation of UB^A in complex with UBE2R1 remained elusive for years. Previous investigations, based on molecular modelling, have identified UB^A's R54 and UBE2R's D143, as residues playing a role in acceptor

UB binding[116]. Modelled location of UB^A in this study confirms the importance of these amino acids (Figure 3.29 a).

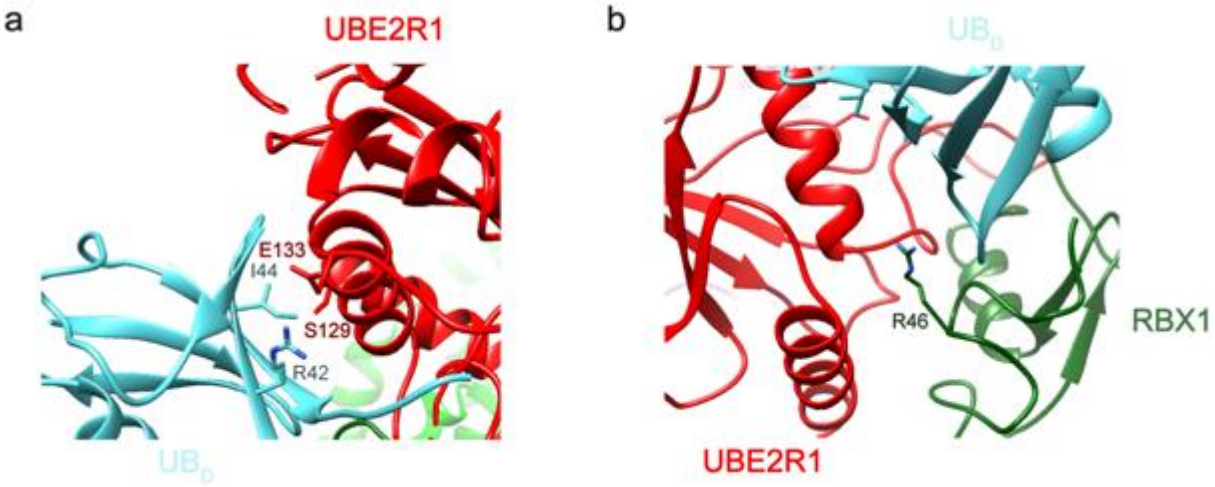


Figure 3.28 UBE2R-UB^D interactions. (a) Key residues playing a role in UBE2R1-UB^D binding. (b) Linchpin arginine sandwiched between UBE2R1 and UB^D.

Even though, the exact positions of side chains of interest can not be determined, they definitely belong to UBE2R1-UB^A interface (Figure 3.29 a). Close inspection of the model allowed to identify two more residues potentially impacting UB^A binding to UBE2R1.

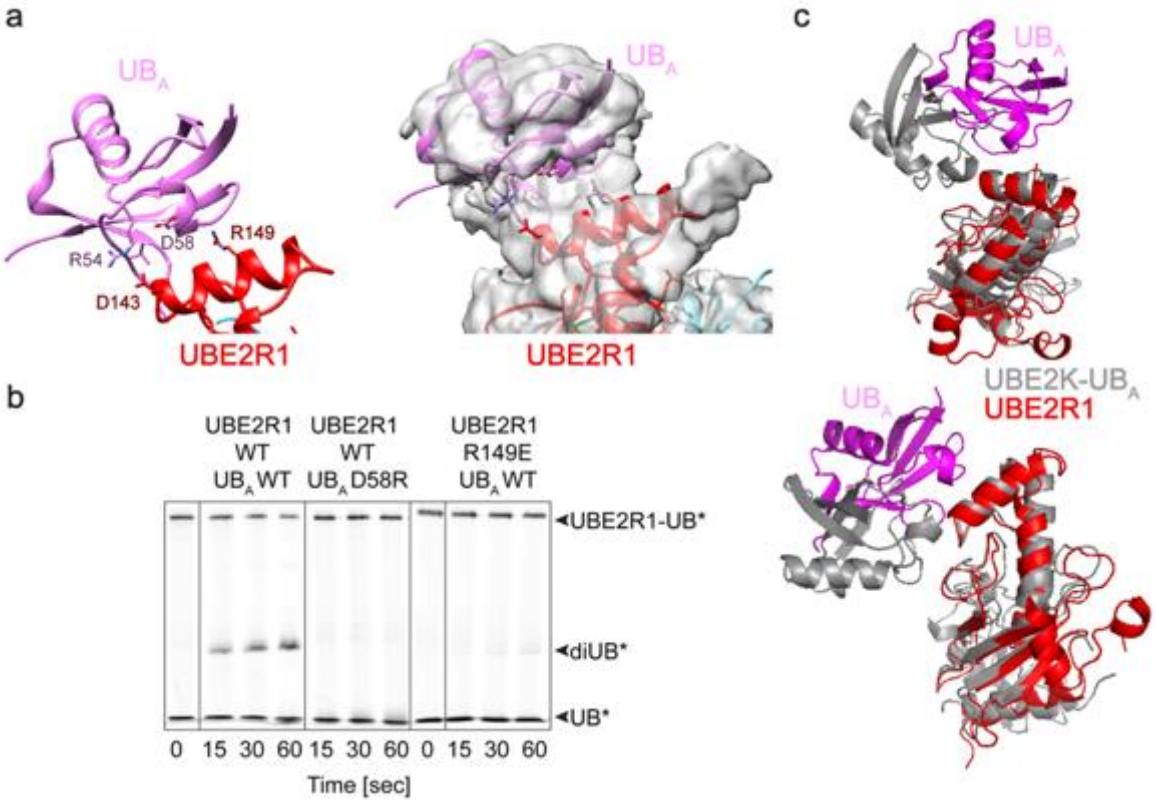


Figure 3.29 UBE2R-UB^A interactions. (a) Key residues, previously and newly identified, playing a role in UBE2R1-UB^A binding[58,116]. (b) diUB pulse-chase assay showing the importance of UB^A's D58 and UBE2R1's R149 in UB^A recruitment. (c) Superposition of UBE2K-UB^A crystal structure (PDB 7OJX) with UBE2R1-UB^A model obtained from Cryo-EM data.

As such, UB^A's D58R and UBE2R1's R149E mutations were tested in a diUB formation assay (Figure 3.29 b). The experiment indicated importance of D58 and R149 in polyubiquitylation, as both mutants almost completely abolished diUB formation. Indeed, R149 was previously suggested to play a role in acceptor UB binding[58]. Interestingly, the position of UB^A in UBE2R1-UB^A complex differs from the one identified in a recent crystal structure of K48-linked UB chain former UBE2K (Figure 3.29 c)[53]. It appears that the same patch of UB^A is bound by a different part of UBE2R1, suggesting diverse mode of action. It is worth to mention, that UBE2K contains an additional UB^A domain, which is known to bind and orient acceptor UB[53]. Furthermore, the model discussed in this study was resolved in the presence of substrate and the whole CRL2 assembled complex.

Acceptor UB binding to UBE2R1 determines its chain specificity. The above discussed interface ensures that K48 points towards the active site. However, precise position of UB^A's K48, catalytic cysteine and UB^D's C-terminus remain to be resolved. The obtained map does not provide sufficient information to conclude the catalytic site conformation. Thus, additional structural studies need to be performed to accurately evaluate UBE2R1's mode of action.

3.2.3.3 UBE2R1's acidic loop

UBE2R1's acidic loop plays an essential role in chain formation. Despite many biochemical studies revealing its importance, the catalytically active position of the loop was never resolved[31,116,118-120]. Chemical trapping of UBE2R1 in the presence of CRL2, mimicking the native intermediate, enabled acidic loop visualization (Figure 3.30 a). The loop wraps around the active site, suggesting its role in acceptor K48 positioning and/or catalysis, as it was suggested by previous studies[31,118]. As far as the loop

resolution limits its precise modelling, it is worth speculating the position of conserved H98. The loop begins with three bulky residues: H98, P99 and P100. The map's density pointed out in Figure 3.30 a suggests HPP position. Moreover, the loop seems to contact RBX1's R91, which was previously proposed (R91 RBX1- D103 UBE2R1) to play a role in loop binding (Figure 3.30 b)[116]. All in all, further structural studies and/or modelling needs to be proceeded to unveil the position of the loop's amino acids.

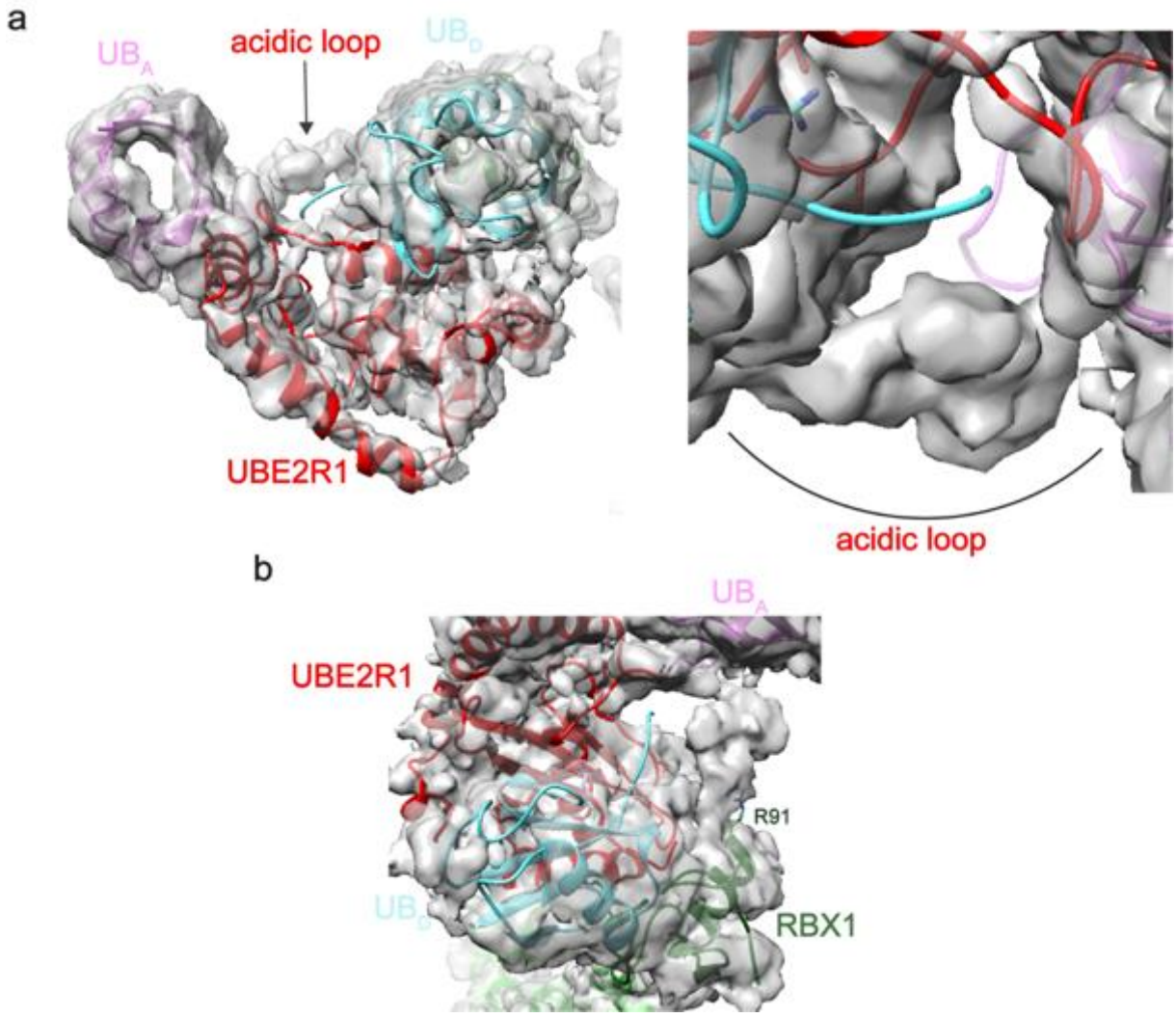


Figure 3.30 Acidic loop of UBE2R1. (a) Pointed out density belonging to UBE2R1's acidic loop. (b) Previously identified RBX1's R91 contacts acidic loop density[116].

3.2.3.4 UBE2R1's acidic tail

UBE2R's acidic, C-terminal extension plays a role, not only in catalysis of UB chain formation, but also in binding to CRL, thus it is essential for its function[52,109,124].

However, the main postprocessed map, generated in DeepEMhancer, does not contain any tail density binding CUL2's C/R domain (Figure 3.31 a) and therefore, additional particle sorting was performed. As such, the additional map was resolved with the tail density visible in CUL2's CTD (Figure 3.31 b).

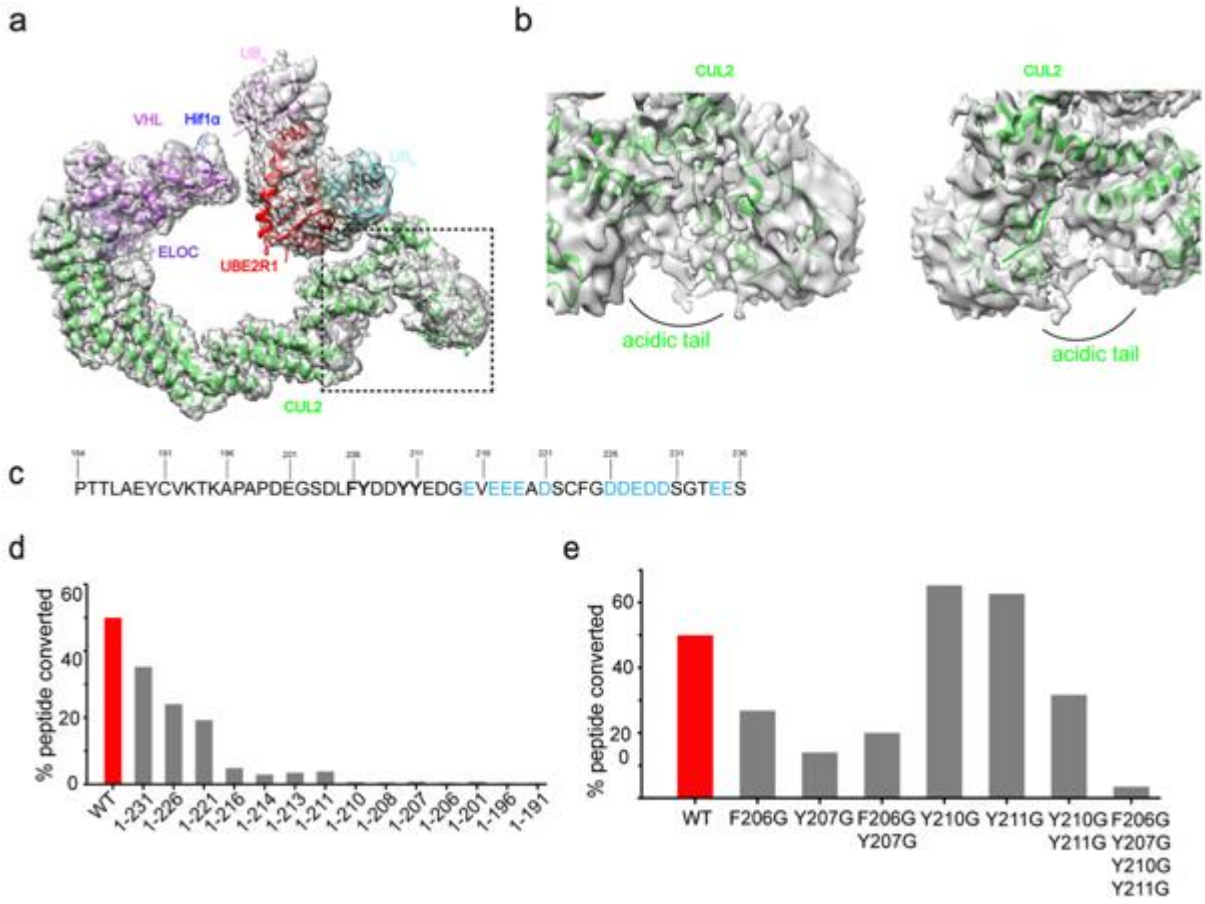


Figure 3.31 Acidic tail of UBE2R1. (a) Cryo-EM map with fitted crystal structures and marked area of interests shown in this figure. (b) Pointed out density belonging to UBE2R1's acidic tail. (c) UBE2R1's acidic tail sequence. Acidic residues are colored in blue. Important hydrophobic residues are bolded. (d) Steady state assay examining UBE2R1's tail truncations. (e) Steady state assay examining UBE2R1's hydrophobic residues mutations.

Yet again, the resolution limited the tail modelling and CUL2's residues determination, responsible for tail-E3 interaction. Nonetheless, it seems that the tail not only binds to the previously investigated basic canyon (Figure 3.31 b)[125]. Earlier studies have identified the importance of UBE2R's tail acidic residues (Figure 3.31 c) in the context of SCF[52]. To ascertain the role of acidic patches in the presence of CRL2, UBE2R1 tail truncations were generated and assayed in steady-state conditions. Consistently with

previous results, UBE2R lacking acidic residues in the tail (1-216) failed to catalyze Hif1α ubiquitylation (Figure 3.31 d)[52,124,125]. Moreover, other investigations have suggested significance of the tail's hydrophobic residues (F206, Y207, Y210, Y211, Figure 3.31 c)[82,126]. Thus, F206G, Y207G, Y210G and Y211G single, double and quadruple mutants were made and tested. The single mutations did not show significant deficiency, expect Y210G, however the quadruple mutant impaired UBE2R's catalysis substantially. The position of hydrophobic residues, their binding mode to CUL2, as well as acidic patch interactions await further structural studies.

3.2.3.5 UBE2R1's NEDD8 activation

CUL neddylation is required for efficient ubiquitylation[145]. Unneddylated CUL remains in an autoinhibited state, where WHB contacts RBX1's RING domain and prevents E2/E3 binding[146]. Upon NEDD8 conjugation to WHB, this autoinhibition is released, thus active complex is formed, ready to engage with E2/E3~UB^{D}. Recent cryo-EM structures revealed mechanisms of NEDD8 activation in CRL1-UBE2D3, CRL1-ARIH1 and CRL5-ARIH2 complexes[50,82,97]. For CRL1-UBE2D3 and CRL1-ARIH1 complexes direct NEDD8 engagement was required with E2/E3 and for CRL5-ARIH2, complex allosteric activation mechanism was revealed.

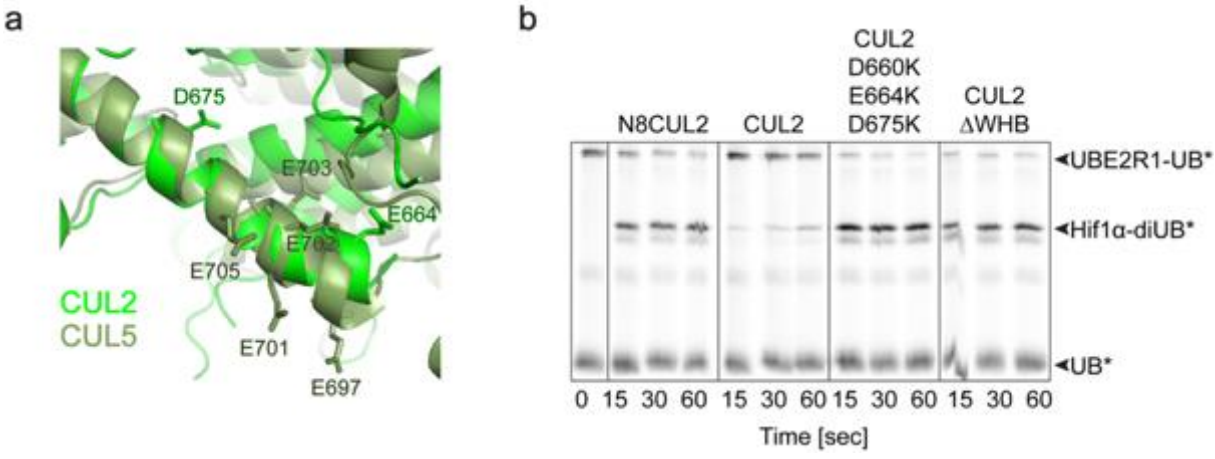


Figure 3.32 The role of NEDD8 in UBE2R1's activation. (a) CUL2 (PDB 5N4W) and CUL5 (7ONI) superposition with helix 29's acidic residues depicted. (b) Pulse-chase compering UBE2R's activity to form Hif1α-diUBs in the presence of various CUL2 mutants.

These assemblies are well-known to catalyze substrate monoubiquitylation[50,82,97]. However, the role of NEDD8 in UBE2R-dependent chain formation remains unknown. Importantly, CUL2 neddylation highly stimulates Hif1α polyubiquitylation (Figure 3.32 b, lane 1,2) in the presence of UBE2R1. The WHB-NEDD8 are not visible in the obtained maps, suggesting its high flexibility. To unveil the mechanism of NEDD8 stimulation for CRL2-UBE2R1 complex, following mutants were design. First, based on the closest CUL2 homolog, CUL5 and its allosteric NEDD8 activation, three acidic residues in helix 29 were mutated to the positive ones (Figure 3.32 a). Helix 29 connects CUL's CTD with WHB. The same strategy applied for CUL5 rescued lack of NEDD8 by repelling helix 29 from C/R domain and thus releasing CUL5 from autoinhibition. Second, the CUL2 version lacking WHB domain was tested. Surprisingly, both mutants were fully active despite lack of NEDD8, suggesting that no direct interaction is required between NEDD8 and the catalytic core (Figure 3.32 b). Careful kinetic experiments with quantitive data need to performed to fully characterize above mentioned mutants.

3.2.3.6 Sample improvement efforts

Lack of high resolution data in crucial parts of the complex, especially catalytic core, limited the analysis and modelling. Additionally, many lingering questions regarding UBE2R1 mode of action remained unanswered. Thus, many different strategies were employed to improve the sample. First, various plunging conditions were tried to overcome preferential complex orientation in ice (Figure 3.24 b). FOM, CHAPSO and β-OG detergents were tested. None of the additives improved complex behavior – CHAPSO and β-OG substantially lowered the number of particles in ice, even though higher complex concentrations were tried. Whereas FOM did not varied complex orientation in ice. Furthermore, data collection on golden grids was attempted to minimize beam-induced motions and improve resolution. This strategy did not lead to any map improvement, suggesting not enough rigidity on the sample site. Moreover, couple of different complex variations were tested. First, CUL2 ΔWHB and CUL2 3K mutant complexes were made, as data depicted in Figure 3.32 b suggested full rescue for these CUL2 versions, under qualitive conditions tested. Lack of WHB-N8 in the complex, not visible in the obtained maps, could potentially improve the sample. However, both specimens seemed to be highly unstable: CUL2 ΔWHB generated

greatly heterogenous sample and CUL2 3K + UBE2R1 could not be trapped. Second, various trap versions were made with longer and shorter Hif1α to test for substrate peptide optimal length. The shortest degron variant (Table 2.4) led to the best reconstruction so far. Lastly, apart from glutaraldehyde batch crosslinking, GraFix was also tested[232]. The longer crosslinking followed by desalting, to remove glycerol, caused substantial complex loss and breakdown. All in all, other strategies need to be employed for further sample improvement, which is further described in the Discussion section.

4. DISCUSSION

The UB system, with its myriad components collectively form a UB code, that oversees a tremendous number of distinct signal transduction events without producing undesirable cross-talk, is thus exquisitely dependent on specificity. Indeed, there has been great focus on how specificity is determined over large scales – tens to hundreds of angstroms – in terms of how substrates are placed relative to the active sites of ubiquitylating enzymes. Major effort has also been applied to deciphering the combined catalytic roles of many conserved residues beyond active site cysteines in UB carrying enzymes[85,100,101,103,233]. The roles of acceptor residues have been also studied in terms of different nucleophilic atoms within natural protein amino acid side chains. Indeed, ubiquitin carrying enzymes display remarkable ranges of reactivities with distinct acceptors within various contexts, including discharge of their thioester-bonded UB to nonspecific amines such as hydroxylamine, free lysine amino acids or even free cysteine or threonine[20,59,214,225]. Moreover, UB can be forcibly enzymatically linked to serine or lysine replacements of E2 or E3 catalytic cysteines, as well as to such residues in the vicinity of the active site to enable structural biology[17,21,234]. Thus, in light of the previous findings, it was not surprising that an E2 conjugating enzyme, UBE2S, could tolerate $^{K11}UB_{C2\text{-}C5}$ acceptors[101]. Unexpectedly, however, data in the first part of this thesis shows that multiple E2s and E3s are sensitive to the acceptor lysine side chain length. This sensitivity occurs on the ångstrom length scale, as shortening or extending the acceptor lysine by a single methylene group had a profound effect on the chain formation. The study included many K48 and K63-linked UB chain forming enzymes like: UBE2N/UBE2V1, UBE2R2, UBE2G1, NEDD4 and Rsp5p (Figure 3.4; 3.5; 3.7). Their activity was examined with or without partner E3 enzymes as well as in the presence of the substrate peptides. Striking preference for the native lysine geometry was not only proved for the E2 enzymes but also for the members of the HECT E3 ligase family (Figure 3.7). Careful kinetics analysis demonstrated that K_M, k_{cat} and pK_a can be influenced by UB_{C5}'s (Table 3.1). Additionally, MD simulations revealed pleiotropic structural effects of UB_{C5}: increased acceptor side chain and UB backbone dynamics and flexibility, along with additional degrees of freedom. Moreover, MS experiments with promiscuous E2 enzyme, UBE2D3, revealed that distribution of diUB chain linkages is impacted by the location of C5 on acceptor UB's surface (Figure 3.9; 3.10).

It might be easily reasoned, why UB_{C1-C3} analogs were defective, as the side chain was too short to reach the active site and attack E2~UB^D. Nonetheless, the fact that the longer acceptor in UB_{C5}'s influenced most of the tested enzymes suggested additional roles of the lysine hydrocarbon.

The mild observed effects on K_M and/or apparent pK_a for UBE2R2 and UBE2N/UBE2V1 were insufficient to justify the k_{cat} defects in the presence of UB_{C5}s analogs (Table 3.1). Performed MD simulations indicated a few aspects of the lysine side chain length, which may be most favorable for UB chain formation catalyzed by E2s and E3s. For instance, for enzymes where lysine positioning and/or substrate binding are rate-limiting, it appears that elevated entropy, caused by extra hydrocarbon in the acceptor side chain could influence the frequency of catalytic encounter (Figure 3.11; 3.12). Intriguingly, this mechanism would contrast from UBE2W mode of action, where disordered, flexible substrate's N-terminus is guided by a non-canonical, disordered UBE2W's C-terminal extension for UB conjugation[60]. Rather than demanding disorder, E2 and E3 enzymes tested in this study appear to benefit from calibrated nucleophile length with the restrained degrees of freedom.

Furthermore, the acceptor lysine length would not only affect entropy but also catalysis. Previous computational studies done for UBE2N (modelled without UBE2V1 and RING E3) has supported a model, where the E2 contains a precise "hole" to fit the lysine amine. Additionally, the data supported the claim that the nucleophilic attack on the thioester carbonyl is rate-limiting[86]. As such, data included in this study may suggest that the acceptor lysine length is not only optimal for the ubiquitylation chemistry but also for accessing the amine hole. Certainly, UBE2N~UB^D/UBE2V1/UB^A's MD simulations suggested multiple ways how the catalysis could be impacted by the acceptor lysine side chain length. First, the native lysine approaches the active site in an optimal way. Second, four hydrocarbon-long acceptor can conformationally influence the active site gate loop in the E2~UB^D intermediate. This conclusion goes in agreement with previous investigations suggesting the role of active site gate loop in formation of the transition state by closing around the acceptor $K63^{223}$. Importantly, MD simulations performed in this study showed active site gate loop distortion in the presence of UB_{C5} analog. This would agree with a UB discharge experiment to free

amino acid, where all nucleophiles served as a good acceptor irrespective of side chain length and great impact on k_{cat} in the context of UB^A analogs oriented by UBE2V1 (Figure 3.14). In such a case, little impact on K_M could be predicted. However, lack of K_M effect might be due to acceptor K63 not playing a role in UB^A recruitment to UBE2V1's UB-binding domain.

In contrast, the acceptor side chain length impacted K_M in the presence of HECT E3 Rsp5p (Table 3.1). This implied the role of the acceptor lysine in effective binding to Rsp5p E3. Potential local interactions – awaiting structural studies – might dominate UB^A recruitment[235]. On the other hand, acceptor lysine placement in the active site might allosterically stabilize conformation of enzyme-UB, which binds the acceptor[145]. Even though, the first part of this thesis relied on installing the acceptor side chain chemical variants, it seems possible that in the cellular environment many natural components might impact presentation of the acceptor lysine. Binding to the protein partners or linkage within a chain might result in E2 and E3 enzyme's specificity. Additionally, strong preference for the native lysine geometry may play a role in robust and sufficient ubiquitylation for proteasomal degradation, even when the favored targeting lysines are inaccessible. Such features, may also have an effect on the success or failure of molecular glues and PROTACs. As targeted protein degradation strategies strictly rely on small molecule binding and recruitment of proteins of interest to the ubiquitylation machinery for therapeutic effect[236,237].

CRLs, being responsible for more than 20% of all degradation events[132] in the eukaryotic system remain not fully understood. Moreover, the full success of the emerging field of targeted protein degradation, employing CRL systems, depends on a mechanistical understanding of ubiquitylation reactions. First, finding favored CRL-E2 pairs awaits further investigation. Recent studies have shown exquisite preference of UBE2R for CRL2, not only in the presence of native substrate but also in the context of neo-substrate[112,192]. Second, the mechanism of K48-linked UB chain formation, catalyzed by UBE2R, remains elusive. Interestingly, there are many long-standing questions, regarding UBE2R mode of action, UB field awaits to answer: (1) How does UBE2R achieve specificity? (2) What is the role and position of the acidic loop? (3) How does the acidic tail bind to CUL? (4) How is UBE2R activated by CRL? The

second part of this study attempted to answer some of the lingering mysteries by proceeding structural studies of UBE2R-UB^D-UB^A-$CRL2^{VHL}$ complex.

To explore the recent hints from the literature, respecting UBE2R preference of CRL2 over CRL1, both systems were compared (Figure 3.17). Furthermore, extensive kinetics studies were performed in collaboration with Gary Kleiger from the Department of Chemistry and Biochemistry, University of Nevada, Las Vegas, NV, USA (data not shown), showing powerful 28-fold decrease in UBE2R2's K_M with CRL2, in comparison to CRL1. As the biochemical, cellular and kinetics data pointed towards the advantage of CRL2 over CRL1, this system was chosen for structural investigations of UBE2R-mediated chain formation. Even more importantly, proceeding $CRL2^{VHL}$ complex with Hif1α as a substrate is to the great interest of many fields. The complex is associated with von Hippel-Lindau syndrome (VHL), a disease linked to multiple tumors growing in many organs, especially developing kidney cancer and pancreatic cancer[156,157]. Hif1α was identified as an important transcription factor responsible for oxygen homeostasis in the cellular environment[161-163]. All in all, the mechanism of Hif1α polyubiquitylation to this date is not known, despite the importance of the complex and its role in tumorigenesis.

Even though, the UB code writers function on millisecond time-scale, it is possible to visualize the active conformations with the help of chemical biology. As such, electrophilic traps were designed (Figure 3.18) and validated (Figure 3.21) to mimic native intermediate of K48-linked diUB formation. Apart from the probes, additional stabilization strategies, such as the CC0651 inhibitor and chemical crosslinking with glutaraldehyde, were used to stabilize the complex for cryo-EM. As a result, a 4.2 Å map was solved giving an insight into the mechanism of UBE2R1-mediated polyubiquitylation in the context of $CRL2^{VHL}$ and bound Hif1α. The obtained density allowed to fit crystal structures, nonetheless due to the lack of high resolution data in crucial parts of the complex, especially catalytic core, the structure could not be build and extensively analyzed on side chain level. However, many interesting features arose from the cryo-EM data. First, the CUL2 CR3, 4HB and C/R domains with RBX1 undergo major rearragmenets in comparison to apo crystal structure (PDB 5N4W) to place UBE2R1-UB^D in proximity of UB^A and acceptor lysine for efficient catalysis[155]. Second, the position of UB^A bound to UBE2R1 helix α3 could be determined. UBE2R1

seems to orient UB^A to achieve specificity via a couple of interactions: previosuly identified UB^A R54-UBE2R1 D143 pair[116], recently suggested UBE2R1's R149[58] and newly determined UB^A D58 (Figure 3.28). Interestingly, the relative position of UB^A in respect to E2 differs from the lately solved structure of UBE2K-UB^A (Figure 3.28 c)[53]. The difference might be due to the additional UBA domain of UBE2K, which binds and orients UB^A[53]. Furthermore, the model discussed in this study was resolved in more native environment, in the presence of substrate and the whole CRL2 assembled complex. Third, essential for chain formation, the acidic loop could be visualized, wraping around the active site (Figure 3.29). As far as the loop resolution limits its precise modelling, it is worth speculating the position of conserved H98, playing a role in lowering K48's pK_a. Moreover, the loop seems to contact RBX1's R91, which was previously proposed (R91 RBX1- D103 UBE2R1) to play a role in loop binding[116]. Fourth, required for UBE2R catalysis and binding, acidic tail density was visible in CUL2's CTD (Figure 3.30). Yet again, the resolution limited the tail modelling and CUL2's residues determination, responsible for tail-E3 interaction. Nonetheless, it seems that tail not only binds to the previously investigated basic canyon[125]. Consistently with previous results, UBE2R lacking tail's acidic residues (1-216) failed to catalyze Hif1α ubiquitylation and quadruple hydrophobic mutant (F206G, Y207G, Y210G, Y211G) impaired UBE2R's catalysis substantially[52]. Lastly, the collected data suggested the role of UBE2R1's NEDD8 activation, which presence is required for efficient chain formation. WHB-NEDD8 is not visible in the obtained maps, suggesting its high flexibility. Surprisingly, WHB domain removal or helix 29 mutations, releasing CUL2 from autoinhibition, fully rescued the lack of NEDD8 (Figure 3.31). Thus, an allosteric mechanism of UBE2R1's NEDD8 activation might be proposed.

As far as the captured model gives more insight into UBE2R1's mode of action, many questions remain unanswered. Precise position of UB^A's K48, catalytic cysteine and UB^D's C-terminus remain to be resolved. The obtained map does not provide sufficient information to argue catalytic site conformation. Moreover, the resolution limits exact determination of UBE2R1-UB^A interactions, acidic loop position and the role of CUL2-UBE2R1's acidic tail binding mode. Thus, many different complex stabilization and plunging strategies were tested to improve the map and resolution (section 3.2.3.6), which, so far, did not lead to any map improvements. Hence, additional structural efforts, either by cryo-EM or crystal studies, need to be proceeded to obtain high

resolution data confirming many of the conclusions and giving insight into the parts not yet visible. Different stabilization strategies can be used for cryo-EM, like changing chemical crosslinker or chemical probe. Additional complex purification step can be included, where the FLAG-tag on a substrate peptide would be used, enabling enrichment for $CRL2^{VHL}$-UBE2R1-trapped sample. Crystal studies of UBE2R1 UBC domain with crosslinked UB^D and UB^A can be advanced. First, UBE2R1 A140K C191I 1-202 – UB^D A46V – UB^A K48C crystals will be improved. Throughout this project, the intermediate mimicking chain formation was made, as described in Streich and Lima, 2016[238]. Briefly, UB^D A46V was linked via an isopeptide bond to an UBE2R1's A140K and UB^A K48C was attached to an UBE2R1's catalytic cysteine via bis-disulfide. Crystalized complex diffracted to 5 Å, thus further optimization rounds have to be performed, like seeding in larger drops. Moreover, one more sample will be prepared following similar strategy, as recently published[53]. UB^D A46V will be reacted with UBE2R1's C93K to form isopeptide bond and UB^A's R54C will form a BMOE crosslink with UBE2R1's D143C. The approach is based on the natural interaction between acceptor UB and UBE2R1, thus potentially placing K48 in the active site[116]. To stabilize UBE2R1-UB^D close conformation, RBX1's RING domain will be used. Possibly, CUL2 CTD with peptide tail crystal trails will be proceeded to visualize acidic tail binding mode. Furthermore, extensive kinetics studies need to be performed to verify the role of many mutants identified in this study. Quantitive experiments determining K_{Ms} and k_{cats} will help to understand UBE2R1-$CUL2^{VHL}$-catalyzed polyubiquitin chain formation. The interesting concept of CUL-E2 pair specificity, could be further investigated in the context of cell biology. Finally, having in mind the importance of targeted protein degradation and $CUL2^{VHL}$ PROTACs, a similar strategy could be used to attempt visualization of neo-substrate-bound complexes.

5. 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 Mukhopadhyay, D. & Riezman, H. Proteasome-Independent Functions of Ubiquitin in Endocytosis and Signaling. *Science* **315**, 201-205, doi:doi:10.1126/science.1127085 (2007).

3 Cruz Walma, D. A., Chen, Z., Bullock, A. N. & Yamada, K. M. Ubiquitin ligases: guardians of mammalian development. *Nature Reviews Molecular Cell Biology*, doi:10.1038/s41580-021-00448-5 (2022).

4 Kim, W. *et al.* Systematic and Quantitative Assessment of the Ubiquitin-Modified Proteome. *Mol Cell* **44**, 325-340, doi:https://doi.org/10.1016/j.molcel.2011.08.025 (2011).

5 Emanuele, Michael J. *et al.* Global Identification of Modular Cullin-RING Ligase Substrates. *Cell* **147**, 459-474, doi:https://doi.org/10.1016/j.cell.2011.09.019 (2011).

6 Kwon, Y. T. & Ciechanover, A. The Ubiquitin Code in the Ubiquitin-Proteasome System and Autophagy. *Trends Biochem Sci* **42**, 873-886 (2017).

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

8 Komander, D. & Rape, M. The Ubiquitin Code. *Annual Review of Biochemistry* **81**, 203-229 (2012).

9 Vijay-Kumar, S., Bugg, C. E. & Cook, W. J. Structure of ubiquitin refined at 1.8Åresolution. *Journal of Molecular Biology* **194**, 531-544, doi:https://doi.org/10.1016/0022-2836(87)90679-6 (1987).

10 Deshaies, R. J. & Joazeiro, C. A. P. RING Domain E3 Ubiquitin Ligases. *Annual Review of Biochemistry* **78**, 399-434, doi:10.1146/annurev.biochem.78.101807.093809 (2009).

11 Schulman, B. A. & Wade Harper, J. Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nature Reviews Molecular Cell Biology* **10**, 319-331, doi:10.1038/nrm2673 (2009).

12 Ye, Y. & Rape, M. Building ubiquitin chains: E2 enzymes at work. *Nature Reviews Molecular Cell Biology* **10**, 755-764, doi:10.1038/nrm2780 (2009).

13 Dikic, I., Wakatsuki, S. & Walters, K. J. Ubiquitin-binding domains — from structures to functions. *Nature Reviews Molecular Cell Biology* **10**, 659-671 (2009).

14 Sloper-Mould, K. E., Jemc, J. C., Pickart, C. M. & Hicke, L. Distinct Functional Surface Regions on Ubiquitin*. *Journal of Biological Chemistry* **276**, 30483-30489, doi:https://doi.org/10.1074/jbc.M103248200 (2001).

15 Hu, M. *et al.* Crystal Structure of a UBP-Family Deubiquitinating Enzyme in Isolation and in Complex with Ubiquitin Aldehyde. *Cell* **111**, 1041-1054, doi:https://doi.org/10.1016/S0092-8674(02)01199-6 (2002).

16 Reyes-Turcu, F. E. *et al.* The Ubiquitin Binding Domain ZnF UBP Recognizes the C-Terminal Diglycine Motif of Unanchored Ubiquitin. *Cell* **124**, 1197-1208, doi:https://doi.org/10.1016/j.cell.2006.02.038 (2006).

17 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 (2009).

18 Shih, S. C., Sloper-Mould, K. E. & Hicke, L. Monoubiquitin carries a novel internalization signal that is appended to activated receptors. *The EMBO journal* **19**, 187-198, doi:10.1093/emboj/19.2.187 (2000).
19 Jin, L., Williamson, A., Banerjee, S., Philipp, I. & Rape, M. Mechanism of Ubiquitin-Chain Formation by the Human Anaphase-Promoting Complex. *Cell* **133**, 653-665, doi:https://doi.org/10.1016/j.cell.2008.04.012 (2008).
20 Stewart, M. D., Ritterhoff, T., Klevit, R. E. & Brzovic, P. S. E2 enzymes: more than just middle men. *Cell Research* **26**, 423-440, doi:10.1038/cr.2016.35 (2016).
21 Zheng, N. & Shabek, N. Ubiquitin Ligases: Structure, Function, and Regulation. *Annu Rev Biochem* **86**, 129-157 (2017).
22 Harrigan, J. A., Jacq, X., Martin, N. M. & Jackson, S. P. Deubiquitylating enzymes and drug discovery: emerging opportunities. *Nature Reviews Drug Discovery* **17**, 57-78, doi:10.1038/nrd.2017.152 (2018).
23 Wang, H. *et al.* Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873-878, doi:10.1038/nature02985 (2004).
24 Bentley, M. L. *et al.* Recognition of UbcH5c and the nucleosome by the Bmi1/Ring1b ubiquitin ligase complex. *The EMBO Journal* **30**, 3285-3297, doi:https://doi.org/10.1038/emboj.2011.243 (2011).
25 Vissers, J. H. A., Nicassio, F., van Lohuizen, M., Di Fiore, P. P. & Citterio, E. The many faces of ubiquitinated histone H2A: insights from the DUBs. *Cell Division* **3**, 8, doi:10.1186/1747-1028-3-8 (2008).
26 Haglund, K. *et al.* Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation. *Nature Cell Biology* **5**, 461-466, doi:10.1038/ncb983 (2003).
27 Wieduwilt, M. J. & Moasser, M. M. The epidermal growth factor receptor family: Biology driving targeted therapeutics. *Cellular and Molecular Life Sciences* **65**, 1566-1584, doi:10.1007/s00018-008-7440-8 (2008).
28 Waterman, H. & Yarden, Y. Molecular mechanisms underlying endocytosis and sorting of ErbB receptor tyrosine kinases. *FEBS Lett* **490**, 142-152, doi:10.1016/s0014-5793(01)02117-2 (2001).
29 Yau, R. & Rape, M. The increasing complexity of the ubiquitin code. *Nat Cell Biol* **18**, 579-586 (2016).
30 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).
31 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 (2005).
32 Stewart, G. S. *et al.* The RIDDLE syndrome protein mediates a ubiquitin-dependent signaling cascade at sites of DNA damage. *Cell* **136**, 420-434, doi:10.1016/j.cell.2008.12.042 (2009).
33 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, doi:10.1016/s0092-8674(00)00126-4 (2000).
34 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).
35 Hofmann, R. M. & Pickart, C. M. Noncanonical MMS2-encoded ubiquitin-conjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair. *Cell* **96**, 645-653, doi:10.1016/s0092-8674(00)80575-9 (1999).

36 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).
37 Brown, N. G. *et al.* Dual RING E3 Architectures Regulate Multiubiquitination and Ubiquitin Chain Elongation by APC/C. *Cell* **165**, 1440-1453 (2016).
38 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).
39 Ikeda, F. *et al.* SHARPIN forms a linear ubiquitin ligase complex regulating NF-κB activity and apoptosis. *Nature* **471**, 637-641, doi:10.1038/nature09814 (2011).
40 Gerlach, B. *et al.* Linear ubiquitination prevents inflammation and regulates immune signalling. *Nature* **471**, 591-596, doi:10.1038/nature09816 (2011).
41 Ordureau, A. *et al.* Defining roles of PARKIN and ubiquitin phosphorylation by PINK1 in mitochondrial quality control using a ubiquitin replacement strategy. *Proceedings of the National Academy of Sciences* **112**, 6637-6642, doi:doi:10.1073/pnas.1506593112 (2015).
42 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, doi:10.1074/jbc.270.29.17442 (1995).
43 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).
44 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).
45 Kaiho-Soma, A. *et al.* TRIP12 promotes small-molecule-induced degradation through K29/K48-branched ubiquitin chains. *Mol Cell* **81**, 1411-1424.e1417, doi:10.1016/j.molcel.2021.01.023 (2021).
46 Lin, Y., Hwang, W. C. & Basavappa, R. Structural and functional analysis of the human mitotic-specific ubiquitin-conjugating enzyme, UbcH10. *J Biol Chem* **277**, 21913-21921, doi:10.1074/jbc.M109398200 (2002).
47 Özkan, E., Yu, H. & Deisenhofer, J. Mechanistic insight into the allosteric activation of a ubiquitin-conjugating enzyme by RING-type ubiquitin ligases. *Proceedings of the National Academy of Sciences* **102**, 18890-18895, doi:doi:10.1073/pnas.0509418102 (2005).
48 Eddins, M. J., Carlile, C. M., Gomez, K. M., Pickart, C. M. & Wolberger, C. Mms2–Ubc13 covalently bound to ubiquitin reveals the structural basis of linkage-specific polyubiquitin chain formation. *Nature Structural & Molecular Biology* **13**, 915-920 (2006).
49 Brzovic, P. S., Lissounov, A., Christensen, D. E., Hoyt, D. W. & Klevit, R. E. A UbcH5/ubiquitin noncovalent complex is required for processive BRCA1-directed ubiquitination. *Mol Cell* **21**, 873-880, doi:10.1016/j.molcel.2006.02.008 (2006).
50 Baek, K. *et al.* NEDD8 nucleates a multivalent cullin–RING–UBE2D ubiquitin ligation assembly. *Nature* (2020).
51 Choi, Y.-S. *et al.* Differential ubiquitin binding by the acidic loops of Ube2g1 and Ube2r1 enzymes distinguishes their Lys-48-ubiquitylation activities. *J Biol Chem* **290**, 2251-2263 (2015).
52 Kleiger, G., Hao, B., Mohl, D. A. & Deshaies, R. J. The acidic tail of the Cdc34 ubiquitin-conjugating enzyme functions in both binding to and catalysis with ubiquitin ligase SCFCdc4. *J Biol Chem* **284**, 36012-36023, doi:10.1074/jbc.M109.058529 (2009).

53 Nakasone, M. A. *et al.* Structure of UBE2K–Ub/E3/polyUb reveals mechanisms of K48-linked Ub chain extension. *Nat Chem Biol* **18**, 422-431, doi:10.1038/s41589-021-00952-x (2022).
54 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~ubiquitin conjugate. *Biochemistry* **50**, 1624-1633, doi:10.1021/bi101913m (2011).
55 Plechanovová, 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 (2012).
56 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 (2012).
57 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 (2012).
58 Welsh, K. A. *et al.* Functional conservation and divergence of the helix-turn-helix motif of E2 ubiquitin-conjugating enzymes. *Embo j* **41**, e108823, doi:10.15252/embj.2021108823 (2022).
59 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 (2011).
60 Vittal, V. *et al.* Intrinsic disorder drives N-terminal ubiquitination by Ube2w. *Nat Chem Biol* **11**, 83-89 (2015).
61 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).
62 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).
63 Zheng, N. & Shabek, N. Ubiquitin Ligases: Structure, Function, and Regulation. *Annual Review of Biochemistry* **86**, 129-157, doi:10.1146/annurev-biochem-060815-014922 (2017).
64 Thien, C. B. & Langdon, W. Y. c-Cbl and Cbl-b ubiquitin ligases: substrate diversity and the negative regulation of signalling responses. *Biochem J* **391**, 153-166, doi:10.1042/bj20050892 (2005).
65 Deshaies, R. J. SCF and Cullin/Ring H2-based ubiquitin ligases. *Annu Rev Cell Dev Biol* **15**, 435-467, doi:10.1146/annurev.cellbio.15.1.435 (1999).
66 Zimmerman, E. S., Schulman, B. A. & Zheng, N. Structural assembly of cullin-RING ubiquitin ligase complexes. *Curr Opin Struct Biol* **20**, 714-721, doi:10.1016/j.sbi.2010.08.010 (2010).
67 Chang, L. & Barford, D. Insights into the anaphase-promoting complex: a molecular machine that regulates mitosis. *Curr Opin Struct Biol* **29**, 1-9, doi:10.1016/j.sbi.2014.08.003 (2014).
68 Baer, R. & Ludwig, T. The BRCA1/BARD1 heterodimer, a tumor suppressor complex with ubiquitin E3 ligase activity. *Curr Opin Genet Dev* **12**, 86-91, doi:10.1016/s0959-437x(01)00269-6 (2002).
69 Rotin, D. & Kumar, S. Physiological functions of the HECT family of ubiquitin ligases. *Nature Reviews Molecular Cell Biology* **10**, 398-409 (2009).
70 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).

71 Huang, L. *et al.* Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade. *Science* **286**, 1321-1326, doi:10.1126/science.286.5443.1321 (1999).
72 Verdecia, M. A. *et al.* Conformational flexibility underlies ubiquitin ligation mediated by the WWP1 HECT domain E3 ligase. *Mol Cell* **11**, 249-259, doi:10.1016/s1097-2765(02)00774-8 (2003).
73 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).
74 Maspero, E. *et al.* Structure of a ubiquitin-loaded HECT ligase reveals the molecular basis for catalytic priming. *Nature Structural & Molecular Biology* **20**, 696-701 (2013).
75 Kamadurai, H. B. *et al.* Mechanism of ubiquitin ligation and lysine prioritization by a HECT E3. *eLife* **2**, e00828-e00828 (2013).
76 Trempe, J. F. *et al.* Structure of parkin reveals mechanisms for ubiquitin ligase activation. *Science* **340**, 1451-1455, doi:10.1126/science.1237908 (2013).
77 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).
78 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).
79 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).
80 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).
81 Scott, D. C. *et al.* Two Distinct Types of E3 Ligases Work in Unison to Regulate Substrate Ubiquitylation. *Cell* **166**, 1198-1214.e1124 (2016).
82 Kostrhon, S. *et al.* CUL5-ARIH2 E3-E3 ubiquitin ligase structure reveals cullin-specific NEDD8 activation. *Nat Chem Biol* **17**, 1075-1083, doi:10.1038/s41589-021-00858-8 (2021).
83 Yunus, A. A. & Lima, C. D. Lysine activation and functional analysis of E2-mediated conjugation in the SUMO pathway. *Nature Structural & Molecular Biology* **13**, 491-499 (2006).
84 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).
85 Berndsen, C. E., Wiener, R., Yu, I. W., Ringel, A. E. & Wolberger, C. A conserved asparagine has a structural role in ubiquitin-conjugating enzymes. *Nat Chem Biol* **9**, 154-156 (2013).
86 Jones, W. M., Davis, A. G., Wilson, R. H., Elliott, K. L. & Sumner, I. A conserved asparagine in a ubiquitin-conjugating enzyme positions the substrate for nucleophilic attack. *J Comput Chem* **40**, 1969-1977 (2019).
87 Mattiroli, F. & Sixma, T. K. Lysine-targeting specificity in ubiquitin and ubiquitin-like modification pathways. *Nat Struct Mol Biol* **21**, 308-316 (2014).
88 Varshavsky, A. Naming a targeting signal. *Cell* **64**, 13-15, doi:10.1016/0092-8674(91)90202-a (1991).
89 Wenzel, D. M. & Klevit, R. E. Following Ariadne's thread: a new perspective on RBR ubiquitin ligases. *BMC Biology* **10**, 24, doi:10.1186/1741-7007-10-24 (2012).

90 Winston, J. T. *et al.* The SCFbeta-TRCP-ubiquitin ligase complex associates specifically with phosphorylated destruction motifs in IkappaBalpha and beta-catenin and stimulates IkappaBalpha ubiquitination in vitro. *Genes Dev* **13**, 270-283, doi:10.1101/gad.13.3.270 (1999).
91 Orlicky, S., Tang, X., Willems, A., Tyers, M. & Sicheri, F. Structural basis for phosphodependent substrate selection and orientation by the SCFCdc4 ubiquitin ligase. *Cell* **112**, 243-256, doi:10.1016/s0092-8674(03)00034-5 (2003).
92 Hon, W. C. *et al.* Structural basis for the recognition of hydroxyproline in HIF-1 alpha by pVHL. *Nature* **417**, 975-978, doi:10.1038/nature00767 (2002).
93 Skaar, J. R. & Pagano, M. Control of cell growth by the SCF and APC/C ubiquitin ligases. *Curr Opin Cell Biol* **21**, 816-824, doi:10.1016/j.ceb.2009.08.004 (2009).
94 Fischer, E. S. *et al.* The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation. *Cell* **147**, 1024-1039, doi:10.1016/j.cell.2011.10.035 (2011).
95 Petroski, M. D. & Deshaies, R. J. Context of multiubiquitin chain attachment influences the rate of Sic1 degradation. *Mol Cell* **11**, 1435-1444, doi:10.1016/s1097-2765(03)00221-1 (2003).
96 Tang, X. *et al.* Suprafacial orientation of the SCFCdc4 dimer accommodates multiple geometries for substrate ubiquitination. *Cell* **129**, 1165-1176, doi:10.1016/j.cell.2007.04.042 (2007).
97 Horn-Ghetko, D. *et al.* Ubiquitin ligation to F-box protein targets by SCF–RBR E3–E3 super-assembly. *Nature* **590**, 671-676, doi:10.1038/s41586-021-03197-9 (2021).
98 Stawiecka-Mirota, M. *et al.* Targeting of Sna3p to the endosomal pathway depends on its interaction with Rsp5p and multivesicular body sorting on its ubiquitylation. *Traffic* **8**, 1280-1296, doi:10.1111/j.1600-0854.2007.00610.x (2007).
99 Chen, Z. & Pickart, C. M. A 25-kilodalton ubiquitin carrier protein (E2) catalyzes multi-ubiquitin chain synthesis via lysine 48 of ubiquitin. *J Biol Chem* **265**, 21835-21842 (1990).
100 Branigan, E., Plechanovová, A., Jaffray, E. G., Naismith, J. H. & Hay, R. T. Structural basis for the RING-catalyzed synthesis of K63-linked ubiquitin chains. *Nature Structural & Molecular Biology* **22**, 597-602 (2015).
101 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 (2011).
102 Kim, H. C. & Huibregtse, J. M. Polyubiquitination by HECT E3s and the Determinants of Chain Type Specificity. *Molecular and Cellular Biology* **29**, 3307 (2009).
103 Stieglitz, B. *et al.* Structural basis for ligase-specific conjugation of linear ubiquitin chains by HOIP. *Nature* **503**, 422-426 (2013).
104 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. *The EMBO journal* **31**, 3833-3844 (2012).
105 Goebl, M. G. *et al.* The yeast cell cycle gene CDC34 encodes a ubiquitin-conjugating enzyme. *Science* **241**, 1331-1335, doi:10.1126/science.2842867 (1988).

106 Schwob, E., Böhm, T., Mendenhall, M. D. & Nasmyth, K. The B-type cyclin kinase inhibitor p40SIC1 controls the G1 to S transition in S. cerevisiae. *Cell* **79**, 233-244, doi:https://doi.org/10.1016/0092-8674(94)90193-7 (1994).
107 Skowyra, D., Craig, K. L., Tyers, M., Elledge, S. J. & Harper, J. W. F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. *Cell* **91**, 209-219, doi:10.1016/s0092-8674(00)80403-1 (1997).
108 Verma, R., Feldman, R. M. & Deshaies, R. J. SIC1 is ubiquitinated in vitro by a pathway that requires CDC4, CDC34, and cyclin/CDK activities. *Mol Biol Cell* **8**, 1427-1437, doi:10.1091/mbc.8.8.1427 (1997).
109 Mathias, N., Steussy, C. N. & Goebl, M. G. An essential domain within Cdc34p is required for binding to a complex containing Cdc4p and Cdc53p in Saccharomyces cerevisiae. *J Biol Chem* **273**, 4040-4045, doi:10.1074/jbc.273.7.4040 (1998).
110 Mathias, N. *et al.* Cdc53p acts in concert with Cdc4p and Cdc34p to control the G1-to-S-phase transition and identifies a conserved family of proteins. *Mol Cell Biol* **16**, 6634-6643, doi:10.1128/mcb.16.12.6634 (1996).
111 Wu, K., Kovacev, J. & Pan, Z. Q. Priming and extending: a UbcH5/Cdc34 E2 handoff mechanism for polyubiquitination on a SCF substrate. *Mol Cell* **37**, 784-796 (2010).
112 Hill, S. *et al.* Robust cullin-RING ligase function is established by a multiplicity of poly-ubiquitylation pathways. *eLife* **8**, e51163 (2019).
113 Gazdoiu, S., Yamoah, K., Wu, K. & Pan, Z. Q. Human Cdc34 employs distinct sites to coordinate attachment of ubiquitin to a substrate and assembly of polyubiquitin chains. *Mol Cell Biol* **27**, 7041-7052, doi:10.1128/mcb.00812-07 (2007).
114 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).
115 Williams, K. M. *et al.* Structural insights into E1 recognition and the ubiquitin-conjugating activity of the E2 enzyme Cdc34. *Nat Commun* **10**, 3296, doi:10.1038/s41467-019-11061-8 (2019).
116 Hill, S., Harrison, J. S., Lewis, S. M., Kuhlman, B. & Kleiger, G. Mechanism of Lysine 48 Selectivity during Polyubiquitin Chain Formation by the Ube2R1/2 Ubiquitin-Conjugating Enzyme. *Mol Cell Biol* **36**, 1720-1732 (2016).
117 Chong, R. A. *et al.* Pivotal role for the ubiquitin Y59-E51 loop in lysine 48 polyubiquitination. *Proc Natl Acad Sci U S A* **111**, 8434-8439, doi:10.1073/pnas.1407849111 (2014).
118 Ziemba, A. *et al.* Multimodal mechanism of action for the Cdc34 acidic loop: a case study for why ubiquitin-conjugating enzymes have loops and tails. *J Biol Chem* **288**, 34882-34896 (2013).
119 Pitluk, Z. W., McDonough, M., Sangan, P. & Gonda, D. K. Novel CDC34 (UBC3) ubiquitin-conjugating enzyme mutants obtained by charge-to-alanine scanning mutagenesis. *Mol Cell Biol* **15**, 1210-1219, doi:10.1128/mcb.15.3.1210 (1995).
120 Spratt, D. E., Wu, K., Kovacev, J., Pan, Z. Q. & Shaw, G. S. Selective recruitment of an E2~ubiquitin complex by an E3 ubiquitin ligase. *J Biol Chem* **287**, 17374-17385, doi:10.1074/jbc.M112.353748 (2012).

121 Kolman, C. J., Toth, J. & Gonda, D. K. Identification of a portable determinant of cell cycle function within the carboxyl-terminal domain of the yeast CDC34 (UBC3) ubiquitin conjugating (E2) enzyme. *Embo j* **11**, 3081-3090 (1992).
122 Silver, E. T., Gwozd, T. J., Ptak, C., Goebl, M. & Ellison, M. J. A chimeric ubiquitin conjugating enzyme that combines the cell cycle properties of CDC34 (UBC3) and the DNA repair properties of RAD6 (UBC2): implications for the structure, function and evolution of the E2s. *Embo j* **11**, 3091-3098 (1992).
123 Wu, K., Chen, A., Tan, P. & Pan, Z. Q. The Nedd8-conjugated ROC1-CUL1 core ubiquitin ligase utilizes Nedd8 charged surface residues for efficient polyubiquitin chain assembly catalyzed by Cdc34. *J Biol Chem* **277**, 516-527, doi:10.1074/jbc.M108008200 (2002).
124 Kleiger, G., Saha, A., Lewis, S., Kuhlman, B. & Deshaies, R. J. Rapid E2-E3 assembly and disassembly enable processive ubiquitylation of cullin-RING ubiquitin ligase substrates. *Cell* **139**, 957-968, doi:10.1016/j.cell.2009.10.030 (2009).
125 Sandoval, D. *et al.* Ubiquitin-conjugating enzyme Cdc34 and ubiquitin ligase Skp1-cullin-F-box ligase (SCF) interact through multiple conformations. *J Biol Chem* **290**, 1106-1118, doi:10.1074/jbc.M114.615559 (2015).
126 Choi, Y. S. *et al.* The human Cdc34 carboxyl terminus contains a non-covalent ubiquitin binding activity that contributes to SCF-dependent ubiquitination. *J Biol Chem* **285**, 17754-17762, doi:10.1074/jbc.M109.090621 (2010).
127 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).
128 Nalepa, G., Rolfe, M. & Harper, J. W. Drug discovery in the ubiquitin-proteasome system. *Nat Rev Drug Discov* **5**, 596-613, doi:10.1038/nrd2056 (2006).
129 Ceccarelli, D. F. *et al.* An allosteric inhibitor of the human Cdc34 ubiquitin-conjugating enzyme. *Cell* **145**, 1075-1087, doi:10.1016/j.cell.2011.05.039 (2011).
130 St-Cyr, D. *et al.* Identification and optimization of molecular glue compounds that inhibit a noncovalent E2 enzyme–ubiquitin complex. *Science Advances* **7**, eabi5797, doi:doi:10.1126/sciadv.abi5797 (2021).
131 Harper, J. W. & Schulman, B. A. Cullin-RING Ubiquitin Ligase Regulatory Circuits: A Quarter Century Beyond the F-Box Hypothesis. *Annu Rev Biochem* **90**, 403-429, doi:10.1146/annurev-biochem-090120-013613 (2021).
132 Soucy, T. A. *et al.* An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature* **458**, 732-736, doi:10.1038/nature07884 (2009).
133 Bennett, E. J., Rush, J., Gygi, S. P. & Harper, J. W. Dynamics of cullin-RING ubiquitin ligase network revealed by systematic quantitative proteomics. *Cell* **143**, 951-965, doi:10.1016/j.cell.2010.11.017 (2010).
134 Bai, C. *et al.* SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. *Cell* **86**, 263-274, doi:10.1016/s0092-8674(00)80098-7 (1996).
135 Kamura, T. *et al.* The Elongin BC complex interacts with the conserved SOCS-box motif present in members of the SOCS, ras, WD-40 repeat, and ankyrin repeat families. *Genes Dev* **12**, 3872-3881, doi:10.1101/gad.12.24.3872 (1998).

136 Xu, L. *et al.* BTB proteins are substrate-specific adaptors in an SCF-like modular ubiquitin ligase containing CUL-3. *Nature* **425**, 316-321, doi:10.1038/nature01985 (2003).
137 Pintard, L. *et al.* The BTB protein MEL-26 is a substrate-specific adaptor of the CUL-3 ubiquitin-ligase. *Nature* **425**, 311-316, doi:10.1038/nature01959 (2003).
138 Zhuang, M. *et al.* Structures of SPOP-substrate complexes: insights into molecular architectures of BTB-Cul3 ubiquitin ligases. *Mol Cell* **36**, 39-50, doi:10.1016/j.molcel.2009.09.022 (2009).
139 Angers, S. *et al.* Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. *Nature* **443**, 590-593, doi:10.1038/nature05175 (2006).
140 Higa, L. A. *et al.* CUL4-DDB1 ubiquitin ligase interacts with multiple WD40-repeat proteins and regulates histone methylation. *Nat Cell Biol* **8**, 1277-1283, doi:10.1038/ncb1490 (2006).
141 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).
142 Goldenberg, S. J. *et al.* Structure of the Cand1-Cul1-Roc1 complex reveals regulatory mechanisms for the assembly of the multisubunit cullin-dependent ubiquitin ligases. *Cell* **119**, 517-528, doi:10.1016/j.cell.2004.10.019 (2004).
143 Enchev, R. I. *et al.* Structural basis for a reciprocal regulation between SCF and CSN. *Cell Rep* **2**, 616-627, doi:10.1016/j.celrep.2012.08.019 (2012).
144 Cavadini, S. *et al.* Cullin-RING ubiquitin E3 ligase regulation by the COP9 signalosome. *Nature* **531**, 598-603, doi:10.1038/nature17416 (2016).
145 Scott, D. C. *et al.* Structure of a RING E3 trapped in action reveals ligation mechanism for the ubiquitin-like protein NEDD8. *Cell* **157**, 1671-1684 (2014).
146 Duda, D. M. *et al.* Structural insights into NEDD8 activation of cullin-RING ligases: conformational control of conjugation. *Cell* **134**, 995-1006 (2008).
147 Yamoah, K. *et al.* Autoinhibitory regulation of SCF-mediated ubiquitination by human cullin 1's C-terminal tail. *Proceedings of the National Academy of Sciences* **105**, 12230-12235, doi:doi:10.1073/pnas.0806155105 (2008).
148 Liakopoulos, D., Doenges, G., Matuschewski, K. & Jentsch, S. A novel protein modification pathway related to the ubiquitin system. *The EMBO journal* **17**, 2208-2214, doi:10.1093/emboj/17.8.2208 (1998).
149 Osaka, F. *et al.* A new NEDD8-ligating system for cullin-4A. *Genes Dev* **12**, 2263-2268, doi:10.1101/gad.12.15.2263 (1998).
150 Pan, Z. Q., Kentsis, A., Dias, D. C., Yamoah, K. & Wu, K. Nedd8 on cullin: building an expressway to protein destruction. *Oncogene* **23**, 1985-1997, doi:10.1038/sj.onc.1207414 (2004).
151 Saha, A. & Deshaies, R. J. Multimodal activation of the ubiquitin ligase SCF by Nedd8 conjugation. *Mol Cell* **32**, 21-31 (2008).
152 Mosadeghi, R. *et al.* Structural and kinetic analysis of the COP9-Signalosome activation and the cullin-RING ubiquitin ligase deneddylation cycle. *eLife* **5**, e12102, doi:10.7554/eLife.12102 (2016).
153 Faull, S. V. *et al.* Structural basis of Cullin 2 RING E3 ligase regulation by the COP9 signalosome. *Nature Communications* **10**, 3814, doi:10.1038/s41467-019-11772-y (2019).
154 Liu, J., Furukawa, M., Matsumoto, T. & Xiong, Y. NEDD8 modification of CUL1 dissociates p120(CAND1), an inhibitor of CUL1-SKP1 binding and SCF ligases. *Mol Cell* **10**, 1511-1518, doi:10.1016/s1097-2765(02)00783-9 (2002).

155 Cardote, T. A. F., Gadd, M. S. & Ciulli, A. Crystal Structure of the Cul2-Rbx1-EloBC-VHL Ubiquitin Ligase Complex. *Structure* **25**, 901-911.e903, doi:10.1016/j.str.2017.04.009 (2017).
156 Linehan, W. M., Lerman, M. I. & Zbar, B. Identification of the von Hippel-Lindau (VHL) gene. Its role in renal cancer. *Jama* **273**, 564-570 (1995).
157 Latif, F. *et al.* Identification of the von Hippel-Lindau disease tumor suppressor gene. *Science* **260**, 1317-1320, doi:10.1126/science.8493574 (1993).
158 Kishida, T., Stackhouse, T. M., Chen, F., Lerman, M. I. & Zbar, B. Cellular proteins that bind the von Hippel-Lindau disease gene product: mapping of binding domains and the effect of missense mutations. *Cancer Res* **55**, 4544-4548 (1995).
159 Duan, D. R. *et al.* Inhibition of transcription elongation by the VHL tumor suppressor protein. *Science* **269**, 1402-1406, doi:10.1126/science.7660122 (1995).
160 Maxwell, P. H. *et al.* The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature* **399**, 271-275, doi:10.1038/20459 (1999).
161 Ohh, M. *et al.* Ubiquitination of hypoxia-inducible factor requires direct binding to the beta-domain of the von Hippel-Lindau protein. *Nat Cell Biol* **2**, 423-427, doi:10.1038/35017054 (2000).
162 Epstein, A. C. *et al.* C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* **107**, 43-54, doi:10.1016/s0092-8674(01)00507-4 (2001).
163 Jaakkola, P. *et al.* Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. *Science* **292**, 468-472, doi:10.1126/science.1059796 (2001).
164 Zhang, Q. & Yang, H. The Roles of VHL-Dependent Ubiquitination in Signaling and Cancer. *Front Oncol* **2**, 35-35, doi:10.3389/fonc.2012.00035 (2012).
165 Tomita, M. *et al.* Activation of hypoxia-inducible factor 1 in human T-cell leukaemia virus type 1-infected cell lines and primary adult T-cell leukaemia cells. *Biochem J* **406**, 317-323, doi:10.1042/bj20070286 (2007).
166 Kaelin, W. G., Jr. Molecular basis of the VHL hereditary cancer syndrome. *Nat Rev Cancer* **2**, 673-682, doi:10.1038/nrc885 (2002).
167 Anderson, K. *et al.* Regulation of cellular levels of Sprouty2 protein by prolyl hydroxylase domain and von Hippel-Lindau proteins. *J Biol Chem* **286**, 42027-42036, doi:10.1074/jbc.M111.303222 (2011).
168 Mikhaylova, O. *et al.* The von Hippel-Lindau tumor suppressor protein and Egl-9-Type proline hydroxylases regulate the large subunit of RNA polymerase II in response to oxidative stress. *Mol Cell Biol* **28**, 2701-2717, doi:10.1128/mcb.01231-07 (2008).
169 Na, X. *et al.* Identification of the RNA polymerase II subunit hsRPB7 as a novel target of the von Hippel-Lindau protein. *Embo j* **22**, 4249-4259, doi:10.1093/emboj/cdg410 (2003).
170 Zhou, L. & Yang, H. The von Hippel-Lindau tumor suppressor protein promotes c-Cbl-independent poly-ubiquitylation and degradation of the activated EGFR. *PLoS One* **6**, e23936, doi:10.1371/journal.pone.0023936 (2011).
171 Starostina, N. G., Simpliciano, J. M., McGuirk, M. A. & Kipreos, E. T. CRL2(LRR-1) targets a CDK inhibitor for cell cycle control in C. elegans and actin-based motility regulation in human cells. *Dev Cell* **19**, 753-764, doi:10.1016/j.devcel.2010.10.013 (2010).

172 Starostina, N. G. *et al.* A CUL-2 ubiquitin ligase containing three FEM proteins degrades TRA-1 to regulate C. elegans sex determination. *Dev Cell* **13**, 127-139, doi:10.1016/j.devcel.2007.05.008 (2007).
173 Wadelin, F. *et al.* Leucine-rich repeat protein PRAME: expression, potential functions and clinical implications for leukaemia. *Mol Cancer* **9**, 226, doi:10.1186/1476-4598-9-226 (2010).
174 Vasudevan, S., Starostina, N. G. & Kipreos, E. T. The Caenorhabditis elegans cell-cycle regulator ZYG-11 defines a conserved family of CUL-2 complex components. *EMBO Rep* **8**, 279-286, doi:10.1038/sj.embor.7400895 (2007).
175 Li, X. S., Trojer, P., Matsumura, T., Treisman, J. E. & Tanese, N. Mammalian SWI/SNF--a subunit BAF250/ARID1 is an E3 ubiquitin ligase that targets histone H2B. *Mol Cell Biol* **30**, 1673-1688, doi:10.1128/mcb.00540-09 (2010).
176 Liu, Y. V. *et al.* RACK1 competes with HSP90 for binding to HIF-1alpha and is required for O(2)-independent and HSP90 inhibitor-induced degradation of HIF-1alpha. *Mol Cell* **25**, 207-217, doi:10.1016/j.molcel.2007.01.001 (2007).
177 Nguyen, H. C., Yang, H., Fribourgh, J. L., Wolfe, L. S. & Xiong, Y. Insights into Cullin-RING E3 ubiquitin ligase recruitment: structure of the VHL-EloBC-Cul2 complex. *Structure* **23**, 441-449, doi:10.1016/j.str.2014.12.014 (2015).
178 Liu, J. & Nussinov, R. Flexible cullins in cullin-RING E3 ligases allosterically regulate ubiquitination. *J Biol Chem* **286**, 40934-40942, doi:10.1074/jbc.M111.277236 (2011).
179 Onel, M., Sumbul, F., Liu, J., Nussinov, R. & Haliloglu, T. Cullin neddylation may allosterically tune polyubiquitin chain length and topology. *Biochem J* **474**, 781-795, doi:10.1042/bcj20160748 (2017).
180 Békés, M., Langley, D. R. & Crews, C. M. PROTAC targeted protein degraders: the past is prologue. *Nature Reviews Drug Discovery* **21**, 181-200, doi:10.1038/s41573-021-00371-6 (2022).
181 Sakamoto, K. M. *et al.* Protacs: Chimeric molecules that target proteins to the Skp1–Cullin–F box complex for ubiquitination and degradation. *Proceedings of the National Academy of Sciences* **98**, 8554-8559, doi:doi:10.1073/pnas.141230798 (2001).
182 Zengerle, M., Chan, K.-H. & Ciulli, A. Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. *ACS Chemical Biology* **10**, 1770-1777, doi:10.1021/acschembio.5b00216 (2015).
183 Lu, J. *et al.* Hijacking the E3 Ubiquitin Ligase Cereblon to Efficiently Target BRD4. *Chem Biol* **22**, 755-763, doi:10.1016/j.chembiol.2015.05.009 (2015).
184 Huang, X. & Dixit, V. M. Drugging the undruggables: exploring the ubiquitin system for drug development. *Cell Research* **26**, 484-498, doi:10.1038/cr.2016.31 (2016).
185 Deshaies, R. J. Protein degradation: Prime time for PROTACs. *Nat Chem Biol* **11**, 634-635, doi:10.1038/nchembio.1887 (2015).
186 Bondeson, D. P. *et al.* Catalytic in vivo protein knockdown by small-molecule PROTACs. *Nat Chem Biol* **11**, 611-617, doi:10.1038/nchembio.1858 (2015).
187 Sievers, Q. L. *et al.* Defining the human C2H2 zinc finger degrome targeted by thalidomide analogs through CRBN. *Science* **362** (2018).
188 Lu, G. *et al.* The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins. *Science* **343**, 305-309 (2014).
189 Kronke, J. *et al.* Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. *Science* **343**, 301-305 (2014).

190 Gadd, M. S. *et al.* Structural basis of PROTAC cooperative recognition for selective protein degradation. *Nat Chem Biol* **13**, 514-521, doi:10.1038/nchembio.2329 (2017).
191 Lu, G. *et al.* UBE2G1 governs the destruction of cereblon neomorphic substrates. *eLife* **7**, e40958 (2018).
192 Mayor-Ruiz, C. *et al.* Plasticity of the Cullin-RING Ligase Repertoire Shapes Sensitivity to Ligand-Induced Protein Degradation. *Mol Cell* **75**, 849-858.e848, doi:https://doi.org/10.1016/j.molcel.2019.07.013 (2019).
193 Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**, 343-345, doi:10.1038/nmeth.1318 (2009).
194 Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, doi:10.7554/eLife.42166 (2018).
195 Kimanius, D., Dong, L., Sharov, G., Nakane, T. & Scheres, S. H. W. New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochem J* **478**, 4169-4185, doi:10.1042/bcj20210708 (2021).
196 Zhang, K. Gctf: Real-time CTF determination and correction. *J Struct Biol* **193**, 1-12, doi:10.1016/j.jsb.2015.11.003 (2016).
197 Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* **14**, 290-296, doi:10.1038/nmeth.4169 (2017).
198 Sanchez-Garcia, R. *et al.* DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Communications Biology* **4**, 874, doi:10.1038/s42003-021-02399-1 (2021).
199 Rappsilber, J., Ishihama, Y. & Mann, M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem* **75**, 663-670 (2003).
200 Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N. & Mann, M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat Methods* **11**, 319-324 (2014).
201 El Oualid, F. *et al.* Chemical synthesis of ubiquitin, ubiquitin-based probes, and diubiquitin. *Angew Chem Int Ed Engl* **49**, 10149-10153 (2010).
202 Liu, M. *et al.* Improved WATERGATE Pulse Sequences for Solvent Suppression in NMR Spectroscopy. *Journal of Magnetic Resonance* **132**, 125 (1998).
203 Piotto, M., Saudek, V. & Sklenar, V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR* **2**, 661-665, doi:10.1007/BF02192855 (1992).
204 Sklenar, V., Piotto, M., Leppik, R. & Saudek, V. Gradient-tailored water suppression for 1H-15N HSQC experiments optimized to retain full sensitivity. *Journal of Magnetic Resonance* **102**, 241-245 (1993).
205 Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687-696, doi:10.1002/prot.20449 (2005).
206 Leman, J. K. *et al.* Macromolecular modeling and design in Rosetta: recent methods and frameworks. *Nat Methods* **17**, 665-680 (2020).
207 Renfrew, P. D., Choi, E. J., Bonneau, R. & Kuhlman, B. Incorporation of Noncanonical Amino Acids into Rosetta and Use in Computational Protein-Peptide Interface Design. *PLoS One* **7**, e32637 (2012).
208 Amaro, R., Daliwal, B. & Luthey-Schulten, Z. Parameterizing a novel residue. (2014).

209 Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J Comput Chem* **26**, 1781-1802 (2005).
210 Brooks, B. R. *et al.* CHARMM: the biomolecular simulation program. *J Comput Chem* **30**, 1545-1614 (2009).
211 R. J. Gowers *et al.* MDAnalysis: A Python package for the rapid analysis of molecular dynamics simulations. . *Proceedings of the 15th Python in Science Conference*, 98-105 (2016).
212 Liu, P., Agrafiotis, D. K. & Theobald, D. L. Fast determination of the optimal rotational matrix for macromolecular superpositions. *J Comput Chem* **31**, 1561-1563 (2010).
213 Shapovalov, M. V. & Dunbrack, R. L., Jr. A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure* **19**, 844-858 (2011).
214 Pao, K. C. *et al.* Activity-based E3 ligase profiling uncovers an E3 ligase with esterification activity. *Nature* **556**, 381-385 (2018).
215 McKenna, S. *et al.* Noncovalent interaction between ubiquitin and the human DNA repair protein Mms2 is required for Ubc13-mediated polyubiquitination. *J Biol Chem* **276**, 40120-40126 (2001).
216 Skaar, J. R., Pagan, J. K. & Pagano, M. Mechanisms and function of substrate recruitment by F-box proteins. *Nature Reviews Molecular Cell Biology* **14**, 369-381 (2013).
217 Bremm, A., Freund, S. M. V. & Komander, D. Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne. *Nature structural & molecular biology* **17**, 939-947 (2010).
218 Saville, M. K. *et al.* Regulation of p53 by the ubiquitin-conjugating enzymes UbcH5B/C in vivo. *J Biol Chem* **279**, 42169-42181 (2004).
219 Wauer, T. *et al.* Ubiquitin Ser65 phosphorylation affects ubiquitin structure, chain assembly and hydrolysis. *Embo j* **34**, 307-325 (2015).
220 Brzovic, P. S. & Klevit, R. E. Ubiquitin transfer from the E2 perspective: why is UbcH5 so promiscuous? *Cell Cycle* **5**, 2867-2873 (2006).
221 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 (2017).
222 Garg, P. *et al.* Structural and Functional Analysis of Ubiquitin-based Inhibitors That Target the Backsides of E2 Enzymes. *J Mol Biol* **432**, 952-966 (2020).
223 Rout, M. K. *et al.* Stochastic gate dynamics regulate the catalytic activity of ubiquitination enzymes. *J Am Chem Soc* **136**, 17446-17458 (2014).
224 Willems, A. R. *et al.* Cdc53 targets phosphorylated G1 cyclins for degradation by the ubiquitin proteolytic pathway. *Cell* **86**, 453-463, doi:10.1016/s0092-8674(00)80118-x (1996).
225 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 (2011).
226 Sievers, Q. L., Gasser, J. A., Cowley, G. S., Fischer, E. S. & Ebert, B. L. Genome-wide screen identifies cullin-RING ligase machinery required for lenalidomide-dependent CRL4(CRBN) activity. *Blood* **132**, 1293-1303, doi:10.1182/blood-2018-01-821769 (2018).
227 Li, G., Liang, Q., Gong, P., Tencer, A. H. & Zhuang, Z. Activity-based diubiquitin probes for elucidating the linkage specificity of deubiquitinating

enzymes. *Chem Commun (Camb)* **50**, 216-218, doi:10.1039/c3cc47382a (2014).
228 Liwocha, J. *et al.* Linkage-specific ubiquitin chain formation depends on a lysine hydrocarbon ruler. *Nat Chem Biol* **17**, 272-279, doi:10.1038/s41589-020-00696-0 (2021).
229 Min, J. H. *et al.* Structure of an HIF-1alpha -pVHL complex: hydroxyproline recognition in signaling. *Science* **296**, 1886-1889, doi:10.1126/science.1073440 (2002).
230 Lv, Z., Williams, K. M., Yuan, L., Atkison, J. H. & Olsen, S. K. Crystal structure of a human ubiquitin E1-ubiquitin complex reveals conserved functional elements essential for activity. *J Biol Chem* **293**, 18337-18352, doi:10.1074/jbc.RA118.003975 (2018).
231 Kidmose, R. T. *et al.* Namdinator - automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *IUCrJ* **6**, 526-531, doi:10.1107/s2052252519007619 (2019).
232 Kastner, B. *et al.* GraFix: sample preparation for single-particle electron cryomicroscopy. *Nat Methods* **5**, 53-55, doi:10.1038/nmeth1139 (2008).
233 Victor Bernier-Villamor, D. A. S., Michael J.Matunis, Christopher D.Lima. Structural Basis for E2-Mediated SUMO Conjugation Revealed by a Complex between Ubiquitin-Conjugating Enzyme Ubc9 and RanGAP1. *Cell* **108**, 345-356 (2002).
234 Cappadocia, L. & Lima, C. D. Ubiquitin-like Protein Conjugation: Structures, Chemistry, and Mechanism. *Chem Rev* **118**, 889-918 (2018).
235 Wang, M., Cheng, D., Peng, J. & Pickart, C. M. Molecular determinants of polyubiquitin linkage selection by an HECT ubiquitin ligase. *Embo j* **25**, 1710-1719 (2006).
236 Burslem, G. M. & Crews, C. M. Proteolysis-Targeting Chimeras as Therapeutics and Tools for Biological Discovery. *Cell* (2020).
237 Verma, R., Mohl, D. & Deshaies, R. J. Harnessing the Power of Proteolysis for Targeted Protein Inactivation. *Mol Cell* **77**, 446-460 (2020).
238 Streich, F. C., Jr. & Lima, C. D. Capturing a substrate in an activated RING E3/E2-SUMO complex. *Nature* **536**, 304-308 (2016).