# Development of Optochemical Strategies to Study Proteome-Wide Ubiquitination Dynamics in a Linkage-Specific Manner and Target Protein-Specific Ubiquitination Events

Dissertation

Submitted for the degree of Doctor of Natural Sciences (Dr. rer. nat)

to the

Faculty of Chemistry and Chemical Biology

Technical University Dortmund

By

Banerjee, Sudakshina

সুদক্ষিণা

Dortmund, 2025

*“Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less.”*

– Marie Curie

This work was prepared from August 2019 to April 2024 in the group of Prof. Daniel Summerer of the Faculty of Chemistry and Chemical Biology at the Technical University Dortmund.

The research projects presented in this work are funded by the European Research Council (ERC) under the ERC consolidator grant EPICODE No. 723863 to Prof. Dr. Daniel Summerer.

Plasmid pASB654 – SE323_wtPylRS_4xPylT (p2970) was kindly provided by Dr. Heinz Neumann.

The compounds used in UPS inhibitor studies (MLN7243, MLN4924 and NMS873) were kindly provided by the lab of Dr. Malte Gersch.

The initial tests of ubiquitination mutants of mTET2CD (K1212N/E) were performed by Shubhendu Palei, in the Summerer group.

# Acknowledgements

The journey of PhD has been one of the most challenging and also the most rewarding journeys of my life. It has largely shaped the person I am today, not only from a professional point of view, but also on a personal level. I could tread on this tricky path and come out successfully largely because of the unwavering support and guidance I received throughout.

First and foremost, I extend my sincere appreciation to my supervisor, Prof. Dr. Daniel Summerer, for making me a part of his wonderful working group and involving me in exciting projects in the field of chemical biology and genetic code expansion. I am grateful for his invaluable scientific supervision and guidance, and even more for his continued support and encouragement even during challenging times. His trust in me and my capabilities has given me the confidence to go forward and the freedom to explore the unknown biological questions.

I would like to express my gratitude to Dr. Malte Gersch who has been an integral part of my PhD projects as my collaborator. I am grateful for his invaluable guidance and expertise in the field of ubiquitin. I thank him for letting me use his resources and lab space for the initial experiments. I thank him for his continuous support and contribution as a member of my thesis advisory committee, and also for agreeing to be the second examiner for this thesis. I am also grateful to his group, especially, Kai Gallant, also for their support with ubiquitination experiments.

I also want to express my cordial thanks to Dr. Leif Dehmelt for his support and guidance as a member of my thesis advisory committee.

I am more than grateful to have been a part of the wonderful Summerer Group, being surrounded by brilliant colleagues who have always been open to all kinds of discussions, both scientific and non-scientific. I have received valuable insights from random after-work chat sessions and huge mental support. I would not have been able to adjust in a foreign land with so much ease if it weren't for my colleagues, who have become more of friends. Former and current Summerer group members who I was lucky to have worked with, Dr. Álvaro Muñoz-López, Dr. Anna Witte, Dr. Anne Jung, Dr. Benjamin Buchmuller, Brinja Kosel, Dr. Damian Schiller, Dr. Jan Wolfgramm, Katrin Bigler, Kotryna Keliuotyté, Lejla Maksumic, Lena Engelhardt, Nadine Schmidt, Sayan Sil, Dr. Shubhendu Palei, Simone Eppmann, Dr. Tzu-Chen Lin and Zeyneb Vildan Çakil. I would particularly like to thank Shubhendu who started the TET ubiquitination project; Zeyneb for her strong support during the revision process of the publication; Jan who really supported me in my initial days and warmed me up to the rest of the group; Damian for being an excellent office-mate; Sayan for being a huge support during my last phase and last but not the least, Anne, Tzu and Lena who have become integral parts of my life as invaluable friends. I also am very grateful to Lena for proof-reading my thesis.

My sincere gratitude goes to the IMPRS-LM PhD program for providing all the valuable workshops/lectures/trainings, travel fundings, and social events that bring the students together. My heartfelt acknowledgements to the IMPRS-LM coordination office, Dr. Lucia Sironi and Christa Hornemann, for organizing all the wonderful events that enriched my PhD journey. Christa has also been a huge support and I have received important guidance in terms of living life in Germany from her. I am immensely thankful to her for printing my thesis and coordinating the entire submission process.

I would also like to thank Maria Sergani, Martina Reibner, Petra Alhorn and Ulrich Schoppe from the TU Dortmund CCB Sekretariat for their help with all kinds of general management and bureaucracy.

My sincere appreciation also goes to Dr. Devrani Mitra, my Masters supervisor and mentor, without whom I don't think I would be here pursuing a PhD. She has ignited the scientist in me and taught me a lot about the life of a scientist. Also, my base of molecular biology and biochemistry was laid in her lab of structural optogenetics at Presidency University, Kolkata.

An integral part of my PhD journey were my friends, Saheli Roy, Dr. Gargi Das and Shayan Sarkar, who, although were living far away from me, were always there in spirit, supporting me in my tough times and always giving me things to laugh about. I am thankful to my friends and family in Germany, Dr. Rida Ahmad, Prasun Banerjee, Shreyasi Banerjee and little Pratyusha Banerjee who made home feel closer than it really was.

I also want to thank my fiancé, Rajarshi Sinha, who has been standing by me like a rock since the day we got to know each other, supporting me in everything I do.

Last but not the least, nothing would ever be possible without the support of my parents, my late grandparents and my maternal family (Gautam Mukherjee, Sonali Mukherjee, Bosudhara Mukherjee, Arindam Mukherjee and Sutapa Ray Mukherjee). My parents, Susanta Kumar Banerjee and Piyali Banerjee, have seen it all, the good and the bad, and have been my biggest support in everything, consoling and celebrating with me at every step. I am nothing without them.

And of course, I am thankful to God for blessing me with the life I have.

Thank you all for being part of this journey.

# Publications

Part of this work has been published in:

**Banerjee, S.**, Cakil, Z. V., Gallant, K., van den Boom, K., Palei, S., Meyer, H., Gersch, M.* and Summerer, D. (2024). Light-Activatable Ubiquitin for Studying Linkage-Specific Ubiquitin Chain Formation Kinetics. *Adv. Sci.* 2024, 2406570. doi: 10.1002/advs.202406570

Other publications:

Jung, A., Muñoz-López, Á., Buchmuller, B., **Banerjee, S**. and Summerer, D. (2023). Imaging-based in situ Analysis of 5-Methylcytosine at Low Repetitive Single Gene Loci with Transcription-Activator-Like Effector Probes. *ACS Chem. Biol.,* 2023, *18*(2), 230-36. doi: 10.1021/acschembio.2c00857

**Banerjee, S.** and Mitra, D. (2019). Structural basis of design and engineering for advanced plant optogenetics. *Trends in Plant Science*. 25, 35-65. doi: 10.1016/j.tplants.2019.10.002

# Table of Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| °C | Degree Celsius |
| α-KG | α-Ketoglutarate |
| 5caC | 5-Carboxylcytosine |
| 5fC | 5-Formylcytosine |
| 5hmC | 5-Hydroxymethylcytosine |
| 5mC | 5-Methylcytosine |
| A | Adenine |
| aaRS | Aminoacyl-tRNA synthetase |
| AM | Active modification |
| AML | Acute myeloid leukemia |
| AR | Active removal |
| BER | Base excision repair |
| BLUF | Blue light using FAD |
| BODIPY | Borondipyrromethene |
| BONCAT | Bio-orthogonal non-canonical amino acid tagging |
| BSA | Bovine serum albumin |
| C | Cytosine |
| Carb | Carbenicillin |
| CC | Coiled coil |
| CD | Catalytic domain |
| CGIs | CpG islands |
| ChIP | Chromatin immunoprecipitation |
| ChR-2 | Channelrhodopsin-2 |
| CID | Chemical inducers of dimerization |
| Co-IP | Co-immunoprecipitation |
| CRY | Cryptochrome |
| CXXC | Cys-x-x-Cys |
| Cys-rich | Cysteine-rich |
| DDMs | DNA demethylases |
| DMEM | Dulbecco's Modified Eagle Medium |
| DMNB | 4, 5-Dimethoxy-2-nitrobenzyl |
| DMRs | Differential methylated regions |
| DNA | Deoxyribonucleic acid |
| DNMTs | DNA-methyltransferases |
| dNTP | Deoxynucleoside triphosphate |
| DPBS | Dulbecco's phosphate-buffered saline |

| | |
|---|---|
| DSBH | Double-stranded β-helix |
| DUB | Deubiquitinase |
| E. coli | Escherichia coli |
| EDTA | Ethylenediaminetetraacetic acid |
| ESCs | Embryonic stem cells |
| FACS | Fluorescence-activated cell sorting |
| FAD | Flavin adenine dinucleotide |
| FBS | Fetal bovine serum |
| FCM | Flow cytometry |
| FL | Full length |
| FMN | Flavin mononucleotide |
| G | Guanidine |
| H(s) | Hour(s) |
| HATs | Histone acetyltransferases |
| HDAC | Histone deacetylase |
| HDMs | Histone demethylases |
| HEK293T | Human embryonic kidney 293T |
| HEPES | 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid |
| HMTs | Histone methyltransferases |
| HP1 | Heterochromatin protein 1 |
| JAMM | JAB1/MPN/MOV34 metalloenzymes |
| K0 | Lysine-less |
| KLF4 | Krüppel-like factor 4 |
| LeuRS/LRS | Leucyl tRNA synthetase |
| LINE | s Long interspersed nuclear elements |
| LOV | Light-oxygen-voltage |
| LTR | Long terminal repeat |
| MBD | Methyl-CpG binding domain |
| MeDIP | Methylated DNA immunoprecipitation |
| mESCs | Mouse embryonic stem cells |
| MFI | Mean fluorescence intensities |
| Min(s) | Minute(s) |
| MINDY | Motif Interacting with Ub-containing DUB family |
| MJD | Machado-Joseph domain containing proteases |
| mRNA | Messenger RNA |
| ncAA | Non-canonical amino acids |
| NMR | Nuclear magnetic resonance |
| NuRD | Nucleosome remodeling and histone deacetylation |

| | |
|---|---|
| OTUs | Ovarian Tumour Proteases |
| pcK | Photocaged Lysine |
| PCR | Polymerase chain reaction |
| PEI | Polyethylenimine |
| PGCs | Primordial germ cells |
| PHY | Phytochrome |
| PPARγ | Peroxisome proliferator-activated receptor-γ |
| PPGs | Photocleavable protecting groups |
| PTL | Photoswitched tethered ligand |
| Pyl | Pyrrolysine |
| PylRS | Pyrrolysyl tRNA synthetase |
| RNA | Ribonucleic acid |
| RT | Room temperature |
| RTT | Rett syndrome |
| SAH | S-adenosyl-L-homocysteine |
| SAM | S-adenosyl-L-methionine |
| SeC | Selenocysteine |
| SELEX | Systematic evolution of ligands by exponential enrichment |
| SOC | medium Super Optimal Broth medium |
| T | Thymine |
| TALE | Transcription factor-like effector |
| TDG | Thymine DNA glycosylase |
| TETs | ten-eleven-translocation dioxygenases |
| TFs | Transcription factors |
| TRD | Transcriptional repressor domain |
| Tris | Tris(hydroxymethyl)aminomethane |
| tRNA | Transfer RNA |
| TSSs | Transcription start sites |
| TyrRS | Tyrosyl tRNA synthetase |
| U | Uracil |
| Ub | Ubiquitin |
| UCHs | Ub C-terminal hydrolases |
| USPs | Ub-specific proteases |
| Wt | Wild type |
| ZUP-1 | Zinc-finger containing Ub peptidase-1 |

# 1. Abstract

Proteins are the primary drivers of cellular processes and normal cellular functioning depends on their precise balance of homeostasis and degradation. Protein degradation is largely brought about by ubiquitination, a post-translational modification that adds multiple molecules of the protein Ubiquitin (Ub) to substrates and targets them for proteasome-mediated degradation. Ubiquitination is a highly conserved, complex phenomenon involving multiple enzymatic reactions that covalently adds Ub to substrates via an isopeptide linkage between the ε-amino group of a substrate lysine (Lys, K) and the carboxyl group of the terminal glycine (Gly, G) of Ub. Ub itself has seven Lys on which more Ub molecules can be added giving rise to polyUb chains. Initially discovered as an important mediator of protein degradation and thus turn over, subsequent studies have revealed that ubiquitination also plays important roles in cellular signal transduction. Linkage-specificity of Ub chains is derived from the specific Lys on which chains are built and their specific structure, giving rise to a complex code of Ub chains Consequently, this linkage-specificity decides the biological outcome of the ubiquitination event. Due to the complexity and inter-twined dynamics of ubiquitination, the study of its linkage specific kinetics becomes largely significant. However, due to its highly transient behaviour, studying protein ubiquitination *in vivo* has proved to be challenging and is often met with background noise from endogenous Ub and the various possibilities of Ub chain structure formation. In this study, a light-activated Ub has been developed by incorporating a photocaged lysine at specific sites, through amber codon suppression, for the monitoring of proteome-wide linkage-specific polyubiquitination and to gain insights into the kinetics of the same. This study reveals rapid, minute-scale ubiquitination kinetics for Lys11 (K11), Lys48 (K48) and Lys63 (K63) linkages. Also, the roles of individual components of the ubiquitin-proteasome system have been studied in K48-initiated chain synthesis by small molecule inhibition. This approach expands the repertoire of current cellular ubiquitination perturbation strategies with the ability to control linkage-specific ubiquitination at high temporal resolution and is a promising tool for studying ubiquitinome dynamics. Furthermore, in this study, advances were made towards looking into the kinetics of a monoubiquitination event in Ten-Eleven translocation (TET) proteins.

# 1. Zusammenfassung

Proteine sind die treibenden Kräfte zellulärer Prozesse. Die normale Zellfunktion beruht auf einem präzisen Gleichgewicht zwischen Homöostase und Abbau. Der Proteinabbau erfolgt größtenteils durch Ubiquitinierung, einer posttranslationalen Modifikation, bei der mehrere Moleküle des Proteins Ubiquitin (Ub) an Substrate angehängt werden, wodurch diese für den proteasomalen Abbau markiert werden. Die Ubiquitinierung ist ein hochkonserviertes, komplexes Phänomen, an dem mehrere enzymatische Reaktionen beteiligt sind, wodurch Ub kovalent an Substrate angehängt wird. Das geschieht durch eine Isopeptidbindung zwischen der ε-Aminogruppe eines Lysins (Lys, K) des Substrats und der Carboxylgruppe des terminalen Glycins (Gly, G) des Ub. Ub selbst hat sieben Lys an die weitere Ub-Moleküle angehängt werden können, so dass Poly-Ub-Ketten entstehen. Ursprünglich wurde die Ubiquitinierung als wichtiger Vermittler des Proteinabbaus, und damit des Proteinumsatzes, entdeckt, darauffolgende Studien haben dann gezeigt, dass Ubiquitinierung auch in der zellulären Signaltransduktion eine wichtige Rolle spielt. Die Verknüpfungsspezifität der Ub-Ketten ergibt sich aus dem spezifischen Lys, an dem die Ketten gebildet werden und der spezifischen Struktur, was zu einem komplexen Ub-Kettencode führt. Folglich bestimmt die Verknüpfungsspezifizität den biologischen Effekt des Ubiquitinierungsereignisses. Aufgrund der Komplexität und der verflochtenen Dynamiken der Ubiquitinierung ist die Untersuchung ihrer bindungsspezifischen Kinetik von großer Bedeutung. Die Untersuchung der Ubiquitinierung von Proteinen hat sich aufgrund ihres kurzlebigen Vorkommens in vivo jedoch als schwierig erwiesen, unter anderem durch Hintergrundrauschen von endogenem Ub und den vielfältigen Möglichkeiten der Ub-Kettenformation. In dieser Studie wurde ein Licht-aktivierbares Ub entwickelt, indem ein photocagiertes Lysin in spezifischen Seiten durch Amber-Codon Unterdrückung integriert wurde, sodass verknüpfungsspezifische Polyubiquitinierung proteomweit beobachtet und Einblick in deren Kinetik gewonnen werden konnte. Diese Studie zeigt minutenschnelle Ubiquitinierungskinetiken für Lys11 (K11)-, Lys48 (K48)- und Lys63 (K63)- Verbindungen. Zusätzlich wurden die Rollen der individuellen Komponenten des Ubiquitin-Proteasom Systems in K48 initiierter Kettensynthese durch Hemmung durch kleine Moleküle untersucht. Dieser Ansatz erweitert das Repertoire der aktuellen Störmethoden zellulärer Ubiquitinierung mit der Möglichkeit verknüpfungsspezifische Ubiqutinierung mit hoher zeitlicher Auflösung zu kontrollieren und ist damit ein vielversprechendes Werkzeug, um Ubiquitindynamiken zu untersuchen. Außerdem wurden in dieser Studie Fortschritte in Bezug auf die Untersuchung der Kinetik eines Monoubiquitinierungsereignisses in Ten-Eleven Translokation (TET) Proteinen gemacht.

# 2. Ubiquitination: The Universal PTM

Ubiquitination (also referred to as ubiquitylation or ubiquitinylation) is a protein post-translational modification (PTM) that covalently adds molecules of the protein Ubiquitin (Ub) in single or multiple copies on protein substrates, and majorly targets them for proteasomal degradation, which together forms the ubiquitin-proteasome system (UPS). Such Ub-mediated degradation of regulatory proteins has implications in controlling numerous processes, including cell-cycle progression, signal transduction, transcriptional regulation, receptor down-regulation, and endocytosis. However, other than degradation, ubiquitination has also been found to be involved in non-proteolytic pathways, including DNA damage response, epigenetic regulation, immune responses and many more (**Figure 1**). The function of ubiquitination is essentially dependent on the "ubiquitin code" that a substrate protein harbors. True to its nomenclature, Ub is indeed ubiquitously present in all eukaryotic life-forms, owing to its central role in regulation of cellular functions. In this chapter, the basis of Ub structure and function and how the "ubiquitin code" is written, read and erased has been delineated. Later on, the focus will also be on approaches that have been employed in regulating this universal biological phenomenon in order to selectively study Ub chain formation and their kinetics.

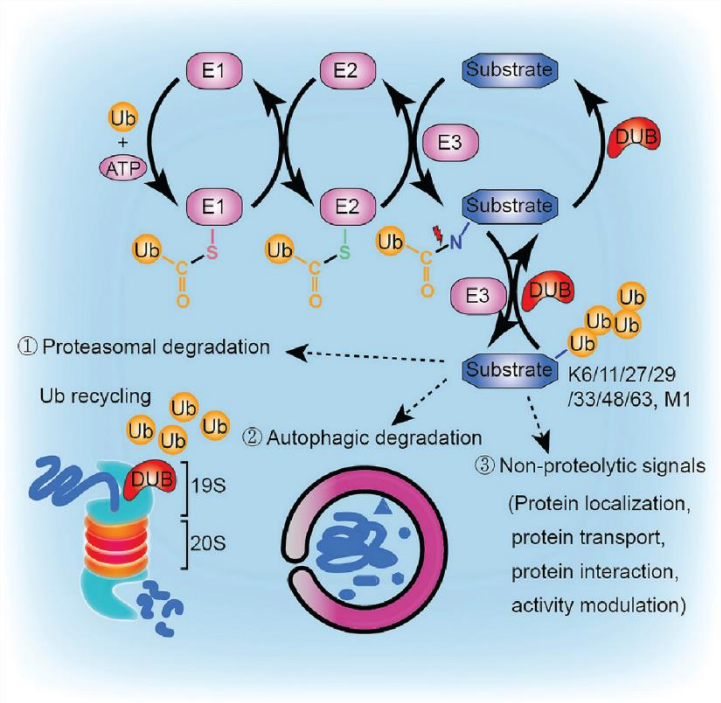



**Figure 1: Schematic diagram showing the overview of the Ubiquitin system.** Ubiquitin (Ub) molecules can be covalently attached to target proteins through a series of enzymatic steps involving Ub-activating (E1), Ub-conjugating (E2), and Ub-ligase E3 enzymes. The process of ubiquitination can be reversed by deubiquitinases (DUBs). Different types of ubiquitin linkages, including K6, K11, K27, K29, K33, K48, K63 and M1, play various roles in the regulation of cellular processes, such as

proteasomal degradation, autophagic degradation, and non-proteolytic signaling. Figure has been adapted from (Xue et al., 2023).

## 2.1 The Ubiquitin System

Ubiquitination involves conjugation of Ub to target proteins via an isopeptide bond that connects the carboxyl group of the Ub C-terminal glycine (Gly, G) with the $N^{e}$-amino group of lysine (Lys, K) residues on target proteins or adjacent Ub molecules. This process is orchestrated by a network of enzymes that involves the activating E1 enzyme, different conjugating E2 enzymes and ligating E3 enzymes (discussed in following sections). This chain of enzymatic reactions catalyze the formation of an isopeptide bond between the C-terminal Gly of Ub and usually a substrate Lys, initially leading to monoubiquitination (**monoUb**). Monoubiquitination can occur at a defined residue, or it might be confined to a domain (Komander & Rape, 2012). More Ub molecules are added either to the same substrate-linked monoUb or other Lys on the substrate to give rise to different polyUb topologies. Over the last few decades, it has been well established that the different forms of mono- and polyUb trigger distinct outcomes in the cell, suggesting that ubiquitination acts as a code to store and transmit information, rightfully referred to as the "ubiquitin code" by Komander and Rape. This code is read by many proteins that bind to the different structural configurations provided by the different Ub conjugations and is translated into different downstream biological outcomes.

### 2.1.1 Structure of Ubiquitin: Gene and Mono- and Oligomeric Chain Structure

Ub is a 76-amino acid protein that is highly conserved in all eukaryotic organisms and the differences between animal, plant, and fungal Ub are two or three residues (Callis and Vierstra, 1989). Ub derives from four genes that encode linear fusion proteins incorporating one or more Ub molecules, from which free Ub is generated by cleavage of the peptide bond (Callis, 2014; Clague et al., 2019; Grou et al., 2015). This *de novo* synthesis of Ub can be as two different types of translational fusions: homomeric fusions and heteromeric fusions. Homomeric fusions are multimers of Ub coding regions repeated head-to-tail with no intervening amino acids (called poly-Ub) with the last Ub terminating with one to a few additional amino acids (**Figure 2a**). In mammals, genes UBB and UBC are poly-Ubs with 3–4 and 7–10 Ubs, respectively, depending on the organism, followed by a variable C-terminal extension that goes from a single amino acid to a few dozen (Larsen et al., 1998). For heteromeric fusions, the 76 amino acid Ub is followed by a different protein in-frame. In mammals, UBA52 and UBA80 comprise a single Ub molecule C-terminally fused to ribosomal proteins L40 and S27A, respectively (Redman

and Rechsteiner, 1989; Baker and Board, 1991) (**Figure 2a**). Ub with even one additional residue at the C-terminus is unable to be recognised by the conjugation pathway. This means that the initial translation products of Ub are non-functional until cleaved to release the active 76-amino acid protein. Several types of Ub-specific proteases or deubiquitinases (DUBs) are capable of processing initial Ub-fusion translation products. These hydrolytic enzymes cleave specifically after Gly76. DUBs have been discussed in more details in the following sub-section. Heteromeric Ub fusions are mostly processed post-translationally whereas processing of poly-Ubs probably occurs through a combination of co- and post-translational mechanisms (Grou et al., 2015). UCHL3 is one of the major enzymes involved in processing heteromeric Ub fusions. Not only UCHL3, Grou et al. also suggested that USP9X (in mouse liver) and USP7 (in HeLa cells) are also highly active in processing these Ub fusions. For processing polyUbs on the other hand, USP5 has been implicated (Falquet et al., 1995). They also have been found to cleave off C-terminal extensions that often follow polyUbs (Grou et al., 2015). The other DUB displaying a large activity towards poly-Ubs is Otulin. Otulin cleaves both anchored and unanchored linear ubiquitin chains and thus, unlike USP5, also processes ribosome-associated nascent polyUbs, provided that they contain two adjacent folded ubiquitin moieties (Keusekotten et al., 2013).

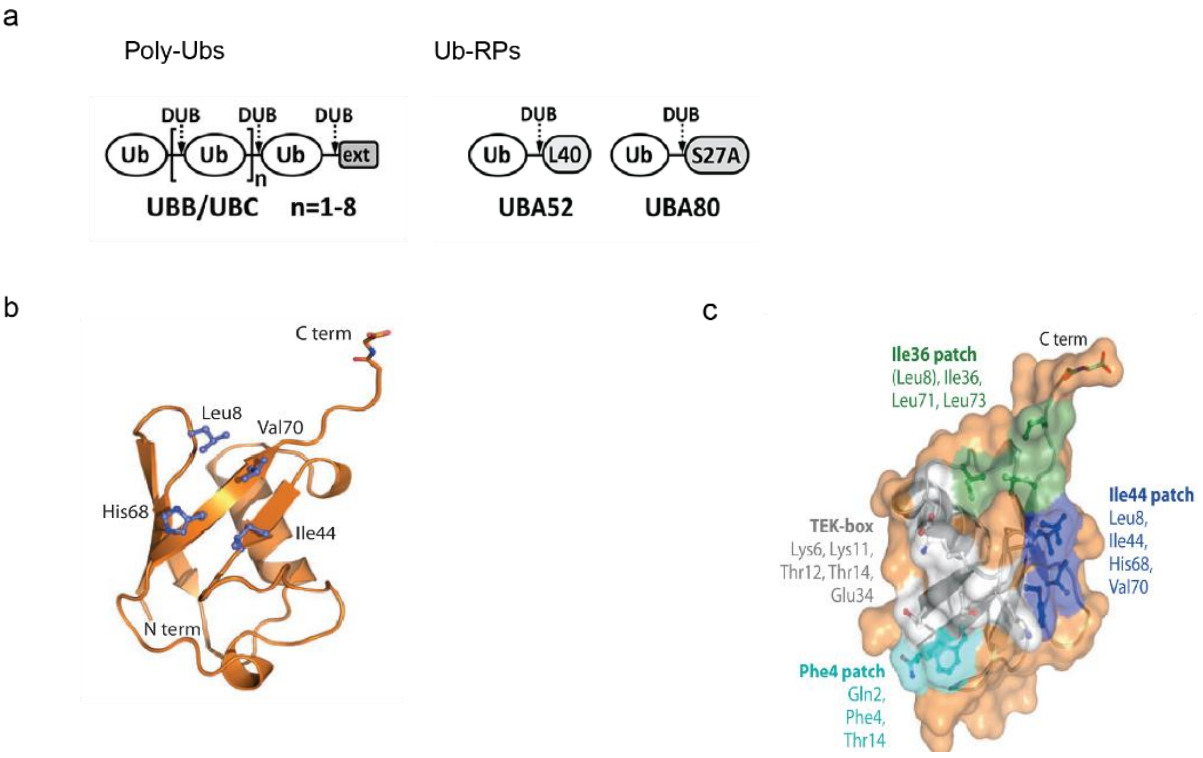



**Figure 2: Ubiquitin structure a.** *De novo* synthesis of Ub as homomeric and heteromeric fusions, respectively. Figure has been created based on (Grou et al., 2015) **b.** Tertiary structure of Ub indicating the residues of the Ile44 patch [Protein Data Bank (PDB) ID **1UBQ**] (Vijay-Kumar et al., 1987). Figure has been adapted with permission from (Komander & Rape, 2012) (A.4) **c.** The Ub surface highlighting

Ile44 (*blue*), Ile36 (*green*), Phe4 patches (*cyan*), and TEK-box (*white*). Figure has been adapted with permission from (Komander & Rape, 2012) (A.4).

After being translated, Ub adopts a highly conserved and stable, compact β-grasp fold with a flexible six-residue C-terminal tail (**Figure 2b**) (Vijay-Kumar et al., 1987). Most of the core residues of the β-grasp are rigid, but the β1/β2 loop containing leucine (Leu, L) 8 shows flexibility and is important for recognition by Ub-binding proteins (**Figure 2b, c**) (Lange et al., 2008). Ub has a hydrophobic surface that consists of isoleucine (Ile, I) 44, Leu8, valine (Val, V) 70, and histidine (His, H) 68 (**Figure 2c**) (Dikic et al., 2009). The Ile44 patch is bound by the proteasome and most Ub binding domains (UBDs). Another hydrophobic surface is around Ile36 and involves Leu71 and Leu73 of the Ub tail (**Figure 2c**). This Ile36 patch is involved in interactions between Ub molecules in chains, and it is recognized by certain E3 enzymes (like the HECT E3s (Kamadurai et al., 2009), DUBs (Hu et al., 2002), and also UBDs (Reyes-Turcu et al., 2006). The Ub face surface comprising phenylalanine (Phe, F) 4 might function in trafficking (Sloper-Mould et al., 2001), and it interacts with the UBAN domain (Rahighi et al., 2009) and the ubiquitin-specific protease (USP) domain of DUBs. Differences in Phe4 patches of Ub and its closest homolog Nedd8 enables distinction by DUBs (Ye et al., 2011). In higher eukaryotes, the TEK-box of Ub, a three-dimensional motif that includes threonine (Thr, T) 12, Thr14, glutamic acid (Glu, E) 34, Lys6, and Lys11, is required for mitotic degradation (**Figure 2c**) (Jin et al., 2008). The most important features of Ub in terms of polyUb chain formation are its N-terminal Met and its seven Lys (Lys-6, 11, 23, 27, 33, 48, 63), which are the attachment sites of adjacent Ubs for chain assembly. These residues cover all surfaces of Ub and point into distinct directions. Lys6 and Lys11 are located in the most dynamic region of Ub and as Lys27 is buried, linkage assembly through this residue requires localized changes in Ub structure (Komander & Rape, 2012).

As mentioned earlier, monoubiquitination can occur at a defined residue, or it might be confined to a domain of the substrate. Alternatively, it is also possible that multiple substrate Lys are decorated with one Ub each giving rise to multimonoubiquitination (**multimonoUb**) (**Figure 3a**). Modification of the N-terminus or one of the seven Lys residues of a substrate-attached Ub leads to formation of polymeric chains (polyubiquitination, **polyUb**). These polyUb chains can be short with only two Ub molecules or long with more than ten Ub, and can be **homogenous**, if the same residue is modified during elongation, like Met1- (or linear), Lys11-, Lys48- or Lys63-linked chains, where Ub molecules are added in a linkage-specific manner. On the other hand, chains can be **heterogenous** and have mixed or branched topology if new Ub molecules are added at different sites of the preceding Ub molecule (**Figure 3a**). Ub chains can adopt either "compact" conformations, where adjacent molecules are in proximity and can form interactions, or "open" conformations, where no interfaces of interaction are present except at the site where the isopeptide bond is formed. The most abundant homogenous Lys48-linked chains adopt compact conformations (**Figure 3b**). In Lys48-linked diUb, the Ub

moieties interact via their Ile44 patches (Tenno et al., 2004), and two such diUbs pack tightly into a tetraUb (Eddins et al., 2007). This structural flexibility allows readers of Lys48-linked chains to access the Ile44 patch, that is major binding interface for many UBDs, as mentioned above. Similar to Lys48 linkages, Lys6- and Lys11- linked chains also have compact conformations, with Lys11-linked chains also displaying structural flexibility and allowing readers to access the Ile44 patch (Bremm et al., 2010). In contrast to these linkages, Met1- (linear) and Lys63-linked chains mostly display open conformations, resembling beads on a string (**Figure 3b**), as shown by NMR analysis of Lys63-linked Ub (Tenno et al., 2004) and crystal structures of both chain types (Komander et al., 2009; Datta et al., 2009). The extended open conformation gives these linkage types high conformational freedom. Most readers of these linkage types utilise the distance and flexibility between the Ub molecules, rather than recognizing assemblies of surface residues. The structural diversity of the various modifications, therefore, forms the foundation of the Ub code.

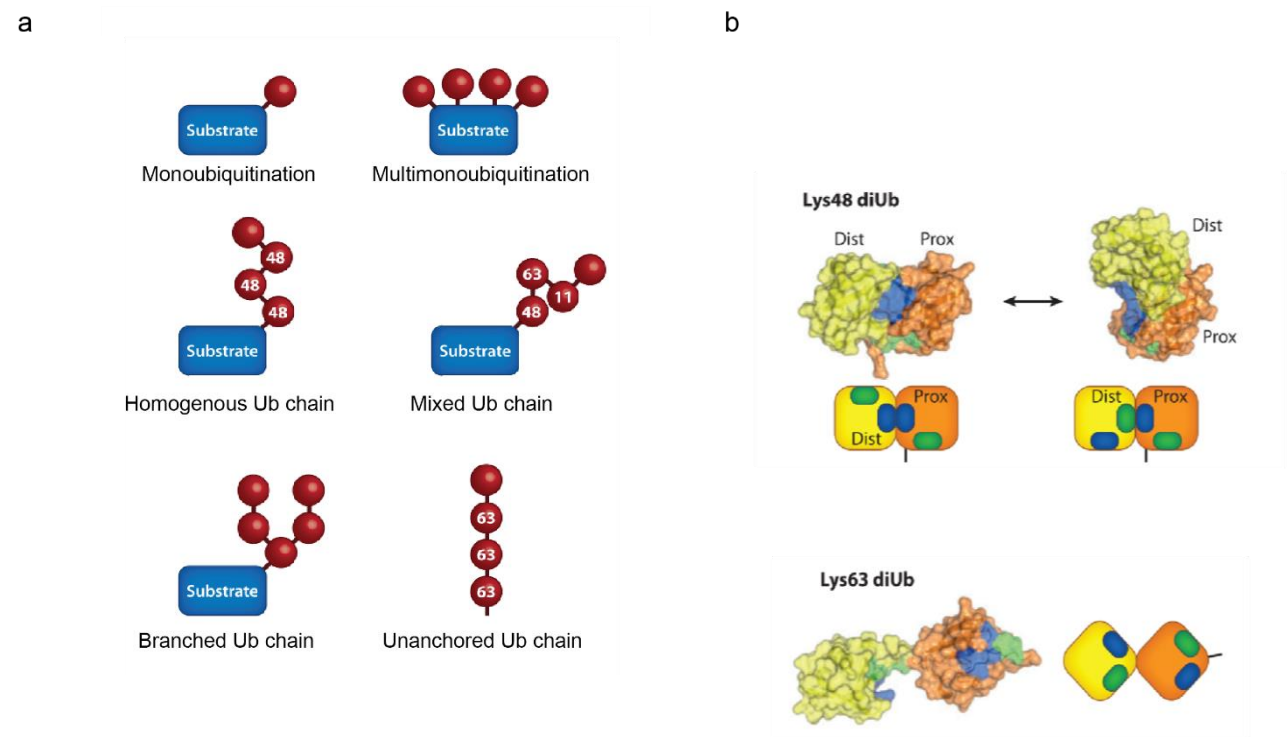



**Figure 3: Ubiquitin chain structure a.** Types of Ubiquitination: mono- and polyubiquitination. **b.** "compact and "open" Ub chain conformations that affect accessibility for reader binding. Figure has been adapted with permission from (Komander & Rape, 2012) (A.4).

### 2.1.2 Ubiquitin Code: Writers, Readers and Erasers

Like any other functional code, the Ub code can induce its specific outcomes only if properly assembled. And this requires not only enzymes that "write" the code, but also proteins that can

"read" the code and eventually aid in the intended biological outcomes in the cells, as well as, enzymes that can "erase" it, to be able to maintain homeostasis. The responsible enzymes catalyze the formation of largely the same product each time they act on their substrate. In this section, the process of how Ub code is written, erased and read by different modules, has been described.

#### 2.1.2.1 Writers of the Ub Code

As mentioned briefly before, Ub is added on to its substrates by the concerted action of Ub activating (E1), conjugating (E2) and ligating (E3) enzymes.

The first enzyme of the Ub conjugation cascade is called Ub activating enzyme (UAE) or E1. In the first step, E1 binds adenosie triphosphate (ATP), and then an adenosine monophosphate (AMP) is ligated to the Ub C-terminus, forming a Ub adenylate that remains noncovalently bound to E1 and releasing pyrophosphate (**Figure 1, 4**). Secondly, the Ub is transferred to an active-site cysteine (Cys, C) of E1, exchanging the acyl phosphate anhydride linkage for a thioester bond, making the AMP leave the enzyme. When one Ub remains thioester-linked to E1, another Ub adenylate is formed, filling the Ub adenylate binding site (Haas and Rose, 1982; Haas et al., 1982a; Haas et al., 1982b). The Ub-activation enzyme 1 (UBA1) and the Ub-like modifier activating enzyme 6 (UBA6) are the two E1 enzymes found in vertebrates, including humans. After being activated by E1, Ub is transferred to the Cys residue of a Ub conjugation enzyme or E2, also through a thioester linkage (**Figure 4**). In all eukaryotes, E2s for Ub and most Ub-like proteins (UBLs) contain a conserved region of approximately 140 to 200 amino acids called the UBC domain that harbors the required Cys. Thereafter in the Ub conjugation cascade, Ub is transferred from the E2 to the substrate protein by the help of E3 ligases (**Figure 4**). E3s are a large and diverse group of proteins and can be divided into two mechanistic categories based on how the Ub is transferred from the E2 to the substrate (Callis, 2014) (**Figure 4**).

A large family of proteins that contain an approximately 350-amino-acid C-terminal region homologous to that of E6-AP, referred to as the "Homologous to E6AP Carboxyl Terminus" (HECT)-domain family, are present in many eukaryotic organisms (Hershko & Ciechanover, 1998). All HECT proteins contain a conserved active site Cys residue near the C-terminus, and the N-terminal regions being highly variable, involved in the recognition of specific protein substrates (Huibregtse et al., 1995). HECT family is one of the largest and earliest studied E3 ligases (Yang et al., 2021). Due to the difference of N-terminal domain, HECT E3 ligases can be classified into three groups: the NEDD4 family (9 members), the HERC family (6 members) and another HECTs (13 members).

The second category of E3 ligases, “Really Interesting New Gene” (RING) and U box types of ligases, involves a noncovalent interaction of the E2-Ub with the E3 via a conserved domain and followed by Ub transfer as part of an E2-E3-substrate complex to the substrate, called the RING domain [reviewed in (Yang et al., 2021)]. There are more than 600 different RING type ligases expressed in human cells. During ubiquitination, the RING domain binds the E2 conjugation enzyme. RING E3 ligases are classified into monomeric and multi-subunit families. Monomeric RING ligases, such as COP1, MDM2, and TRAF6, possess substrate-binding, ubiquitination, and autoubiquitination functions. Multi-subunit ligases, like cullin-RING ligases (CRLs) and APC/C, are diverse. CRLs consist of a cullin scaffold with a RING-box protein at the N-terminus, an adaptor protein, and a C-terminal substrate receptor. APC/C comprises 19 subunits, including a RING subunit (APC11) and a cullin-like subunit (APC2) (Metzger et al., 2014). SCF E3 ligases, the largest complex, include SKP1, Cullin1, and F-box proteins, each with distinct roles: F-box recognizes substrates, SKP1 connects the catalytic core to the F-box, and Cullin1 links SCF components. RING E3s are regulated by modifications like autoubiquitination, neddylation, phosphorylation, and interactions with small molecules. U-box E3 ligases, a smaller family, control posttranslational protein quality via a conserved 70-residue U-box domain, structurally similar to the RING finger domain, essential for enzymatic activity (Ryu et al., 2019).

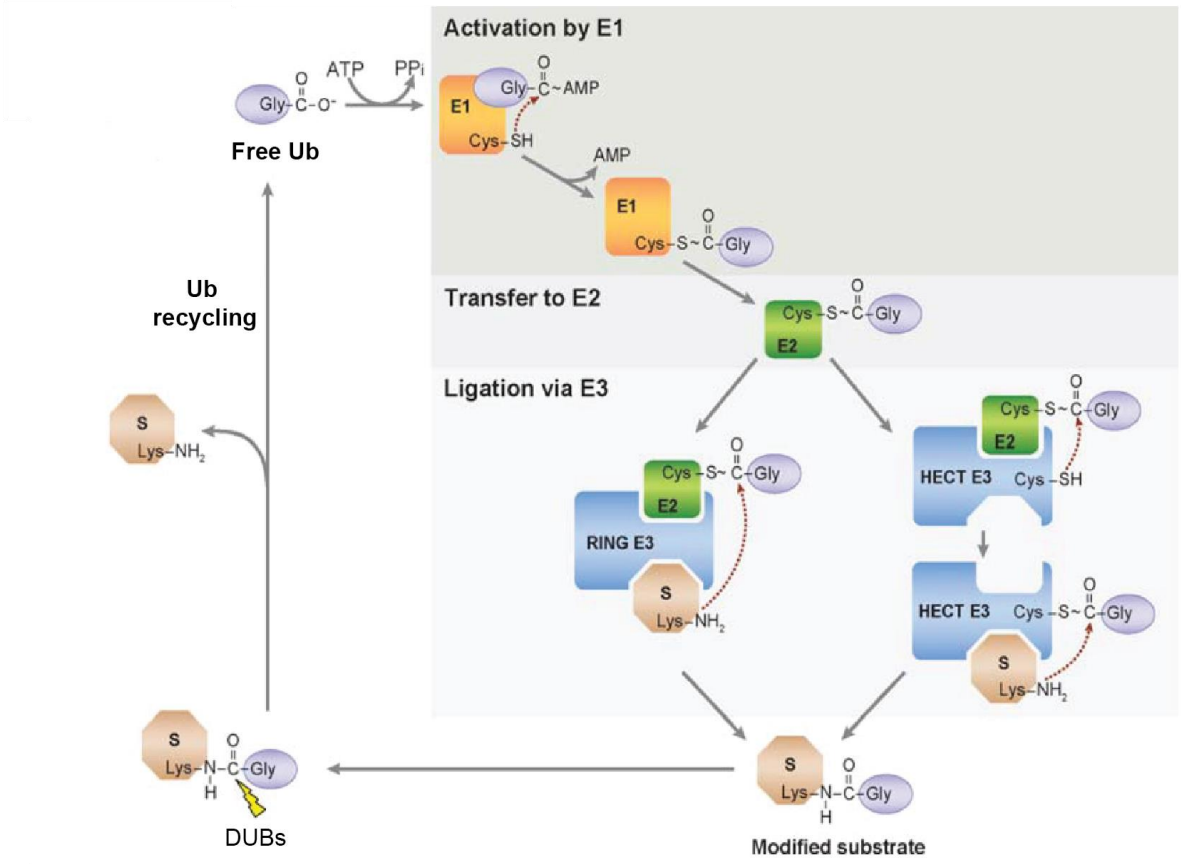


**Figure 4: A schematic representation of the Ub-conjugation pathway.** A monomeric (free) Ub is activated with ATP by E1. The E1 adenylates the Ub C-terminal carboxyl group, forming a high-energy Ub-AMP intermediate (the "~" represents a high-energy bond), which is attacked (*dashed red arrow*) and covalently bound by the catalytic Cys of the E1, creating a thioester linkage and releasing AMP. Ub is the transferred to the catalytic Cys of the E2, or Ub-conjugating enzyme, via a transthiolation reaction and is then ligated to a substrate with the help of E3-ligase. The adaptor-like RING E3s catalyze modification by binding simultaneously the Ub-E2 thioester complex and the substrate to be modified. HECT E3s catalyze substrate ligation in two steps, where the Ub is first transferred to a catalytic Cys of the HECT E3 via transthiolation. Then, the E3-Ub thioester complex transfers Ub to the substrate. The DUBs can remove Ubs from substrates. Figure has been adapted and modified with permission from (Kerscher et al., 2006) (A.4)

Lastly, a third hybrid class of E3 ligases have also been described. This class is the RING between RING (RBR) type where the E3 interacts with the E2-Ub as in the RING/U box type, but then transfer the Ub to a RBR Cys residue prior to transfer to proteins as in the HECT-type mechanism. (Callis, 2014) The RBR ligases are specialized by a conserved catalytic region, including a RING1, a central in between-RINGs (IBR) and a RING2 domain (Yang et al., 2021), that gives its name. RING1 can recruit the E2-Ub, and RING2 domain contains the catalytic Cys. The IBR domain can adopt the same fold as the RING2 domain only lacking the catalytic Cys. Additionally, different RBR E3 ligases also contain specific domains to distinguish from each other. RBR E3 ligases can be involved in intermolecular interactions to keep the proteins in an autoinhibited state. Such state is regulated by different kinds of mechanisms, such as phosphorylation or protein-protein interactions. Examples of these type of ligases include human homolog of Ariadne (HHARI) and Parkin, that tend to ubiquitinate substrates through linear (Met1-) Ub chains.

The histone code, in its entirety, is determined essentially by the E2, E3 or the substrate-E2 complex (Komander & Rape, 2012). The enzymes catalyzing monoubiquitination have to recognize substrate Lys, and not of Ub itself and thus, prevent Ub chain elongation. For instance, the polycomb E3 ligase complex BMI1-RING1 monoubiquitinates histone H2A on Lys119, even though it collaborates with UBE2D/UBCH5, a nonspecific E2 that is also implicated in modifying multiple substrates and Ub lysine residues (Bentley et al., 2011), indicating that the specificity comes from the E2-substrate interaction. Another mechanism of specificity is that E3 enzymes can block the ability of E2s to catalyze chain formation, as seen with RAD18 and its E2 RAD6. Although Rad6 can synthesize mixed or Lys48-linked chains (Hibbert et al., 2011), it leads to the monoubiquitination of PCNA when interacting with RAD18 (Kim et al., 2009). Similar to other E2s, in RAD6, a noncovalent Ub-binding site is required for chain formation. RAD18 has been found to occupy this site, thereby blocking chain formation by RAD6.

Enzymes facilitating Ub chain formation face a different specificity challenge where there is the need to modify specific lysine residues of Ub. For E3s containing a RING or U-box domain,

this linkage specificity is determined by the E2 as E3s can synthesize different chain types depending on the E2. For instance, BRCA1-BARD1 or MURF catalyzes formation of Lys63-linked chains with the heterodimeric E2 enzyme UBE2N-UEV1A, but Lys48 linkages when bound to UBE2K (Christensen et al., 2007; Kim et al., 2007). Multi-subunit E3 ligases of the SCF family form Lys48-linked chains on substrate proteins by using the Lys48-specific E2 UBE2R1 (Petroski and Deshaies, 2005), GP78, a regulator of endoplasmic reticulum–associated degradation (ERAD), assembles Lys48-linked chains with the Lys48-specific E2 UBE2G2 (Li et al., 2007); and the anaphase promoting complex (APC/C) produces Lys11-linked chains using the Lys11-specific UBE2S (Williamson et al., 2009). RING E3s and their specific E2s initiate chain formation on a substrate Lys, at random positions or in preferred sequence environments, referred to as chain initiation motifs (Williamson et al., 2011). The E2s can be either the "chain initiating" E2s that often assemble short chains which can then be connected non-specifically as for UBE2D (Wu et al., 2010), contain a favoured linkage, as seen for the Lys11-preferring UBE2C (Kirkpatrick et al., 2006), or can be homogenous, as for Lys48-specific UBE2R1 (Petroski and Deshaies, 2005). In most cases, chain the initiating E2s cooperate with a specific chain-elongating E2. This allows for assembly of Lys11-linked chains by UBE2S (Williamson et al., 2009), Lys48-linked chains by UBE2K/UBC1 or UBE2R1 (Christensen et al., 2007; Petroski and Deshaies, 2005; Rodrigo-Brenni et al., 2010), and Lys63-linked chains by UBE2N-UEV1A (Christensen et al., 2007; Eddins et al., 2006). Selection of the specific Lys on which the chain is built requires recognition of a specific acceptor Ub surface by the E2-donor Ub complex.

HECT E3s also display a wide range of linkage specificities. For instance, yeast RSP5 and human NEDD4 assemble Lys63-linked chains, E6AP (E6-associated protein) synthesizes Lys48 linkages (Kim et al., 2009), KIAA10/UBE3C promotes formation of Lys29 and Lys48 linkages (Wang and Pickart, 2005), and HUWE1 appears to be nonspecific (Kim et al., 2009). As the acceptor lysine attacks the thioester between the Cys of the E3 and Ub, linkage specificity, in this case, is determined by the HECT E3 and not the E2. Indeed, RSP5 and E6AP assemble Lys63- or Lys48-linked chains, respectively, despite using the nonspecific E2 Ube2D. Many HECT E3s also facilitate Ub chain elongation with UBE2L3/UBCH7, a thiol-reactive E2 that does not modify lysine residues (Wenzel et al., 2011), and HECT domain swaps are sufficient to change linkage specificity despite using the same E2s (Kim et al., 2009).

RBR E3s function with UBE2L3, an E2 that does not have any lysine specificity (Wenzel et al., 2011), thus, the linkage specificity of chain formation must be determined by the E3. Expectedly, RBR E3s can synthesize chains that differ from the inherent specificity of a cooperating E2. Whereas UBE2K usually assembles Lys48-linked chains (Haldeman et al. 1997), LUBAC and UBE2K produce Met1-linked chains (Tokunaga et al., 2009).

Sometimes it also so happens that E3 ligases recognize their substrates through some degradative signals or degrons that facilitate their (poly)ubiquitination to ultimately target them for proteasomal degradation. This is termed as the "N-end rule" and proteins having this N-end degrons typically have basic (Type I) or bulky-hydrophobic (Type II) N-terminal amino acid residues (Hershko & Ciechanover, 1998) (**Figure 5**). Specific residues called primary ($1^o$) destabilizing residues, such as the basic amino acids arginine (Arg) and Lys, are de-stabilizing directly by virtue of recognition by a specific E3 or E3s. Other amino acids at the N-terminus require modification prior to interaction with an E3, for example, glutamine (Gln, Q) and asparagine (Asn, N) are considered tertiary de-stabilizing ($3^o$) residues because they must be first deaminated to Glu and aspartic acid (Asp, D), respectively, and then Glu and Asp (secondary destabilizing, $2^o$) are substrates of an arginyl-tRNA protein arginyl transferase (R-transferase) activity, finally resulting in a $1^o$ destabilizing residue, Arg, at the N-terminus (Callis, 2014).

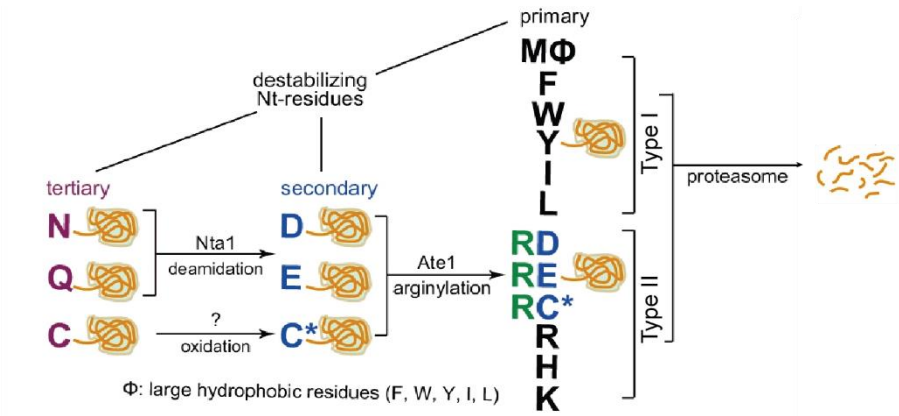



**Figure 5: The N-end rule pathway.** Nta1 N-terminal amidase deamidates tertiary destabilizing N-terminal residues Asn and Gln into secondary destabilizing N-terminal residues Asp and Glu, respectively. Subsequently, Ate1 arginyltransferase attaches Arg to the N-terminal Asp or Glu. Eventually, Ubr1 recognizes the type-1 or type-2 primary destabilizing N-terminal residues of the target substrates for polyubiquitination and subsequent proteasomal degradation. Figure has been created on the basis of (Nguyen et al., 2018).

### 2.1.2.2 Interpretation of the Ub Code by Readers

The information carried on the Ub code is deciphered by "reader" proteins that selectively bind to these modifications, as is common with every biological code. Proteins containing Ub-binding domains (UBDs) read the code and target the ubiquitinated substrates towards desired outcomes. Here, the several functions of the Ub code and reader proteins that act as the mediators, have been delineated. The biological outcomes are essentially dependent on the

architecture of (poly)Ub and its complexity. For example, amongst the homotypic chains, K48-linked chains are the most abundantly occurring polyUb marks and are mostly responsible for targeting substrates to the proteasome for degradation. And recently, importance of branched chains has also been uncovered where they have been found to be involved in these processes along with the linkage specific ones. Here, we focus on the most predominant functions of the predominant chain types.

## Proteolytic Functions

Targeting protein substrates for their cellular turnover has been the most predominant outcome of (poly)ubiquitination. (Poly)ubiquitinated proteins having homotypic (Lys48 and Lys11-linked, for example) chains or heterotypic branched chains (like, K48/K11 branches) are signalled for proteasomal as well as non-proteasomal degradation.

- **Proteasome-dependent Degradation**

Lys48-linked Ub chains was the first polyUb linkage to be assigned a role in proteasomal targeting of protein substrates (Chau et al., 1989). As we also read before, many E3s like the SCF, GP78, or E6AP, are responsible for constructing these chains, making Lys48-linked Ub chains the most abundant chain type and their levels increase rapidly when the proteasome is inhibited. However, some studies indicated that other linkages could also be recognized by the proteasome (reviewed in Komander & Rape, 2012). Lys11-linked chains are the second most abundant chain types with roles in proteasomal targeting. As the APC/C complex triggers formation of these Lys11-linked chains during anaphase in mitosis, the chains bind proteasomal receptors and trigger degradation of cell cycle regulators during mitosis (Jin et al., 2008; Williamson et al., 2009). Some other chain types have also been implicated in proteasomal degradation, albeit, less frequently. Lys29-linked chains contribute to protein turnover in the ubiquitin-fusion-degradation pathway (Koegl et al., 1999), and in a few cases, Lys63-linked or mixed chains were seen to be responsible for triggering substrate degradation (Kim et al., 2007; Kirkpatrick et al., 2006). Lys48- and Lys11-linked chains trigger degradation more frequently than other modifications, most likely, owing to the fact that the enzymes synthesizing Lys48- and Lys11-linked chains are less likely to introduce branches, which can impede degradation (Kim et al., 2007). In addition, Lys48- and Lys11-specific enzymes, such as the APC/C or SCF, often interact with the proteasome to efficiently couple ubiquitination and degradation (Seeger et al., 2003). For some E3s, binding to the proteasome is required for targeting substrates to degradation, for instance, deletion of a proteasome-binding domain in the E3 ligase UFD4 does not affect ubiquitination, but inhibits substrate turnover (Xie and Vershavsky, 2002). The

mechanism of recognition and processing of poly-ubiquitinated substrates at the proteasome has been described in details in **Box 1**.

Apart from the regular, cytosolic protein turnover in proteasomes, ERAD is responsible for the clearance of misfolded proteins in the ER through proteasomal degradation (**Figure 6**). The SEL1L-HRD1 protein complex represents the most conserved ERAD machinery in mammals with SEL1L being the cofactor for the E3 ligase HRD1. ERAD substrates are ubiquitinated by the action of HRD1 and SEL1L followed by their extraction from the ER membrane to the cytosol, a process referred to as 'retrotranslocation' or 'dislocation'. This process is facilitated by the energy derived from ATP hydrolysis by the AAA+ ATPase vasolin-containing protein (VCP)/p97 (CDC48 in yeast) (Hwang & Qi, 2018). However, p97 does not act alone but instead associates with a set of adaptor proteins that act to recruit or modify the substrates. Many of the adaptors contain UBDs, for instance, the best-characterized p97 adaptor is the heterodimer of NPLOC4/NPL4 and UFD1L/UFD1 (UN), each bind Ub chains, with UFD1L showing specificity for K48-linked chains (Locke et al., 2014). They interact with p97 at separate sites to form a complex with the stoichiometry of one UN heterodimer per p97 hexamer (Blythe et al., 2017). The accumulation of misfolded proteins in the ER triggers UPR by activating three main UPR sensors, IRE1α, PERK and ATF6, on the ER membrane. They subsequently initiate signaling pathways to enhance protein folding and degradation, and stall protein translation (Hwang & Qi, 2018) (**Figure 6**).

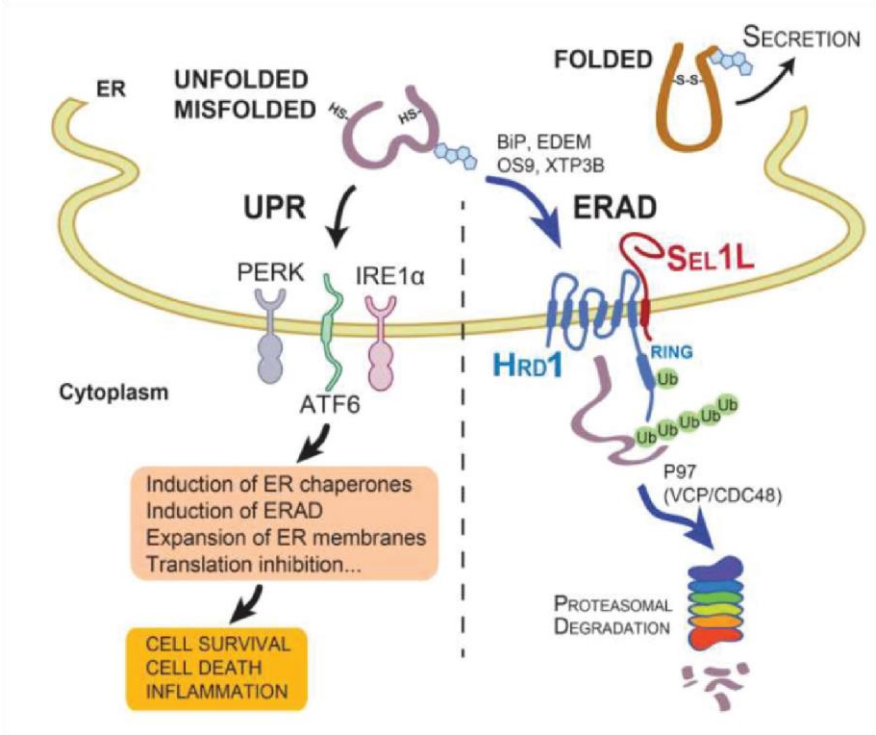



**Figure 6: The ERAD quality control system.** Misfolded proteins are recruited to the ERAD complex via the activity of various ER chaperones (right) such as BiP, EDEM, OS9 and XTP3B for cytosolic

degradation. The Sel1L-Hrd1 protein complex represents the most conserved ERAD complex in mammals. Following retrotranslocation into the cytosol, substrates are ubiquitinated and, with the help of p97 (VCP/CDC48), degraded by the proteasome in the cytosol. Terminally unfolded or misfolded proteins in the ER (left) also activate UPR via three sensors IRE1α, PERK and ATF6, which initiate multiple signaling pathways leading to the induction of ER chaperones and ERAD. Figure has been adapted with permission from (Qi et al., 2017) (A.4).

Although linear Ub chains, mainly K48- and K11-linked chains are the most intensively studied for their roles in proteasomal degradation, recent findings on branched Ub chains, which are possibly generated in collaboration between pairs of E3s with distinct linkage specificities (R) or by a single E3 that recruit different E2s with varying linkage specificities, have been found to have proteolytic functions as well (French et al., 2021). In a seminal study, the APC/C was found to attach branched K11/K48 chains to mitotic cell cycle regulators, such as cyclin A and NEK2A leading to their proteasomal degradation (Lopez-Mosqueda and Dikic, 2014). Branched K11/K48 chains have also been shown to be involved in the degradation of newly synthesized misfolded proteins and cytoplasmic aggregates (Yau et al., 2017), many of which seem to require the activity of VCP/p97 (Yau et al., 2017; van den Boom and Meyer, 2018). Similarly, branched K29/K48 and K48/K63 chains have been shown to target a diverse array of proteins, including substrates of the UFD pathway, ERAD substrates, and apoptotic regulators, for proteasomal degradation (Liu et al., 2017; Ohtake et al., 2018).

As there are a variety of Ub signals that target substrates for degradation, what is the threshold? As Ub molecules are added to the substrate and the chain is elongated, the avidity of the conjugate to the proteasome increases. Once the avidity reaches a certain threshold, and a stable binding of the adduct to the proteasome is achieved, it is detached from the conjugating machinery and in turn, bound stably to the proteasome to be degraded processively and efficiently. With larger protein targets that may require a longer processing time, a longer polyUb chain may be necessary to generate the required avidity. For substrates that are polyubiquitinated, the spatial arrangement of a large enough number of single Ub moieties that bind to multiple points in the proteasomal Ub receptors assures the strong binding necessary for processive degradation (Ciechanover & Stanhill, 2014). It has been shown that a minimum length of n=4 is in many cases required for K48-linked chains to constitute an effective degradation signal (Thrower et al., 2000). In a recent study, it was also shown that K29- and K33-linked chains with six or more Ub moieties seem to have an increased tendency to bind metabolite interconversion enzymes (like hydrolases, transferases, or oxidoreductases), while shorter chains (Ub2, Ub4) prefer to interact with protein-modifying enzymes like kinases, proteases, or Ub ligases (Lutz et al., 2020).

- **Proteasome-independent Degradation: Autophagy/ Lysosomal Degradation**

Autophagy is a fundamental cellular process responsible for eliminating subcellular components, including nucleic acids, proteins, lipids and organelles, via lysosome-mediated degradation to promote homeostasis, differentiation, development and survival. The flux of substrates through the autophagy pathway is a multistep process that relies on protein-protein and protein-lipid–driven interactions. In "macroautophagy" (or simply, autophagy), cellular material is engulfed by a double membrane structure, the autophagosome which then elongates and matures before fusion with the lysosome for cargo degradation and recycling. This is the major response to cellular stress, maintaining metabolic building blocks during periods of nutrient deprivation, and eliminating unwanted cellular contents such as toxic protein aggregates, damaged organelles, or intracellular pathogens. Ub acts as a tag for autophagic degradation. Ubiquitinated protein substrates are recognized by the autophagy machinery via autophagy adaptors (also referred to as autophagy receptors), which contain both a ubiquitin-binding domain and an autophagy-related protein 8 (ATG8) binding region (commonly called the LC3-interacting region), which binds ATG8 family proteins on autophagosomes (Mizushima, 2024) (**Figure 7**). So as is evident, despite their distinct degradation mechanisms, the UPS and autophagy share molecular determinants and substrates. There is a dynamic cross-talk between the two pathways, with Ub playing a central role in both. As both systems use Ub as a degradation signal for substrates, cellular Ub concentration as well as accessibility of Ub receptors both appear crucial. For instance, the most studied Ub receptor, p62/SQSTM1, binds to ubiquitinated substrates via its Ub-associated (UBA) domain and escorts ubiquitinated substrates to the proteasome and also acts as an autophagy receptor recruiting Ub-conjugated substrates to autophagosomes. The p62 dimer recognizes substrates with K48-linked polyUb chains via its UBA domain and shuttles them to the proteasome. On the other hand, binding to K63-linked polyUb chains induces oligomerization of p62 and liquid-liquid phase separation, resulting in the formation of p62 bodies that subsequently recruit autophagic membranes (Pohl & Dikic, 2019).

## Non-proteolytic Functions

As mentioned earlier, monoubiquitination or even polyubiquitination of protein substrates is involved in signal transduction in cellular signalling cascades. Major examples that have been discussed in more details include DNA repair and even NF-κB signalling.

**Proteasome Assembly and Recognition of PolyUb Chains [BOX 1]:**

The substrates decorated with Lys48- or other above-mentioned polyUb chains are targeted to the proteasome for ATP dependent degradation. The 26S proteasome is the only known Ub ATP-dependent protease consisting of a compartmentalized cylinder-shaped complex attached to an AAA+ type ATPase with an 'unfoldase' activity required for substrate processing. The protease complex is composed of a 20S catalytic core particle (CP) and two 19S regulatory particle (RP) subunits (**Figure B1a**). The 20S particle is barrel-shaped, composed of four stacked heptagonal rings consisting of the two inner β catalytic rings (β 1–7) that harbour the protease active sites, and two outer α rings (α 1–7) of which β 1, 2, and 5 are active subunits with varying catalytic specificities (**Figure B1b**). The outer α rings serve as an interface for RP binding through a pocket created between two adjacent α subunits. Upon RP engagement, the α subunit N-termini that obstruct entry into the catalytic cavity are displaced and the entrance into the chamber's interior is dilated. One of the functions of the RP is therefore to open the narrow entrance of the latent 20S CP (Ciechanover & Stanhill, 2014).

The 19S regulatory particle of the proteasome is a highly controlled gatekeeper for most of the proteasome substrates, comprising two main subcomplexes: the 'base' and the 'lid' (reviewed in Ciechanover & Stanhill, 2014). The base consists of six RP Triphosphatases (Rpt1-6) that are paralogous AAA+ ATPases forming a hexameric ring, the RP non-ATPases RPN1, RPN2, and two Ub receptors, RPN10/S5A and RPN13. The lid includes the remaining of the 13 RPN subunits, with only RPN11 known to have catalytic DUB activity. Recent structural studies reveal that the ATPases are organized as a trimer of dimers, and are positioned towards the 20S catalytic particle in a 6-fold rotational symmetry. For each RPT, the ATPase domain is facing the α ring, whereas a second oligosaccharide binding (OB) domain faces upwards (**Figure B1c**). All the six OB domains are thus generating an additional ring. The AAA sub-domains appear to be arranged in a spiral staircase with RPT3 and RPT2 at the highest and lowest positions in the ring, respectively, bridged by RPT6. Their arrangement suggests that substrates traverse a 'staircase' of ATPase domains, ultimately unfolding and entering the 20S catalytic core. RPN1 and RPN2 are positioned at distinct locations on the regulatory particle—RPN1 on the sides and RPN2 at the ends—interacting with specific RPT ATPases. These positions allow them to act as platforms for various proteasome adaptors and Ub receptors. RPN1 interacts with RPT1 and RPT2, while RPN2 interacts with RPT3 and RPT6. The RPN subunits of the 'lid', all contain a proteasome-cyclosome initiation factor (PCI) domain (**Figure B1d**). RPN5 and RPN6 are directly connected to the α1 and α2 subunits of the 20S particle. The PCI cluster extends all along the 19S RP, all the way from the catalytic particle interface to the distal end of the RP, and adjacent to the RPT hexamer in the opposing side of the assigned RPN1. This connection links the 'lid' component (RPN5 and RPN6) to the catalytic particle. The catalytically inactive DUB RPN8

forms a dimer with active RPN11, both MPN domain containing metalloproteases (see below, section 2.1.2.3), positioning RPN11 near RPT3, which may explain its dynamic role in substrate processing. RPN10/S5A and RPN13 are well-characterized Ub receptors. RPN10/S5A binds Ub through two Ub-interacting motifs (UIMs), while RPN13 uses a Pleckstrin-like receptor for Ub (PRU) domain. Both receptors interact with the hydrophobic patch of Ub. Additionally, Ub-like (UBL) domains on shuttle factors, which contain UBL–UBA domains, facilitate polyUb substrate delivery to the proteasome. RPN1 also mediates UBL-proteasome interactions. RPN10/S5A and RPN13 are positioned about 70-100 Å apart, setting a minimal length for polyUb chains needed for effective substrate binding. RPN11 is centrally located between these receptors, ensuring proper timing of deubiquitination and substrate commitment to degradation. NMR studies suggest that RPN13 preferentially binds the proximal Ub moiety in a polyUb chain, with both RPN10/S5A and RPN13 potentially binding a single K48 triUb chain. This implies that a minimal polyUb chain length of four Ub units is necessary for effective simultaneous binding and processing, aligning with reports on required chain lengths for proteasomal degradation (Ciechanover & Stanhill, 2014). UBP6, a sub-stoichiometric DUB, is located near RPN1, facilitating the processing of polyubiquitinated substrates. This positioning, supported by cross-linking studies, suggests that UBP6 can interact with substrates brought by shuttling factors, which may require extended polyUb chains to bridge the distance to RPN11 and the RPT ATPases.

Assembly of the CP, on the other hand, is initiated with the formation of the heptameric α ring (**Figure B1e**). Although the exact order of α-subunit assembly is unknown, there are possibilities of aberrant subunit arrangements (Yao et al., 1999). The β subunits, several of which are synthesized as precursors with N-terminal propeptides, then add to the α ring in an ordered fashion, creating a half-proteasome (Nandi et al., 1997). Upon dimerization of two half-proteasomes to form the preholoproteasome (PHP), the active site bearing β subunits undergo autocatalytic cleavage of their propeptides, generating the mature catalytic sites and completing CP assembly (**Figure B1e**, Tomko & Hochstrasser, 2013).

It is clear from the available structural data of RP that the placement of subunits and proteasome-associated factors within the RP is organized in a way that facilitates movement of the substrates from the outer edges of the RP toward the CP for degradation. The subunits responsible for each successive step in substrate degradation, i.e., recognition, deubiquitylation, unfolding, and translocation into the CP are positioned sequentially closer to the CP pore. After a substrate binds the proteasome, its polyUb tag must be removed either prior to or potentially during substrate degradation. The proteasome harbors two non-integral DUBs UBP6 (USP14 in humans) and UCH37 (not in all species) that trim polyUb chains from the distal end, and the RPN11 subunit of the RP lid that removes whole Ub chains at their point of attachment to substrate. Additionally, several E3 ligases, most notably HUL5,

are found associated with proteasomes as well, strongly suggesting that the polyUb chains of substrates can undergo remodeling on the proteasome, perhaps as a substrate triage or quality control step (Crosas et al., 2006). Compiling knowledge from other hexameric ATPases, it is possible that, instead of, or in addition to, the mechanical forces from the ATPases being exerted through small, local movements of pore loops, unfolding may be driven by movements of the entire large domains in which they proceed through the conformations observed in the staircase arrangement (Tomko & Hochstrasser, 2013).

Overall, these findings refine the understanding of the structural and functional dynamics of the 26S proteasome, emphasizing the complex interactions between its subunits and substrates. (Grice & Nathan, 2016; Osei-Amponsa & Walters, 2022).

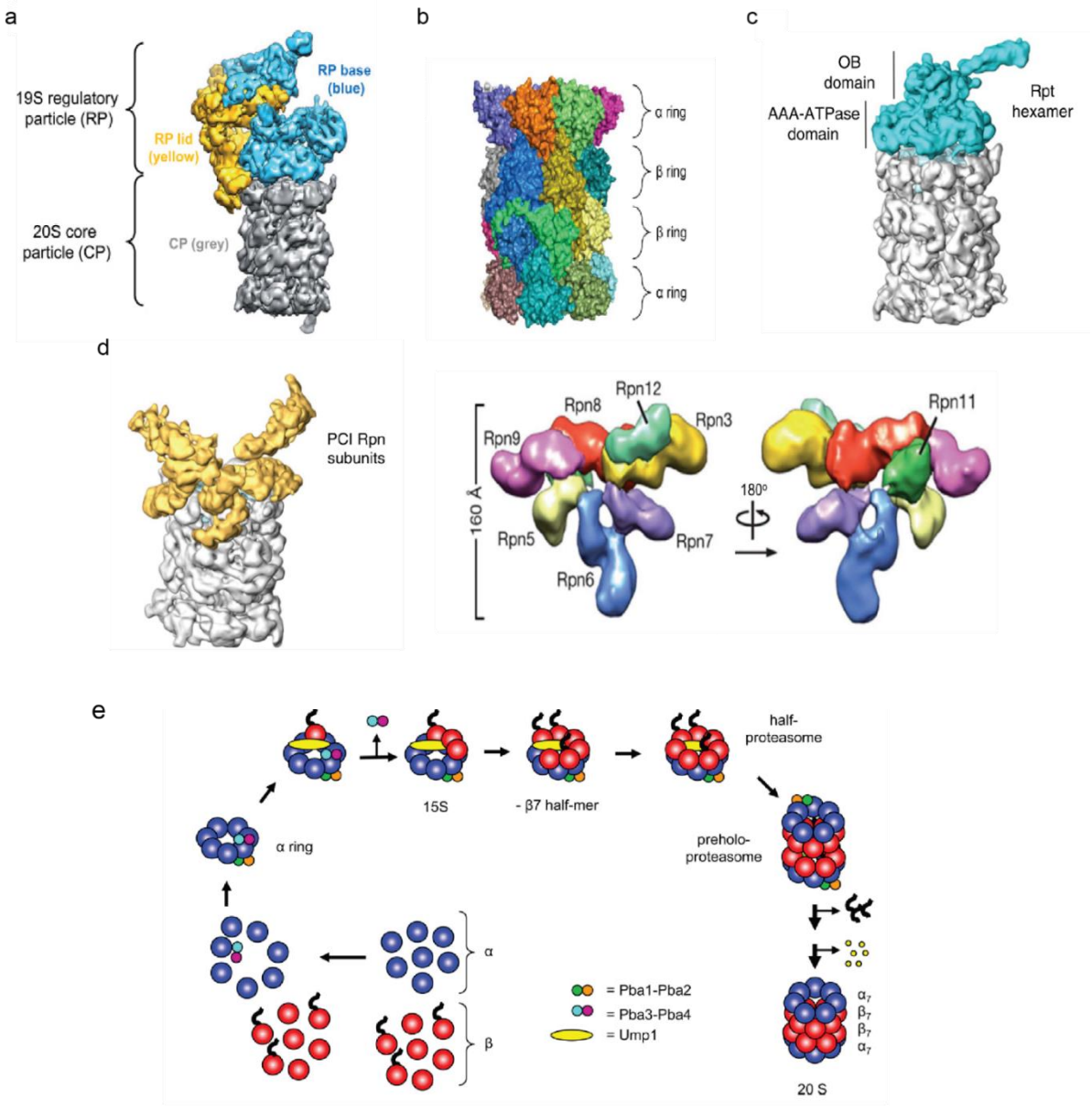


**Figure B1: Proteasome assembly**. **a.** The cryo-EM density of the 26S proteasome. The 19S regulatory particle (RP) lid subcomplex is displayed in yellow, the RP base subcomplex in blue, and the 20S core particle (CP) in grey. **b.** Space-filling model of the 20S CP atomic structure from yeast (PDB ID **1RYP**). **c.** Subunit architecture of the hexameric RPTs (blue) of the RP base assembled on top of the 20S particle (gray). The AAA+ ATPase and OB domains of the hexameric RPT ring have been shown. **d.** Subunit architecture of the RP lid with the PCI-winged helix domain RPN subunits (gold) that resembles a horseshoe-shaped arrangement with close proximity to the 20S particle (gray) (left), EM structure of the purified lid complex (right). **e.** The 20S CP particle assembly pathway. Panels **a, b, d** (right) and **e** have been adapted with permission from (Tomko & Hochstrasser, 2013) (A.4), Panels **c, d** (left) are adapted with permission from (Ciechanover & Stanhill, 2014) (A.4).

- **DNA Repair**

In response to DNA damage, PCNA is monoubiquitylated, recruiting Y family DNA polymerases that recognize PCNA through a PCNA-interaction motif, the PIP-box, and Ub through UBZ or UBM domains, leading to a high-affinity interaction that replaces replicative

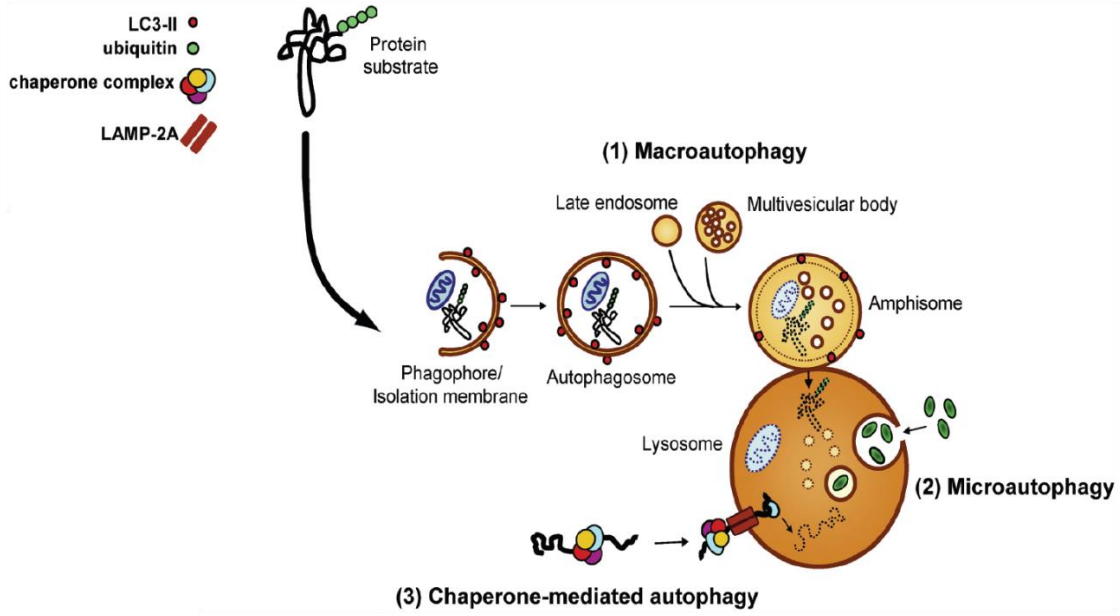



**Figure 7: Autophagy-lysosomal system.** This system is comprised of (1) macroautophagy, in which cytosolic components are engulfed and delivered to the lysosome in bulk, (2) microautophagy, in which small volumes of cytosol are directly engulfed by lysosomes, and (3) chaperone-mediated autophagy (CMA), in which soluble substrates associated with a specific chaperone complex are translocated into the lysosome through the LAMP-2A lysosomal receptor. LC3-II is a protein that associates with the inner and outer surfaces of autophagic membranes and provides a histological marker of autophagic vacuoles. Figure has been adapted and modified with permission from (Nedelsky et al., 2008) (A.4).

polymerases from PCNA. Although this leads to error-prone incorporation of nucleotides, such a polymerase switch rescues stalled replication forks from collapsing. Amongst the polyUb chains, the major roles of Lys63-linked chains, which are the second-most abundant in cells

and are implicated in cellular signal transduction, include scaffolding during DNA damage, an event that is dependent on a series of E3 enzymes. The recruitment of several E3 ligases to the sites of the damaged DNA depends on Lys63-linked chains that are probably attached to histone proteins. The E3s RNF8 or RNF168 directly recognize Lys63 linkages through MIU or UBZ domains, whereas BRCA1 binds together with RAP80. Together, these E3s generate Ub-rich foci that act as stable recruitment platforms for DNA repair enzymes and for checkpoint proteins that inhibit cell cycle progression (reviewed in Komander & Rape, 2012). It was also shown in a study that, K63-linked polyUb chains bind DNA directly during DNA repair and that Ub mutations impairing this process may facilitate tumorigenesis in cancer patients. It was observed that K63-linked Ub chains interact with DNA in a Ub chain length–dependent manner and through a DNA-interacting patch (DIP) region in Ub that critically includes Thr9, Lys11, and Glu34. Notably, this identified DIP motif localizes spatially distant from the Ile44 patch that mediates the interaction of K63-linked ubiquitin chains with RAP80, indicating that the DNA binding and the protein-binding function of the K63-linked polyUb chains could contribute synergistically to govern DNA damage repair (Liu et al., 2018).

- **NF-κB Signalling**

Ubiquitination is also an important mediator of the NF-κB signalling pathway that is a proinflammatory signalling pathway that leads to the expression of proinflammatory genes like those of cytokines, chemokines and adhesion molecules. Activation of NF-κB usually occurs in response to external stimuli, such as the interleukin 1 (IL-1) or tumor necrosis factor α (TNFα). Binding of these stimuli to the respective membrane receptor initiates a variety of linkage-specific ubiquitination events by respective, specific E3 ligases. For example, formation of Lys63-linked chains by TRAF6 (Deng et al., 2000), mixed Lys11/Lys63-linked chains by the RING E3 ligase cIAP1 (Dynek et al., 2010), or Met1-linked chains by LUBAC (Tokunaga et al., 2009). Recognition of these polyUb marks is subsequently responsible for activation of the pathway. LUBAC, for instance, modifies NEMO, a subunit of the IκBα kinase (IKK) complex. The Met1-linked chains on NEMO are recognized by the UBAN domain of NEMO itself, which may cause a conformational change in the intertwined helices of NEMO dimers (Rahighi et al., 2009). As NEMO is a core regulatory subunit of IKK, these conformational changes might lead to allosteric activation of IKK. NEMO also associates with Lys63-linked chains (Laplantine et al., 2009) and with mixed Lys11/Lys63-linked chains that are detected on the receptor interacting protein 1, RIP1 (Dynek et al., 2010). Lys63-linked chains activate IKK by promoting its binding to the upstream TAK1 kinase complex (Wang et al., 2001). Unanchored Lys63-linked ubiquitin chains have also been shown to mediate the activation of TAK1 and IKK kinases and also the RIG-I antiviral protein (Zeng et al., 2010). RIG-I binds Lys63-linked chains through tandem CARD domains (Zeng et al., 2010), which

leads to RIG-I dimerization and facilitates the downstream events that ultimately leads to the release of active NF-κB and regulation of gene expression (Komander & Rape, 2012). DUBs (like CYLD and A20) that disassemble K63-linked chains have been found to inhibit the activation of TAK1 and subsequent downstream events. Branched K48/K63 chains have also been found to be having a stabilizing effect on NF-κB signaling by being resistant to such DUBs, leading to stabilization of K63 linkages. Branched M1/K63 chains, which are reported to be resistant to cleavage by A2020, show similar stabilizing effects on NF-κB signaling (French et al., 2021).

**Additional Functions:**

Lys6-linked chains are thought to be involved in DNA repair as they were found to be associated with the BRCA1-BARD1 E3 ligase complex (**Figure 8a**). Several findings suggest a role for Lys27 linkages in mitochondrial quality control. Mitochondrial damage leads to the activation of the RBR E3 ligase Parkin, which assembles Lys27-linked polyubiquitin on several mitochondrial proteins, including voltage-dependent anion-selective channel protein 1 (VDAC1). Ubiquitinated VDAC1 is recognized by the autophagy adaptor p62, thus triggering clearance of damaged mitochondria by mitophagy (**Figure 8b**). Lys27-linked chains have also been found to be upregulated by DNA damaging reagents that induce DNA double-strand breaks, and have been shown to recruit DNA damage repair factors, such as p53 binding protein 1 (53BP1), to DNA damage foci. Lys33-linked ubiquitination of TCRζ leads to reduced phosphorylation of the receptor subunit and inhibits association of the activating kinase ZAP70 (ζ-chain-associated protein of 70 kDa) with the receptor, thus dampening TCR signalling through a non-degradative mechanism (**Figure 8c**). Several studies report roles for both Lys29- and Lys33-linked chains in different processes, such as the regulation of the AMP-activated protein kinase (AMPK)-related protein kinases which are polyubiquitinated with Lys29 linkages and Lys33 linkages but not targeted for degradation by the same. Lys29-linked chains are also implicated in epigenetic regulation as the specific DUB for this linkage, TRABID, was found to have the histone demethylase JMJD2D as a target, which it appears to deubiquitinate and stabilize, so that JMJD2D can act on the IL gene promoters to release repression (Kulathu & Komander, 2012; Swatek & Komander, 2016).

As we see, the Ub code is a universal code and plays a central role in dynamic biological processes.

### 2.1.2.3 Erasers of the Ub Code (DUBs)

As there are writers of the Ub code, there must be erasers of the same in order to maintain homeostasis. In this case, the erasers that actively oppose the function of E3 ligases, are the superfamily of isopeptidases that are collectively called "deubiquitinases" (DUBs) (Komander et al., 2009) (**Figure 9**). This superfamily can be subdivided into seven families: Ub C-terminal hydrolases (UCHs), Ub-specific proteases (USPs), Ovarian Tumour Proteases (OTUs), Machado-Joseph domain containing proteases (MJDs, also called Josephins), Motif Interacting with Ub-containing DUB family (MINDY) (Abdul Rehman et al., 2016), Zinc-finger containing Ub peptidase-1 (ZUP1) (Kwasna et al., 2018) and JAB1/MPN/MOV34 metalloenzymes (JAMMs, also known as MPN family).

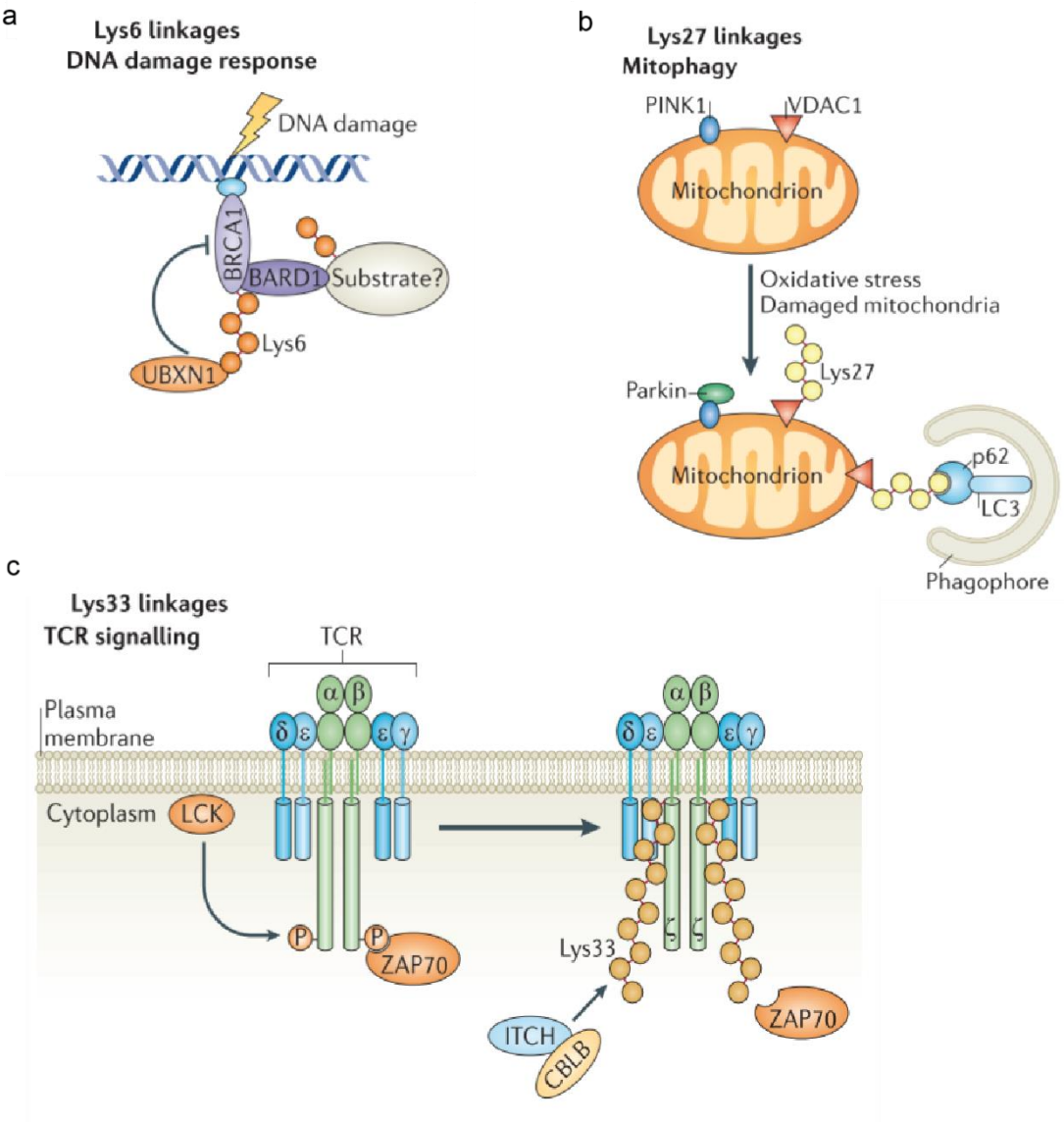


**Figure 8: Functional significance of atypical Ub chains. a.** The DNA repair-associated E3 ligase complex consisting of BRCA1-BARD1 assembles Lys6-linked ubiquitin chains on itself or its substrates. UBXN1 binds to Lys6 ubiquitylated BRCA1 to inhibit the ligase activity of the complex. **b.** Mitochondrial damage leads to the activation of the RBR E3 ligase Parkin, which assembles Lys27-linked Ub chains on several mitochondrial protein ultimately triggering clearance of damaged mitochondria by mitophagy. **c.** TCR stimulation results in the phosphorylation of tyr (Tyr, Y) in the cytoplasmic tail of the ζ-chain by the kinase LCK, which then leads to recruitment and amplification of TCR signalling by the kinase ZAP70. The ligases ITCH and CBLB negatively regulate TCR signalling by ubiquitinating the ζ-chain of the TCR with Lys33-linked Ub chains, which leads to reduced association of ZAP70 with the receptor. UBXN1= ubiquitin regulatory X (UBX) domain-containing protein 1, TCR = T cell receptor, LCK = lymphocyte cell-specific protein Tyr kinase, ZAP70 = ζ-chain-associated protein of 70 kDa, CBLB = casitas B lineage lymphoma proto-oncogene B. Figure has been adapted and modified with permission from (Kulathu & Komander, 2012) (A.4).

Out of the seven families, the UCHs, USPs, OTUs, MJDs, MINDY and ZUP-1 are cysteine proteases, whereas the JAMM family members are zinc metalloproteases. Except for the MJDs, each DUB family is conserved from yeast to humans (Clague et al., 2019). Out of the 99 family members, 11 are considered to be "pseudoenzymes" as these enzymes have lost residues critical for DUB activity but still can perform vital functions (Walden et al., 2018). Pseudo-DUBs are also common in the 12-member JAMM family, which contains five such members. The relatively newly discovered MINDY family has two members in *Saccharomyces cerevisiae* and five in humans, including one pseudo-DUB and has been found to show specificity for Lys48-linked Ub chains, indicating roles in protein homeostasis (Kristariyanto et al., 2017). As for the ZUP1 family, humans have one representative, found to be specific for Lys63-linked chains and has been linked to genome maintenance pathways (Haahr et al., 2018).

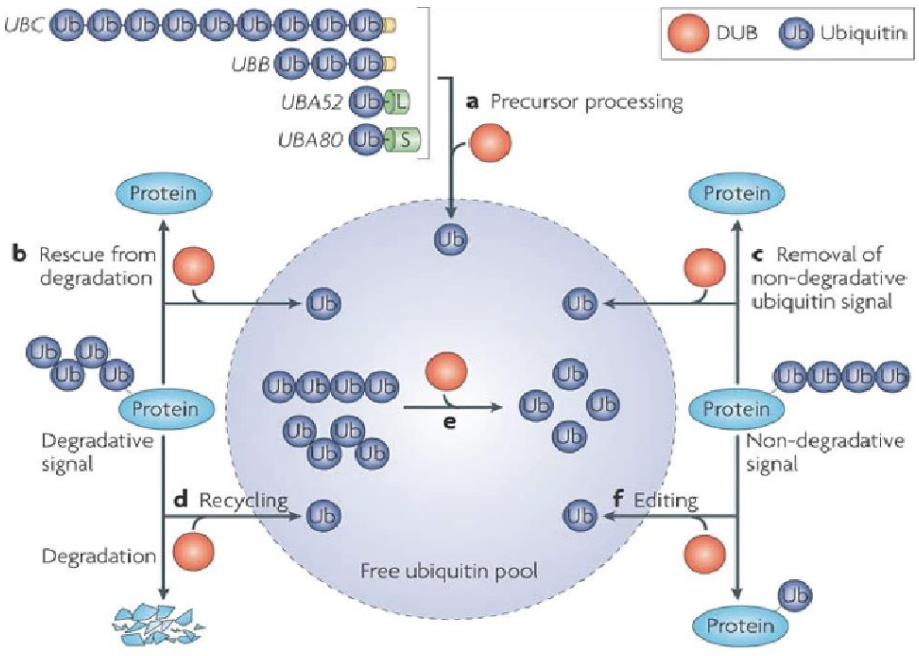


**Figure 9: General roles of DUBs in cells. a.** Generation of free Ub from precursors poly-Ub or Ub-RP genes. **b.** Rescuing protein substrates from degradation. **c.** Removal of non-degradative Ub signal. **d.** Maintaining Ub homeostasis and preventing degradation of Ub together with substrates of the proteasomal and lysosomal pathways (recycling of Ub). **e.** Disassembly of Ub chains generated by *en bloc* removal from substrates. **f.** Edit Ub chains and thereby help to exchange one type of ubiquitin signal for another. Figure has been adapted with permission from (Komander et al., 2009).

Most DUBs cleave between the ε-amino group of Lys and the carboxyl group at the C terminus of Ub. The Cys protease DUB families have two to three crucial residues, that together form a catalytic diad or triad, respectively (Komander et al., 2009). In these enzymes, a nearby His side chain lowers the pKa of the catalytic Cys, enabling it to make a nucleophilic attack on isopeptide linkages. A third amino acid (mostly Asn or Asp) aligns and polarizes the catalytic His, however this is not always essential for DUB activity (Komander and Barford, 2008). Then there is formation of a catalytic acyl intermediate, in which the carboxyl group is covalently bound to the catalytic Cys after the amino group has been cleaved. The negatively charged transition state is stabilized by hydrogen-donating residues, which form an oxy-anion hole in the vicinity. Next, a water molecule hydrolyzes the acyl-Cys intermediate to complete the catalytic cycle (**Figure 10a**). All Cys protease DUB families have divergent folds but their catalytic residues are conserved with only small deviations when bound to the Ub C-terminus (Komander and Barford, 2008). The JAMM family of metalloproteases, on the other hand, contain the JAMM/MPN+ motif that coordinates two zinc ions, one of which activates a water molecule to attack the isopeptide bond. The amino group is subsequently released from the charged catalytic intermediate, by a mechanism similar to that of cytidine deaminase (Tran et al., 2003; Maytal-Kivity et al., 2002). It is likely that the zinc ion in the Jab1/MPN domain activates a water molecule to form a hydroxide ion that in turn makes a nucleophilic attack on the carboxyl group. This then leads to the elimination of ammonia. A conserved glutamate residue that is positioned near to the active site zinc functions both as a proton acceptor and donor during the deamination reaction (Tran et al., 2003) (**Figure 10b**).

The different aspects of DUB specificity and therefore, the activity, depend how the DUBs recognize their substrate (Mevissen & Komander, 2017). All DUBs have at least one Ub-binding site, referred to as the S1 site, that guides the C-terminus and the isopeptide linkage into the DUB active site, where hydrolysis can then take place. This active site cleft specifically binds a linear epitope of the Ub C-terminus, consisting of the sequence 5'-RLRGG-3', collectively referred to as the P5-P1 sites. This linear sequence is highly conserved throughout eukaryotes, hinting towards a role in confering specificity for Ub conjugation or deconjugation, or both (Drag et al., 2008). In the case of cleaving a diUb, the distal Ub occupies the S1 site and the proximal Ub occupies a second S1' site. In addition, some DUBs are found to bear

additional Ub-binding sites (S2, S3, S2', or S3', and so on). These sites enable interaction between the enzyme and long Ub chains and may contribute to linkage specificity. An Ub mutant was even generated that is resistant towards cleavage by DUBs (Békés et al., 2013).

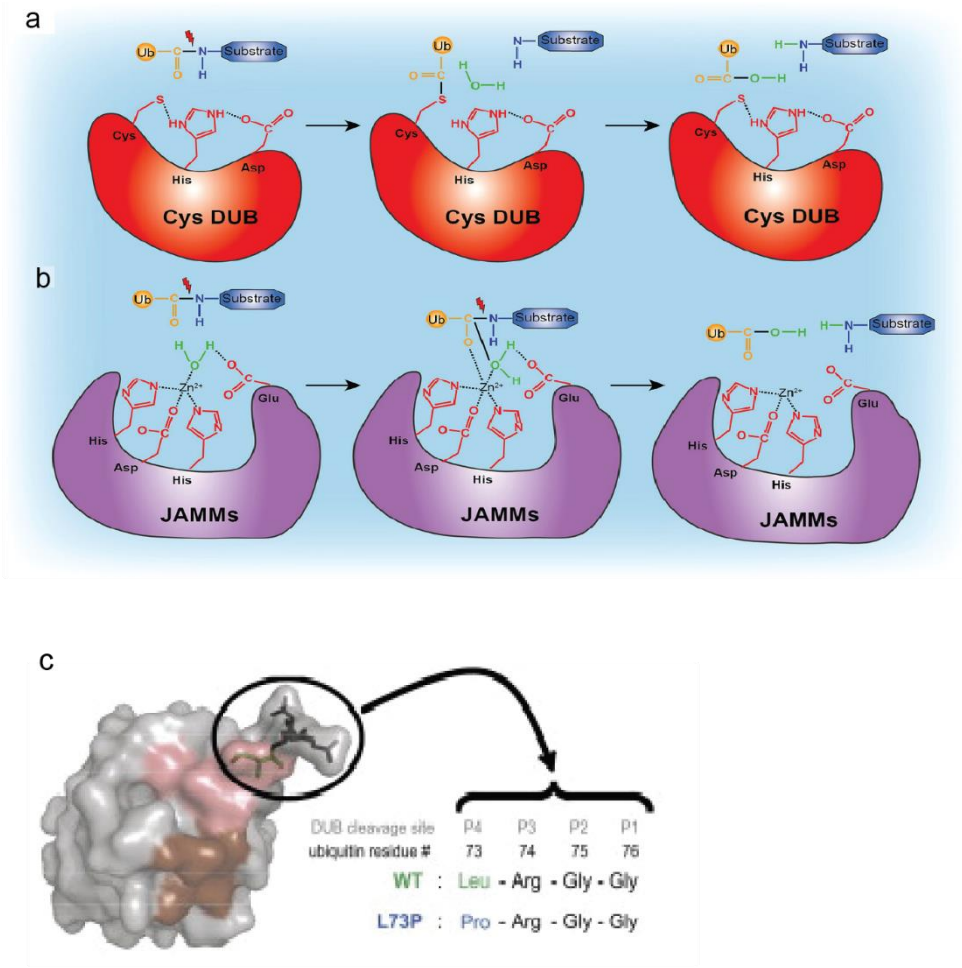



**Figure 10: Catalytic mechanisms of DUB cleavage. a.** Cys DUBs employ a canonical catalytic triad composed of Cysteine-His-Asp/Asn residues to catalyze the hydrolysis of peptide or isopeptide bonds. **b.** The catalytic reaction of JAMM DUBs is facilitated by a polarized water molecule bound to the $Zn^{2+}$ ion, which attacks the carbonyl carbon of the isopeptide bond, leading to the formation of a noncovalent intermediate and cleavage of the scissile isopeptide bond. Figures have been adapted from (Xue et al., 2023). **c.** Surface representation of Ub (PDB ID **1UBQ**) with the C terminus shown in details. The DUB cleavage site (P4-P1, amino acids 72–76 of Ub) is shown in sticks (in black), with Leu73 shown in green. The Ile44 hydrophobic patch is coloured brown on the surface, whereas the Ile36 hydrophobic patch is coloured pink. Figure has been adapted with permission from (Békés et al., 2013) (A.4).

In this study the Leu73, that is the P4 site of Ub for DUB cleavage, was mutated to Pro resulting in loss of DUB binding, rendering a stable Ub variant (**Ub L73P**) (**Figure 10c**). DUBs vary in their substrate preferences according to complexity of Ub chain architectures. Adjacent Ub molecules within a chain are not equivalent. Aside from differentiating through the linkage

type, DUBs may choose between cleaving from the distal end (the Ub which presents its C-terminal Gly to the DUB active site), gradually chewing down the chain (exo-DUB activity), or cleaving within chains (endo-DUB activity). Chain length provides another parameter for discrimination, with some DUBs preferring longer chain types (for example, MINDY, YOD1 (also known as OTUD2 or OTU1) and ATXN3) (Abdul Rehman et al., 2016; Winborn et al., 2008; Mevissen et al., 2013). Others will specialize in cleaving monoUb from specific protein substrates (for example, histone- directed DUBs like USP3, USP16 etc.) or clipping off an intact Ub chain (en bloc cleavage, for example, USP14). The members of the OTU family display diverse chain specificities and thus studying them revealed factors that might be involved in determining the same (Mevissen et al., 2013). As mentioned above, S1' sites within the catalytic domains of DUBs are involved in conferring linkage specificity. Understandably so, as the position of the proximal Ub with respect to the active site determines which lysine residue is presented to the catalytic cleft. Some OTU family members like OTULIN (specific for Met1-linked chains) and Cezanne (specific for Lys11-linked chains) are autoinhibited in the absence of Ub. Only when the proximal Ub with the specific modification on Met1 or Lys11 is bound, is this autoinhibited state released and the DUB becomes active. OTUs have also been found to use a S2 site to bind longer polyUb chains in a linkage specific manner (Mevissen et al., 2013). By contrast, studies of USP family members revealed only limited Ub chain preferences but the catalytic rate constants vary largely across the family (Faesen et al., 2011).

As regulators of the major PTM in cells, DUBs have many implications in cellular processes. Specific DUB functions are often associated with their linkage-specificity and the specificity of products they generate (Clague et al., 2019). There are DUBs within the category of "core fitness genes" are essential across multiple cell types (Meyers et al., 2017). These essential DUBs are widely expressed in high copy numbers (Clague et al., 2015). An essential DUB module comprising a MPN+–MPN combination, PSMD14 and PSMD7 (Rpn11 and Rpn8 in yeast), is involved in substrate processing by the proteasome. USP5 is also a core fitness protein as it is involved in the suppression of the accumulation of unattached Ub chains and maintaining levels of free Ub, the essential currency of the Ub economy. USP36 is also a prominent nucleolar DUB and most likely contributes to cell viability by governing the stability of RNA polymerase I and consequent ribosome biogenesis (Richardson et al., 2012).

Within the 26S proteasome, three catalytically active DUBs from distinct families, USP14 (Ubp6 in yeast), UCHL5 (also known as UCH37) and PSMD14, are associated with the lid of the 19S regulatory particle and coordinate essential proteasomal substrate preprocessing. A large fraction of DUBs can also be found within the nuclear space, where they can influence genome surveillance and repair pathways, epigenetic modifications and chromatin organization as well as transcription. As Ub is post-translational component of the histone landscape, influences epigenetic regulation, DUBs directed to histone deubiquitination are prevalent. For

example, USP22 is a component of the Spt–Ada–Gcn5–acetyl-transferase (SAGA) complex responsible for deubiquitination of histone H2B. USP1 deubiquitinates PCNA in order to prevent the error-prone translesion synthesis repair pathway (Huang et al., 2006). Global proteomics studies have revealed a massive increase in Lys6 and Lys33-linked polyUb chains in response to DNA damage from ultraviolet and ionizing radiation (Elia et al., 2015). Accordingly, many DUBs like USP22 have also been associated with this response (Nishi et al., 2014). OTUB1, on the other hand, although is majorly a Lys48 linkage- specific DUB, it has also been found to limits Lys63 chains in the DSB repair pathway by binding to and inhibiting transfer from the Ub-E2 conjugate of UBC13 (also known as UBE2N).

DUBs are also implicated in the regulation of innate immune receptor signaling through the NF-κB pathway. OTULIN, a stringent Met1 linkage- specific DUB, binds to the LUBAC component of the E3 ligase HOIP (also known as RNF31) (Fiil et al., 2013; Keusekotten et al., 2013). CYLD, a USP family member specific for Lys63-linked and linear Ub chains (Komander et al., 2008), can also indirectly bind to the same domain on HOIP through an adaptor protein, SPATA2 (Kupka et al., 2016). Both DUBs can restrict NF-κB signaling by removing Ub chains from the adaptors. In addition, DUBs are also involved in the regulation of the cell cycle by being active towards Lys11-linked chains that are accumulated on protein targets by the APC/C during anaphase. The OTU family member Cezanne has been shown to control the degradation kinetics of some APC/C substrates like cyclin B and securing. All this emerging knowledge on the specificity and mechanism of action of DUBs is invaluable and adds to the toolbox for studying and understanding the Ub system and hence to various cellular processes they regulate.

## 2.2 Small Molecule Control of UPS

As we have seen, ubiquitination is implicated in diverse biological processes including cell-cycle progression, apoptosis, oncogenesis, protein-quality control, and angiogenesis. Expectedly, the inhibition of the Ub system, and especially the proteasome-dependent degradation of substrates, presents numerous attractive targets for the development of not only therapeutic treatments for many diseases, but also tools for biological research. Each class of enzymes, ranging from the E1 enzymes to the E3 ligases and the multiple subunits of proteasomes, have been tested as drug targets and some have proved to be promising while others have not. In this section, I will discuss how the different modules of the Ub system have been brought under small molecule control and how they have been useful in therapeutics and research.

### 2.2.1 Inhibitors of the E1 Ubiquitin-Activating Enzyme

As we have seen above, the E1 Ub-activating enzyme (UAE) carries out its function in two discrete steps. It first adenylates the C terminus of Ub and then forms a covalent thioester bond with the active-site Cys. Inhibition at the E1 level would prevent ubiquitination and thus globally disrupt the UPS. The first cell-permeable inhibitor was the pyrazone PYR-41, followed by another PYZD4409, both of which were found to covalently inhibit UAE in cells and thus prevent the ubiquitination and degradation by the proteasome of target proteins. Besides the pyrazones, nitric oxide prodrug JS-K and the disulfide NSC624206, both of which act by modification of the catalytic Cys, were reported. The natural product largazole, was also found to block Ub adenylation and hence inhibit UAE (reviewed in Buckley & Crews, 2014; Laplante & Zhang, 2021). Recently, another first-in-class inhibitor of UAE was identified. TAK-243 (MLN7243, **Figure 11a**) is a mechanism-based inhibitor that potently inhibits UAE via formation of a TAK-243–Ub adduct. Treatment of cells *in vitro* with TAK-243 led to the desired loss of cellular Ub conjugates, resulting in defective Ub-dependent protein turnover and signaling, impaired cell cycle progression and defective DNA repair, increased proteotoxic stress, and ultimately cancer cell death (Hyer et al., 2018).

### 2.2.2 Inhibitors of the E2 Ubiquitin-Conjugating Enzymes

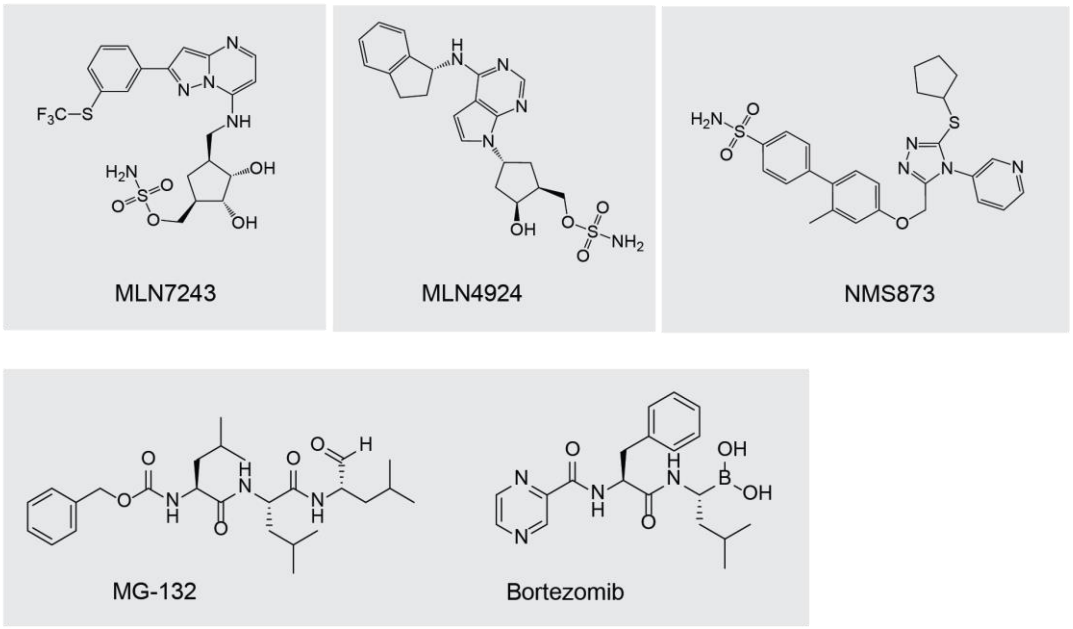
The E2 proteins are responsible for accepting Ub from the E1 enzyme and transferring them to the substrate in conjunction with the E3 ligase (RING type E3 ligases) or for the transfer of Ub to HECT E3 ligases. Despite the existence of a catalytic Cys, the development of E2 inhibitors has lagged behind and only two E2 inhibitors have been reported. First, Ceccarelli et al. discovered a compound CC0651 that inhibited p27Kip1 (involved in cell-cycle progression and a possible target for cancer therapy) ubiquitination using a high-throughput screen of inhibitors of the same. The activity of CC0651 was found to be due to allosteric inhibition of hCdc34, the E2 protein for ubiquitinating p27Kip1. This was confirmed by X-ray crystallography (Ceccarelli et al., 2011). The inhibition of a distinct E2 protein, Rad6 (essential for postreplication DNA repair), was reported to be caused by the compound TZ9. It was developed by the use of a virtual screen and designed to bind to the catalytic site of Rad6. TZ9 was successful at inhibiting histone ubiquitination *in vitro* and inhibited cell proliferation. Unlike CC0651, which acts through an allosteric mechanism, TZ9 is predicted to block thioester formation; this mechanism of action makes TZ9 the first competitive E2-ligase inhibitor (Sanders et al., 2013). Another compound NSC697923 was discovered as an inhibitor of the E2 enzyme UBC13-UEV1a. It also impedes the formation of the UBC13-Ub thioester

conjugate and was found to suppress constitutive NF-κB activity in ABC-DLBCL cells (Pulvino et al., 2012).

### 2.2.3 Inhibitors of the E3 Ubiquitin-Ligating Enzymes

As we have seen earlier, as E3 ligases catalyze the addition of Ub to their target proteins. Since the substrate specificity of the UPS is derived primarily from the selectivity of this interaction, the E3 ligases are attractive targets for the development of targeted therapeutics. However, most E3 proteins lack any enzymatic activity, with the exception of the HECT E3s, which form a thioester bond with Ub before transferring it to their substrates. Therefore, the inhibition of E3 ligases requires the targeting of protein–protein interactions (Buckley & Crews, 2014).

a




b

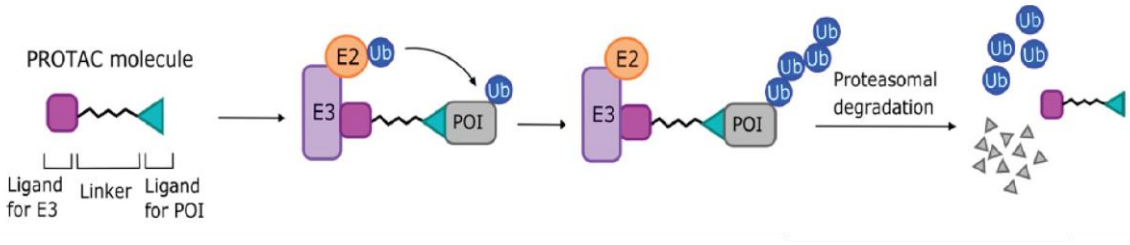


**Figure 11: Small molecules designed to control different steps of the of UPS. a.** Chemical structures of inhibitors that have also been used in this study (as described in later chapters): MLN7243 (UAE inhibitor, MLN4924 (NAE inhibitor), NMS873 (p97 inhibitor), MG-132 and Bortezomib (proteasomal inhibitors). **b.** Schematic diagram representing the mode of action of PROTACs. They contain a ligand for an E3 ligase, a linker, and a ligand for a protein that is to be targeted for degradation, and can then bind to both the E3 ligase and the target, inducing the formation of a ternary complex. This leads to the polyubiquitination of the target protein, followed by its degradation by the proteasome. Figure is adapted from (Laplante & Zhang, 2021).

The first successfully targeted E3 was MDM2, which ubiquitinates the tumor suppressor p53. The company Roche reported the discovery of Nutlins, cis-imidazoline inhibitors of the MDM2–p53 interaction that stabilize p53 in cells and inhibit the growth of tumor xenographs in nude mice (Vassilev et al., 2004). Since then, an orally administered Nutlin derivative, RG7112 (Gonzalez-Lopez de Turiso et al., 2013), has advanced to phase I clinical trials for the treatment of solid and hematological tumors. An additional MDM2 inhibitor, the tryptamine JNJ-26854165, is also orally administered and in phase I clinical trials but appears to act instead by blocking the interaction of the p53– MDM2 complex with the proteasome. Numerous other classes of MDM2 inhibitors have been developed, including the spirooxindoles, which were discovered through structure-based design and include MI-219 and MI-888 (reviewed in Buckley & Crews, 2014). However, although p53 is an important tumor suppressor, attempts to stabilize it will have little effect on the large percentage of cancers with mutated p53. Although the development of potent MDM2 inhibitors remained long relevant in the p53 field, the importance of therapeutic targeting MDM4 was recognized later. Inhibitors such as SJ-172550, WK298 and stabilized peptide SAH-p53-c8 have been developed against MDM4. In addition, simultaneous inhibition of both MDM2 and MDM4 enzymes, by means of dual inhibitors, enhances p53 activation. Stapled peptide ALRN-6924 (Aileron, Inc.) is an example of dual inhibitor that is currently undergoing clinical trials against solid tumors and lymphomas (reviewed in Bulatov et al., 2018).

The inhibitor of apoptosis protein (IAP) family of E3 ligases has also been extensively targeted, usually with largely peptidic or peptidomimetic inhibitors inspired by the natural protein inhibitor of the IAPs, SMAC/DIABLO, and numerous of these compounds are in clinical trials (Cohen and Tcherpakov, 2010).

Next, numerous SCF members have been targeted with small molecule inhibitors. $SCF^{Skp2}$ ubiquitinates numerous proteins involved in cell-cycle control, such as $p27^{Kip1}$. $SCF^{Skp2}$ inhibitors were developed through virtual screening to target the $SCF^{Skp2}$–p27 interface. Both the rhodanine C1 and the pyrrolinone C2 inhibited the ubiquitination of p27 in vitro. Recently, a class of chromones was also reported to inhibit $SCF^{Skp2}$ by preventing the binding of Skp2 to the remainder of the SCF complex. The lead, compound 25, was found to have antitumor activity in animal models (Chan et al., 2013). SMER3 was discovered through the use of a

chemical-genetics screen of enhancers of rapamycin and found to inhibit Met4 ubiquitination by $SCF^{Met30}$ through blockage of the interaction between Met30 and the core of the SCF complex (Aghajan et al., 2010). Furthermore, small molecules that can bind to the VHL protein, the substrate-recognition subunit of a CRL, in a competitive mechanism were reported. But these compounds were only able to inhibit VHL *in vitro*, not in cell-based assays. Thalidomide, that was originally developed for its sedative properties before it was infamously discovered to be a potent teratogen that causes serious birth defects, is still in use today for the treatment of serious disorders, such as leprosy and multiple myeloma, despite its serious side effects. A study has found that thalidomide binds to and inactivates the E3 ligase cereblon (CRBN), a component of a CRL important to limb development. This finding strongly indicates a mechanism for the side effects of thalidomide, and could possibly enable the development of thalidomide derivatives that do not target CRBN (Cohen and Tcherpakov, 2010; Ito et al., 2010)

In addition, it is reasonable to mention that the auxiliary protein NEDD8 that serves as a crucial regulator of CRL functions (Soucy et al., 2009). The ubiquitin-like protein NEDD8 gets covalently conjugated to a specific conserved Lys at C-terminal domain of Cullin scaffold (NEDDylation). NEDD8 conjugation enhances substrate ubiquitination by promoting CRL dynamics and inducing conformational shift of the protein ensemble that brings together E3-bound substrate and the E2 (Duda et al., 2008, 2011). Pevonedistat (also known as MLN4924, **Figure 11a**) is a first-in-class selective inhibitor of NEDD8-activating enzyme (NAE) that mediates NEDDylation (Soucy et al., 2009). It covalently binds NEDD8, forming a complex that prevents NAE's conjugation of NEDD8 to CRL and causes accumulation of CRL substrates. Pevonedistat is currently undergoing several clinical trials against different types of cancer (Swords et al., 2015).

### 2.2.4 Inhibitors of Proteasomal Degradation (Proteasome and Accessories)

Numerous classes of proteasome inhibitors have been described, (Rentsch et al., 2013) including peptide aldehydes, peptidyl boronates, epoxyketones, vinyl sulfones, β-lactones, hydroxy-ureas, and α-ketoaldehydes. Many contain an electrophile that targets the key nucleophilic $Thr^{1}$ residues in the catalytic β1, β3, and β5 subunits of the proteasome. These electrophiles are often attached to a linear or cyclic peptide chain that mimics the substrate protein. Additionally, noncovalent inhibitors have been reported, such as TMC-95A and the noncompetitive imidazoline class (reviewed in Buckley & Crews, 2014).

The first class of proteasome inhibitors identified were the peptide aldehydes, of which MG132 (or Z-LLL, by Myogenics, **Figure 11a**) is the most studied (Palombella et al., 1994). These

compounds are potent ($K_i$ = 4 nM) covalent inhibitors of the chymotryptic-like activity of the β5 subunit. However, the lack of selectivity towards other Cys proteases (such as calpains and cathepsins), rapid oxidation, and the rapid reversibility of the formation of a hemiketal with Thr 1 prevented these compounds from being useful therapeutically. Nonetheless, the fact that it is highly potent and widely available at relatively low costs, has made MG132 one of the most used proteasome inhibitors in biological studies. Other than synthetic compounds, natural products like fellutamide B, has proved to be a proteasomal inhibitor (Hines et al., 2008). Potency of peptide aldehydes as proteasome inhibitors led to the discovery of the peptidyl boronates. It was found that the converting the aldehyde group of MG132 into a boronic acid (MG262 or Z-LLL-boronate) led to a vast improvement in its $K_i$ ( ~ 18 pM) (Kisselev and Goldberg, 2001) and also enabled the development of smaller, dipeptidyl boronic acids, such as **bortezomib** (developed by Myogenics/ProScript, **Figure 11a**). This class of compounds has an added advantage as the boronic acid moiety is far less capable of reacting with Cys proteases, conferring them with increased specificity (Adams et al., 1998).

Another class of proteasome inhibitors are the epoxyketones. The first epoxyketones described were eponemycin and epoxomicin (Hanada et al., 1992), both of which are natural products. Crews and colleagues showed that the epoxyketones reacted with Thr 1 of the proteasome to first form a hemiketal, and that the subsequent attack of the terminal amine on the epoxide led to a stable morpholine ring (Groll et al., 2000). This two-step nucleophilic attack on the epoxyketone not only led to irreversible inhibition of the proteasome (as opposed to the slowly reversible inhibition by peptidyl boronates, such as bortezomib), but also leads to specificity, owing to the uniqueness of the N-terminal Thr residue amongst proteases. Further optimization for specificity towards the chymotrypsin-like activity of the proteasome yielded YU101 (later circulated as carfilzomib). Although bortezomib/velcade and carfilzomib/kyprolis are both FDA-approved proteasome inhibitors at the moment, the development of improved proteasome inhibitors to improve their bioavailability. Currently there are multiple orally bioavailable proteasome inhibitors in clinical trials, such as ixazomib citrate (MLN9708) from Millennium Pharmaceuticals and oprozomib (ONX 912) from the company Onyx (Huber and Groll, 2012).

Recently there are efforts in developing inhibitors that specifically target the immunoproteasome, which is predominantly expressed in lymphocytes. Targeting the immunoproteasome has been proposed as a potential therapeutic approach for autoimmune and neurodegenerative disorders. Selective inhibitors like ONX 914 demonstrate a preference for the immunoproteasome over the constitutive proteasome.

Despite the success in the discovery and development of proteasome inhibitors, one concern remains. The inhibition of protein degradation leads to the formation of protein aggregates, that is linked to neurodegenerative disorders, such as Alzheimer's disease. However, bortezomib and carfilzomib do not readily cross the blood–brain barrier, and their effect on the central

nervous system should therefore be limited (Demo et al., 2007; Gilardini et al., 2008), although, as noted above, peripheral neuropathy is a serious side effect of bortezomib treatment.

In addition to directly inhibiting proteasomal degradation of ubiquitinated substrates by targeting the proteasome, other proteins involved in degradation pathways can also be targeted to bring about the same outcome. For instance, as we have seen in section 2.1.2.2, VCP/p97 aids in ERAD by providing the energy for the retrotranslocation and dislocation of ER substrates within the membranes, into the cytosol, finally being targeted for degradation in the proteasome. A covalent, allosteric inhibitor of p97 was developed, called NMS-873 (Magnaghi et al., 2013, Figure 11a), which was found to activate unfolded protein response in cells, interfered with autophagy, ultimately leading to cancer cell death.

### 2.2.5 DUB Inhibitors

DUBs, as we already know, reverse the action of the E3 ligases by editing or removing Ub chains via their proteolytic activity. Thus, inhibiting DUBs leads to selective degradation of substrates as it prevents the removal of their Ub tags. DUBs are attractive targets for developing small-molecule inhibitors against as they have substrate specificities and defined catalytic sites.

USPs are the largest family of DUBs and are characterized by a catalytic USP domain of variable size containing six conserved motifs. USP7 deubiquitinates various substrates implicated in numerous cellular processes, most importantly, it has a paradoxical role in the p53–MDM3 axis. On one hand, USP7 deubiquitinates and stabilizes p53, reversing the action of MDM2 (Li et al., 2002). However, USP7 can also stabilize MDM2 by opposing its auto-ubiquitination. Interestingly, when cellular USP7 levels are decreased, p53 is destabilized. But when USP7 activity is abolished, MDM2 is downregulated, stabilizing p53 (Li et al., 2004). The first generation of small-molecule USP7 inhibitors included HBX 41108, HBX 19818, HBX 25258, P5091, P50429, and P22077 (reviewed in Laplante & Zhang, 2021). Out of these P50429 and P22077 binds irreversibly to the catalytic C223 of USP7. However, none of these first-generation USP7 inhibitors had optimal potency or specificity as they were developed using an activity-based approach. Almost a decade later, several potent and selective allosteric USP7 inhibitors were discovered through combination of fragment-based screens and structure-guided optimizations. Gavory and colleagues first identified ‘Compound 1’ (Gavory et al., 2018) and subsequently improved on it to produce ‘Compound 4’, which noncompetitively binds an allosteric site distal to the USP7 catalytic site, exhibiting ~2000-fold increased potency and better cell permeability. Other allosteric modulators include GNE-6640 and GNE-6776 which bind at the catalytic domain 12 °A away from the catalytic Cys, preventing Ub binding (reviewed in Laplante & Zhang, 2021). FT671 and FT827 are more

allosteric inhibitors with high affinity exhibited both *in vitro* and in human cells. A non-covalent inhibitor XL188 was also discovered.

USP1 is also an important DUB whose catalytic activity is stimulated upon complex formation with UAF1 (Cohn et al., 2007). Two small molecules, pimozide and GW7647, were identified in 2011 and found to non-competitively inhibit USP1. USP14 is one of three proteasome-associated DUBs and IU1 was discovered to be its selective inhibitor, identified via a small-molecule HTS strategy (reviewed in Laplante & Zhang, 2021). Structural analysis revealed that this family of molecules functions by binding a previously unknown steric site on USP14. Analogs of IU1 have been further developed to increase potency, membrane permeability. Another IU1 analog, 1B10, also has heightened potency, as well as better membrane permeability. The antirheumatic drug auranofin (also known as Ridaura) also has inhibitory action against USP14 (Li et al., 2021). Another molecule, b-AP15, was discovered to be an inhibitor of USP14 and another proteasome-associated DUB, UCHL5. It selectively blocks 19S regulatory particle DUB activity and was shown to induce tumor cell apoptosis in various solid tumor types. Rpn11 is the third proteasome-associated DUB, belonging to the JAMM family of metalloproteases. Several small-molecule inhibitors of Rpn11 have been identified, including 8TQ and Capzimin. Non-specific inhibitors of Rpn11 include SOP11, an epidithiodiketopiperazine molecule, and thiolutin, a naturally-occurring antibiotic, which both also inhibit other JAMM family DUBs. Recently, UCHL1 became a popular target for inhibition and a fluorometric cell lysate-based assay called AlphaLisa was used to develop UCHL1 inhibitors that yielded a series of inhibitory compounds, the most selective being celasterol and mangaferin (reviewed in Laplante & Zhang, 2021).

Furthermore, dual USP9X/USP24 inhibitors EOAI3402143 (aka G9) were developed (Peterson et al., 2015). And so were inhibitors against OTUB2 (otubain-2), that preferentially cleaves Lys63-linked poly-Ub chains but also has activity against Lys11- and Lys48-linked chains (Resnick et al., 2019).

In addition to these specific DUB inhibitors, many broad-spectrum DUB inhibitors, such as chalcone molecules G5 and F6 (Aleo et al., 2006), PR-619 (Altun et al., 2011), and betulanic acid (Reiner et al., 2013) have been discovered, but unspecific inhibition of DUBs can result in undesirable off-target effects and high toxicities, making them unideal for clinical application but have found applications in biological studies.

Moreover, chemical reagents that are routinely used in biological assays for broad range inhibition of DUBs include N-ethylmaleimide (NEM) and 2-chloroacetamide (2-CAM), both of which block the sulfhydryl group of cysteines, thus blocking the Cys protease families of DUBs. Being a chelating agent, ethylenediaminetetraacetic acid (EDTA) is often used to block metalloproteases like the JAMM family DUBs.

### 2.2.6 PROTACs

Discussion about small molecule regulators of UPS is incomplete without the mention of Proteolysis Targeting Chimeras (PROTACs). PROTACs, are heterobifunctional molecules (Corson et al., 2008) that contain a ligand for an E3 ligase, a linker, and a ligand for a protein that is to be targeted for degradation (**Figure 11b**). The molecule can then bind to both the E3 ligase and the target and thus induce the formation of a ternary complex. This hijacking of the E3 ligase can lead to the polyubiquitination of the target protein, followed by its degradation by the proteasome.

Early PROTACs were peptide-based, which had limitations due to instability, poor cell permeability, and low activity. The first PROTAC (PROTAC-1) was developed by Sakamoto et al. in 2001. It contained a phosphopeptide derived from IκBα, which binds to the E3 ligase $SCF^{\beta TrCP}$, as well as a moiety derived from ovalicin, which covalently binds MetAP-2. PROTAC-1 was able to specifically and covalently bind to MetAP-2, which was then recruited to $SCF^{\beta TrCP}$ and ubiquitinated *in vitro* (Sakamoto et al., 2001). The first cell-permeable PROTACs (PROTAC-4 and PROTAC-5) were developed by the incorporation of a peptide derived from HIF (ALAPYIP) that binds to VHL, the substrate recognition portion of a CRL E3 ligase. A poly-D-arginine chain was also added to aid cell permeability. This was, in turn, linked to a ligand for FKBP12 (F36V) to give PROTAC-4, which was able to degrade GFP–FKBP12 (F36V) in cells efficiently (at 25 μM). PROTAC-5 contained a DHT moiety and was used to degrade an androgen receptor/GFP fusion protein (at 25 μM) in cells. Since then, numerous PROTACs have been developed for use in therapeutics and research and have been reviewed extensively (Buckley & Crews, 2014; Laplante & Zhang, 2021; Békés et al., 2022).

Subsequently, variants of PROTACs have also been developed, like the specific and non-genetic IAP-dependent protein erasers (SNIPERs). The first of these capitalized on the discovery that methyl bestatin (MeBS) interacts with cIAP1 and induces its autoubiquitination. Researchers added a ligand for the target protein, cellular retinoic acid binding proteins (CRABPs) I and II, to the methyl residue of MeBS. Consequently, the first SNIPERs induced the simultaneous degradation of the recruited E3 and the target proteins. Additionally, PROTACs that can be activated by UV or visible light, termed photo-PROTACs, opto-PROTACs, or PHOTACs have also been developed. These could be applied clinically to directly target tumours via photodynamic therapy (PDT). Other variants include homo-PROTACs and CLIPTACS (reviewed in Laplante & Zhang, 2021)..

## 2.3 Methods to Study Proteome-Wide Ub Dynamics

Although ubiquitination had been discovered way back in the 1950s, methods to elucidate ubiquitination dynamics on a proteome-wide level have been established rather recently, to gain insights into Ub turnover rates, protein-ubiquitination dynamics and cellular Ub pools. However, the challenge of controlling cellular ubiquitination, especially in a linkage-specific manner on rapid time-scales, still persists. These challenges mainly stem out of the transient nature of ubiquitinated species and the highly ubiquitous nature of Ub making it difficult to selectively control one target or one Ub linkage type. Nonetheless, a number of strategies have been developed to date that deepen our understanding of proteome-wide ubiquitination kinetics, that rely on effective perturbation of cellular ubiquitination and that enable controlling and tracking of ubiquitination events at high temporal resolution.

### 2.3.1 Isotope-Labeling and Mass-Spectrometry based Approaches

Proteomic studies provide important insights into the molecular mechanisms involved in cellular protein homeostasis during physiological processes. Metabolic turnover and protein half-life measurements are generally based on "pulse-chase" approach, which requires the introduction of radioactive, biochemical, or stable isotope-labelled tracers into target proteins. These tracers are introduced into the target cell's metabolism (e.g., $^{13}$C-labeled carbohydrates, $^{15}$N-labeled ammonia salts, or $^{13}C_6$-lysine or $^{13}C_6$ $^{15}N_4$-arginine). These tracers can then be followed using a suitable detection system (Wilkinson, 2018). The "pulse" is then followed by a "chase" period, in which the labeled tracer is replaced or "chased" away by an excess of the same unlabeled compound after a given time-point. Depending on the experiment, monitoring the labeled tracer over time measures its incorporation (protein production) and/or its loss (protein degradation).

Mass-spectrometry based detection methods have been used since more than 80 years to analyze protein synthesis and degradation (Foster et al., 1939; Schoenheimer, 1946). Shotgun proteomics approaches in combination with translation inhibition have been used to quantify the differences in subcellular proteome turnover (Larance et al., 2013), and are used along with techniques such as Western blotting or immunofluorescent imaging to determine subcellular localization. Since early 21st century, the combination of high-resolution liquid chromatography (LC), nano electrospray ionization (ESI), and ultrahigh-resolution tandem mass spectrometry with fast MS/MS cycles enabled the quantitative analysis of thousands of peptides and proteins in a few hours. Amino acids with a defined and selectable number of

stable isotopes are introduced into culture media or food sources and used for quantitative studies (stable isotope labeling of amino acids in cell culture [SILAC]), and the isotopic

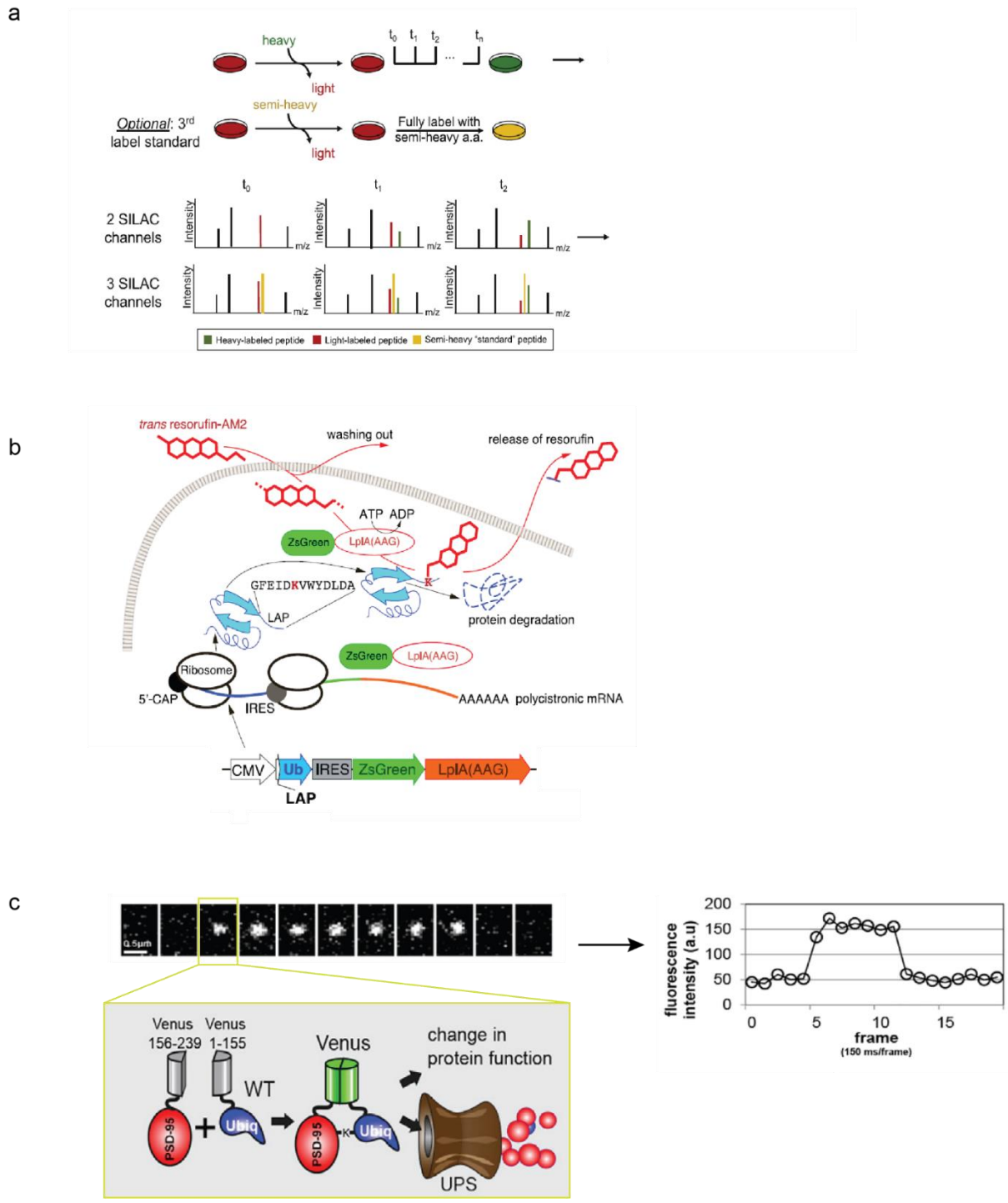



**Figure 12: Methods to study proteome-wide Ub dynamics. a.** Dynamic SILAC workflow: Sample preparation followed by data acquisition. In a standard two-label dynamic SILAC experiment, cultures are plated in an unlabeled media and then the media is swapped for the one containing stable isotope–labeled amino. Samples are then collected over a time course, with a separate culture harvested at each time point. A fully labeled sample using a third stable isotope, typically a "semi-heavy /medium-heavy" isotope, can also be generated as normalization standard for data analysis. LC-MS/MS enables direct monitoring of "light" (red) and "heavy" (green) peptide signals, which correspond to pre-existing and

newly synthesized proteins, respectively. In three-channel designs, the signal from the constant "semi-heavy"-labeled sample (yellow), spike-in provides an internal normalization standard between different MS measurements, allowing for relative signal from "light" and "heavy" channels to be quantitated. Figure adapted and modified from (Ross et al., 2021). **b.** PRIME methodology adopted for Ub turnover studies. Expression of resorufin ligase LplA(AAG) and the Ub fused with with LAP is synchronized due to the IRES. Cells are pulsed with resorufin, and further resorufin is washed out due to protein degradation. Figure adapted and modified with permission from (Kudriaeva et al., 2021) (A.4). **c.** Visualization of *de novo* ubiquitination events using SM-UbFC. Raw data montage showing one reconstituted split-Venus molecule detected. Upon reconstitution, fast folding of Venus promotes its fast detection with single-molecule resolution. The quantification of fluorescence intensity over time indicates one-step appearance and one-step disappearance, characteristic for single-molecule imaging. Figure has been adapted and modified from (Ifrim et al., 2022a).

envelopes of “light” and “heavy” peaks are separated by predefined shifts (e.g., $^{13}C_6$-lysine or $^{13}C_6$ $^{15}N_4$-arginine), greatly facilitating data acquisition and interpretation. Coupling lysine/arginine labeling with trypsin as a protease for sample preparation guarantees that nearly each peptide has a labeled amino acid.

This quantitative approach has been used to study proteome turnover in “pulsed-only” experiments. In these so-called “dynamic SILAC” experiments (**Figure 12a**), cells are switched from unlabeled medium to a medium containing isotopically labeled lysine and/or arginine. Samples are then measured via LC-MS/MS over a time course. The rate of appearance of a heavy amino acid–labeled peptide signal corresponds to its rate of synthesis, whereas the rate of disappearance of a light amino acid–containing peptide corresponds to its rate of degradation. And the ratio of heavy to light peptide signal thus directly reflects protein turnover. Schuman and Tirell labs developed bio-orthogonal non-canonical amino acid tagging (BONCAT) to overcome difficulties in detecting low-abundant, nascent proteins (Dieterich et al., 2006). BONCAT makes use of natural amino acid surrogates, typically methionine mimetics, that usually carry an azido or alkyne functional group and can be immobilized on a solid phase using click chemistry and affinity purification. BONCAT can also be used to visualize overall proteome synthesis in cells using fluorescent non-canonical amino acid tagging (Dieterich et al., 2010) or to measure synthesis of target proteins in a spatially resolved manner using a proximity ligation assay (tom Dieck et al., 2015). It can also be combined with pSILAC to improve the accuracy of nascent proteome analyses by the incorporation of heavy isotope–labeled amino acids (reviewed in Ross et al., 2021) .

To study Ub turnover, many approaches have been utilized that primarily perturb the cellular ubiquitination and enable tracking of ubiquitination events. One such approach was the development of lysine-less Ub (K0-Ub) in order to study proteome-wide ubiquitination sites (Oshikawa et al., 2012). K0-Ub is unable to form Ub chains because of the lack of the ε-amino group of lysine, but it can be linked to lysine residues of target proteins via an isopeptide bond, trapping substrates at a mono-ubiquitinated state. Two-step digestion of K0-Ub-modified

proteins with lysyl endopeptidase (Lys-C) and trypsin coupled with affinity purification allows the efficient enrichment of Ub signature peptides derived from target proteins and the precise identification of ubiquitination sites. These signature marks are those of the Gly-Gly (GG, or diGly) remnant attached to the modified lysine left after trypsin cleavage. This GG remnant is widely used to identify lysine residues harboring these PTMs by MS (Prus et al., 2024). Very recently, a combination of GG remnant profiling, partial chemical modification, and serial dilution SILAC (SD-SILAC) methods was used to systematically quantify site-specific ubiquitination. Using NHS-Gly-Gly-Boc, SILAC heavy-labeled proteins were partially modified with GG remnant and the degree of partial chemical GG modification (PC-GG) was measured using MS. A known amount of PC-GG-modified proteins (SILAC-heavy) was introduced into native proteins (SILAC-light). Following trypsin proteolysis, GG-modified peptides were enriched and quantified by MS and the site occupancy was calculated. Using this approach, it was concluded that global ubiquitination occupancy is low, but spans a large dynamic range. Furthermore, to determine site-specific deubiquitination kinetics, cells were treated with TAK-243 (E1 inhibitor) for 5, 10, 30, and 60 min, and the change in ubiquitination was quantified using SILAC-based MS. Ubiquitination sites were classified into four categories: ‘‘very fast,’’ ‘‘fast,’’ ‘‘slow,’’ and ‘‘very slow,’’ corresponding to half-lives of <5, 5–15, 15–60, and >60 min, respectively. Over 22% of sites exhibited a half-life of <5 min, 45% had a half-life of <10 min, and 67% had a half-life of <30 min.

Another approach, synthetic peptide absolute quantification (AQUA) MS, has enabled precise quantification of mono- and polyubiquitination of purified substrates and relative quantification of affinity-captured polyUb species from cell and tissue lysates . (Kaiser et al., 2011) developed a method that combines differential affinity chromatography and protein standard absolute quantification (PSAQ) MS11 to enable the precise measurement of cellular molar concentrations of Ub. This strategy, termed Ub-PSAQ, uses stable isotope–labeled free Ub and Ub conjugates as recovery standards, which are ‘spiked’ into cell lysates and captured with affinity reagents that are selective for free Ub or chains. Using this approach the steady-state distribution of Ub pools was determined in tissue culture cells before and after inducing proteasome stress and in lysates of mouse and human brain. Tissue samples were divided into two equal portions. One portion was used to determine total Ub concentration after converting all Ub conjugates to free Ub using the DUB USP2 catalytic domain (USP2CC12). Free Ub species were isolated from this portion and from the portion not treated with DUB using the zinc-finger domain from isopeptidase T (BUZ) as it can bind specifically to the free C-terminal diGly motif of Ub. PolyUb chains were isolated by affinity capture with the human PLIC2 Ub-association domain (hP2 UBA). Next the affinity-captured material was washed and digested with trypsin and analyzed by LC-ESI TOF MS, and the sample-derived Ub species were measured.

### 2.3.2 Fluorescence based Approaches

The complexity of Ub turnover studies also arises from its unique structure, which restricts fusion with fluorescent proteins and tags longer than 20 amino acids, the difficulties with immunoprecipitation of polyUb conjugates, and the impossibility of quantitative analysis by blotting techniques. (Kudriaeva et al., 2021) determined fundamental parameters of Ub metabolism utilizing its real-time fluorescent tracking. The group used PRIME (probe incorporation mediated by enzymes, **Figure 12b**) technique based on mutated *E.coli* lipoic acid ligase (LplA) that conjugates the ε-amino group of lysine in a short 13-amino acid peptide tag with chemical fluorophores (Cohen et al., 2011; Uttamapinant et al., 2010) and used for real-time and endpoint monitoring of Ub metabolism in physiological conditions. A triple mutant E20A/F147A/H149G LplA ligase (herein after referred as LplA(AAG)) was used for intracellular labeling of recombinant proteins containing ligase acceptor peptide (LAP) with a derivative of the low molecular-weight fluorophore resorufin, emitting at 595 nm (Liu et al., 2014). Stably transduced HEK293T cells expressing LAP fused N-terminally to Ub (LAP-Ub) and ZsGreen-LplA (AAG) were pulsed (i.e., incubated with resorufin-AM2). The stained cells were treated by DMSO, proteasome inhibitor and DUB inhibitors for 6 h and further labelled with resorufin-AM2. Immediately after washing, cells were lysed and subjected to PAGE. Analysis of in-gel resorufin fluorescence showed classic traces representing polyUb chains conjugated to numerous intracellular proteins. Using this methodology and Ub profiling revealed that Ub intensity decreased twice in ~4 h, and unlabeled Ub completely replace resorufin-labeled Ub in polyUb chains in 10 h. The pool of unconjugated Ub is stabilized after 6 h of exposure to the PS-341, which means a balance is established between Ub ligases and DUBs.

Furthermore, (Ifrim et al., 2022b) exploited the covalent bond formation between a Lys of the target protein and Ub and applied it in the widely used bimolecular fluorescence complementation (BiFC) (Hu et al., 2002) to visualize ubiquitinated proteins by ubiquitin mediated fluorescence complementation (Ub-FC) (Fang and Kerppola, 2004). Their method, "Single-Molecule Ubiquitination Mediated Fluorescence Complementation" (SM-UbFC, **Figure 12c**), uses single-molecule imaging of reconstituted split Venus generated by Ub mediated fluorescence complementation (Fang and Kerppola, 2004). SM-UbFC combines the advantages of single-molecule detection of Venus fluorescence protein in live neurons (Tatavarty et al., 2012) with those of UbFC (Fang and Kerppola, 2004). However, the study makes use of SM-UbFC, which images the global ubiquitination rate of a protein of interest in live cells. SM-UbFC cannot, however, distinguish between mono- and polyubiquitination, between different types of polyUb chains, or between different sites of ubiquitination on the

protein of interest. Using this approach, the authors detected ubiquitination sites in dendrites with high spatial, temporal and molecular resolution.

So, as is evident, there have been seminal studies that provide important insights into proteome-wide Ub turnover at high temporal resolution, but information about linkage-specific Ub dynamics remains to be scarce. It remains challenging to control the fast and complex dynamics of the Ub system that demands rapid control of the individual components and modification events. Controlling ubiquitination using a tool that confers both spatial and temporal resolution, as well as, be able to confer linkage-specificity is still unknown.

# 3. Light as a Fundamental Tool to Control Biological Processes

Light is a central driver of natural biological processes like photosynthesis and photomorphogenesis in plants, regulating circadian rhythms across all kingdoms of life and ofcourse, helping us to visually perceive our surroundings. In all of these pathways, photo-responsive proteins, or photoreceptors, absorb specific wavelengths of light and bring about the desired effector functions, finally activating or inhibiting a cellular pathway (Fisk et al., 2018; Partch et al., 2014; Möglich et al., 2010). Light enables fast switch on/off kinetics of the pathways it regulates. Also, with its easy and focussed availability, light can be applied to targets at high spatial and temporal resolution, in a non-invasive set-up. All these attributes make light an excellent tool to control cellular pathways, enabling studies of dynamic processes.

Currently, there are two main methods available for applying photocontrol in living systems: optochemical genetics and optogenetics. These methods differ in their receiver modules that bear core chromophores to initiate light-responsive reactions (reviewed in Kneuttinger, 2022).

Optogenetic tools fuse protein effector domains with naturally occurring sensory or signaling photoreceptors, which absorb light through an organic, non-protein component known as a chromophore, and undergo reversible conformational changes to translate these photo responses into the controls of the hybridized protein effectors (**Figure 14**) (Banerjee & Mitra, 2020; Möglich et al., 2010). Optochemical genetics, on the other hand, exploits chemically synthesized small molecules as the core chromophore in accordance with its name, and it can be further categorized into two approaches: the photopharmacology approach (**Figure 15a**) and the codon-specific protein modification approach (**Figure 15b**).

A suitable strategy for photocontrol is selected for use in specific studies on the basis of the principles to implement the above-mentioned methods. In this section, the principles of implementation of photocontrol methods have been discussed.

## 3.1 Optogenetic Tools

Light-sensitive proteins or photoreceptors sense light of respective wavelengths and activate the appropriate behavioral responses in photosynthesis, photomorphogenesis and more in plants, visual perception and circadian rhythms in animals and even phototropic movements in bacteria.

Optogenetic tools fuse protein effector domains with naturally occurring sensory or signaling photoreceptors (**Figure 13a**), to engineer light-sensitive proteins (reviewed in Banerjee & Mitra, 2020). The photoreceptor modules of these engineered proteins absorb light through chromophore of the photoreceptor and undergo conformational changes to translate these photo responses into the controls of the hybridized protein effectors. For example, rhodopsins, as discussed above, are naturally occurring light-gated ion channels and have been used as powerful optogenetic tools that alter ion flow in response to a light trigger. Optogenetics was developed as a tool to study responses in neuropathological conditions. In the seminal experiment, that marked the beginning of this diverse field of optogenetics, the gene encoding channel channelrhodopsin-2 (ChR-2) from *Chlamydomonas reinhardtii* (ChR-2) (Nagel et al., 2003) was inserted into mammalian neuronal cells that selectively secrete the neurotransmitter glutamate, by transfection using a benign lentiviral vector. Following expression and using blue-light pulses, Deisseroth and colleagues (Boyden et al., 2005) could attain reliable and conditional control of neuronal spiking and transfer information from one neuron to another, at a millisecond timescale. Since then, optogenetics has been expanding and currently accommodate even synthetic and engineered photoreceptors that have been used in a wide range of applications to decode cellular dynamics beyond neuroscience.

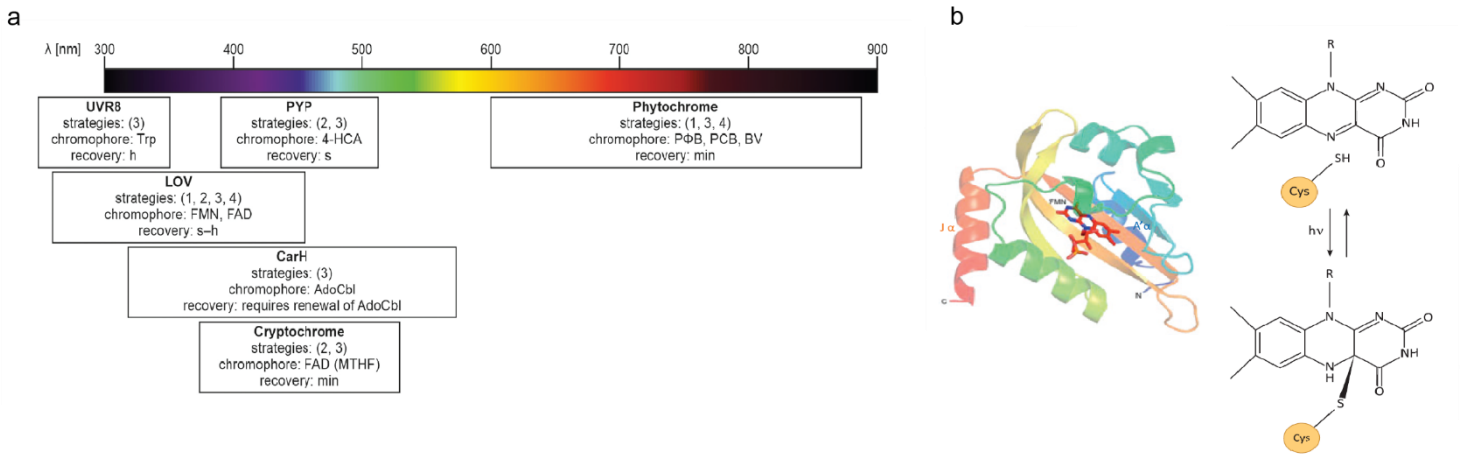



**Figure 13: Use of naturally occurring photoreceptors as optogenetic tools. a.** Representative examples of photosensor modules and their respective absorption spectrum. Figure adapted from (Kneuttinger, 2022). **b.** Structure and photochemistry of Light-oxygen-voltage (LOV) photosensor domain from Aureochrome1 of Vaucheria frigida (PDB ID: 3UE6) comprises a conserved Per-Arnt-Sim (PAS)/LOV fold. Upon absorption of 450 nm blue light, the dark state conformation converts to a light state conformation by intersystem crossing. Within microseconds, a covalent, thioether bond is formed between the C4a atom of flavin and the cysteinyl sulfur atom of the protein. This S390 state constitutes the LOV signaling state. The dark state is again generated in the reverse step of the photocycle, with the loss of light signal and disruption of the covalent linkage. Figure adapted from (Banerjee & Mitra, 2020).

Using light as a control trigger, optogenetics has many advantages, including its high spatio-temporal resolution, the easy availability of light, the ability to be conditional by quickly turning light ON and OFF, non-invasive delivery of the light signal, the potential for **reversibility on a biologically meaningful timescale**, minimal off-target effects, and the resulting ability to dissect complex signaling pathways. Additionally, the optogenetic tools are **modular** in nature, meaning each photosensor module can be connected to diverse effector modules and vice versa. And by virtue of these advantages, optogenetics was selected as one of the ten breakthroughs of the decade (INSIGHTS OF THE DECADE, 2010) and the method of the year in 2010 and 2014 (Rusk, 2011; *Science.346.6209.568*, n.d.).

Amongst the most commonly used photoreceptors as optogenetic tools, the blue-light-controlled light–oxygen–voltage (LOV) family of protein domains is one (**Figure 13b**). LOV-domains rely on non-covalently bound flavin mononucleotide (FMN) or flavin adenine dinucleotide (FAD) chromophores and have the smallest size among currently available photosensory modules used for optogenetics. Light irradiation induces the formation of a transient bond between a Cys in the LOV domain and the C4α position of flavin. As a result, the conformation of the LOV domain changes, enabling the LOV domain to function as a photoswitch (Manoilov et al., 2021; Möglich et al., 2010). Using this property, many optogentic tools have been developed to study various biological processes.

### 3.1.1 Design Strategies and Applications

The resulting conformational change in natural photoreceptor proteins following light absorption has been one of the principles of engineering optogenetic proteins. For instance, the LOV2 protein domain of *Avena sativa* Phototropin1 (AsLOV2) was used to make a light activatable *Escherichia coli* Trp repressor (TrpR), which they named LOV-TAP, in order to control its DNA-binding activity (Strickland et al., 2008). In this synthetic LOV-TAP photoreceptor, the LOV2 domain is attached to the N-terminus of TrpR through the J$\alpha$-helix connector from LOV2 (**Figure 14a**). Upon irradiation with blue light, the Jα-helix unfolds, leading to a conformational change in TrpR and restoring its active form, rendering it capable of binding its target DNA sequence. Such conformational changes can also be exploited to inhibit protein functions. One such example is when Dagliyan and co-workers (Dagliyan et al., n.d.) inserted AsLOV2 into loops within effector proteins (such as Src kinase, Rac1 GTPase, and Vav2) that otherwise do not interfere with normal protein function. Upon light irradiation, LOV2-Jα unfolds, perturbing the structure of the effector protein, which leads to its inactivation. Using a similar principle, proteins can also be activated upon light illumination if

the light sensor is inserted within the autoinhibitory domain of the protein; perturbation of this domain leads to the activation of the protein.

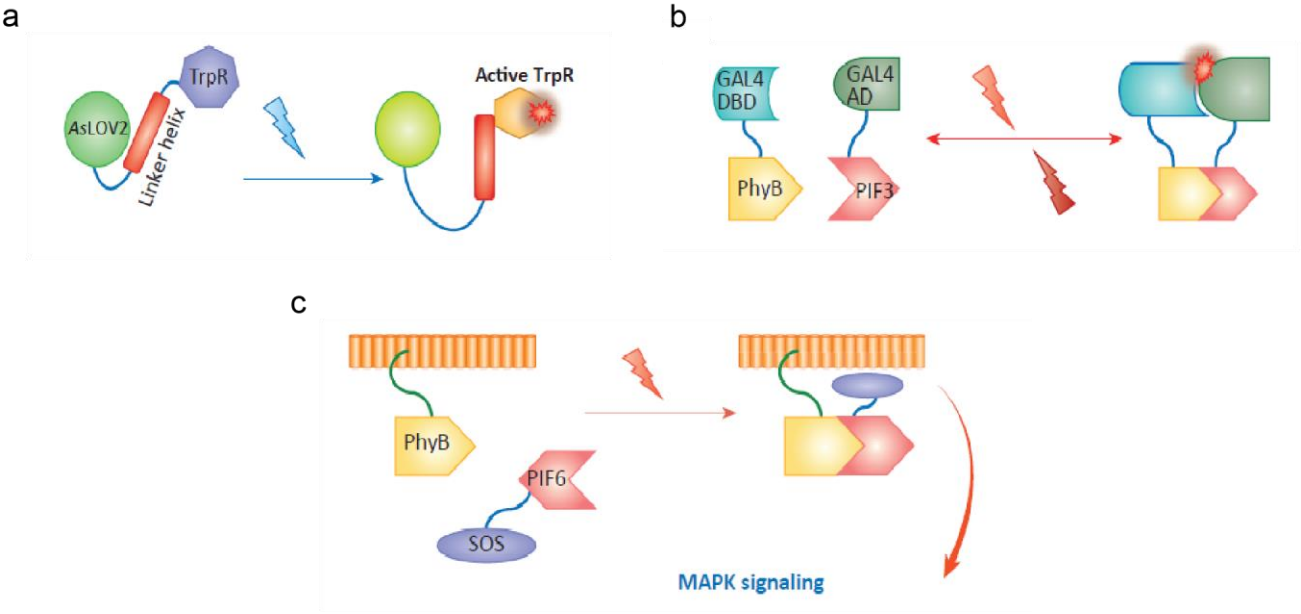



**Figure 14: Examples of optogenetic designs. a.** LOV-TAP. TrpR is connected to the C terminus of the LOV2 domain via a linker helix. In the dark state, the TrpR is distorted. In blue light, the linker dissociates from the LOV core, restores TrpR structure, and increases its affinity for DNA binding. **b.** Light-induced promoter system. PhyB and its interacting partner PIF3 are connected to the DBD and AD of GAL4, respectively. In red light, PhyB and PIF3 interact with each other, thus bringing the GAL4 domains together and rendering them functional. **c.** Opto-SOS. PhyB is anchored to the plasma membrane via a membrane localization peptide. Its interacting partner, PIF6, is fused to the SOS protein of the MAPK signaling pathway. In red light, PIF6 interacts with PhyB, bringing SOS close to the plasma membrane, where it can interact with its downstream signaling molecules and mediate the signaling pathway. LOV = Light-Oxygen-Voltage; TrpR = Trp repressor; Phy = Phytochrome; PIF = Phy Interacting Factor; SOS = Son of Sevenless; MAPK = Mitogen-Activated Protein Kinase. Figure has been adapted and modified from (Banerjee & Mitra, 2020).

Light-induced dimerization of naturally occurring photoreceptors with respective binding partners have also been utilised in the construction of optogenetic tools. The interaction between CRY and its binding partner, CRY-interacting basic Helix-Loop-Helix (CIB), has been applied in the development of artificial split proteins that aid in the dimerization of two separate proteins or domains, in response to blue light (Kennedy et al., 2010). This design principle has been exploited to study the propagation of protein activity pulses across cells (Aoki et al., 2013), live-imaging of *Drosophila* cells (Boulina et al., 2013), the control of phosphoinositide metabolism (Idevall-Hagren et al., 2012), and the development of light-inducible transcriptional effectors (Konermann et al., 2013). Cry–CIB dimerization has also been used to study protein–protein interactions (Taslimi et al., 2014). Recent works have further utilized

the CRY2-CIB1 dimerization pair to control the site-specific epigenome editing with de novo DNA methylation writer DNMT3A and methylation eraser TET1. The target specificity was achieved by a CIB1-fused transcription activator-like effector (TALE) targeting the Ascl1 gene promoter, which recruits the CRY2-fused DNMT3A or TET1 upon blue light irradiation and alters the local methylation pattern as well as gene expression (Choudhury et al., n.d.; Lo et al., 2017). Similarly, interaction of PHY with its binding partner PHY-interacting factor (PIF) has been utilised by Shimizu-Sato and colleagues, who fused the GAL4 DNA-binding domain to PhyB and the GAL4 transactivation domain to PIF3 (**Figure 14b**) to produce a red light-activatable promoter system (Shimizu-Sato et al., 2002). In another instance, the Opto-SOS system (Toettcher et al., 2013) utilises PhyB which is anchored to the plasma membrane through a membrane localization peptide. Its interacting partner, PIF6, is fused to the SOS protein of the mitogen-activated protein kinase (MAPK) signaling pathway. Upon light irradiation, PIF6 interacts with PhyB, bringing SOS close to the plasma membrane, where it can interact with its downstream signaling molecules and mediate the signaling pathway (**Figure 14b**).

### 3.1.2 Disadvantages and Possible Solutions

Although optogenetics has proven to be one of the most advantageous toolboxes of protein engineering, there are still certain limitations of such optogenetic tools. A significant example is their intrinsic reversibility. Photosensors return to their dark state conformation either by relaxation or light-triggered scission (Pudasaini et al., 2015). Reversibility is one of the virtues of optogenetic tools but also is a double-edged sword since continuous illumination might be necessary in certain cases to lock the “light” state conformation to obtain the desired effector function (Nihongaki et al., 2015). Although efforts have been made on structural optimizations to attenuate the decay, for instance, a ‘constitutively lit’ and a ‘constitutively dark’ state variant of AsLOV2 were generated by substituting a conserved glutamine (Gln513 in AsLOV2) residue with asparagine, or by altering the chromophore-binding residue cysteine to alanine/serine, respectively (Yee et al., 2015). But more careful evaluation and selection of suitable photosensor modules with lifetime and decay kinetics matching the timescale of targeted biological processes is desirable. Furthermore, the tunability and versatility of CRY-CIB system could be improved by generating CRY variants with improved half-lives (Taslimi et al., 2016). Background leakage (residual dark-state function) and off-target effects are also noteworthy drawbacks of optogenetic tools, especially the off-target enzymatic activity when the binding of an enzyme effector is regulated by photosensor dimerization. Some other limitations also include photostability, phototoxicity, and photochemical properties *in vivo*.

Optogenetic applications on humans are limited by technical difficulties of gene introduction and the necessity of invasive methods for light waves to reach the target location. Therefore, optimization is required for every application in order to achieve a satisfying result.

## 3.2 Optochemical Genetics Tools

Optochemical genetics exploit chemically synthesized small molecules as the core chromophore in accordance with its name, and it can be further categorized into two approaches: the photopharmacology approach (**Figure 15a**) and the protein modification approach using light-controlled amino acids (**Figure 15b**).

The photopharmacology approach involves protein ligands decorated with light-sensitive functional groups as receivers to influence protein function. Not only the pharmacologically active chemical compounds but also the chemical inducers of dimerization (CID), such as the commonly used rapamycin or the TMP/DHFR conjugation system, can be the protein ligands in this approach (Ankenbruck et al., 2018). However, the CID system would require additional genetical engineering to equip target proteins with responsible ligand-binding domains. It is also noteworthy that the photo-pharmacological ligands can either be used alone or tethered onto the target protein to proximately control the protein function. The most well-known example of the tethered ligand approach is the successful application of PTL (photoswitched tethered ligand) in neurons reported by Trauner and Kramer in 2004 (Banghart et al., 2004; Fehrentz et al., 2011) (**Figure 15c**).

On the other hand, the protein modification approach using light-controlled amino acids, enables chemical modification of photo-responsive functional groups on amino acid residues, and the user-defined site-specific incorporation of the non-canonical amino acid (ncAA) into the target proteins to directly control protein function (**Figure 15b**) (Ankenbruck et al., 2018; Fehrentz et al., 2011).

Considering that the application of photopharmacology is more restricted to the availability of protein ligands, this section will only focus on the protein modification approach.

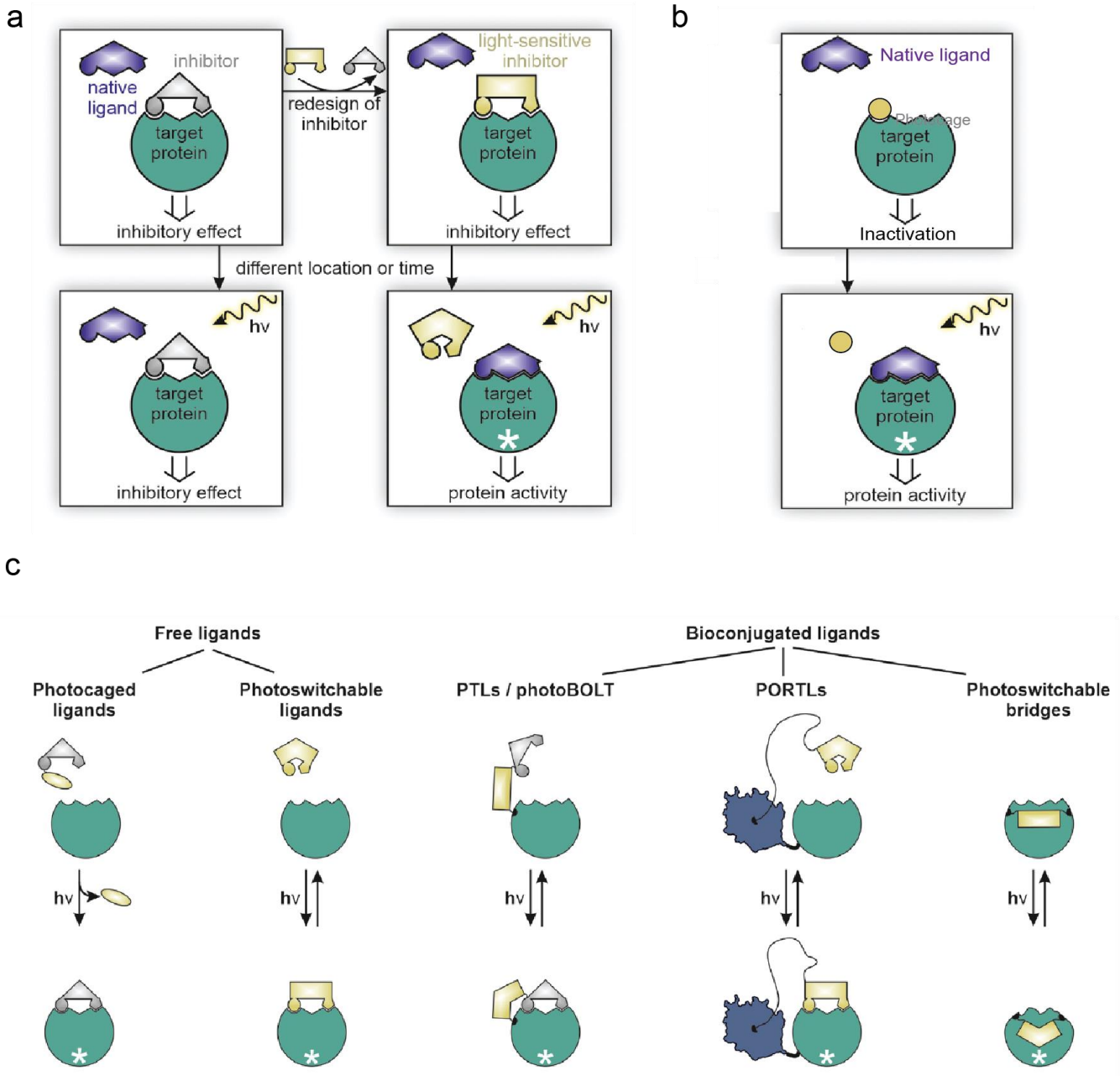



**Figure 15: Working principle of optochemical genetics tools. a.** Photopharmacology. **b.** Protein modification approach using light-controlled amino acids. **c.** Main strategies of photopharmacology. PTLs = photoswitchable tethered ligands; photoBOLT = photoswitchable biorthogonal ligand tethering; PORTLs = photoswitchable orthogonal remotely tethered ligands. White asterisk: protein activation/inhibition. Figure adapted and modified from (Kneuttinger, 2022).

### 3.2.1 Protein Modification with Light-controlled Amino Acids

Instead of using light-responsive proteins, chemists take advantage of light-responsive chemical functional groups to expand the optical toolbox for living systems. One of the most important applications of such chemical tools is the codon-specific protein modification to render a target protein light-responsive, which enables precise control of the targeted protein function or activity at single residue resolution by site-selective side chain modification. The available light responsive side chain moieties can be grouped by their applications in either photocage or photoswitch approach.

Approaches to impart photocontrol to biological targets share two pivotal strategies: photoswitch and photocage. Photoswitch employs reversible conformational change of chemical functional groups or photoreceptor modules to enable the switch between on/off state, as seen for optogenetic tools. Photocage, on the other hand, generally inhibits the function or interaction of target proteins via steric interference in the presence of conformational "cage" which can be "uncaged" upon light irradiation. Photoswitch typically allows multiple rounds of on/off control, whereas photocage, particularly the chemical photocage groups, is often limited to irreversible on/off control (Lang & Chin, 2014).

#### 3.2.1.1 Photocaged ncAAs

Photocage groups aim to block the function or activity of target proteins either via inhibiting the enzymatic activity conferred by residues like serine and cysteine, or via interrupting local noncovalent interactions that stabilize the protein-substrate, protein-cofactor, or protein-protein interactions. This type of chromophore utilizes photocleavable protecting groups (PPGs) to protect active chemicals/residues, where the active chemical/residue can be released from the PPG as a leaving group upon light irradiation.

A handful of PPGs have been explored to enable precise control of the release of chemicals in biological systems, including derivatives of coumarin, borondipyrromethene (BODIPY), cyanine, and *ortho*-nitrobenzene (Bardhan & Deiters, 2019; Klan et al., 2013). Among the PPGs, *o*-nitrobenzyl is the most common moiety in caging biomolecules. It is noteworthy that rather than the generic *o*-nitrobenzyl group which has disadvantages of slow cleavage kinetics and short absorption wavelength (< 350 nm), their dimethoxy or methylenedioxy derivatives (**Figure 16**) are extremely useful due to their high quantum yield, rapid cleavage, red-shifted absorption spectrum (> 350 nm and strong absorption around 400 nm), and good solubility in

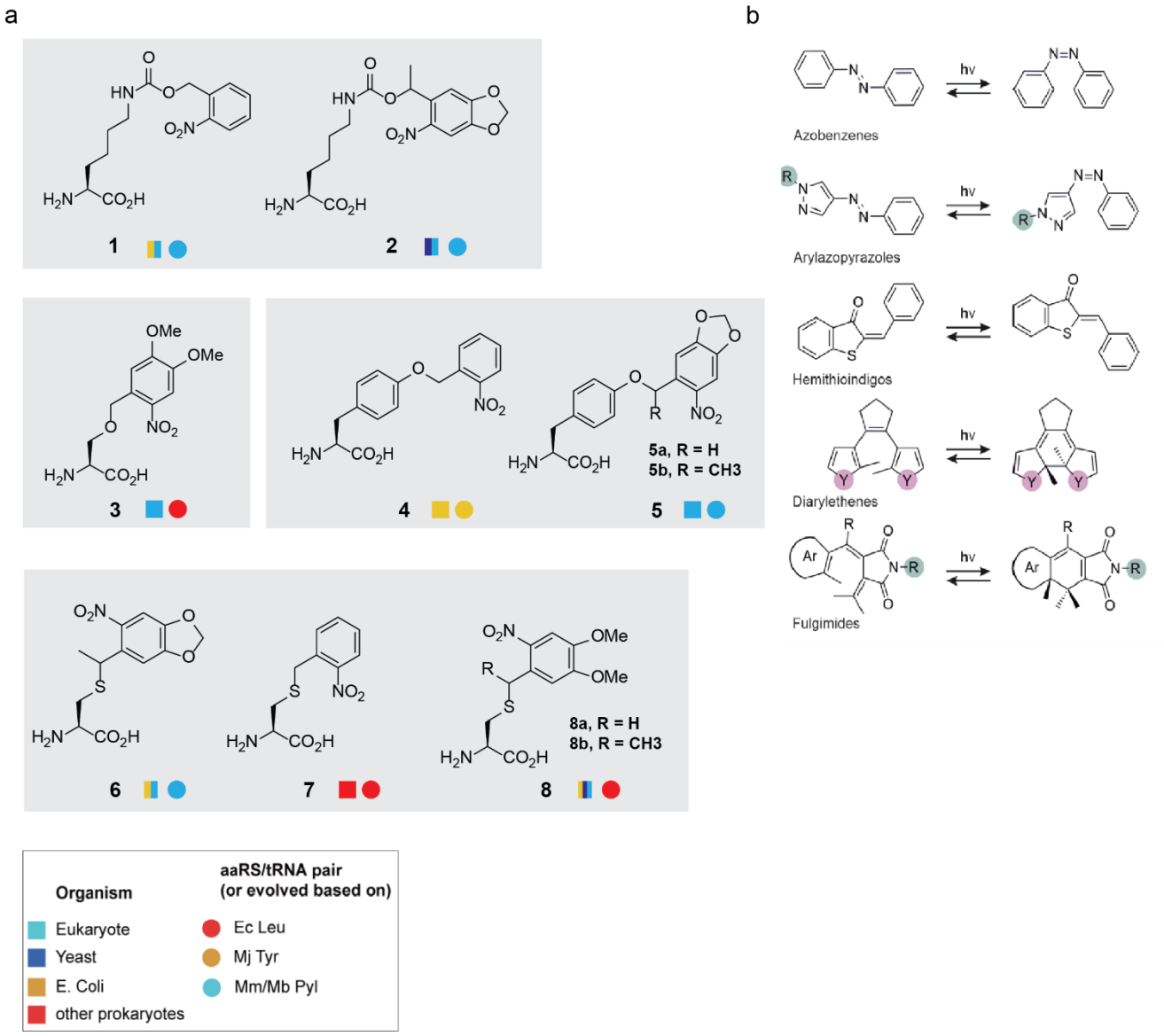



**Figure 16: Examples of photocaged and photoswitch ncAAs**. Chemical structures of **a**. photocaged that have been applied in chemical biological studies. Colour and shape codes indicate the applicable organism and required amber suppressor pair. Figure created based on (Dumas et al., 2015) (A.4). **b**. photoswitch ncAAs. R = H and other substituents; Y = S, O, N. Figure adapted from (Kneuttinger, 2022).

biological media (mostly aqueous solution) (Klan et al., 2013; Patchornik et al., 1970; Walker et al., 1993). *o*-Nitrobenzyl derived PPGs have been used to modify lysine (**1-2**, **Figure 16a**) (Chen et al., 2009; Gautier et al., 2010), serine (**3**, **Figure 16a**) (Lemke et al., 2007), tyrosine (**4-5**, **Figure 16a**) (Deiters et al., 2006), cysteine (**6-8**, **Figure 16a**) (Wu et al., 2004), selenocysteine (Rakauskaite et al., 2015), and threonine (only been applied in solid phase synthesis) (Mainz et al., 2016) as photocaged ncAAs. The canonical amino acids are attached via ester/thioester or carbamate bond to the benzylic position as leaving groups to be released

upon photoreaction of *o*-nitrobenzyl group (Il'ichev & Wirz, 2000; Klan et al., 2013). However, potential drawbacks of these photocaged ncAAs include the concomitant release of toxic nitrobenzaldehyde by-products and the potential condensation reaction between amine residues and the nitrobenzaldehyde by-product. Apart from the *o*-nitrobenzyl derived ncAAs, Luo et al. reported a set of coumarin-based photocaged lysine which can be activated via single-photon or two-photon pathway and provide advantageous three-dimensional resolution (Luo et al., 2014).

#### 3.2.1.2 Photoswitch ncAAs

Since photocaged ncAAs are limited to irreversible activation, photoswitch ncAAs (**Figure 16b**) were developed in order to mimic the reversible biological processes in nature. In this regard, the azobenzene moiety is the most suitable chromophore. Azobenzene is known for its unique ability to isomerize between *trans* and *cis* conformations upon different wavelengths of light. It exists in the thermodynamically favoured *trans* conformation in dark, isomerized to *cis* conformation upon UV-A or deeply violet light (315–380 nm), and can be reverted to *trans* conformation via thermo-relaxation or visible light-induced thermo-isomerization (Zimmerman et al., 1958). Azobenzenes hold many attractive characteristics, including high extinction coefficients and quantum yields, fast photoswitching rate, and relatively high photostability. These characters allow azobenzenes to isomerize with lower light intensity and be switched over multiple cycles (Fehrentz et al., 2011). Phenylalanines equipped with various azobenzene derivatives bearing distinct aromatic substituents constitute the currently available photoswitch ncAAs. The different geometry and dipole of two azobenzene isomers are useful for reversible modulation of the substrate or ligand binding site to impart activity control of target proteins. For instance, the *trans* conformation is long and extended which often does not disrupt protein conformation or the local interactions in the binding pocket, whereas the *cis* conformation is compact and bulky and may bring large steric interference to destabilize substrate binding (Bose et al., 2006). Nonetheless, applications of photoswitch ncAAs are not as well-established as the photocaged ncAAs. It is practically difficult for azobenzene to completely switch to *cis* conformation. In addition, it has the potential of being reduced by thiol reagents in cells and organisms, such as glutathione (Boulegue et al., 2007). More applications and improvements of the method remain to be explored.

### 3.2.2 Principles of Codon-specific ncAA Incorporation via Genetic Code Expansion (GCE)

With the light-responsive ncAA toolbox in hand, a target protein can be engineered at specific positions to render the desired light response. Site-selective protein modification methods, such as solid-phase synthesis, native chemical ligation, and other biosynthetic approaches although offers maximum flexibility in designing caged peptides and proteins, frequently face challenges like size limitations of synthetic peptides, installation of nonspecific caging groups onto expressed proteins, and the necessity to deliver the caged peptides and proteins into cells and organisms for biological studies (C. C. Liu & Schultz, 2010). In comparison, codon-specific ncAA incorporation via genetic code expansion is unparalleled in overcoming these challenges of modifying proteins in live cells (Baker & Deiters, 2014; Lawrence, 2005). Genetic code expansion (GCE) allows simultaneous ncAA incorporation into the polypeptide chain during endogenous protein synthesis in response to specifically inserted unassigned codon(s), the most widely used codon being the amber stop codon (UAG), at a desired site in the gene of interest. In the ribosomal polypeptide synthesis, a canonical amino acid is selectively recognized and aminoacylated ("charged") on the cognate tRNA by the aminoacyl-tRNA synthetase (or aminoacyltRNA transferase, aaRS) and subsequently transferred onto the growing polypeptide chain if the anti-codon of the cognate tRNA matches the mRNA codon presented in the ribosomal A-site (**Figure 17a**). Genetic code expansion, in turn, involves engineering and rewiring protein translation in order to facilitate site-specific ncAA incorporation. This generally requires (1) an orthogonal aminoacyl-tRNA synthetase which binds to the ncAA, (2) an unassigned codon which can be recognized by the orthogonal tRNA, and (3) an orthogonal cognate tRNA which can pair with the orthogonal tRNA synthetase (**Figure 17b**) (Chin, 2017; Liu & Schultz, 2010 ).

To establish the orthogonality, meaning, a system that would not interfere with the endogenous functions of the native machinery, the chemical structure of the ncAA must be distinguishable from the canonical amino acids and not recognized by the endogenous aminoacyl-tRNA synthetases, while the orthogonal aminoacyl-tRNA synthetase must be capable of binding the ncAA but not canonical amino acids (**Figure 17c**). The ncAA must be metabolically stable, permeable through the cell membrane, and be able to be recognised and accepted by the translation elongation factor Tu (EF-Tu) and the ribosome (Kneuttinger, 2022; C. C. Liu & Schultz, 2010). One approach to ensure ncAA recognition by the aaRS that is applicable to living cells involves the use of wild-type aaRSs to incorporate ncAAs that are close structural analogs of canonical amino acids. In this approach, a strain auxotrophic for one of the common 20 amino acids is used to substitute that amino acid with an ncAA analog. Although the resulting wholesale replacement of a common amino acid by an ncAA cannot sustain

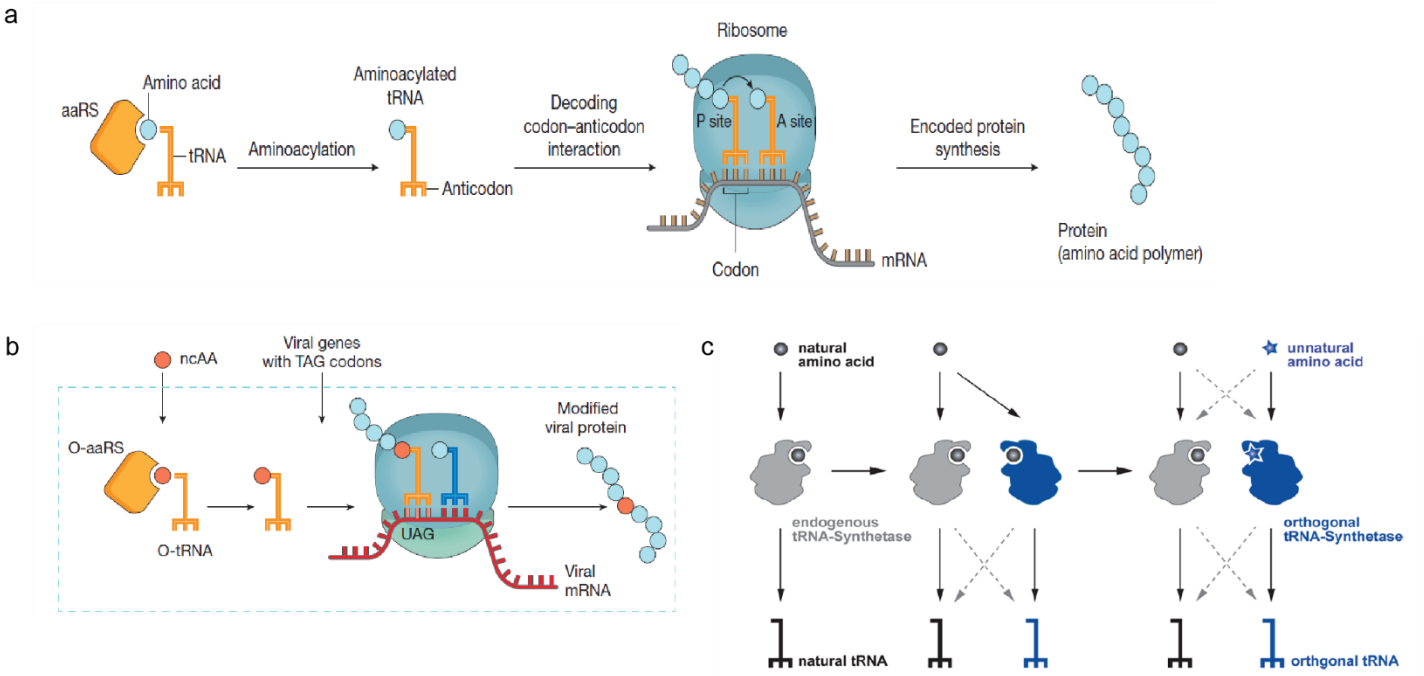



**Figure 17: Overview of the cellular translation machinery and genetic code expansion. a.** An aaRS aminoacylates its cognate tRNA with a specific amino acid. During translation elongation, the aminoacylated tRNA is decoded on the ribosome in response to a cognate codon in the mRNA, leading to the addition of an amino acid to the growing polymer. **b.** ncAA incorporation into proteins. Genes of proteins of interest containing amber (TAG) codons are introduced into cells containing the orthogonal aminoacyl-tRNA synthetase (O-aaRS) and orthogonal tRNA. The orthogonal aaRS transfers the ncAA onto the tRNA, and the modified protein containing the ncAA is translated. Figures are adapted with permission from (Chin, 2017) (A.4). **c.** Orthogonal aaRS/tRNA pairs are evolved by a two-step selection process. A heterologous aaRS/tRNA pair from a different domain of life is first imported into the host of interest, this is followed by selection of a mutated active site in the orthogonal synthetase that specifically recognizes a desired ncAA (blue star). Dashed gray lines connect nonfunctional combinations. Figure is adapted from (Lang & Chin, 2014). (o) aaRS = (orthogonal) aminoacyl-tRNA synthetase; ncAA =non-canonical amino acid.

exponential growth, nondividing cells are still viable and are able to overexpress proteins that contain the ncAA. In these newly synthesized proteins, the canonical amino acid is efficiently replaced by its ncAA analog at all sites. However, most of the orthogonal aminoacyl-tRNA synthetases originally recognize canonical amino acids and require active site evolution to build up the orthogonality, except for the aminoacyl-tRNA synthetase for pyrrolysine (PylRS) (Ibba & Söll, 2004; Polycarpo et al., 2004) which does not recognize any canonical amino acids but can promiscuously bind to a variety of ncAA analogues. Moreover, PylRS–tRNA pairs from certain methanogens (notably *Methanosarcina barkeri* and *Methanosarcina mazei*) are orthogonal in both bacteria and eukaryotic cells. The specificity of ncAA recognition can be further enhanced by engineering the editing domain of the aminoacyl-tRNA synthetase. An editing domain is present in certain aminoacyl-tRNA synthetases to facilitate the correct tRNA aminoacylation by hydrolysing mischarged aminoacylated tRNAs, whereby mutating critical

residues in the editing domain can attenuate the hydrolysis of ncAAs. One example of such strategy is when an orthogonal PylRS/tRNA pair was evolved for the incorporation of a photocaged lysine in response to an amber stop codon. First, a library of 108 mutants of a PylRS from *Methanosarcina barkeri* (*Mb*PylRS) was created, in which 5 positions (M241, A267, Y271, L274, and C313) in the binding pocket of the pyrrolysine ring were randomized to all possible amino acids. After many rounds of selection, the most active synthetase that accommodated the photocaged lysine, contained the mutations M241F, A267S, Y271C, and L274M with respect to wild-type *Mb*PylRS. This synthetase was named photocaged Lysyl-tRNA synthetase (PCKRS) (Gautier et al., 2010).

In nature, all the 64 combinations of mRNA triplet codons are used to encode 20 canonical amino acids and 3 stop signals for translation termination. It is therefore challenging to expand the usage of genetic codons as recoding sense codons (encoding canonical amino acids) will compete with normal protein synthesis, and recoding nonsense codons (stop signals) might disturb the termination machinery. Nonetheless, discovery of deviated codon usage in vertebrate mitochondria and codon reassignments in prokaryote and eukaryote (reviewed in Ambrogelly et al., 2007; Knight et al., 2001) challenges the concept of nonevolving genetic code from the "frozen accident" hypothesis (F. H. C. Crick, 1968). Stop codon reassignments found in all domains of life shed light on the expanded use of genetic codons. The opal codon UGA is ambiguously used to encode selenocysteine (Sec, **Figure 18a**) for synthesizing selenoproteins in all domains of life despite the limited organisms (Johansson et al., 2005; Leinfelder et al., 1988), while the amber codon UAG is reassigned in methanogenic archaea and bacteria to encode pyrrolysine (Pyl, **Figure 18b**) (Hao et al., 2002; Srinivasan et al., 2002). In the light of the natural expansion of genetic codes, the highest degree of orthogonality can therefore be provided by the least frequent amber codon, which is also the most used codon to date to incorporate ncAAs. This is known as the "amber suppression technique". On the other hand, strategies to enhance the orthogonality has been demonstrated using evolved orthogonal ribosomes. The orthogonal ribosomes are used for the translation of orthogonal mRNAs in parallel to natural ribosomal translation (Rackham & Chin, 2005; Wang et al., 2007). Since it is not involved in the synthesis of host proteome, the orthogonal ribosomes are flexible to be re-engineered for decoding new blank or unassigned codons such as the quadruplet codons which are poorly processed by natural ribosomes (Neumann et al., 2010).

The orthogonal aminoacyl-tRNA synthetase–tRNA pair must not cross-react with the host aminoacyl-tRNA synthetase–tRNA pair. Meaning, the orthogonal aminoacyl-tRNA synthetase does not charge host tRNAs, and the orthogonal tRNA is not recognized by host aminoacyl-tRNA synthetases (**Figure 17c**). Since the orthogonality of the aminoacyl-tRNA synthetase–tRNA pair reflects the interactions of tRNA or aminoacyl-tRNA synthetase with the host translational components, it is defined with respect to a specific host organism. This

orthogonality is typically provided by importing an aminoacyl-tRNA synthetase–tRNA pair from one domain of life into a heterologous host. A *Methanocaldococcus janaschii* (formerly *Methanococcus janaschii*)-derived TyrRS–$tRNA_{CUA}$ pair is orthogonal in *Escherichia coli* and other bacteria, *E. coli* TyrRS–$tRNA_{CUA}$ and LeuRS–$tRNA_{CUA}$ pairs are orthogonal in eukaryotic cells, and PylRS–$tRNA_{CUA}$ pairs from certain methanogens (notably *Methanosarcina barkeri* and *Methanosarcina mazei*) are orthogonal in both bacteria and eukaryotic cells. The anticodon loop of the imported tRNA is then mutated to create a blank codon suppressor tRNA (tRNASB), and the orthogonality of the resulting aaRS/tRNASB pair is assessed (Chin, 2017; C. C. Liu & Schultz, 2010). Orthogonality of the aminoacyl-tRNA synthetase–tRNA pair can also be created *de novo* by directed evolution. It can be improved by a two-step process, involving both positive and negative rounds of selection to identify functional optimized tRNASBs that exhibit no cross-reactivity with endogenous aaRSs. Finally, structure-based mutagenesis and a similar two-step selection strategy are used to alter the specificity of the heterologous aaRS so that it uniquely recognizes the UAA of interest.

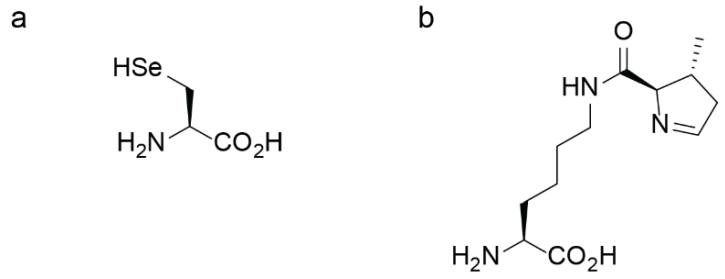



**Figure 18: Chemical structures of naturally occurring non-canonical amino acids (ncAAs). a.** selenocysteine (Sec) and, **b.** pyrrolysine (Pyl).

Some of these orthogonal aminoacyl-tRNA synthetase–tRNA pairs have been evolved to incorporate light-responsive ncAAs in mammalian cells. The evolved PylRS/PyltRNA pairs have been used for encoding photocaged lysines (Chen et al., 2009; Gautier et al., 2010), photocaged cysteine (Nguyen et al., 2014), photocaged tyrosine (Arbely et al., 2012), and photoswitch phenylalanines (Hoppmann et al., 2014). While the evolved LeuRS/LeutRNA pairs have been used for incorporating photocaged serine (Lemke et al., 2007) and cysteine (Kang et al., 2013).

### 3.2.3 Application Strategies

Photocontrol of protein function at single residue, especially via photocaging, has vast applications in cell biology due to its unsurpassed precision, even over optogenetic tools. It has been used to control the activity of caspase 3 (Wu et al., 2004) and β-galactosidase (Deiters et al., 2006), signal transduction (Arbely et al., 2012; Gautier et al., 2011; Lemke et al., 2007; Liaunardy-Jopeace et al., 2017), protein localization (Gautier et al., 2010), gene expression (Chou & Deiters, 2011; W. F. Edwards et al., 2009; Hemphill et al., 2015; Luo et al., 2016), and activating fluorescent or luminescent reporters (Groff et al., 2010; Wilkins et al., 2010; J. Zhao et al., 2013) (Ankenbruck et al., 2018; Baker & Deiters, 2014).

One of the most important recent applications of photocaged ncAAs has been in the control of the epigenetic landscape, mainly to regulate the writing, erasing and reading of DNA methylation marks. The functional studies of epigenetic regulatory elements at chromatin level are often suffered from the challenging druggability and the poor selectivity of small molecule inhibitors. Especially the dynamic and sophisticated nature of chromatin regulation requires tools with high spatial-temporal resolutions to enable kinetic insights uncoupled from background cellular events. In this regard, selective photoactivation of protein function at a user-defined residue provides unparalleled advantages. DNA methylation at the 5'-carbon of cytosine (C) within CpG dinucleotides is the most abundant DNA modification. As expected, there are dedicated enzymes that serve as writers, erasers and readers of this epigenetic mark. In three seminal papers, Summerer and group reported the activation of these three classes of enzymes with light, by virtue of photocaged ncAAs.

A photocaged 4, 5-dimethoxy-2-nitrobenzyl-L-cysteine (**8a**, DMNB-Cys, **Figure 16a**) was genetically encoded into the catalytic site of *de novo* DNA methyltransferase (DNMT) in place of the catalytic Cys710. Light activation of photocaged DNMT3 illustrated differential *de novo* methylation kinetics in mammalian cells influenced by frequent DNMT mutations in cancer (Wolffgramm et al., 2021). Subsequently, Lin *et al.* reported the light activation of MBD reader association to chromatin by genetically encoding a DMNB caged serine (**3**, DMNB-Ser, **Figure 16a**) at the MBD-DNA interface. They incorporated DMNB-Ser at S45 of human MBD1 enabling the expression of MBD protein that is inactive in DNA binding, which could be rapidly activated with a light pulse. The authors also studied the association of MBD1 to mouse pericentromeres, dependent on its CXXC3 and transcriptional repressor domains (TRD) which interact with unmethylated CpG and heterochromatin-associated complexes, respectively (Lin et al., 2024). Light activation of TET dioxygenases, the epigenetic DNA methylation erasers, was also achieved by genetically encoding a DMNB-Ser to replace an active site serine residue S1812 (murine TET2 numbering) responsible for α-KG cofactor binding. The DMNB

protecting group serves as a temporary block in the active site to keep TET dioxygenases inactive in cells, and subsequent DMNB removal with light restores the canonical serine residue as well as enzymatic activity. This allows protein inactivation with minimum structural changes. The light-activatable TET2 dioxygenase has been demonstrated to activate target gene expression and induce differential transcriptome modulation in mammalian cells in a light-dependent manner (Palei et al., 2020). This report has been a basis of the current study as it showed the path towards regulating TET activity with light, as will be discussed more in details in later chapters.

Protein post-translational modifications like phosphorylation have also been regulated with light using ncAAs by modifying the responsible kinases with a photocaged residue. Lemke et al. developed DMNB-Ser and used this ncAA to regulate phosphorylation of Pho4, a transcription factor that is central to the signaling cascade responsible for yeast to grow under different concentrations of inorganic phosphate (Pi). Under Pi-starved conditions, Pho4 is hypophosphorylated and localized in the nucleus, where it activates transcription of phosphate-responsive genes. In the presence of high concentrations of extracellular Pi, Pho4 is phosphorylated by the cyclin-cyclin dependent kinase (cyclin-CDK) complex (Pho80-Pho85), which inactivates Pho4 and triggers its translocation to the cytoplasm. Light triggered phosphorylation of individual serine residues in Pho4 led to the observation of distinct export kinetics for differentially phosphorylated Pho4 mutants, which demonstrated dynamic regulation of Pho4 function (Lemke et al., 2007). In another important study, by incorporating the genetically encoded methyl-o-nitropiperonyl caged lysine, or simply **photocaged lysine** (**2**, pcK, **Figure 16a**), a light-activated kinase MEK1 was developed in which a universally conserved lysine residue within the ATP binding pocket is photocaged (Gautier et al., 2010). MEK1 is a critical component of the Raf/MEK/ERK signaling pathway and is involved in many essential cell functions including cell growth, adhesion, survival, and differentiation. With this approach, the authors could demonstrate the specific, rapid, and receptor independent activation of an artificial subnetwork within the Raf/MEK/ERK pathway (Gautier et al., 2011).

### 3.2.4 Disadvantages and Possible Solutions

Despite the unsurpassed precision to control protein function with light in living systems, codon-specific incorporation of light-responsive ncAAs via genetic code expansion has a few limitations. Firstly, the ability to incorporate multiple different light-responsive residues is limited by the mutually orthogonal aminoacyl-tRNA synthetase–tRNA pairs and the availability of blank codons. Although new blank codons are constantly under study, like development of quadruplet codons and orthogonal ribosomes, the risk of ambiguity in the

correct assignment of orthogonal machinery for each ncAA remains (de la Torre & Chin, 2021). Photoswitchable ncAAs are also met with limitations of structure, being bulky even after isomerization which bring conformational constraints. Moreover, the selection of the site of ncAA incorporation is necessarily based on knowledge of the target system, e.g., residues involved in catalysis, ligand binding or allostery, for which the three-dimensional structure of the protein may be beneficial. Since the effect of the light-sensitive ncAAs is often difficult to anticipate, some computation-based approaches have been developed to facilitate the choice of incorporation sites (e.g., CAGE-prox) (Wang et al., 2019; Kneuttinger, 2022). The most common limitations in this approach are the truncated protein side-products and the potentially mistranslated endogenous proteome due to stop codon suppression (compromised fidelity of incorporation) that lead to reduced yields of modified full-length proteins and potential extended/misfolded endogenous proteins. Efforts have been made to address the low ratio of full length to truncated protein products by expressing these engineered proteins in engineered *E. coli* strains (Kuznetsov et al., 2017) lacking release factors. Despite these shortcomings, modifying proteins with ncAAs by genetic code expansion continue to be useful in diverse applications to implement dynamic control in biological systems.

# 4. Ten-Eleven Translocation Proteins (TET): A Target for Light Control

Ten-Eleven Translocation (TET) proteins are α-ketoglutarate (α-KG) and iron (II) (Fe(II)) dependent dioxygenase enzymes that actively reverse methylation at 5'-carbon of cytosine (5mC) nucleotide bases of DNA.

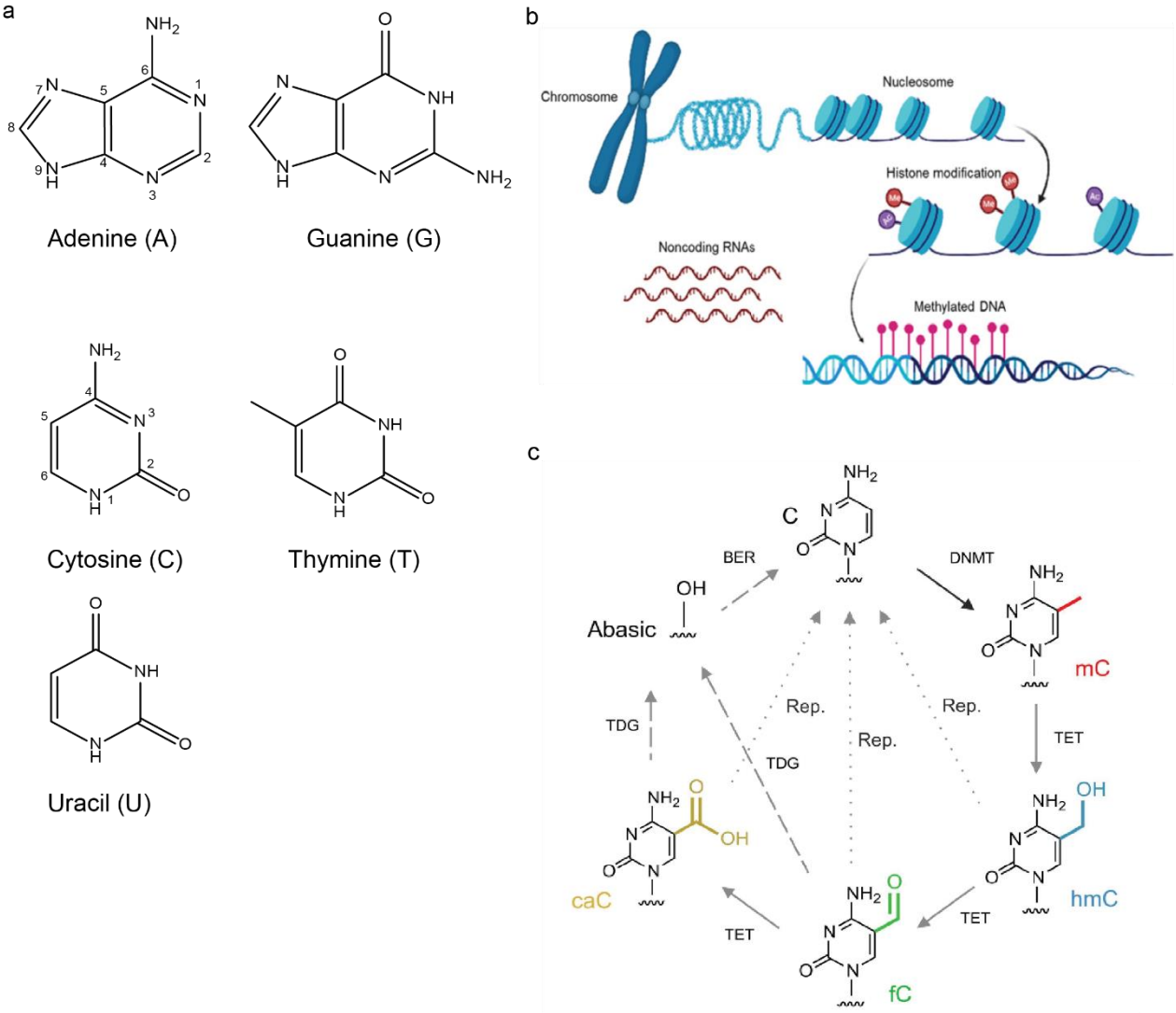



**Figure 19: Basics of epigenetics. a.** Chemical structures of DNA and RNA bases. Purine bases are Adenine (A) and Guanine (G) found in both DNA and RNA; Pyrimidines are Cytosine (C), found also in both DNA and RNA, Thymine (T) that is a DNA letter and Uracil (U) that is found in RNAs. **b.** Epigenetic modifications include histone modifications (methylation, acetylation, etc.), DNA modifications (methylation) and long non-coding RNAs (lncRNAs). Figure has been adapted from (Zhu et al., 2023). **c.** The complete cycle of cytosine methylation. Figure has been adapted with permission from (Muñoz-López & Summerer, 2018) (A.4).

Decoding the information embedded in the nucleotide sequence of DNA (genotype) to build up functional multi-cellular organisms is a process that is highly regulated at multiple points. The central dogma of life explains the decoding process as a sequential flow of genetic information from DNA to proteins with mRNA as the mediator. The four DNA letters, adenine (A), thymine (T), cytosine (C), and guanine (G) are transcribed residue-by-residue into four RNA monomers, A, uracil (U), C, and G (**Figure 19a),** where nucleic acid triplets further constitute the basic information units that are translated into proteins (F. Crick, 1970). However, the phenotypic diversity on the cellular level cannot be explained by the central dogma alone. For instance, the same DNA sequence of a human being results in over 200 different cell types bearing distinct functions and identities. Therefore, studies to discover the mechanisms underlying the genotype to phenotype transformation has become a very interesting subject. Scientists later discovered that this "epigenetic" (outside of genetics) decision-making process involves heritable changes in gene expression and function without altering the DNA sequence. Numerous such epigenetic factors have been found that together contribute towards altered gene expression patterns. These include the post-synthetic modifications on DNA (DNA methylation) and histones (acetylation, methylation, phosphorylation, etc.), as well as non-coding RNAs and long-range chromatin interactions (**Figure 19b**). In particular, DNA methylation on the fifth carbon of cytosine base, 5-methylcytosine (5mC), is a hallmark in epigenetic regulation and the most abundant DNA modification. 5mC plays important roles in transcriptional regulation, differentiation, and development. It is found in a variety of organisms including plants, vertebrates, and fungi. In mammals, about 2-7% of the cytosines are methylated. And nearly all 5mC in mammalian somatic cells occurs in the context of CpG dinucleotides and accounts for 60-80% of the total CpGs (Razin & Riggs, 1980), whereas as much as a quarter of methylation happens in a non-CpG context in embryonic stem cells (B. Jin et al., 2011). Mammalian genomes are globally depleted of CpGs but contain short, interspersed sequences bearing high relative density of non-methylated CpGs termed CpG islands (CGIs) (Deaton & Bird, 2011; Smith & Meissner, 2013), that are protected from methylation and associated with transcription initiation. They either span the promoter regions of housekeeping genes and developmental regulator genes or sit more remote from the annotated transcription start sites (TSSs) but exhibiting regulatory (enhancer) functions (reviewed in Deaton & Bird, 2011). Hypermethylation of CGIs often leads to transcriptional silencing of associated genes and is an important process involved in genomic imprinting and X-chromosome inactivation (Bird, 2002; Deaton & Bird, 2011; Kohli & Zhang, 2013). Apart from CGIs, methylation also silences the expression of endogenous repetitive elements including the pericentromeric satellites and transposable elements. These functions of 5mC suggest its diverse roles in various biological processes and the importance to uncover its dynamic regulation and consequences in the epigenome.

Dynamic regulation of the methylation landscape is enabled by the reversibility of 5mC (**Figure 19c**). The life cycle of 5mC starts with methylation writer-catalyzed active modification (AM), then followed by the replication-dependent passive dilution (PD) or methylation eraser-mediated active removal (AR) to restore the unmodified cytosine. Active modification is catalyzed by the methylation writers, DNA methyltransferases (DNMTs), by transferring a methyl group from the *S*-adenosyl-L-methionine (SAM) cofactor onto the fifth carbon of the cytosine pyrimidine ring (Du et al., 2016; Zangi et al., 2010). It is an important process involved in both the *de novo* creation of methylation and the maintenance of methylation patterns between cell generations to ensure faithful epigenetic information transfer. If methylation patterns cannot be efficiently maintained during DNA replication, 5mCs will be gradually replaced by unmodified cytosines which eventually leads to the PD of methylation patterns after progressive rounds of replication. On the other hand, 5mC can be actively removed from the genome by the methylation erasers, ten-eleven translocation dioxygenases (TETs). TET dioxygenases catalyze the iterative oxidation of 5mC to 5-hydroxymethylcytosine (5hmC) (Kriaucionis & Heintz, 2009; Tahiliani et al., 2009), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) (**Figure 19c**) (He et al., 2011; Ito et al., 2011; Pfaffeneder et al., 2011), whereby 5fC and 5caC are recognized and removed by the thymine DNA glycosylase (TDG) leaving abasic sites for the base excision repair (BER) machinery to restore unmodified cytosines (He et al., 2011; Maiti & Drohat, 2011; Weber et al., 2016).

The TET dioxygenase family has three members: TET1, TET2, and TET3. TET enzymes catalyze the oxidation of 5mC in an iron(II)/α-ketoglutarate (Fe(II)/α-KG)-dependent manner with their conserved core catalytic domain at the C-terminus, which is comprised of a cysteine-rich (Cys-rich) domain and a double-stranded β-helix (DSBH) domain (**Figure 20a**). In complex with DNA, the catalytic domain exhibits a compact globular conformation with a central DSBH core buttressed by flanking segments of both the DSBH domain and the Cys-rich domains. The double-stranded DNA sits above the DSBH core with a methylated cytosine flipping out from the base pairing and inserting into the catalytic pocket (**Figure 20b**). Zooming in the catalytic cavity of TET2, the methyl group is orientated by the residues in the DSBH core in coordination with Fe (II) and α-KG for oxidation (**Figure 20b**). The Fe(II) is chelated in an octahedral coordination by conserved residues H1382, D1384, H1881, oxygen atoms from the 1-carboxylate and 2-keto group of α-KG, and a water molecule which is replaced by $O_2$ in the catalysis (**Figure 20b**). In addition, residue R1261 stabilizes the 1-carboxylate of α-KG, whereas the 5-carboxylate at the other end of α-KG is stabilized by residues H1416, R1896, and S1898. Most importantly, all residues that contribute to the interactions with Fe(II) and α-KG are highly conserved across TET dioxygenases. Zooming out from the catalytic cavity again, the interaction between DNA and the DSBH core is further supported by two loop regions (L1 and L2) of the Cys-rich subdomains (Cys-N and Cys-C) that wrap

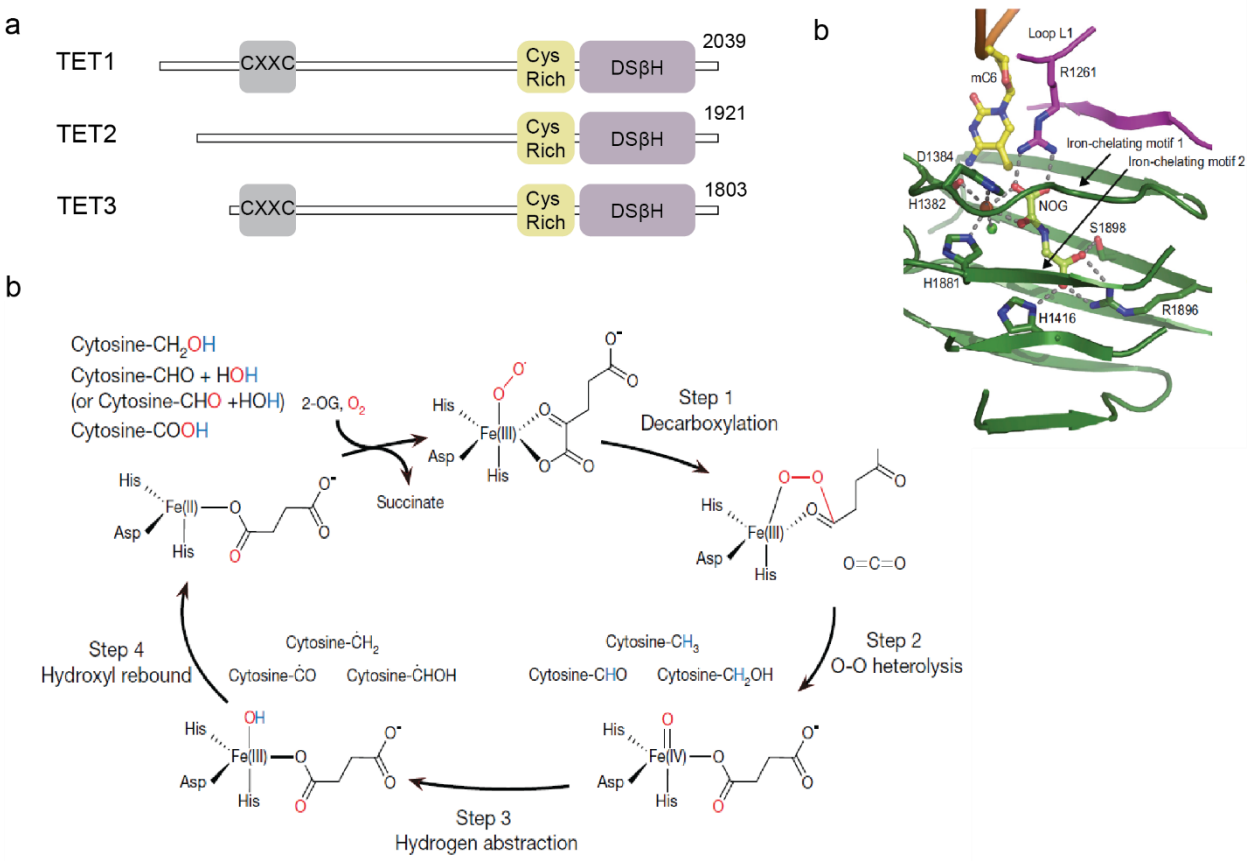



**Figure 20: Structure of Ten-Eleven Translocation (TET) proteins. a.** Domain structure of the three TET isotypes TET1, TET2 and TET3. All share a conserved C-terminal catalytic domain containing the Cys-rich and DSβH domains. **b.** Critical interactions at the DSβH core of human TET2 (PDB ID **4NM6**) with NOG substituting α-KG. Dashed lines: hydrogen bonds and Fe(II) coordination; green ball: crystallographic water. Figure adapted from (Hu et al., 2013) with permission (A.4). **c.** The proposed catalytic cycle of TET dioxygenase-mediated 5mC oxidation. Figure adapted with permission from (Hu et al., 2015) (A.4). DSβH = double stranded β-helix; NOG = N-oxalylglycine; α-KG = α-keto glutarate.

around the DSBH core and stabilize its conformation by zinc coordination (L. Hu et al., 2013, 2015a; Kohli & Zhang, 2013; X. Lu et al., 2015). TET dioxygenases exhibit a substrate preference for 5mC over 5hmC and 5fC. The 5mC to 5hmC conversion has been shown to be 3-5 times faster than that of 5hmC to 5fC and 5fC to 5caC although all of them are possible substrates in the sequential oxidation (L. Hu et al., 2015a; Ito et al., 2011). In fact, hints for this substrate preference can be found in the molecular mechanism of TET-mediated 5mC oxidation which has been simulated with other Fe (II)/α-KG-dependent dioxygenases. The catalytic cycle is initiated with the oxidation of Fe (II) to reduce a dioxygen molecule and form a Fe(IV)-peroxo intermediate. The Fe (IV)-peroxo complex is further converted into a Fe(IV)-oxo (Fe(IV)=O) intermediate following the decarboxylation of α-KG co-factor and the release of

succinate and $CO_2$. The Fe(IV)-oxo intermediate then abstracts a hydrogen from the methyl group of 5mC to activate the C-H bond for subsequent hydroxylation (**Figure 20c**) (L. Hu et al., 2015; X. Lu et al., 2015; Tarhonskaya et al., 2014). Structural analysis has elucidated the different availabilities of hydrogen atom in the C-H activation step between 5mC, 5hmC, and 5fC. In the case of 5fC, the hydrogen atom is relatively far from the Fe (IV) complex due to the intramolecular hydrogen bonding, and the planar conformation of 5fC also restricts the rotation of C-H bond to adopt a preferable orientation for reaction (L. Hu et al., 2015; Münzel et al., 2011). Similar with 5fC, calculations also indicate the tendency of 5hmC to form hydrogen bonds which could prevent the hydrogen abstraction (L. Hu et al., 2015).

## 4.1 Regulation of TET Function by Ubiquitination

TET, being such a highly instrumental player in the gene expression game, is regulated at multiple layers. These include structural regulation (Hu et al., 2016); regulation by co-factors (Hu et al., 2016; Yin and Xu, 2016); cellular turnover and degradation by other protein complexes (Wang and Zhang, 2014) and most extensively, by various post-translational modifications.

Apart from glycosylation (Shi et al., 2013; Zhang et al. 2014; Bauer et al., 2015), phosphorylation (Wu et al., 2018; Zhang et al., 2019; Rao et al., 2020) and acetylation (Zhang et al., 2017), all three forms of TET were recently found to be mono-ubiquitinated at a conserved Lys residue at the C-terminal domain (Nakagawa et al., 2015). Vpr binding protein (VprBP)/DCAF1 is essentially the strongest and the most abundant binding partner of the damaged DNA binding protein 1 (DDB1), an intermediate through which it forms an active CUL4 assembly. CUL4 is a member of the Cullin Ring Ligase family of RING type E3 ligases and is conserved from yeast to humans. In humans both the two paralogs of CUL4 (CUL4A/B) use DDB1 protein as a linker to bind to proteins containing the WD40 domain, that in turn, serve as substrate recognition modules. As expected, VprBP binds to DDB1 with its C-terminal WD40 domain. Together this complex of CUL4A/B with DDB1 and VprBP is annotated as $CRL4^{VprBP}$ (**Figure 21**). In the seminal study, Nakagawa et al. showed that all three TETs, TET 1, TET 2 and TET 3, bind to VprBP through their C-terminal Cys rich domain (**Figure 21**) (Nakagawa et al., 2015). Upon binding VprBP, TET was found to be monoubiquitinated at a conserved Lys residue K1299 in human TET2 and the corresponding residues in the other TETs, both in human and murine cells. Mutating these Lys residues to Asn or Glu or even Arg led to significant reduction in ubiquitination. Also, depletion of even one of the components of $CRL4^{VprBP}$ complex led to a disruption in the monoubiquitination. This monoubiquitination event was found to be directly affecting the catalytic function of TET, and absence of this PTM

leads to loss of TET activity. The authors postulated that this activation results from increased DNA binding upon monoubiquitination which they could observe both *in vitro* and *in vivo*. Indeed, many TET2 mutants have been discovered that are defective in monoubiquitination, being mutated directly at the target Lys or at residues that are involved in VprBP binding. Later, Chen et al. showed that the USP15 DUB reverses this Lys monoubiquitination of TET (Chen et al., 2020). They showed that USP15 deubiquitinates human TET2 and inactivates it by inhibiting its binding to substrate DNA (**Figure 21**).

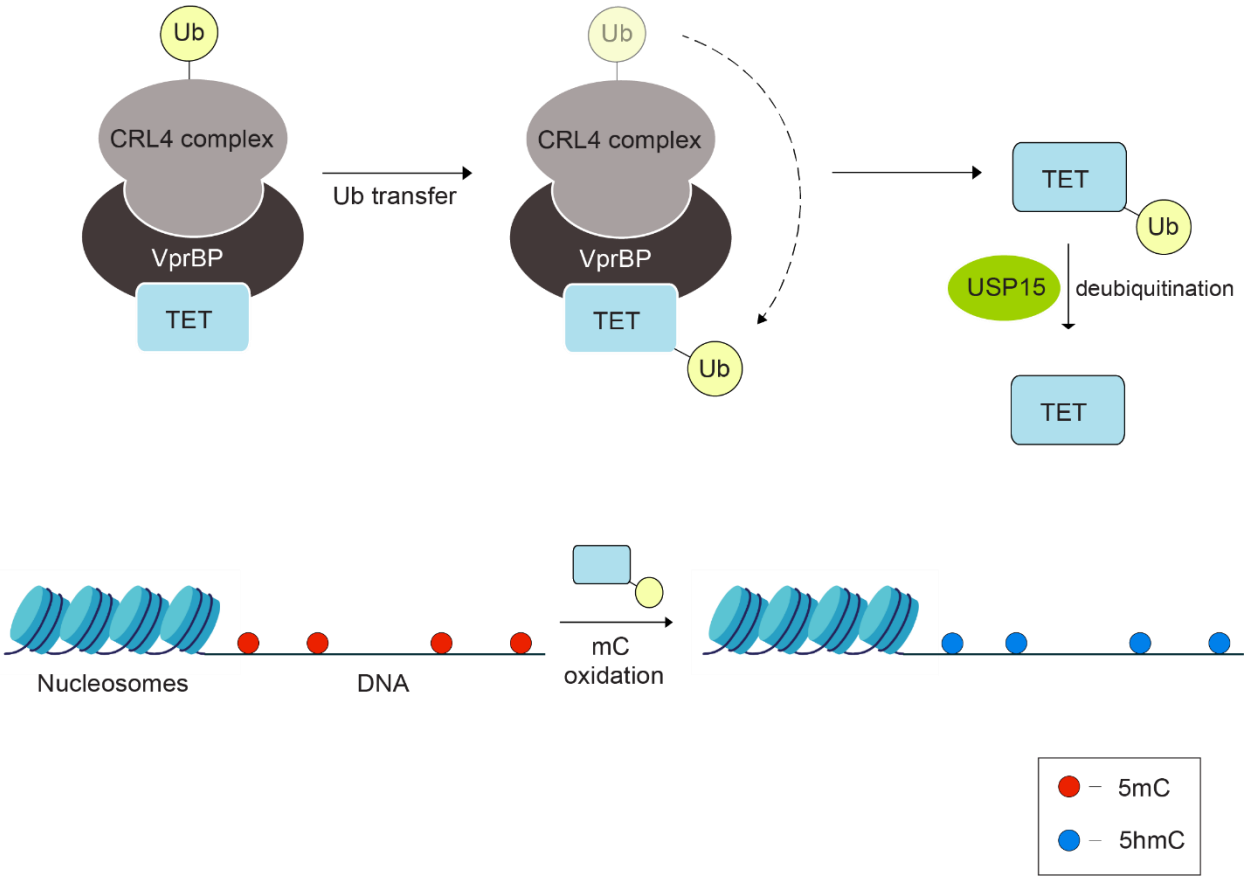



**Figure 21: Monoubiquitination of TET leads to its activity.** TET is monoubiquitinated at a conserved Lys, brought about by the CRL4$^{VprBP}$ complex. This activates TET to carry out its function of oxidizing 5mC to 5hmC and further. This monoubiquitination is reversed by the DUB USP15. Figure created based on (Nakagawa et al., 2015).

In a subsequent study, Cardoso and colleagues took this knowledge further by showing that monoubiquitinated TET1 is recognized by Uhrf1, which is a key epigenetic regulator that has been predominantly found to be binding to hemimethylated DNA and recruiting DNMT1 to ensure faithful DNA methylation. TET1 binds to the DNA binding domain of UHRF1 and is recruited to late replicating heterochromatin. This leads to spreading of 5mC oxidation to

heterochromatin regions, LINE 1 activation and chromatin decondensation. However, this action is inhibited if TET1 monoubiquitination is depleted (Arroyo et al., 2022).

Owing to its indispensability, this monoubiquitination site of TET can serve as a possible target for the external control of TET activity, without perturbing its catalytic centre.

# 5. Aim of the Study

This study mainly aims towards unraveling the dynamics of proteome-wide ubiquitination and study the kinetics of linkage-specific Ub chain formation. In light of the persisting difficulties in targeting cellular ubiquitination events due to the transient and ubiquitous nature of ubiquitination and the obvious background effects from endogenous Ub, this study aims to confer photocontrol to Ub molecules and control initiation of linkage-specific chain formation with a high temporal resolution. Such light-activatable Ub molecules allow us to specifically look into linkage-specific Ub chains in a time-dependent manner.

This work is divided into two parts where the first part majorly intends to investigate ubiquitination dynamics by: (1) Introduction of a photo-responsive ncAA pcK at specific sites in Ub to control linkage-specific chain initiation, (2) unravelling dynamics of proteome-wide K11, K48 and K63-linked chains, and (3) studying the roles of individual components of the UPS in the synthesis of K48-initiated chains. On the other hand, in the second part, the focus is on a protein-specific ubiquitination event where kinetics of TET monoubiquitination is under the spotlight, once again by utilizing the power of light as a stimulus to control protein properties.

Overall, the established system envisions to provide a valuable tool for future studies on dynamics of linkage-specific ubiquitination.

# 6. Proteome-Wide Control of Ubiquitination Reveals Kinetics of Linkage Specific Ub Chain Formation

## 6.1 Expression of Ub Variants Show Different Proteome-Wide Polyubiquitin Smear Patterns in Whole Cell Lysates

As we have seen in Chapter 2.3, strategies such as pulse-chase methods using isotope-labeled amino acids (Ross et al., 2021) or fluorophores, capping with lysine-less Ub (Oshikawa et al., 2011), advances in mass spectrometry (Prus et al., 2024), and many more, have been used to gain more insights on ubiquitination kinetics on a proteome-wide level. However, these strategies can only shed light onto the dynamics at timescales that the respective events or assays work, i.e., the expression rate of labeled proteins or diffusion and association rates of small molecules. Hence, to be able to investigate rapid, linkage-specific polyUb chain formation kinetics, construction of light-activatable Ub molecules was a promising approach. This could be achieved by activating specific Ub lysine residues with light. This strategy could be the first step towards the elucidation of the very complicated dynamics of the ubiquitome.

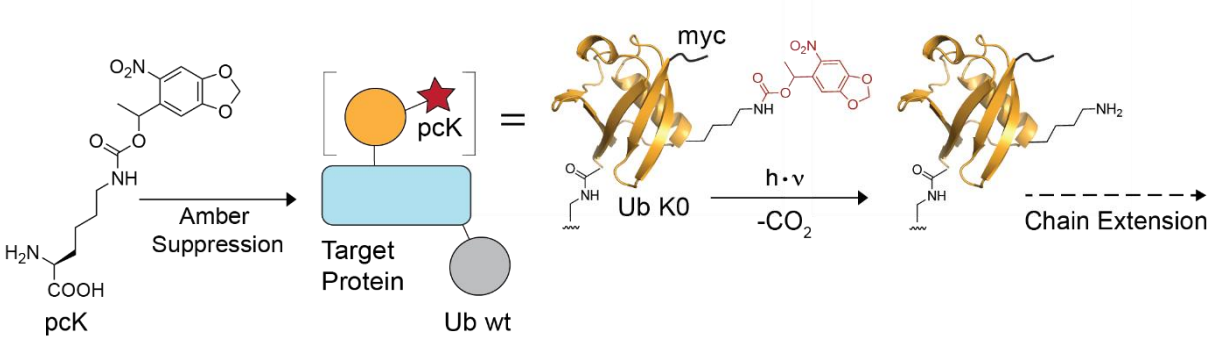



**Figure 22: The concept of light-activated ubiquitination using a genetically encoded pcK.** Incorporation of pcK into Ub via amber suppression leads to its addition on to substrate proteins. Its subsequent decaging on light irradiation leads to polyubiquitination initiated by specific linkages. pcK = photocaged Lys; Ub K0 = Ub with K→R mutations in all lysine sites except for the pcK site. Figure has been adapted from (Banerjee et al., 2024).

To establish the methodology, the focus was put on Ub variants bearing a photocaged ncAA at one of the three lysine positions K11, K48 and K63 as these represent the main ubiquitination sites most abundantly found in cellular Ub chain topologies with diverse and important functions (Dikic and Schulman, 2022). The ncAA pcK (**2**, see Chapter 3.2.1.1) was selected to be incorporated at these lysine positions so that chain elongation could be prevented at these

sites in the absence of light. Upon decaging with 365 nm light, the caging group would fall off, thereby exposing the lysine ε-amine for chain elongation (**Figure 22**). pcK itself has the advantage that it generates a ketone byproduct upon photolysis that, unlike in other simpler ONB-caged Lys, does not undergo an undesired condensation with the ε-amino group of lysine. Also, a red-shifted absorption of pcK makes it a better candidate for cellular applications over other simpler ONB groups.

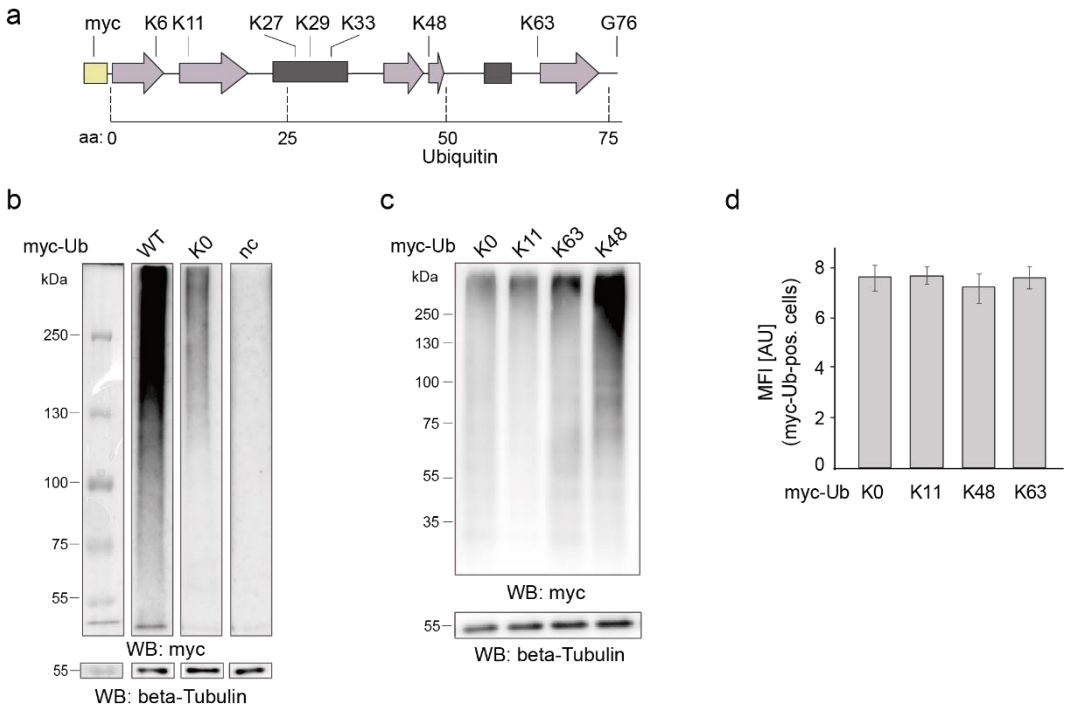



**Figure 23: Analysis of proteome-wide polyubiquitin smear patterns in whole cell lysates from expression of Ub variants. a.** Domain structure of Ub construct used in this study. Pink arrows = β-strands, grey boxes = α-helices. **b.** SDS PAGE/anti-myc blots of proteomes from HEK293T cells expressing indicated Ub variants for 24 h (all lanes are from the same blot with identical exposure time). Cells were treated with 10 μM MG132 5 h before harvesting (also applies to panels c-d). **c.** Experiment as in Fig. 2b with indicated Ub linkage variants. **d.** Anti-myc immunostaining and FACS analysis of cells from figure 23c reveal similar total expression levels of Ub variants (error bars from N=3 replicates). MFI= Mean fluorescence intensity; AU= arbitrary units. Figure has been adapted from (Banerjee et al., 2024).

As the first step towards studying proteome-wide ubiquitination, an assay is required to be able to visualize proteome-wide ubiquitination in cells, which would also be able to differentiate between the Ub readouts at different times, when different Ub variants are expressed. As an initial test, wild type Ub (Ub wt), Ub K0, a Ub mutant where all the seven lysines are mutated to arginine and thus are only capable of ubiquitinating targets with a monoUb or multimonoUb, and a Ub version that cannot be conjugated to lysines because of an altered C-terminus lacking G76 (Ub nc, **Figure 23a**), with N-terminal myc tags, were individually expressed in HEK293T

cells. The proteasome inhibitor MG132 (10 μM) was added to the cells 5 hours (h) before harvesting so that degradation of polyubiquitinated protein substrates in the proteasome could be prevented and a clearer picture of the proteome modified with the overexpressed Ub variant could be achieved. SDS PAGE analyses of isolated whole proteomes and subsequent anti-myc blots revealed the proteome subpopulation modified with our Ub variants as a smear with a broad range of molecular weights (**Figure 23b**). As expected, the proteome fraction modified with Ub K0 showed only a fraction (14 %) of the intensity observed for Ub wt as Ub K0 can only modify substrates with mono and multi-monoUb due to the absence of any Lys on which polyUb chains can be built, and as a result, the overall number of Ub molecules on a target protein are much less. For Ub nc, on the other hand, no smear was observed as expected due to the lack of Ub C-terminal Gly, showing that the observed signal exclusively resulted from protein conjugates of the Ub variants (**Figure 23b**). The polyUb smear patterns obtained for each of the Ub variants coinciding with the expected trends and showing visible difference, led us to think that it would be possible to use this system as a readout for modified Ub molecules.

Next, Ub variants with only single Lys residues at positions of interest, that is, Ub K11, K48 and K63 that each differ from K0 only in the single indicated R→K mutation, were expressed. Owing to the single lysines, these variants allow Ub chain elongation only in a linkage-specific manner. Similar workflow of analysing ubiquitome showed that all three single K variants resulted in similar or higher myc-proteome intensities than Ub K0. K48 showed the highest levels, being in agreement with the generally high abundance of this linkage type in Ub topologies (Dikic and Schulman, 2022) (**Figure 23c**). To rule out the possibility of differential expression patterns of the different Ub variants, FACS analyses of anti-myc immunostained cells were conducted which revealed similar total expression levels for all four Ub variants, indicating that the differences in the formed myc-Ub proteome results from specific linkage preferences of the UPS rather than different expression and availabilities of the Ub variants (**Figure 23d** and **Figure S1**).

## 6.2 Construction of Light-Activatable Ub by Incorporation of pcK

With an assay for readout in place, the light-activatable, photocaged Ub variants were ready to be designed. First, amber mutants of the single Lys containing Ub variants, as described above, were constructed, where the respective lysine positions are mutated to an amber stop codon (TAG) (**Figure 23a**). For instance, in pcK48 Ub, the 48th residue is mutated to TAG. To incorporate pcK at positions K11, K48 and K63, vector encoding the *Methanosarcina mazei* pyrrolysyl-tRNA-synthetase pair pcKRS/tRNAPyl that has been engineered to use pcK as

substrate (explained in more details in section 3.2.2) along with the respective Ub vectors containing the single in-frame amber codon (TAG) at the incorporation target sites, were co-transfected into HEK293T cells. The cells were further cultivated in presence or absence of 0.32 mM of pcK in the growth medium. Upon immunoblot analyses of myc-positive proteome, it was observed that when 1 is not present in the growth medium of cells, no significant polyUb smear is visible (**Figure 24a**) owing to the fact that without any amino acid to be incorporated at the amber site, it is read as a stop codon and gives rise to truncated Ub that cannot be added to any substrate. The truncated Ub generated are ultimately degraded. In presence of the ncAA, a considerable increase in the myc-positive proteome was observed in the cells for all the three amber positions (**Figure 24a, b**). Moreover, as a proof of the specificity of incorporation of 1 by the pcKRS, only a trace amount of myc-positive Ub proteome was observed when pcKRS was replaced with wt PylRS as the aaRS/tRNA pair, for which pcK is a poor substrate (**Figure 24a, b**). This confirmed site-specific incorporation of pcK specifically in the presence of the correct aaRS/tRNA pair, for all three caged Ub variants.

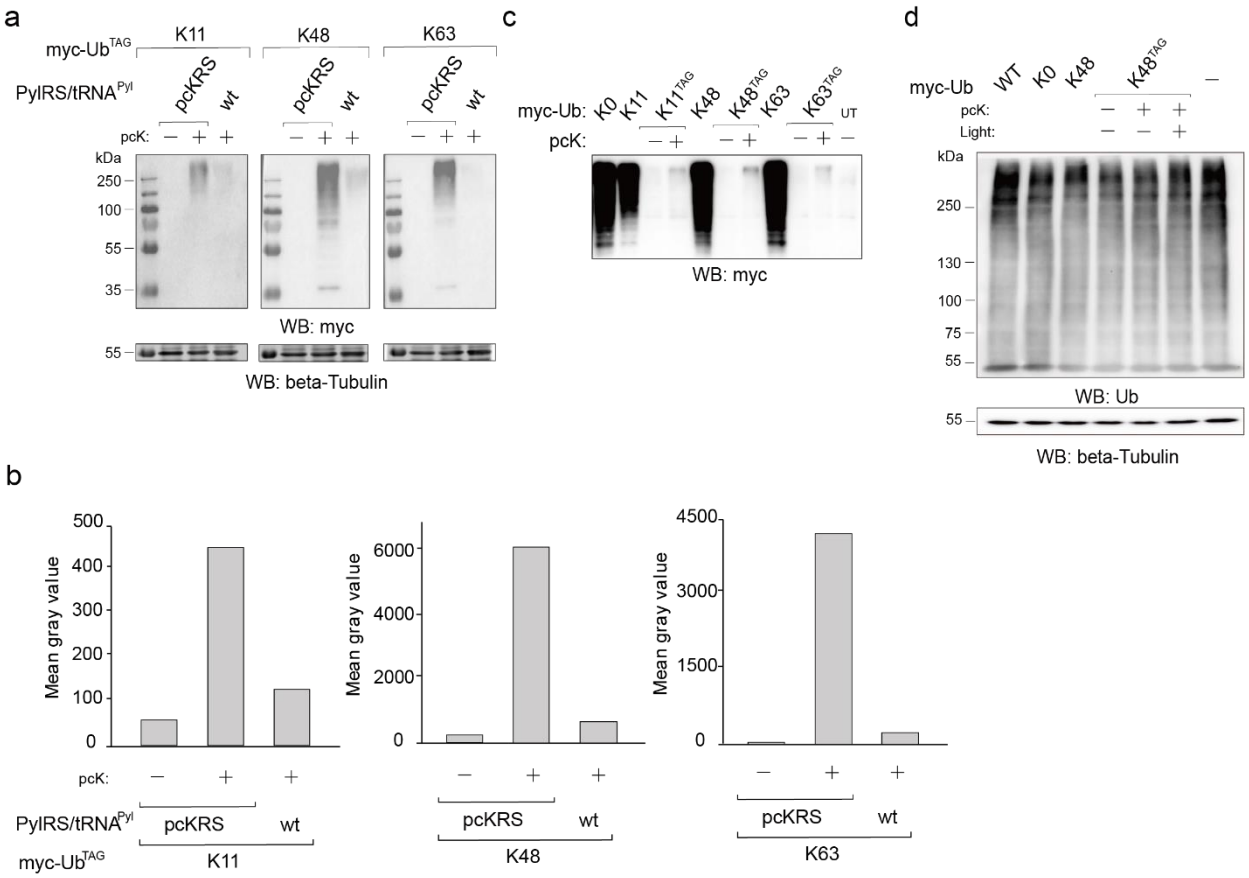



**Figure 24: Analysis of Ub proteome subpopulation upon pcK incorporation. a.** High fidelity incorporation of pcK at three different Ub Lys sites in HEK293T cells expressing the indicated amber myc-Ub variants along with the indicated PylRS/tRNA$^{Pyl}$ pair for 24 h (individual contrast settings for

each linkage). Cells were treated with 10 µM MG132 5 h before harvesting (also applies to panels c-d). **b.** Quantification of blots from (a) using Fiji Image J. **c.** SDS PAGE/anti-myc blots of proteomes from HEK293T cells expressing myc-Ub variants with or without amber codon and in presence or absence of pcK for 24 hours reveal minimal expression levels of caged Ub variants. **d.** SDS PAGE/anti-Ub blots of proteomes from HEK293T cells expressing Ub variants for 24 h with or without amber codon and in presence or absence of pcK reveal minimal expression levels of Ub variants compared to the endogenous Ub pool. Cells expressing non amber Ub variants were harvested directly after 24 h and those expressing amber Ub variants were further treated with light and harvested 24 h later. Figure has been adapted from (Banerjee et al., 2024).

In this (dark) state, the expressed caged Ub is functionally equivalent to the Ub K0. Hence the Ub proteins with caged lysines are able to be added on to substrates as mono- or multimono-ubiquitination marks as well as be added to Ub chains as caps, but they prevent chain extensions in the absence of light. Such priming of the ubiquitome thus allows us to selectively look at the linkage-specific chain extension upon decaging of the lysine. However, a common side effect of amber expression is a decrease in expression due to recognition of the amber as a stop codon leading to the formation of truncated products, as described in section 3.2.4. Incorporation of pcK thereby results in a considerable reduction in the expression levels of caged Ub variants, compared to the non-caged variants (**Figure 24c**). However, when anti-Ub blots were performed, the intensity of the total ubiquitome did not differ between cells expressing pcK-modified myc-Ub or non-amber Ub (**Figure 24d**). This indicates that the population of caged myc-Ub proteome represents only a small proportion of the total ubiquitome, which is less likely to perturb the cell´s UPS by extensive overexpression.

## 6.3 Light Activation and Kinetics

After achieving successful incorporation of pcK in Ub, the focus of the study was shifted towards the light activation of the caged Ub variants to elucidate linkage-specific Ub chain formation kinetics. Initially, the kinetics of *de novo* ubiquitination on the respective lysines (giving rise to the linkage-specific polyUb chains) were observed over a time window of 6 h. Cells were transfected and cultivated in the presence of 0.32 mM of pcK, as described above, and irradiated with 365 nm light using an UV trans-illuminator. For the light treatment, the growth media containing pcK was exchanged with warm DPBS (lacking pcK) and after light irradiation, cells were grown in full media also lacking pcK to terminate expression of the photocaged Ub. This growth media further contained 25 µM MG132 so that the requirement of studying *de novo* ubiquitination kinetics could be fulfilled and ubiquitome synthesis kinetics could be studied uncoupled from simultaneous proteasomal degradation. HEK293T cells were then harvested at corresponding time-points after light and their proteomes extracted. Myc-Ub proteome formations were analysed by SDS PAGE/anti-myc blots as described above (**Figure**

**25a**). *De novo* myc-ubiquitination in both irradiated and nonirradiated cells was observed to increase, with the latter however showing markedly slower increases. This indicated that a sufficiently high fraction of the ubiquitination sites in the proteome was modified with the

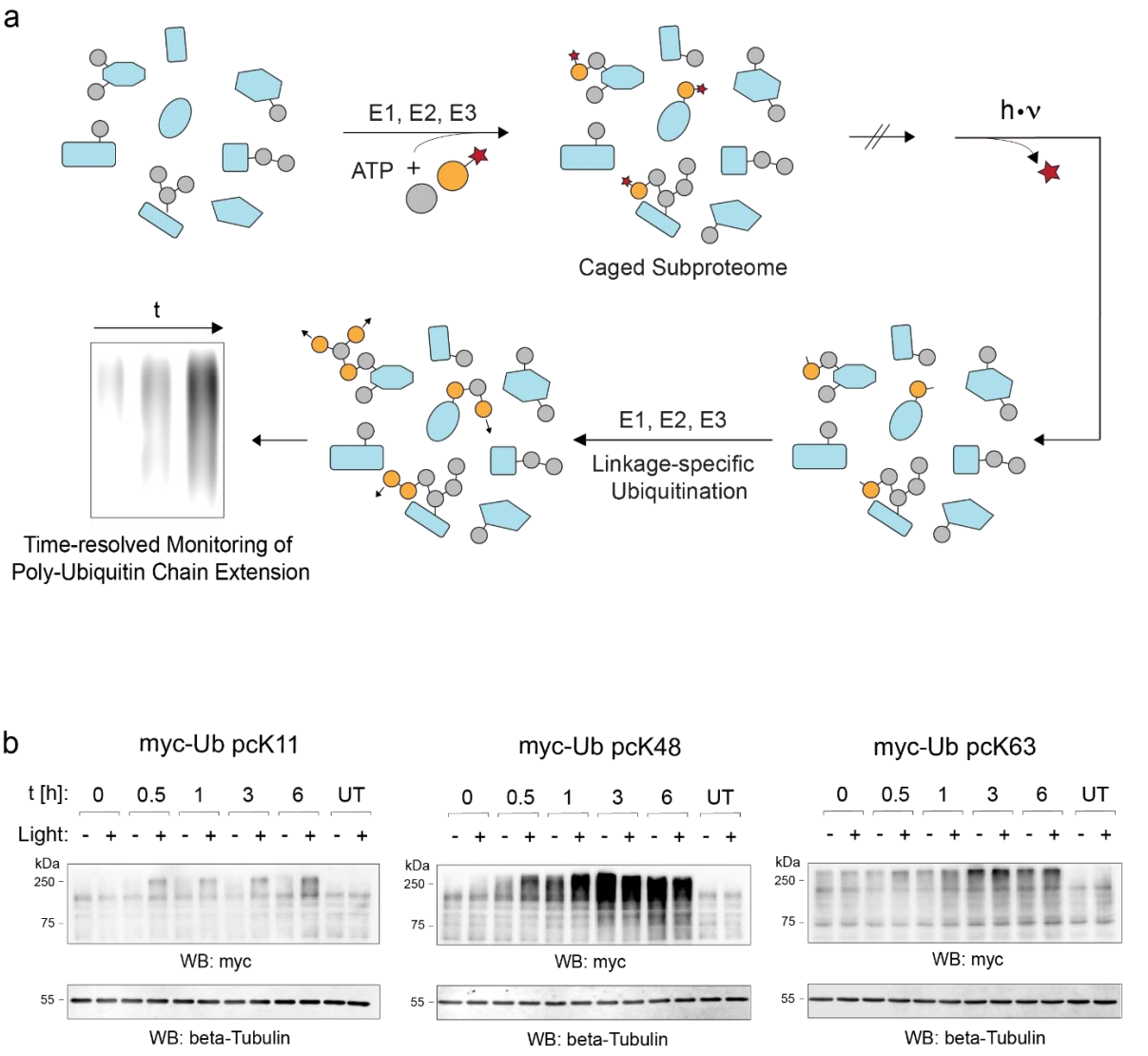



**Figure 25: Light mediated decaging of caged Ub leads to increase in high molecular weight myc-Ub proteome. a.** Schematic diagram showing the biosynthesis and light-activation of a caged Ub subproteome that acts within the ubiquitin proteasome system (UPS) and upon light-activation, leads to Ub chain extension from the specific Lys. **b.** Long-term, linkage-specific ubiquitination kinetics after light activation of caged myc-Ub variants. HEK293T cells were cultivated in presence of 0.32 mM pcK for 24 h after transfection. After light activation, cells were grown in full media containing 25 µM MG132, lacking pcK, and harvested at each corresponding time-point after light. t = time after light. Figure has been adapted from (Banerjee et al., 2024).

photocaged Ub proteins, making this approach suitable for selectively studying the linkage-specific *de novo* ubiquitination after light (**Figure 25b**, **Figure S2**). As a control, cells grown

under similar conditions, but expressing non-amber Ub K48 or K0 did not show any changes in the myc signals after light treatment, confirming that the observed light-dependent increases of high molecular weight myc-Ub proteome were not an unspecific light response of the cells (**Figure S3**). Overall, the observed kinetics were rapid for all three linkage types, with large increases already after 0.5 h. The three linkage types however strongly differed in the abundance, that is, the amount of proteome modified with the myc-Ub in each case, with K48 again showing the highest signals. It was further observed that the myc-Ub signals reached similar levels in both – and + light samples with time. This meant that myc-Ub proteome amounts synthesized from available free lysine residues alone (- light) reached similar levels as the ones from both free and decaged lysines combined (+ light). This could be explained by the concurrent proteasome inhibition that leads to a depletion of monomeric Ub. This further becomes visible by comparing -/+ light for pcK48 and pcK63, whereas for pcK11, no saturation is reached within the observation time window (**Figure 25b, Figure S2**).

To have a closer, more zoomed-in look at the kinetics, measurements were next conducted with a short time window of up to 45 minutes (min) post light. It was observed that the myc-Ub proteome increases already after 15 min, confirming a very rapid build-up of Ub proteome from all three linkages (**Figure 26a, Figure S4**). In order to reveal the light-induced, linkage-specific synthesis kinetics without taking the background synthesis of myc-Ub proteome on free lysines (as seen in -light cells) into account, the myc signals from irradiated cells were normalised to the nonirradiated cells of each time point.

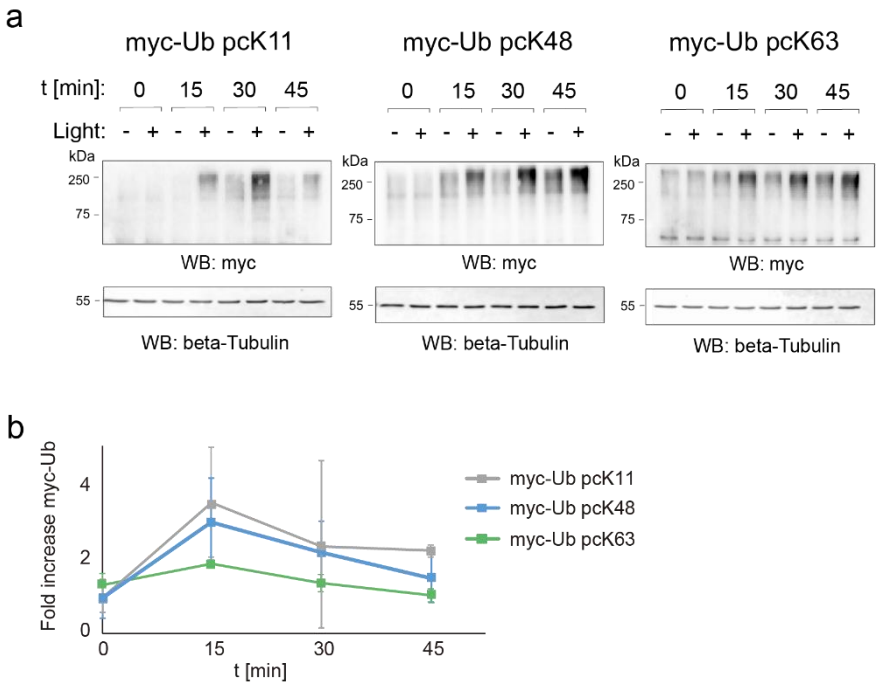



**Figure 26: Short time-course studies of *de novo* Ub chain formation. a.** Rapid, initial ubiquitination kinetics (<15 min after light) are seen for all linkages. HEK293T cells were cultivated in presence of 0.32 mM pcK for 24 h after transfection. After light activation, cells were grown in full media

containing 25 µM MG132, lacking pcK, and harvested at each corresponding time-point after light (also applies to panel b). **b.** Plots of rapid, initial ubiquitination kinetics (myc-Ub proteome intensities normalized to beta-tubulin blot and no light control of each lane; error bars from N≥2). t = time after light. Figure has been adapted from (Banerjee et al., 2024).

These kinetics – though differing between the linkages – generally revealed an initial burst with a maximum at 15 min in all cases, and a subsequent stabilization in relation to the background (**Figure 26b**).

## 6.4 Effect of E1-E2-E3 Inhibitors

As we have seen in details in Chapter 2.2, small molecule inhibitors of UPS components hold great therapeutic potential and have been used as major perturbation tools to study cellular roles of the UPS. Applying such tools in this study could help in elucidating the involvement of the individual UPS components in the observed rapid, linkage-specific de novo synthesis of Ub chains. Thus, the above-mentioned methodology was applied to compare the cellular polyUb syntheses in cellular states with differently inhibited UPS components. Three small molecule inhibitors were chosen that perturb different steps of the ubiquitination/proteasomal degradation pathway: MLN7243 as a mechanism based, AMP-mimicking inhibitor of both mammalian E1 enzymes (UAE and UBA6), MLN4924 as a covalent inhibitor of the NEDD8-activating enzyme (NAE) to inhibit Cullin-Ring-E3 ligase activation, and NMS873 as an inhibitor of the AAA+ ATPase VCP/p97 (**Figure 11a**, **Figure 27**). Under the influence of these inhibitors, the initiation of particularly K48-linked chains, the most abundant linkage type, was measured under conditions of proteasomal inhibition by MG132 (**Figure 28a**).

MLN7243, being an E1 inhibitor, effectively starved the cells of their Ub pools and shut the downstream ubiquitination events off (**Figure 27, Figure 28a**). Even after pre-incubating the HEK293T cells with 1 µM MLN7243 for 6 h, interestingly, a residual *de novo* ubiquitination was however still visible. This hints at the fact that low levels of active E1 and/or activated Ub are still available under these widely used E1 inhibitory conditions (**Figure 28a**).

Next, MLN4924, a covalent inhibitor of the NEDD8-activating enzyme (NAE) was tested for its effect on the system under study. As we have also seen in the previous chapter, NAE-dependent NEDDylation controls Cullin-RING-E3 ligases (CRL), the largest E3 ligase class that regulates proteins with important functions in cancer cell growth and survival pathways (**Figure 27**). However, it was observed that the global K48-linked polyubiquitin kinetics

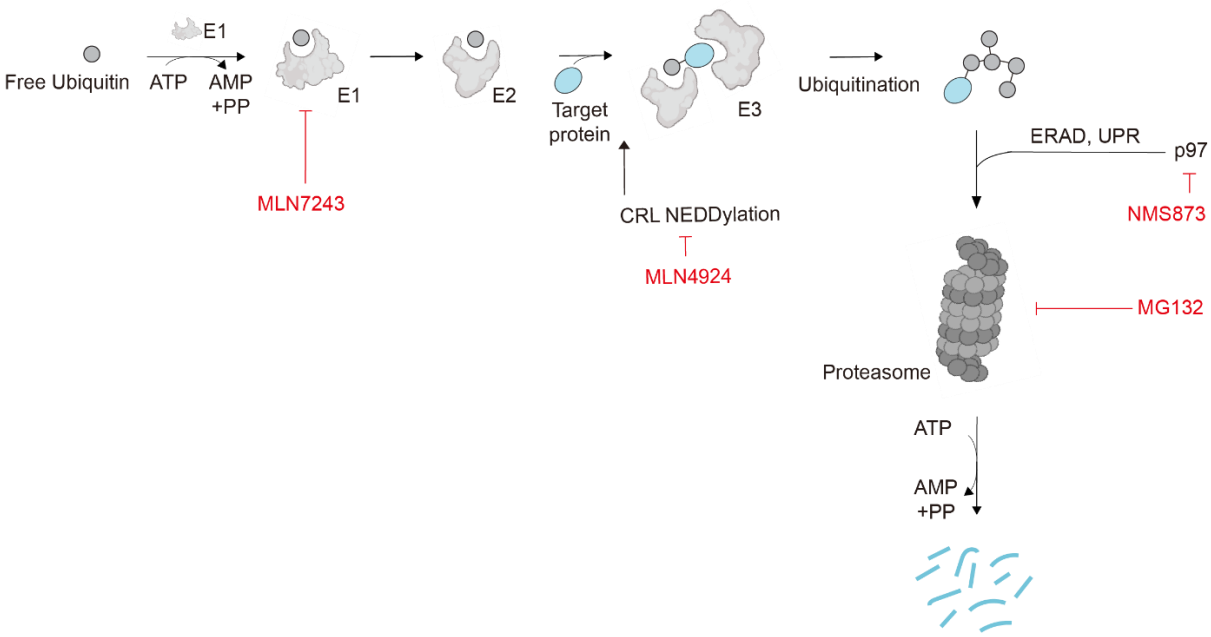



**Figure 27: Schematic diagram showing UPS synthesis and degradation pathway highlighting the UPS factors targeted by inhibitors used in the study.** Figure has been adapted from (Banerjee et al., 2024).

did not change even after a 4 h preincubation with 1 µM MLN4924 (**Figure 28a**). This suggested that the majority of the observed myc-Ub chain extension initiated at K48 occurs independently of CRL activity (Soucy et al., 2009), in agreement with previous observations (Hyer et al., 2018).

Lastly, the effect of NMS873 was studied. As mentioned above, NMS873 is an inhibitor of the AAA+ ATPase VCP/p97 that promotes proteasomal degradation of ubiquitinated substrate proteins by their extraction from membranes or complexes, and possess co-factors that have specificities towards K48-linked Ub chains predominantly (**Figure 27**). Initially, treatment with 5 µM NMS873 for 4 h did not result in a reduced K48-initiated ubiquitome synthesis rate or accumulation of myc-Ub positive proteome as expected, in presence of MG132. This was in agreement with the fact that the proteasome activity is required for client protein degradation, and effects of NMS873 would be overshadowed by the effects of MG132 (**Figure 28a**). Indeed, experiments in the absence of MG132 showed a marked increase of myc-Ub proteome levels in the presence of NMS873, directly revealing the effect of VCP/p97 as an activator or helper of K48-linked *de novo* ubiquitome synthesis on a rapid timescale (**Figure 28b, c, Figure S5**), being consistent with previous reports (Locke et al., 2014).

## 6.5 Conclusion

In conclusion, light activation of Ub chain extension is reported here in mammalian cells. It could be shown that different variants of Ub give rise to characteristic smear patterns of polyubiquitination in whole cell lysates of treated HEK293T cells. Ub variants with single lysine residues also differ in such smear patterns depending on the position of the lysine, K48 polyUb chains generating the most intense smears, followed by K63 and K11. With this strategy, protein polyUb could be controlled in a linkage specific manner at a temporal resolution. Incorporation of a genetically encoded photocaged lysine via an engineered PylRS/tRNAPyl pair enabled the expression of myc-tagged Ub variants that are only able to synthesize chains from specific, individual lysine sites and are then employed by the UPS to

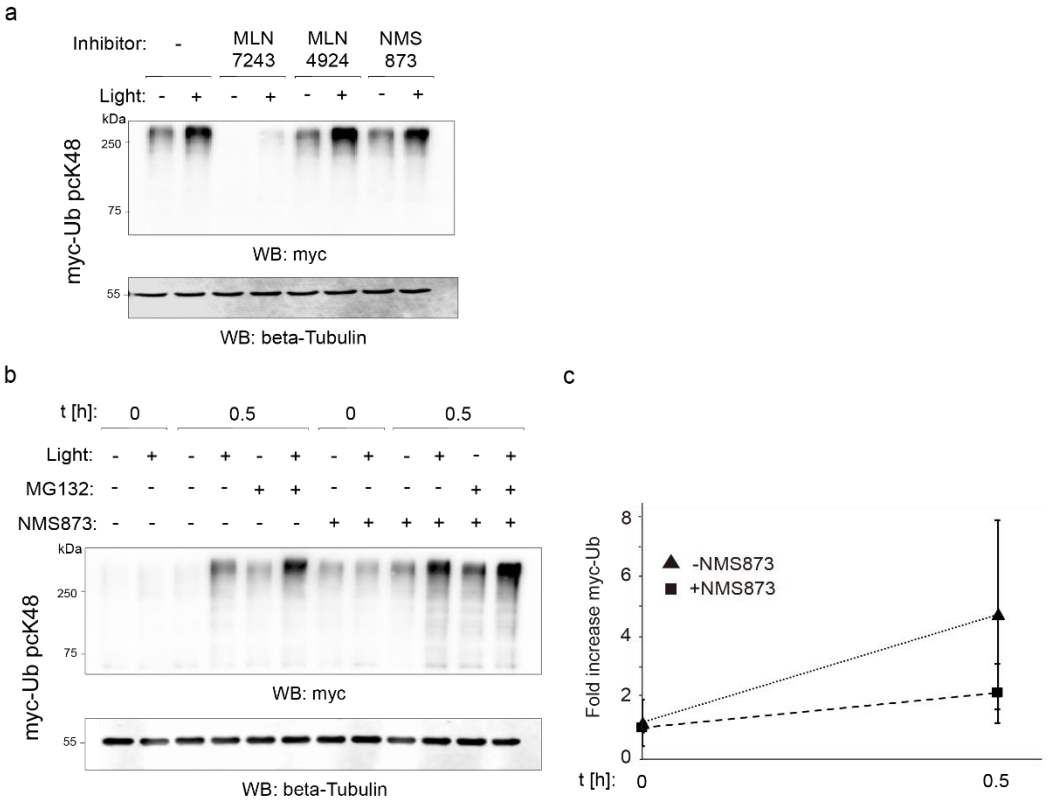



**Figure 28: Effects of UPS inhibitors on early, K48-specific *de novo* ubiquitinome synthesis. a.** HEK293T cells were cultivated in presence of 0.32 mM pcK for 24 h after transfection. The inhibitors were added to the media in concentrations as mentioned. Following the pretreatment times, light activation was performed and cells were grown in full media containing 25 µM MG132 and the respective inhibitors, lacking pcK, and harvested 0.5 h after light (also applies to panels b and c). **b, c.** Effect of NMS873 on early, K48-specific *de novo* ubiquitinome synthesis in presence and absence of MG132 (error bars from N=2). Figure has been adapted from (Banerjee et al., 2024).

synthesize a sub-population of the cellular ubiquitome that is indeed a small percentage and trackable.

It was shown that irradiation with light triggers ubiquitination from the respective lysine sites of interest following decaging, enabling the monitoring of de novo ubiquitination in the proteome initiated by the specific linkage types. Decaging following light activation showed the synthesis of K11, K48 and K63 linked polyUb chains at different levels, with burst-like and rapid kinetics that result in reaching a maximum already after 15 minutes for all linkage types. Furthermore, inhibition of individual components of the UPS that are responsible for Ub activation (E1) or activation of CRL E3 ligases (NAE) or even involved in the unfolding of substrate proteins ubiquitinated with K48-linked chains, for proteasome-dependent degradation by ERAD (VCP/p97), revealed the impact of these components on the synthesis of K48-initiated Ub chains on the low minutes time scale. It could be seen with the E1 inhibitor MLN7243 that even after treating cells with standard inhibitory conditions, there are still low levels of active E1 and/or activated Ub available. The NAE inhibitor MLN4924 showed that the global K48-linked polyubiquitin kinetics did not change even after a 4 h preincubation with 1 μM compound, suggesting that the majority of the observed myc-Ub chain extension initiated at K48 occurs independently of CRL activity. Finally, p97 inhibitor NMS873 proved the effect of VCP/p97 as an activator or helper of K48-linked *de novo* ubiquitome synthesis on a rapid timescale as it showed a marked increase of K48-linked myc-Ub proteome levels in the absence of MG132.

The high temporal resolution of this approach and its ability to selectively zoom into linkage-types adds new abilities to the current toolset of UPS perturbation strategies. The straightforward workflow also makes it compatible with diverse downstream analysis techniques that could be undertaken to study rapid, linkage specific ubiquitination dynamics for individual proteins as well as on global proteomic levels as done here. Given the central role of Ub in regulating protein function, localization and quality control, in a vast number of normal and disease-related processes, this approach is highly anticipated to find wide use for studying the dynamics of linkage-specific ubiquitination.

# 7. Control of Target-specific Ubiquitination in Mouse TET2

## 7.1 Inhibition of TET-Ub Leads to Inactivity

Upon the establishment of an approach to control global cellular ubiquitination in a linkage-specific manner, it was of interest to see whether a similar approach could be applied to study the ubiquitination kinetics of a single protein of interest. As the protein of interest, Ten-Eleven Translocation protein was chosen. As we have seen in Chapter 4, TET proteins are extremely important players in the epigenomic scenario of mammals and are always a very relevant substrate to be studied. And as seen above, all three isotypes of TET undergo monoubiquitination at a conserved Lys residue, that in turn appears to be indispensable for TET activity by facilitating stronger chromatin interactions. This made TET proteins the perfect target on which we could study the kinetics of the monoubiquitination event. This would not only shed light on the in-cell kinetics of TET monoubiquitination, but also confer control over activity without tweaking the catalytic center. And using light-mediated control would overcome system noise and enable uncoupling the kinetics of monoubiquitination itself from potentially rate-limiting upstream processes in the TET life.

Out of the three TET isotypes, TET2 was chosen as the *TET2* gene has been found to be frequently mutated in hematopoietic cancers, of both myeloid and lymphoid lineages. And many of these mutations occur at the conserved lysine or other residues involved in VprBP binding. Although it was shown extensively by Nakagawa et al. (Nakagawa et al., 2015), to start with, it was important to reassess the importance of the said Lys residue of the different TET isoforms on their activity. As the starting point, to establish the importance of the conserved ubiquitinated lysine, this residue was mutated to Aspargine (N) or Glutamate (E) in the catalytic domain (roughly from residues 1042 to 1849 (Hu et al., 2013)) of mouse TET2 (K1212 of mTET2CD). For this, a construct previously generated in the lab (p2392), where mTET2CD with a C-terminal HA tag fusion, was cloned into a mammalian expression vector, was used and the K1212 of TET was accordingly mutated. The activity of the mutated TET2 was analyzed using an immunostaining based FACS assay. To this end, HEK293T cells expressing the mutated mTET2CD were stained for the total cellular 5hmC using anti-5hmC antibody, in addition to staining for the expression tag (HA, in this case). Cells expressing the TET were then analyzed for the total cellular 5hmC levels, that is indicative of TET activity. As the background levels of 5hmC is quite low in HEK293T cells (Ito et al., 2011), expression of a functionally active TET leads to a significant rise in cellular 5hmC levels. When the K1212N or K1212E variants of mTET2CD were expressed in HEK293T cells, a marked

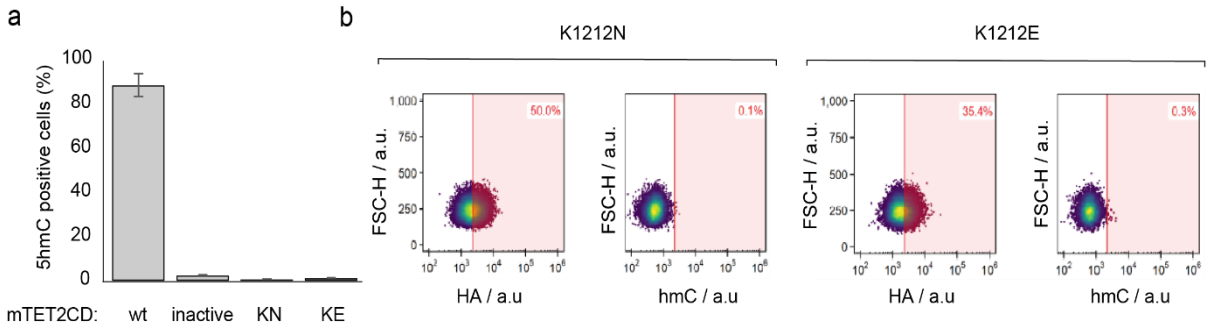



**Figure 29: Mutation of K1212 of mTET2CD leads to its inactivation. a.** Anti-HA and Anti-hmC immunostaining and FACS analysis shows inactivity of K1212 mutated mTET2CD (K1212N/E), comparable to that of a catalytically inactive variant (H1295Y, D1297A) (error bars from N=2). **b.** Representative FACS density plots. Red boundary indicates the intensity thresholds determined by the fluorescence intensity of untransfected cells. a.u. = arbitrary units.

reduction was seen in the 5hmC content of cells expressing the construct, compared to that of cells expressing the wild-type mTET2CD (**Figure 29**). And indeed K1212N/E mutants showed activity similar to another inactive TET mutant where the catalytic residues His1295 and Asp1297 (both involved in stabilizing the α-KG cofactor, also mentioned in Chapter 4) are mutated to Tyr and Ala (H1295Y, D1297A), respectively, leading to complete inactivity (Tahiliani et al., 2009). This indicates that indeed the conserved lysine residue that is the site for monoubiquitination is very much indispensable for TET activity and mutations at this site leads to disruption of 5mC oxidation.

## 7.2 Co-expression of TET, Ub, VprBP Leads to Active TET Function

After the establishment of the importance of K1212 for TET2 activity, the system followed by Nakagawa and colleagues was tried to be reproduced. With an *in vivo* set-up expressing mTET2CD, VprBP and Ub in place, it could later be utilized to express an amber mutant of the mTET2CD with an amber stop codon in place of K1212 that would incorporate pcK upon amber suppression. In the absence of light, pcK1212 would prevent TET from getting monoubiquitinated, hence keeping it in an inactive state. Upon light irradiation, decaging of 1 would reinstate monoubiquitination, making the TET active again. This would allow for the ontrol of TET-monoubiquitination, and hence its activity with light and shed light on the kinetics of monoubiquitination.

So as the next step, HEK293T cells were co-transfected with a wt mTET2CD construct (p2963, described in more details in section 8.2.1) with C-terminal fusions of 3X Flag tags (mTET2CD-Flag), wildtype human VprBP with N-terminal HA tag fusion (Addgene #133586) (HA-VprBP) and wildtype Ub with an N-terminal myc tag fusion as described in Chapter 6 (myc-Ub). The previously HA-tagged TET was changed to a Flag-tagged construct due to a higher sensitivity in immunostaining as well as in immunoblots, of the 3X Flag tag (**Figure S7**) as seen with FACS analysis. The thought behind this was that the overexpression of VprBP along with TET would ensure that there is enough E3 ligase to ubiquitinate the overexpressed TET. Also, heterologously expressing wt myc-Ub generates enough Ub for the ubiquitination of the overexpressed TET. In addition to FACS analysis, after expressing the desired constructs for 24 h, the proteomes of the cells were isolated and subjected to SDS-PAGE followed by immunoblot analysis against each expression tag. As we see in **Figure 30a,** TET2CD-Flag and HA-VprBP are well expressed as indicated by the molecular weights. And expression of myc-Ub was shown by the polyUb smears indicative of the proteome being (poly)ubiquitinated with our myc-Ub.

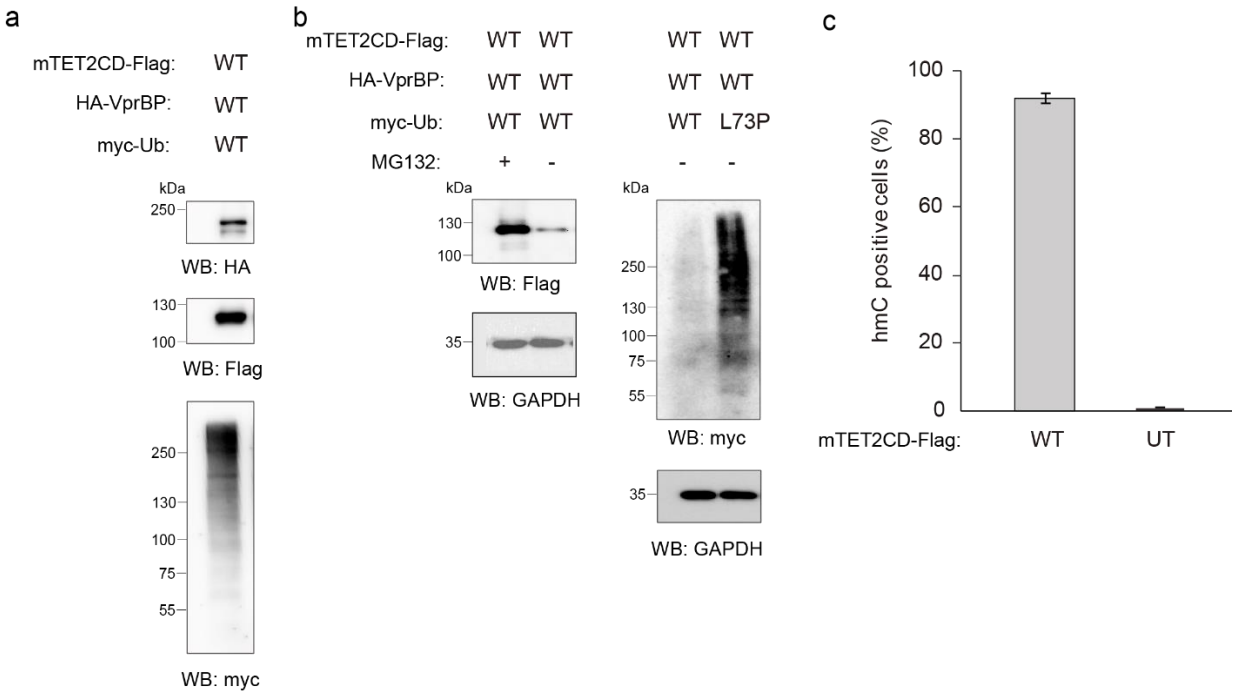



**Figure 30: Co-transfection of HEK293T cells with TET, VprBP and Ub leads to the expression of a functional TET. a.** Immunoblots against indicated expression tags show the successful expression of the three proteins of interest. Cells were treated with 10 µM MG132 5 h before harvesting (also applies to panel b). **b.** Treatment of cells with MG132 helps retain a large population of ubiquitinated proteome (left). However, DUB resistant Ub variant L73P is resistant to the effect of functional proteasomal degradation (right). **c.** Anti-hmC immunostaining and FACS analysis shows the TET expressed in the described set-up is functionally active (error bars from N=2).

For the immunoblotting, before harvesting cells for lysis, they were treated with 10 µM MG132 for 5 h in order to prevent turnover of TET substrate owing to polyubiquitination with endogenous or myc-Ub as in seen in **Figure 30b.** As is observed, in the absence of MG132, the available mTET2CD population is only about 20 % of the total population available in presence of it. Only the DUB resistant Ub mutant L73P (myc-Ub-L73P) seems to show resistance of the proteome decorated with this Ub variant to proteolytic degradation, owing to its ability to oppose Ub chain cleavage.

Next, to reassess TET activity in this established set up, the above-mentioned immunostaining based FACS assay was used on similarly treated cells to look at the percentage of 5hmC positive cells amongst the population expressing all the three constructs. It is worth mentioning that for FACS assay, cells were not treated with MG132 as the final readout is not of an ubiquitinated species, but the catalytic activity. As we see in **Figure 30c** and **Figure S8a**, wildtype TET is active in oxidizing 5mC to 5hmC, where above 91 % of cells expressing all the three constructs are 5hmC positive, showing hmC levels above the threshold set on the basis of untransfected cells. These results provided a good base on which the next step of controlling TET-monoubiquitination with light could be built.

## 7.3 Construction of pcTET-Ub and Successful Control of TET-Ub Through Photocaging as Seen with FACS Analysis

With the wildtype system established, the focus was shifted to the amber mutant of mTET2CD (p2962), where the K1212 is mutated to an amber stop codon (TAG). To incorporate the pcK, in addition to the mTET2CD amber mutant (p2962), the vector encoding the *Methanosarcina mazei* pyrrolysyl-tRNA-synthetase pair pcKRS/$tRNA^{Pyl}$ with a wt myc-Ub fusion (p2956) along with wt VprBP vector (Addgene # 133586) and were co-transfected into HEK293T cells. To reduce the number of plasmids for transfection in the caged TET (pcTET) set-up, the pcKRS/$tRNA^{Pyl}$ vector was modified with a fused ORF for myc-Ub under the CMV promoter (described in more details in section 8.2.1). The cells were further cultivated in presence or absence of 0.32 mM of 1 in the growth medium. Upon FACS analyses, it was observed that when 1 is not present in the growth medium of cells, there's only a basal Flag signal in cells (similar to untransfected cells) (**Figure S8b**) owing to the fact that without any amino acid to be incorporated at the amber site, it is read as a stop codon and gives rise to truncated mTET2CD lacking the C-terminal 3X Flag tags, thus allowing for the discrimination between fully expressed and truncated protein constructs. The truncated TET generated are ultimately degraded. In presence of the ncAA, a considerable increase in the Flag-positive cells was

observed (**Figure S8b**). This confirmed site-specific incorporation of pcK at 1212 position of mTET2CD in the presence of pcKRS/tRNA$^{Pyl}$ pair.

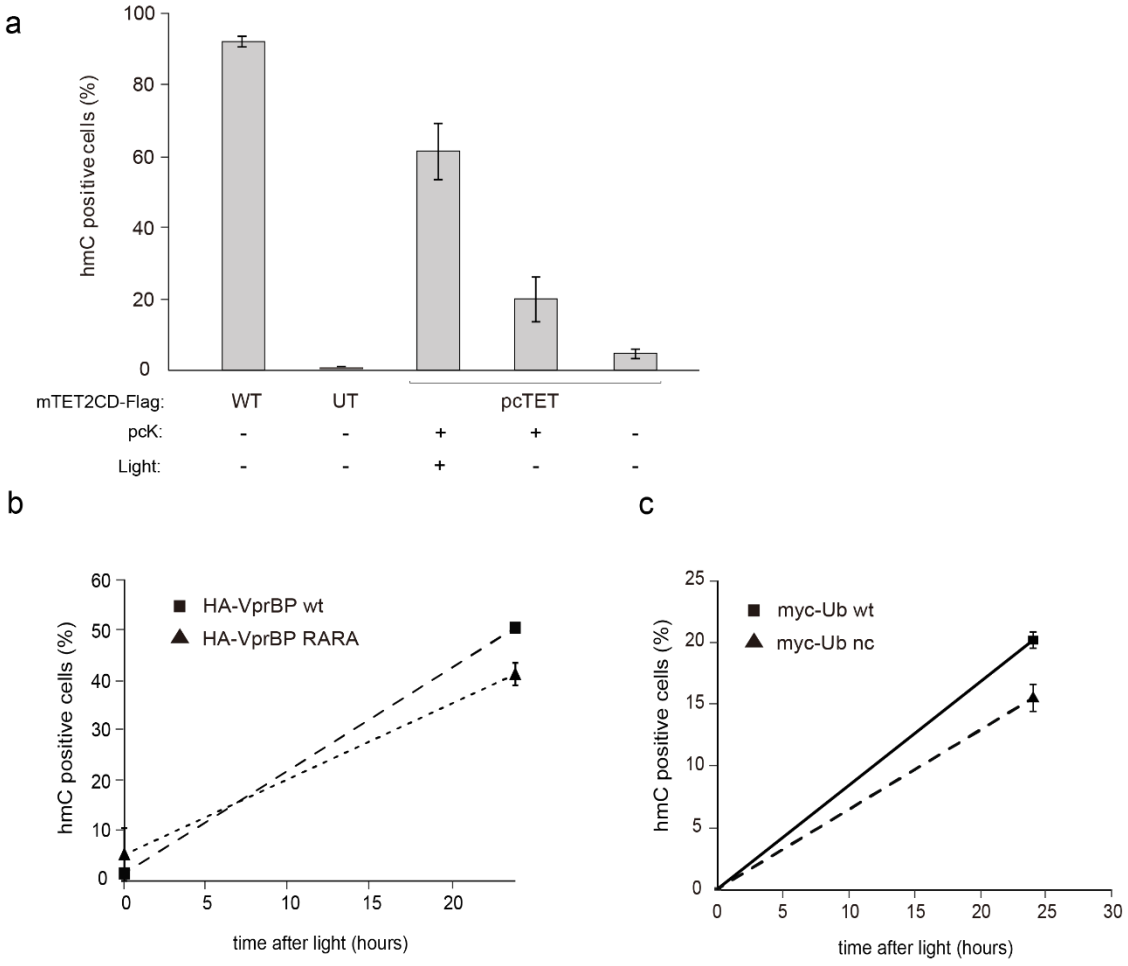



**Figure 31: Succesful incorporation of pcK in mTET2CD and subsequent light-mediated decaging rescues TET activity. a.** Anti-hmC immunostaining and FACS analysis shows a significant rise in the 5hmC levels upon light irradiation that prove successful decaging of pcK and resumption of TET monoubiquitination, rescuing TET activity (error bars from N=2). **b.** Effect of Ub and VprBP mutants on pcTET activity. Co-transfecting HEK293T cells expressing the non-binding variant of VprBP (HA-VprBP RARA) shows a decrease in 5hmC formation kinetics relative to the HA-VprBP wt. Likewise, the inactive myc-Ub nc also shows a significant decrease in 5hmC formation after 24 h, relative to the myc-Ub wt, as seen from anti-hmC immunistaining and FACS analysis (error bars from N=2). WT/wt = wild-type; UT= untransfected.

After achieving successful incorporation of pcK in mTET2CD, as the light-mediated decaging resumes the monoubiquitination, the activity of TET was looked into. Initially, the TET activity was observed after 24 h of pcK decaging. Cells were transfected and cultivated in the presence of 0.32 mM of pcK, as described above, and irradiated with 365 nm light using an UV trans-

illuminator. After following the same workflow as described in section 8.2.4 for the light irradiation, cells were grown in full media lacking pcK to terminate expression of the photocaged TET, for same reasons as mentioned above in Chapter 6.3. FACS assay revealed that in the absence of light irradiation, the 5hmC levels were basal in cells expressing the mTET2CD and other constructs, as seen in inactive K1212N/E mutants, indicative of a catalytically inactive state of TET due to presence of the caging group that prevents the monoubiquitination. Upon light irradiation, a significant rise is seen in the 5hmC levels that prove successful decaging of pcK and resuming of TET monoubiquitination making the TET catalytically active again. However, again the recognition of the amber as a stop codon leading to the formation of truncated products led to a considerable reduction in the expression levels of caged mTET2CD, causing overall lower levels of 5hmC, compared to the non-caged variants (**Figure 31a, Figure S8b**).

As a proof that the TET activity seen here is indeed the outcome of its monoubiquitination, the effect of overexpressing mutants of the TET-monoubiquitination machinery on TET activity was tested. It was observed that when a mutant of VprBP (VprBP-RARA) that is defective in binding to DDB1 (Chapter 4.1) and hence is unable to monoubiquitinated TET, is expressed along with the caged mTET2CD and wt Ub, the activity of TET is significantly reduced as indicated by lower 5hmC levels 24 h after decaging (**Figure 31b)**. On the other hand, expressing the mutant Ub with a deletion of G76 that makes it unable to be added to substrates (myc-Ub nc) also lead to markedly lower levels of 5hmC compared to wt Ub (**Figure 31c**).

All these data indicate that indeed we are able to control TET monoubiquitination and activity with light. However, as we do not directly visualize the monoubiquitination event with the used FACS assay, this proof is rather indirect. And hence, there still remains the necessity to establish an assay with which the direct visualization of TET-monoubiquitination would be possible. In the next section, it is described how an immunoprecipitation followed by immunoblotting assay was tried to be established to allow for the direct study of TET-monoubiquitination and its kinetics.

## 7.4 Detection of TET-Ub

As stated before, an immunoprecipitation followed by immunoblotting workflow was the most suitable assay to study TET-monoubiquitination kinetics as it would allow for the direct visualization of the monoubiquitinated TET species. Immunoprecipitating TET with its Flag tag followed by immunoblotting for the myc tag of Ub or vice versa would enable us to specifically look at the overexpressed mTET2CD-Flag (caged or uncaged) that has been ubiquitinated with the overexpressed myc-Ub (**Figure 32**). As the first step to establish this

workflow, the uncaged mTET2CD was thought to be used for optimization which would later be extrapolated to the caged mTET2CD (pcTET) for ease of handling and avoid issues like low expression that are common for amber suppression.

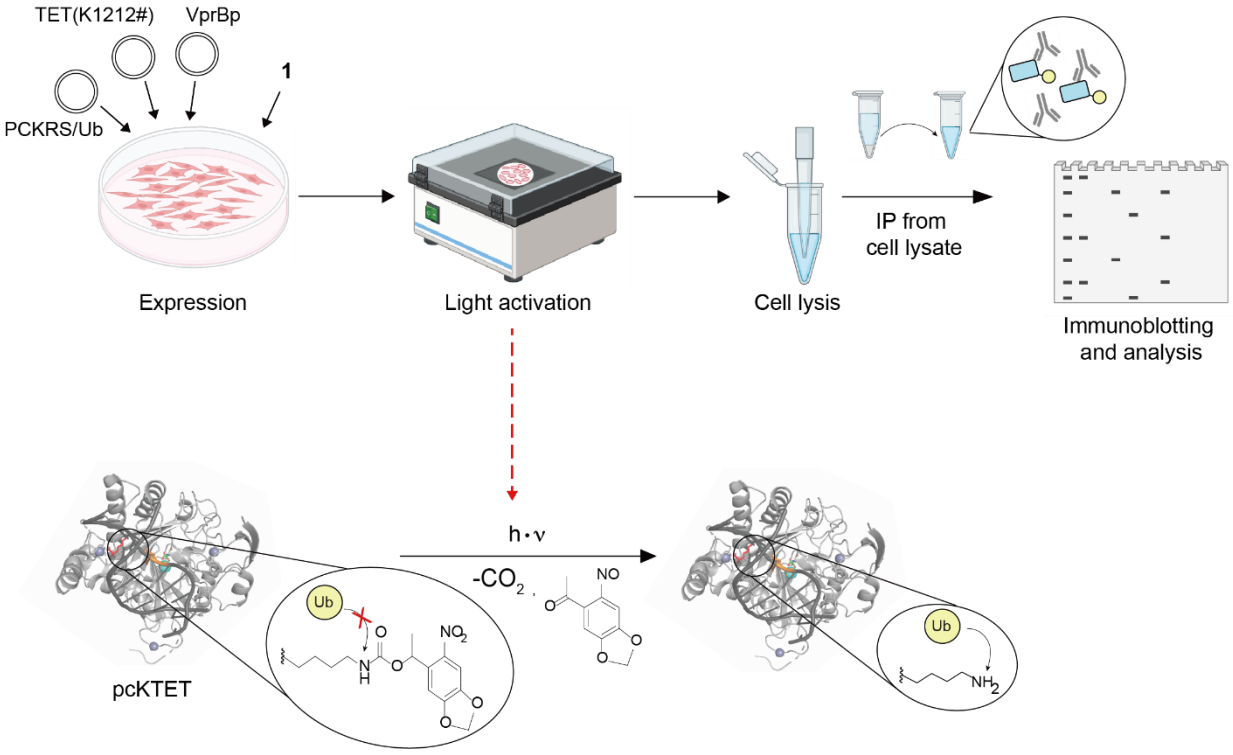



**Figure 32: Workflow for the detection of TET-monoubiquitination (TET-Ub).**

Although a fairly straightforward workflow, the results were not yielding the expected readout. The three constructs were expressed in HEK293T cells as described above for other experiments. After 24 h, cells were harvested following treatment with 10 µM MG132 for 5 h. Cells were lysed in denaturing conditions with 0.1 % SDS in an effort to release all the chromatin bound TET. Then from the whole cell lysates, either the total expressed mTET2CD-Flag was pulled down using an antibody specific for Flag tag or the total expressed and conjugated myc-Ub was pulled down using an antibody specific for myc tag. These immunoprecipitated proteomes were subjected to immunoblots using antibodies specific for myc tag or Flag tag. What was expected were two bands at around 120 kD and 130 kD, corresponding to non-ubiquitinated and mono-ubiquitinated TET respectively on the Flag immunoblots following Flag pulldown. But what was observed was only an intense band corresponding to the overexpressed, non-ubiquitinated TET at 120 kD. On the myc

immunoblots on the other hand, the expected outcome would have been a single band at around 130 kD corresponding to the mono-ubiquitinated TET. But what was observed were smears that are expected for polyubiquitinated protein samples (**Figure S9**). Upon coming across such shortcomings, several steps of the experimental workflow were attempted to be optimized. For example, other cell types like U2OS and HeLa were tested; other TET isotypes like the human TET1 and mouse TET1 were tested (according to (Arroyo et al., 2022)); different stringencies in the lysis methods were tried and each step of the immunoprecipitation and immunoblot were also optimized, starting from DNA amounts of each construct used for transfection to different incubation times and temperatures during pulldown. Although technically the experimental workflow of immunoprecipitation followed by immunoblotting worked (as seen with control constructs, data not shown), for some reason the TET-monoubiquitination could not be directly observed.

## 7.5 Conclusion

In conclusion, in the effort to develop light control over the monoubiquitination of TET to be able to elucidate the kinetics of this PTM, final success could not be achieved. Nonetheless, it was successfully shown that TET could be caged at the conserved Lys residue (K1212 in mTET2) that undergoes monoubiquitination by successful and site-specific incorporation of pcK. Furthermore, it was also proved that indeed the caging of TET at the monoubiquitination site prevents the addition of Ub and in turn, leads to the inactivation of its dioxygenase activity that oxidizes 5mC to 5hmC, as shown by immunostaining based FACS analysis. And the activity was recovered when irradiated with light (365 nm) as the bulky caging group falls off, exposing the Lys at 1212 position to be added with Ub that aids in the DNA binding strength of TET leading to activation. It would be of high interest to get to study the kinetics of such a monoubiquitination event, given the tool required for the same being established. But even after all the groundwork being laid, a direct proof of TET-ubiquitination remained to be shown, possibly due to incompatibility with the cell line used or some other unknown biological factor. Nevertheless, this effort shows a way to use such light activatable ncAAs to control single ubiquitination events and this principle could be applied to study kinetics of such PTM events.

# 8. Summary and Outlook

The present work confers light control to Ub molecules and allows temporal control over Ub chain extension in mammalian cells by irradiating them with light.

The first part of the study demonstrated the construction of a novel tool to perturb cellular ubiquitination with the ability to control linkage-specific ubiquitination with high temporal resolution. Different Ub variants showed characteristic polyubiquitination patterns in whole cell lysates of treated HEK293T cells. Ub variants with single Lys at specific positions differ in such smear patterns depending on the position, and hence proving the proteome-wide abundance of polyUb chains generated from each position. K48 polyUb chains generating the most intense smears, followed by K63 and K11. Genetic code expansion enabled the incorporation of photocaged lysine (pcK) via an engineered PylRS/tRNAPyl pair into myc-tagged Ub constructs at the single Lys sites and are then employed by the UPS to synthesize a trackable sub-population of the cellular ubiquitome. Irradiation with light triggers ubiquitination from the respective Lys following decaging, enabling the monitoring of *de novo* ubiquitination in the proteome initiated by these specific linkage types. Decaging of pcK following light activation allowed synthesis of K11, K48 and K63 linked polyUb chains, indicated by an increase in polyubiquitination signal, which was proved to not be a mere effect of light activation or treatment conditions. Time-course analyses of proteome-wide polyubiquitination patterns revealed burst-like and rapid kinetics of synthesis for all the linkage types. For each linkage type, a maximum of ubiquitination intensity was reached already after 15 minutes. Furthermore, inhibition of individual components of the UPS revealed the impact of these components on the synthesis of K48-initiated Ub chains on the low minutes time scale. Treatment with the E1 inhibitor, MLN7243 (TAK-243), showed that even after treating cells with standard inhibitory conditions, there are still low levels of active E1 and/or activated Ub still available. The NAE inhibitor, MLN4924, inhibits NEDDylation of CRL E3 ligases, keeping them in an inactive state and hence, inhibits Ub ligation by them. Treating cells with MLN4924 showed that the global K48-linked polyUb kinetics did not change even after a 4 h preincubation with 1 µM compound, suggesting that the majority of the observed myc-Ub chain extension initiated at K48 occurs independently of CRL activity. Finally, p97 inhibitor NMS873, that inhibits its interaction with substrates ubiquitinated with K48-linked chains and hence their unfolding and proteasome-mediated degradation by ERAD, proved the effect of VCP/p97 as an activator or helper of K48-linked *de novo* ubiquitome synthesis on a rapid timescale as it showed a marked increase of K48-linked myc-Ub proteome levels in the absence of MG132.

To summarize, the ability of this approach to selectively zoom into linkage-types at a high temporal resolution makes it suitable as a UPS perturbation strategy with added new abilities

compared to the current toolset. The rather straightforward workflow also makes it compatible with diverse downstream analysis techniques. Given the central role of Ub in regulating protein function, localization and quality control, in a vast number of normal and disease-related processes, this approach is highly anticipated to find wide use for studying the global, proteome-wide dynamics of linkage-specific ubiquitination.

The second part of the study demonstrated efforts to look at a protein-specific ubiquitination event in a time-resolved manner. As TETs, the DNA demethylating enzymes, have been found to be monoubiquitinated at a conserved Lys, this approach was aimed at elucidating their kinetics utilising a similar strategy as above. Genetic code expansion was once again employed to incorporate a pcK via an engineered PylRS/tRNAPyl pair into TET constructs at the conserved Lys. Irradiation with light triggers monoubiquitination at the Lys following decaging, enabling the monitoring of TET activity as well as its *de novo* ubiquitination. Although it was proved that the caging of TET at the monoubiquitination site prevents the addition of Ub leading to the inactivation of its dioxygenase activity, which was subsequently recovered when irradiated with light, final success in elucidating the kinetics of this PTM could not be achieved due to the inability in detecting ubiquitinated TET species. Nevertheless, this approach showed a way to utilize the strategy of light activatable ncAAs to control single ubiquitination events and this principle holds the capacity to be successfully applied in kinetics studies of such PTM events.

# 9. Materials and Methods

## 9.1 General Information

**Synthesis of Oligonucleotides**

The oligonucleotides listed in Table S1 were synthesized by Merck KGaA (Darmstadt, Germany). The desalted oligonucleotides were stored as 100 μM stock in TE buffer.

**Sanger Sequencing**

The sequence of constructed plasmids was routinely checked by Sanger sequencing by Microsynth Seqlab GmbH (Göttingen, Germany).

**Purification of Plasmid DNA and Double Stranded Oligonucleotides**

Plasmid DNAs were isolated from bacteria strains via silica column purification using NucleoSpin® Plasmid EasyPure kit (MachereyNagel, Düren, Germany) following the manufacturer's instruction. Large scale purification of endotoxin-free plasmids for mammalian cell transfection was performed with NucleoBond® Xtra Maxi kit (MachereyNagel). Double stranded oligonucleotides amplified by PCR or digested from plasmids were identified or separated by DNA electrophoresis (discussed below), then the crude reaction mixture or agarose gels containing desired products were purified using NucleoSpin® Gel and PCR Clean-Up kit (MachereyNagel).

For DNA electrophoresis, 1% (w/v) of agarose gel in 0.5× TBE buffer was prepared. The DNA samples were resolved using 8 – 12 V/cm. The agarose gels were stained with 0.5 μg/mL ethidium bromide, destained with water, and visualized with UV fluorescence.

**Enzymes**

The enzymes used in in this study that are not indicated in the methods are listed below. Corresponding experiments were conducted following the supplier's instructions.

**Table 1. List of Enzymes**

| Enzyme | Company |
|---|---|
| NotI | NEB (New England Biolabs) |
| DpnI | NEB |
| SalI | NEB |
| Q5 HiFi DNA polymerase | NEB |
| Phusion DNA polymerase | NEB |
| KOD hot start DNA polymerase | Merck Millipore |

| | |
|---|---|
| T5-Exonuclease | NEB |
| Taq-DNA ligase | NEB |
| T4-DNA ligase | NEB |

## Bacteria Strains

**Table 2. *E. coli* strains used in this study.**

| Strain | Genotype | Supplier |
|---|---|---|
| DH10B (Top10™) | F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ- | Invitrogen™ (Thermo Fisher Scientific) |

## Mammalian Cell Lines

**Table 3. Mammalian cell lines used in this study.**

| Cell Line | Origin | Supplier |
|---|---|---|
| HEK293T | Human Embryonic Kidney | MPI Dortmund |
| HeLa | Human Cervical Cancer | MPI Dortmund |
| U2OS | Human Osteosarcoma | Sigma Aldrich |

## Chemicals

The chemicals used in this study that are not listed in the methods are listed below.

**Table 4. List of chemicals.**

| Name | CAS No. | Supplier |
|---|---|---|
| 2-Amino-2-(hydroxymethyl)propane-1,3-diol (Tris), buffer grade | 77-86-1 | Carl Roth |
| 2-Log DNA ladder | | NEB |
| Acetic acid | 64-19-7 | Carl Roth |

| | | |
|---|---|---|
| Agarose LE, molecular biology grade | 9012-36-6 | Biozym Scientific |
| Ammonium acetate | 631-61-8 | Carl Roth |
| Bovine serum albumine (BSA) | 9048-46-8 | Cell Signaling Technology |
| Carbenicillin, disodium salt | 4800-94-6 | Carl Roth |
| Dimethyl sulfoxide (>= 99.7%) | 67-68-5 | Merck |
| dNTPs | | NEB |
| Ethanol, absolute | 64-17-5 | Merck |
| Ethidium bromide | 1239-45-8 | Carl Roth |
| Ethylenediaminetetraacetate (EDTA) | 6381-92-6 | Alfa Aeser |
| Formaldehyde, 37 wt. % in H2O, contains 10-15% Methanol as stabilizer | 50-00-0 | Merck |
| Hydrochloric acid, 37% | 7647-01-1 | Merck |
| Isopropanol | 67-63-0 | Fisher Scientific |
| Opti-MEM® I | | Gibco |
| Sodium chloride | 7647-14-5 | Merck |
| Sodium hydroxide | 1310-73-2 | Merck |
| Triton® X-100 | 9002-93-1 | Fluka Chemika |
| Tween® 20 | 9005-64-5 | Fisher Bioreagents |

## Software

The software used in in this study that are not indicated in the methods are listed below.

**Table 5. List of software.**

| **Name** | **Company** |
|---|---|
| SnapGene v4.3 | Dotmatics |
| NanoDrop 2000 v1.6 | Thermo Fisher Scientific |
| BioDoc Analyze v2.1 | Biometra |
| Adobe Illustrator 2022 v26.4 | Adobe |
| PyMol v2.5.2 | Schrödinger LLC |
| ChemDraw Professional v21.0.28 | PerkinElmer |
| Office 365 | Microsoft |

## 9.2 Methods

### 9.2.1 Generation of Constructs

Firstly, in order for the incorporation of pcK, a pEVOL (Young et al., 2010) based plasmid (p787) with a *M. mazei* PylRS construct, containing the same mutations as reported (Gautier et al., 2010), was constructed. p787 was digested with XbaI and BamHI and ligated into a similarly digested vector backbone obtained from the lab of Prof. Dr. Heinz Neumann, containing four repeats of $tRNA^{Pyl}$ under U6 promoter (pASB654 - SE323_wtPylRS_4xPylT, p2970) to generate the modified PCKRS plasmid (p2946).
A pcDNA3-based Ub plasmid (pcDNA3_HA-Ub, p2972) was constructed by Gibson assembly to insert one copy of Ub with an N-terminal myc-tag (amplified with o4324 and o4327) into plasmid p2932 (amplified with o4325 and o4326) to yield p2945. Quickchange mutagenesis was performed on p2945 to generate the Ub nc variant p2950 (using primers o4421 and o4422). The Ub K0 variant was cloned into the backbone of p2945 using Gibson assembly by first amplifying Ub K0 from p1.9T7 minus1 with primers o4948 and o4949 and amplifying p2945 with primers o4946 and o4947, yielding p3041. The Ub K11, K48 and K63 constructs and Ub $K11^{TAG}$, $K48^{TAG}$ and $K63^{TAG}$ constructs were all obtained by performing Quickchange mutagenesis on p3041. The Ub K11 was generated by Quickchange mutagenesis with o5278 and o5279 to generate p3455, for Ub $K11^{TAG}$, o5280 and o5281 were used to generate p3456, for Ub K48, o5226 and o5227 were used to generate p3439, for Ub $K48^{TAG}$, o5228 and o5229 were used to generate p3440, for Ub K63, o5274 and o5275 were used to generate p3457, and lastly for Ub $K63^{TAG}$, o5276 and o5277 were used to generate p3458.

For the photocaging of K1212 of mTET2CD, firstly a pcDNA3 based mTET2 construct, previously generated in the lab, having the mTET2CD with a C-terminal fusion of HA-tag (p2392) was used to generate K1212E (p2927) and K1212N (p2928) mutants through Quickchange mutagenesis using o4160 and o4161 for KE mutation and o4162 and o4163 for KN mutation. This construct was further used to generate the $K1212^{TAG}$ amber mutant of mTET2CD (p2932), using Quickchange primers o4196 and o4197, still with a C-terminal HA-tag. Later, this TET construct of p2932 was inserted into the backbone of p2945 (amplified by o4336 and o4337) by Gibson assembly after being amplified using primers o4335 and o4338, replacing the *M. mazei* PylRS but retaining the four repeats of $tRNA^{Pyl}$ under U6 promoter, in order to increase the transfection efficiency in co-transfections with the pcKRS plasmid, and also to ensure greater amber suppression and incorporation efficiency of 1. This also brought the TET gene under the EF1a promoter, instead of the previously used CMV. The resulting plasmid was named p2947, carrying the mTET2CD with $K1212^{TAG}$ amber mutation under EF1a promoter and co-expressing 4X $tRNA^{Pyl}$ under U6 promoter. Next, a 3X-Flag tag was

cloned into p2947 using primers o4651 and o4652, replacing the HA-tag at the C-terminus of mTET2CD, yielding the plasmid p2962. And finally, the K1212$^{TAG}$ mutation in p2962 was reverted back to lysine using Quickchange primers o4653 and o4654 to yield the wildtype construct p2963 used in all subsequent studies. The HA-VprBP construct was obtained from Addgene (# 133586) and the non-binding variant (VprBP-RARA) was generated by quickchange mutations on the wildtype construct using primers o4657 and o4658, yielding p2965. The myc-Ub construct under the CMV promoter from p2945 was fused within the PCKRS p2946 construct in order to reduce the number of plasmids for transfection in co-transfection experiments with photocaged TET, VprBP and Ub. The backbone (p2946) was double digested with NotI and SalI and the myc-Ub under CMV promoter from p2945 was amplified with primers o4511 and o4417 and ligated together by Gibson assembly, resulting in p2956. In order to clone similar constructs but with Ub nc, a similar approach was taken. After double digesting p2946 as above, p2950 was amplified with primers o4511 and o4595 and assembled into the backbone by Gibson assembly to yield p2958.

### 9.2.2 Cell Culture and Transfection

Cells were cultivated in DMEM (Dulbecco's Modified Eagle Medium, w/ 4.5 g/L Glucose, w/o: L-Glutamine, w: Sodium pyruvate, w: 3.7 g/L NaHCO3, PAN Biotech, # P04-03600), supplemented with 10% FBS (South America origin, premium grade, PAN Biotech, # P30-3306), 2 mM L-Glutamine (PAN Biotech, # P04-80100), 100 U/mL Penicillin and 0.1 mg/mL Streptomycin (PAN Biotech, # P06-07100) in a sterile, humidified incubator (≥95%) at 37ºC and a $CO_2$ level of 5 %.

For transfection, cells were seeded in 6-well cell culture plates (Sarstedt), a day before so that they reached 70-80% confluency at the time of transfection. Transient plasmid transfection was performed using Lipofectamine-2000 (Thermo Fisher Scientific, # 11668019) following the supplier's guidelines. For transfection, 1 µg of DNA was used in case of non-amber containing constructs and a total of 2 µg of DNA was used in case of the amber containing constructs (except for experiment in Fig. S4 where cells were cultivated in 10 cm dishes and a total of 12 µg DNA was used for transfection).

### 9.2.3 Light Activation

HEK293T cells were grown in 6-well cell culture plates (Sarstedt) transfected with plasmids containing amber Ub variants (Ub K11$^{TAG}$, K48$^{TAG}$ and K63$^{TAG}$) and p2946 (modified PCKRS) (Figure S1. Data from Fig. S4 are from cultivation in 10 cm dishes) as described under 4.2. After adding the transfection mixture, 0.32 mM of pcK (synthesized as describe above or

purchased from Sigma Aldrich, # 915793) was added directly in the growth media and the cells were allowed to express the protein for 24 h. For treatment with light (Palei et al., 2020), the growth medium containing pcK was exchanged with warm DPBS (Dulbecco's Phosphate-Buffered Saline, Mg/Ca free, PAN Biotech, #P04-36500) and placed on a 365 nm UV-transilluminator (Witeg DH. WUV00010, 6x 15 W) for 4 minutes (decaging of this and related noncanonical amino acids is also possible with 405 nm light and higher temporal resolution using suited light sources such as LEDs). Immediately after the light treatment, the DPBS was exchanged and pre-warmed medium (without pcK, but with 25 µM MG132; Enzo Life Sciences, # BML-PI102) was added and the cells were further grown in a humidified incubator (≥95%) at 37ºC and a $CO_2$ level of 5 %.). Cells were harvested at indicated time points.

### 9.2.4 Cell Lysis and Immunoblotting

Cells were trypsinised and collected in microcentrifuge tubes (Sarstedt). The cell pellets were lysed with 200 µL of lysis buffer (0.5% IGEPAL, 50 mM Tris (pH=7.5), 150 mM NaCl, 10% glycerol) containing protease inhibitor (Roche complete™, 11697498001), 0.1 mM PMSF and in addition, 20 mM NEM (Thermo Fisher Scientific, 23030) and 2 mM EDTA for inhibiting DUB activity. After completely dissolving the pellets with lysis buffer, the samples were incubated at 4ºC for 60 min, shaking at 700 rpm. Crude lysates were cleared by centrifugation at full speed at 4ºC for 15 min. The lysates were collected and protein concentrations were determined using Bicinchoninic Acid (BCA) protein assay. Samples were prepared using 4X Laemmli sample buffer and loaded on 12% SDS gels.
Proteins were transferred to 0.2 µm polyvinylidene fluoride (PVDF) membranes using a Trans-Blot Turbo system (1.9 A, 25 V, 7 min; Bio-Rad, # 1704150) and transfer kit (Bio-Rad, # 1704272). The membranes were blocked with 5% (m/v) nonfat milk in TBS-T (TBS + 0.1% v/v Tween-20) buffer and incubated with indicated primary antibodies (rabbit monoclonal anti-myc 71D10, 1:1000 dilution, Cell Signaling, # 2278; rabbit monoclonal anti-ubiquitin E412J, 1:1000 dilution, Cell Signaling, # 43124; mouse anti-β-Tubulin, 1:2000 dilution, Cell Signaling, # 86298) overnight at 4°C with gentle shaking. The membranes were then incubated with the respective secondary antibody (HRP linked rabbit IgG, 1:5000, Merck, # GENA934; Dylight680-conjugated goat anti-mouse secondary antibody, 1:10000 dilution, Invitrogen, # 10797775). The chemiluminescent reaction was initiated using a Clarity Western ECL substrate and enhancer (Bio-Rad # 1705061; # 1705062) and chemiluminescent images were taken on a Bio-Rad ChemiDoc™ imaging system. Fluorescent images of the blots were taken on the Odyssey® DLx imaging system (LI-COR). Typically, the chemiluminescent staining and imaging were done first, followed by the fluorescent staining and imaging of the β-Tubulin loading control. The myc or Ub smears from the individual images of the immunoblots and the

β-Tubulin bands were quantified using Fiji Image J software (Schneider et al., 2012; Rueden et al., 2017).

### 9.2.5 Immunostaining and Flow Cytometry

For analysis of expression of the different non-amber Ub variants, HEK293T cells expressing the respective constructs were harvested 24 h after transfection (adding 10 µM MG132 5 h before harvesting). Cells were washed with DPBS, trypsinized with Trypsin 0.05% / EDTA 0.02% (PAN Biotech, # P10-038100) for 5 minutes at 37ºC and blocked with full DMEM medium. Harvested cells were washed once with DPBS and fixed with medium A (Fix & Perm kit, Thermo Fisher Scientific, # GAS004) for 15 min at RT followed by washing with wash buffer (PBS + 5% FBS + 0.1% $NaN_3$). Then the fixed cells were permeabilized with medium B for 20 min. Thereafter, the cells were blocked with blocking buffer (PBS+1% BSA + 0.05% Tween) overnight at 4°C with gentle shaking. For immunostaining, mouse monoclonal anti-myc (9E10) (Santa Cruz Biotechnology, # sc-40) primary antibody was added to the cells in 1:500 dilution followed by incubation at RT for 1 h. After washing with PBST buffer, Alexa Fluor 488 conjugated goat anti mouse (Thermo Scientific, # A-11001) secondary antibody was incubated in 1:1000 dilutions for 1 h. Cells were then washed with PBST and PBS buffer before being measured by Flow Cytometry (Sony, # LE-SH800SFP). FCM results were analysed using Cell Sorter Software (Sony). Untransfected cells (UT) were used as a control.

### 9.2.6 Cell Lysis and Immunoprecipitation

In the attempt to visualise monoubiquitination on mTET2CD, an immunoprecipitation followed by immunoblot approach was adapted. For this, HEK293T cells expressing the respective TET-Flag constructs along with HA-VprBP and myc-Ub, were harvested as above after respective times, with or without light treatment. The cell pellets were lysed with 200 µL of lysis buffer (see 1.2.5) additionally containing 1% SDS (Ultra-pure, Carl Roth, CAS No. 151-21-3). The solutions were boiled at 95°C for 5 minutes and following this, were diluted 10X with NP-40 lysis buffer (without SDS). The sticky pellets were broken down and dissolved by passing the solutions through 1.20 x 40 mm needles (B. Braun Sterican®) attached to 5ml syringes (B. Braun Omnifix®), until dissolved. The crude lysates were cleared by centrifugation at full speed for 15 mins at 4 degrees, and supernatants were incubated with respective antibodies (mouse monoclonal anti-myc (9E10), Santa Cruz Biotechnology, # sc-40; mouse monoclonal anti-FLAG® M2 antibody, Merck, # F1804) or anti-FLAG® M2 magnetic beads (Merck, # M8823) and precipitated with Dynabeads®-ProteinG complex (Thermo Fisher Scientific, #10004D). The immunocomplexes were washed thrice with wash

buffer (0.1% Triton X-100, 10% glycerol in TBS) and eluted in 1X 4X Laemmli sample buffer (boiling for 5 mins) and loaded on to SDS-PAGE gels. Immunoblots were then performed as described above.

# A. Appendix

## A.1 Supplementary Figures

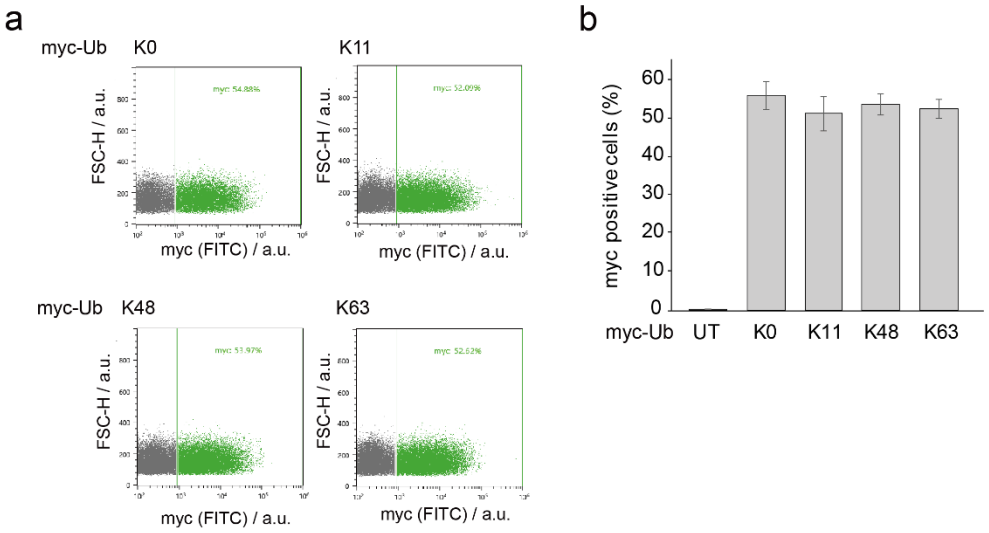



**Figure S1: Representative FACS density plots and alternate analyses showing the similar expression levels of myc Ub K0 and other single lysine Ub variants in HEK293T cells. a.** The myc positive cell population of HEK293T cells expressing myc Ub K0, K11, K48 and K63 for 24 hours, were analyzed by FCM. Green lines indicate the intensity thresholds determined by the fluorescence intensity of untransfected cells. **b.** Bar diagrams showing the mean of the percentage of myc positive cells. Error bars from N=3 biological replicates. UT= untransfected cells. a.u. = arbitrary units. Figure has been adapted from (Banerjee et al., 2024).

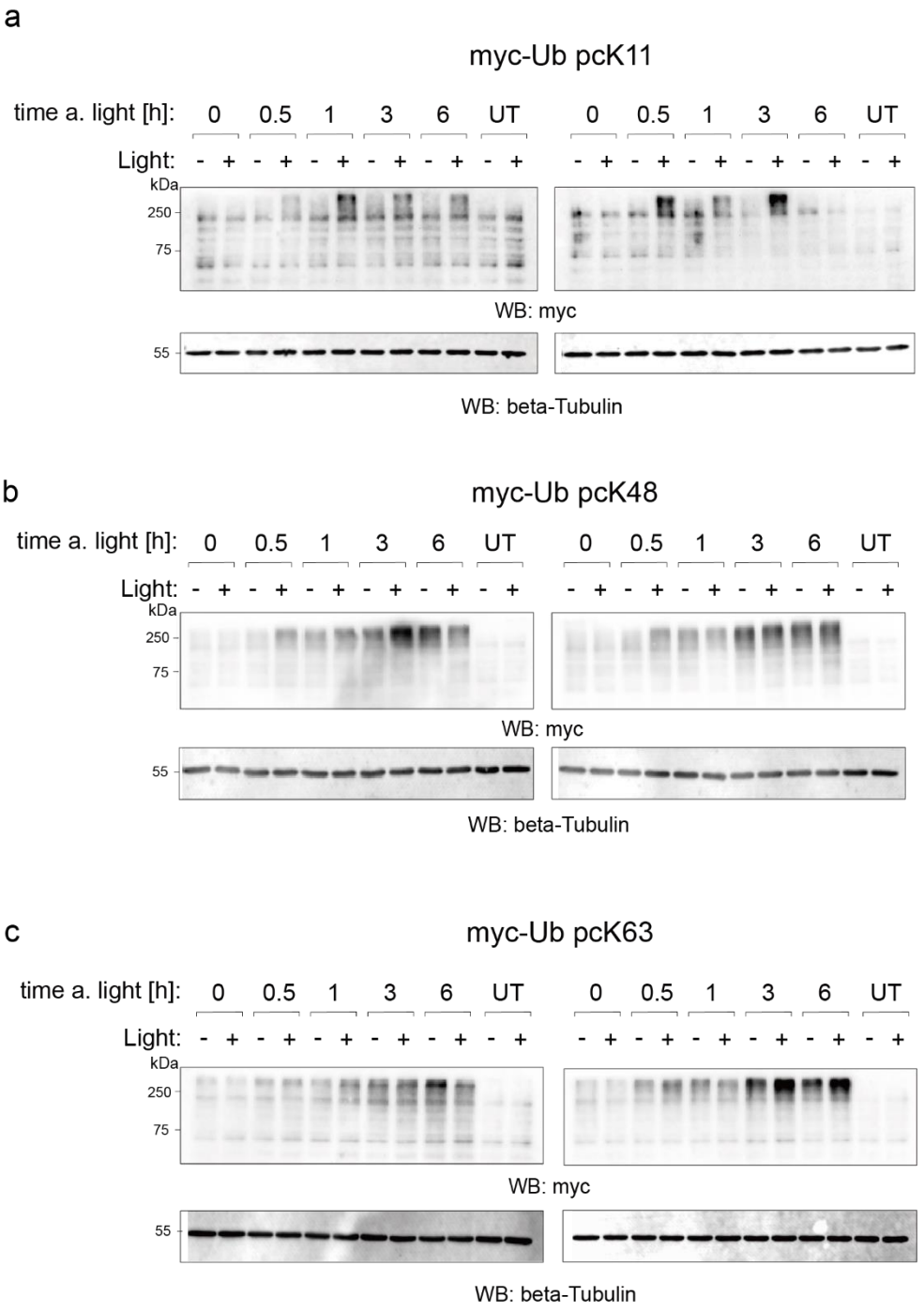



**Figure S2: Additional SDS PAGE/anti-myc blots for the analyses of long-term, linkage-specific ubiquitination kinetics after light activation of caged Ub variants. a.** myc-Ub pcK11; **b.** myc-Ub pcK48; **c.** myc-Ub pcK63. Figure has been adapted from (Banerjee et al., 2024).

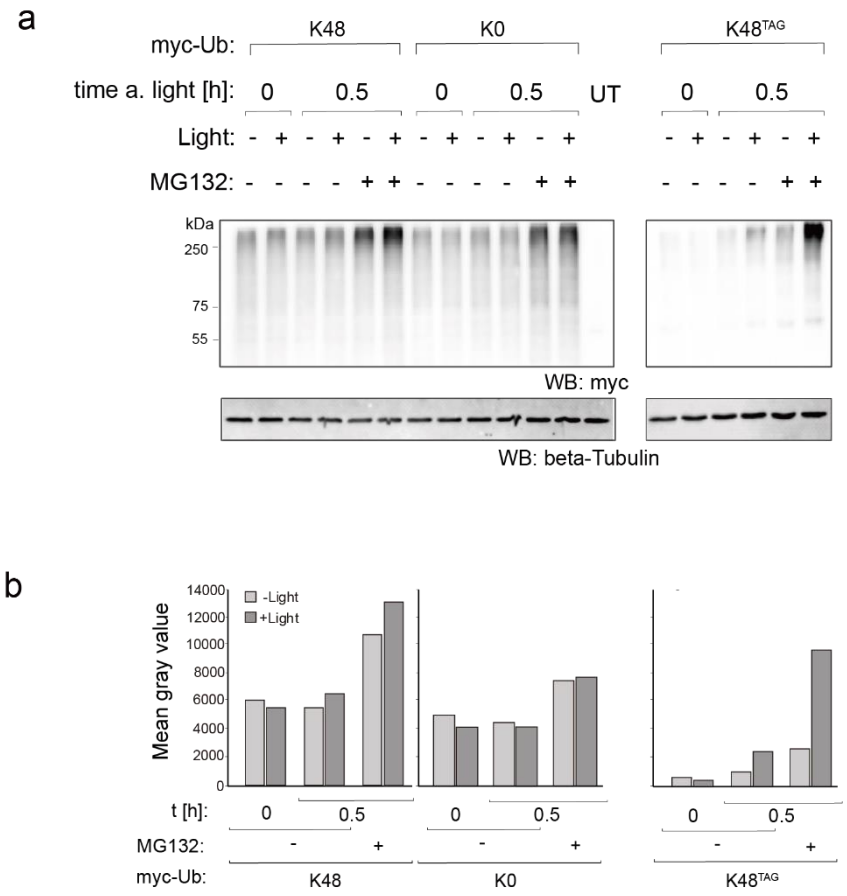



**Figure S3: Light alone or MG132 do not have an effect on the increase in high molecular weight myc-Ub in the proteome. a.** HEK293T cells expressing myc-Ub K0, K48 or pcK48 (K48TAG in presence of pcKRS and pcK) for 24 hours were irradiated or not irradiated with light (365 nm, 4 min), treated or not treated with 25 μM MG132, and harvested immediately or after 30 minutes. Only myc-Ub pcK48 showed an increase in the high molecular weight myc-ubiquitinome upon light irradiation, proving that light or MG132 alone are not responsible for this increase. **b.** Bar diagrams showing the mean gray values (arbitrary units) corresponding to the myc-ubiquitinated proteome, normalized to the loading control (beta-Tubulin). Figure has been adapted from (Banerjee et al., 2024).

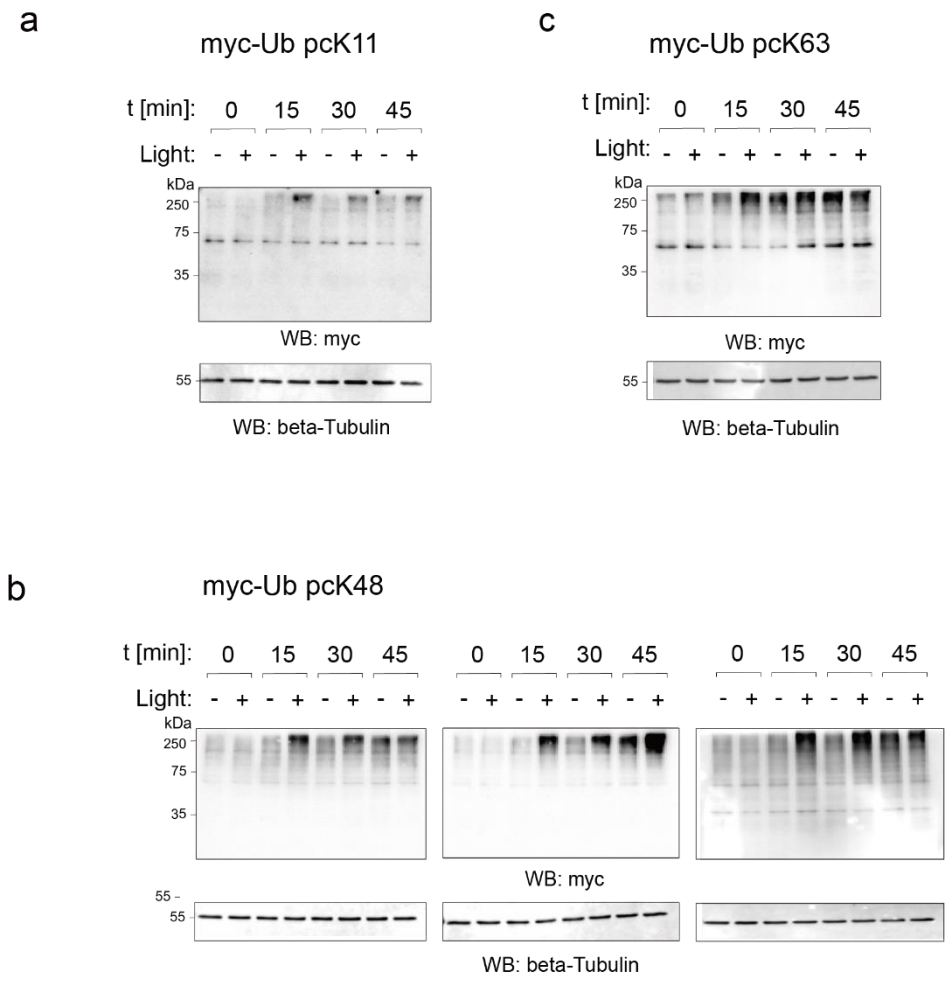



**Figure S4: Additional SDS PAGE/anti-myc blots for the analyses of short-term, linkage-specific ubiquitination kinetics after light activation of caged Ub variants. a.** myc-Ub pcK11; **b.** myc-Ub pcK48; **c.** myc-Ub pcK63. Figure has been adapted from (Banerjee et al., 2024).

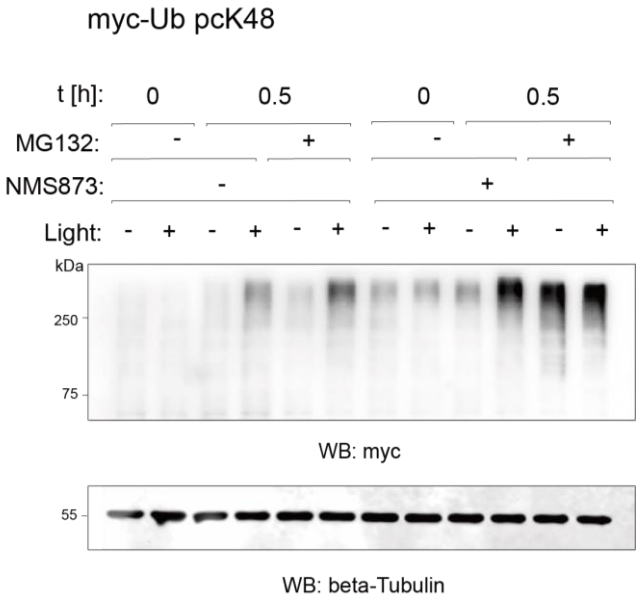



**Figure S5: Additional SDS PAGE/anti-myc blot for the analyses of effect of NMS873 on early, K48-specific *de novo* ubiquitome synthesis (t=0.5 h) in presence of MG132.** Figure has been adapted from (Banerjee et al., 2024).

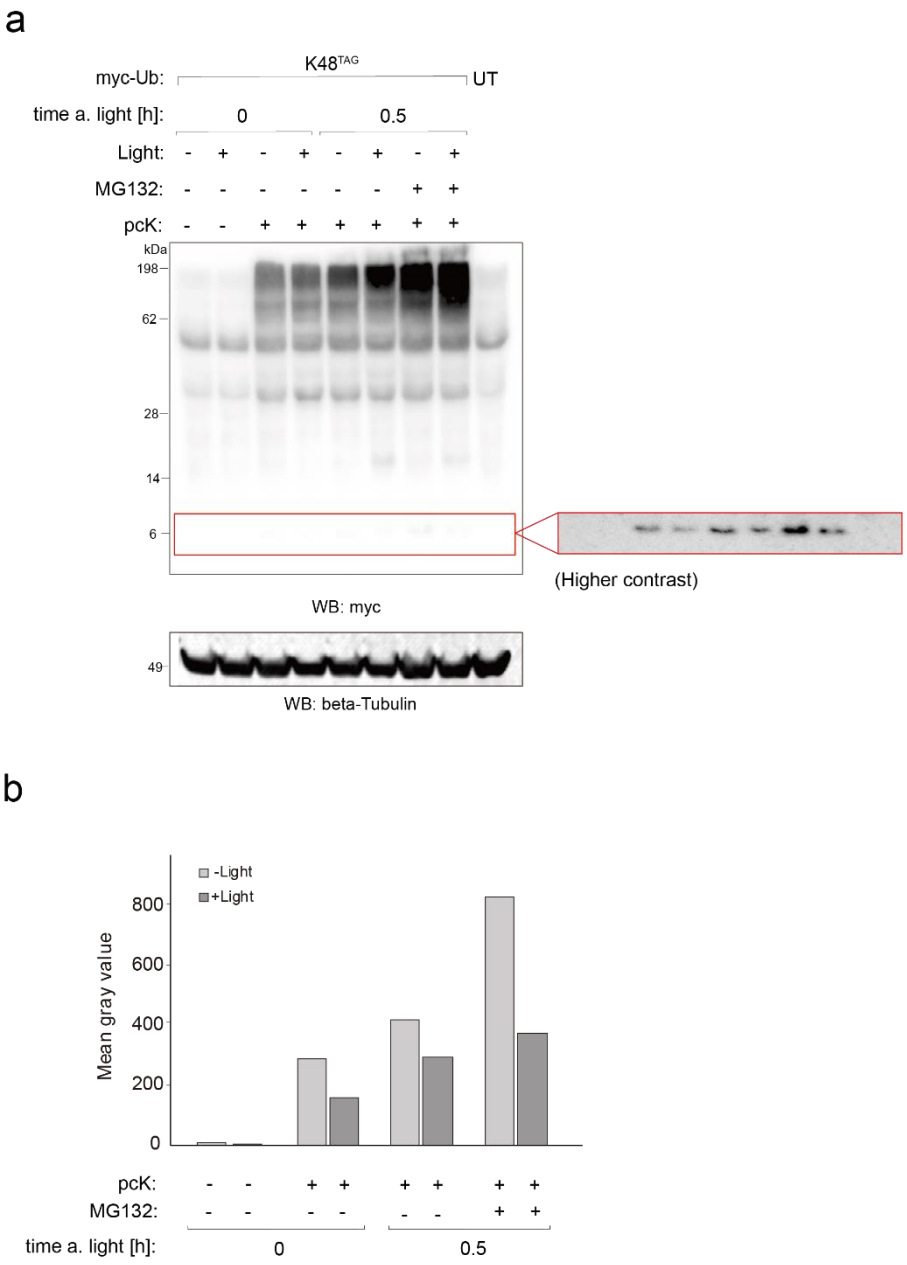



**Figure S6: Changes in mono myc-Ub pool in HEK293T cells expressing myc-Ub pcK48 in presence or absence of pcK, with or without light irradiation. a.** HEK293T cells grown in 10 cm plates, co-transfected with myc-Ub pcK48 and pcKRS (a total of 12 µg DNA) and grown in the absence of pcK do not show mono myc-Ub expression. Only when grown in presence of pcK, expression of myc-Ub pcK48 is observed (compare lanes 1-2 with lanes 3-8 in high contrast panel). In order to visualize mono-Ub, samples were run in 4-12% gradient Bis-Tris gels, transferred to 0.2 µm PVDF membranes and stained. **b.** Bar diagrams showing the mean gray values corresponding to the mono myc-ub, that also quantitatively shows changes in this pool upon light irradiation. Values are normalized to the loading control (beta-Tubulin). UT= untransfected cells. Figure has been adapted from (Banerjee et al., 2024).

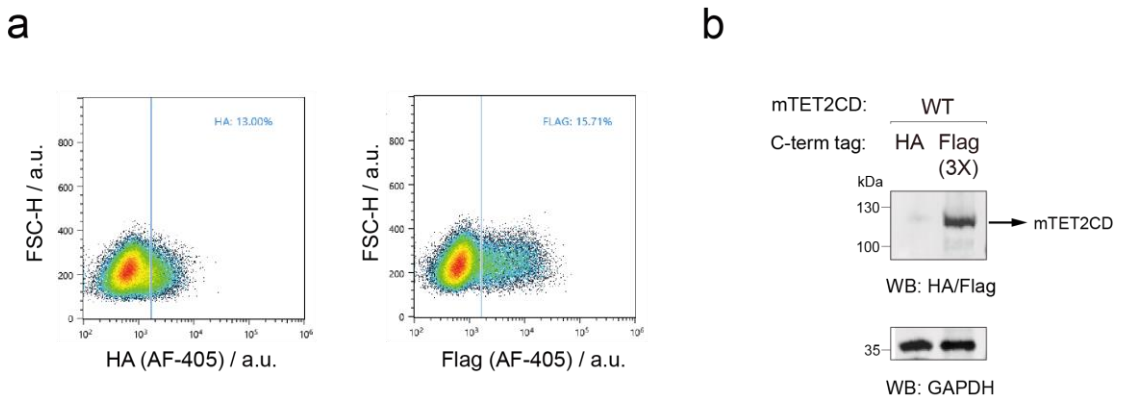



**Figure S7: Comparison of staining efficiency of HA tag and 3X-Flag tag. a.** Representative density plots from FACS analysis of anti-HA and anti-Flag immunostainings. **b.** Immunoblots to compare HA and Flag staining. a.u.= arbitrary units.

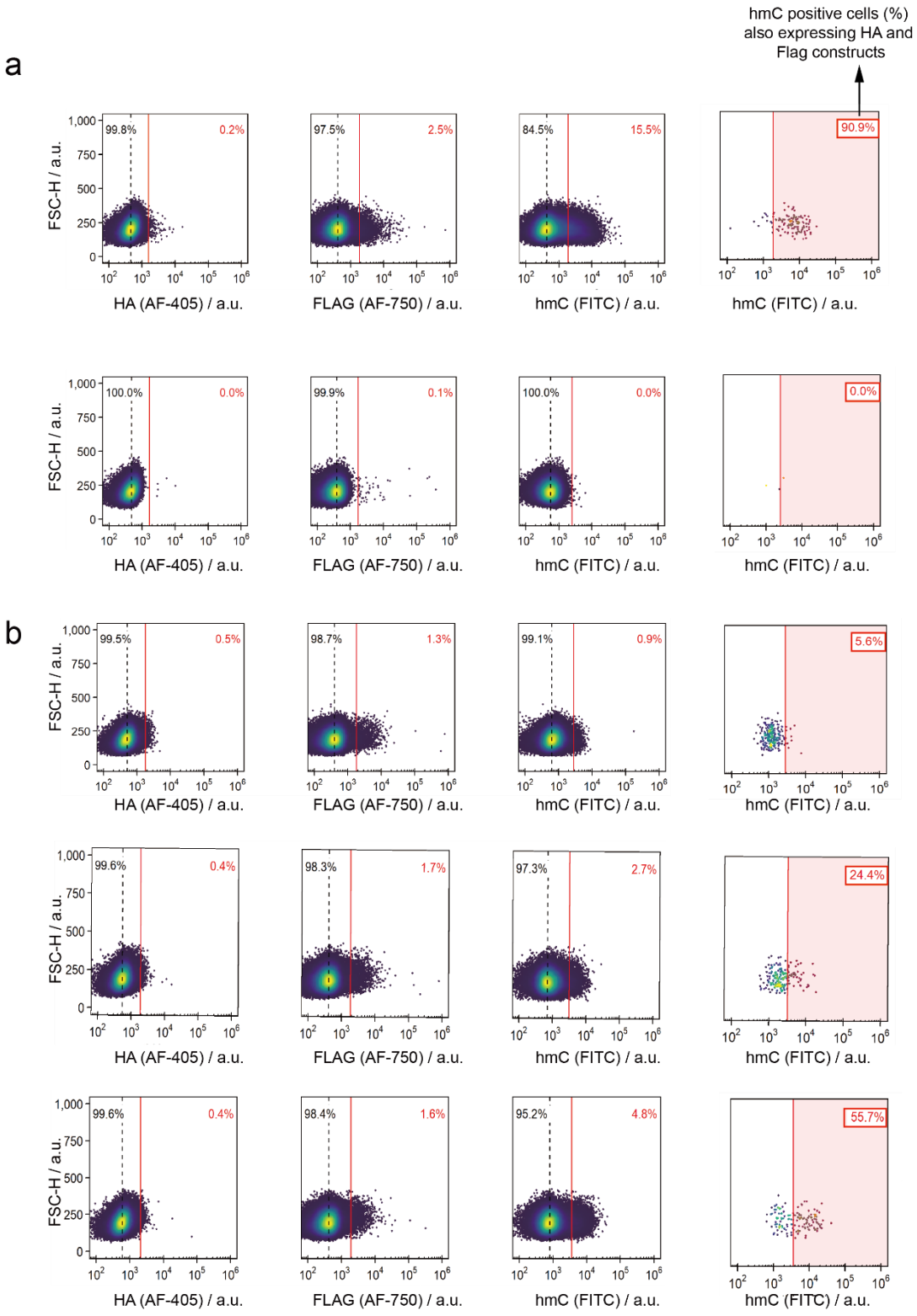



**Figure S8: Representative density blots from FACS analysis**, corresponding to **a.** Figure 30c. Wild-type cells (upper pane) show 91% hmC positive cells, compared to 0% as shown by untransfected cells (lower pane). **b.** Figure 31a. Light irradiated cells (lower pane) shows increase in hmC positive cells compared to those grown without pcK (upper pane) or the ones not irradiated with light (middle pane). a.u.= arbitrary units.

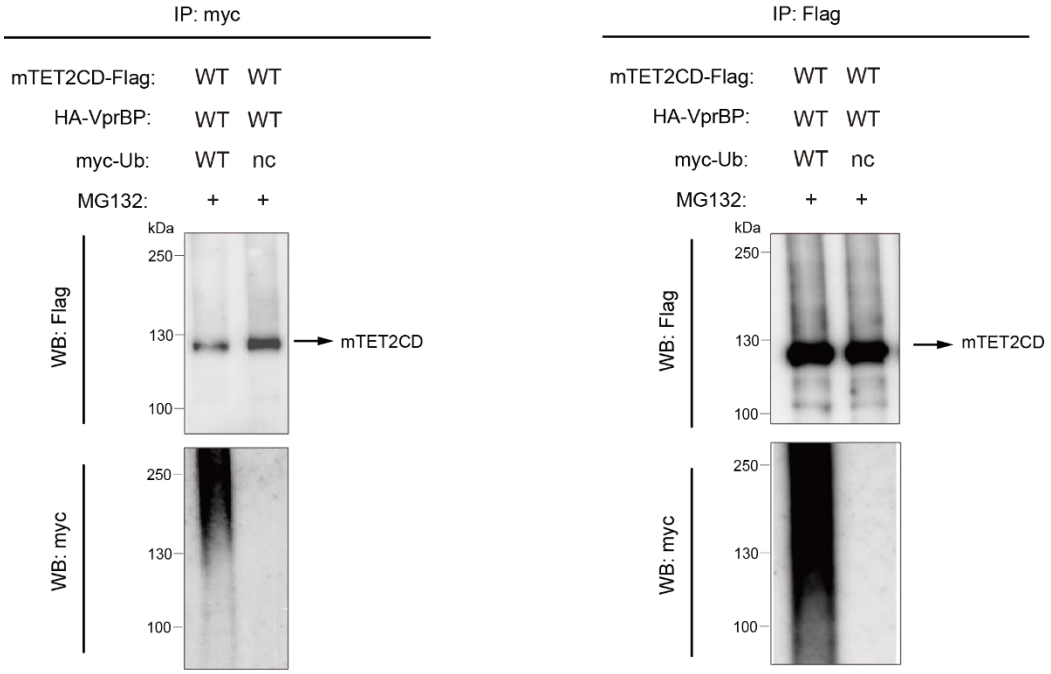



**Figure S9: Anti-Flag and anti-myc immunoblots of cells immunoprecipitated with anti-myc and anti-Flag antibodies**. No band corresponding to TET-Ub could be detected on any of the blots.

## A.2 Supplementary Tables

**Table S1: Oligonucleotides for plasmids construction.**

| Name | Sequence (5'→3') |
|---|---|
| o4160_ShP | CTATAATGGATGTGAGTTTGCCAGAAGC |
| o4161_ShP | GCTTCTGGCAAACTCACATCCATTATAG |
| o4162_ShP | CTATAATGGATGTAACTTTGCCAGAAGCAAG |
| o4163_ShP | CTTGCTTCTGGCAAAGTTACATCCATTATAG |
| o4196_ShP | GTACTATAATGGATGTTAGTTTGCCAGAAGC |
| o4197_ShP | GCTTCTGGCAAACTAACATCCATTATAGTAC |
| o4324_ShP | CTGATCTCAGAGGAGGACCTGAGCCTGTCTAGAGGCATG |
| o4325_ShP | CCTCTGAGATCAGCTTCTGCTCGCCCATGGTGGCTAGC |
| o4326_ShP | GGTGGGATGACTAGCTGAATCGGTAGGAATTC |
| o4327_ShP | GAATTCCTACCGATTCAGCTAGTCATCCCACC |
| o4335_ShP | CTCTAGAGCTAGCACCATGGCCCCCAAGAAG |
| o4336_ShP | CTTCTTGGGGGCCATGGTGCTAGCTCTAGAG |
| o4337_ShP | CGTTCCGGACTACGCTTAATGATAAGGATCCGCG |
| o4338_ShP | CGCGGATCCTTATCATTAAGCGTAGTCCGGAACG |
| o4417_ShP | TGTAATCCAGAGGTTGATTGTCGACTCAGCTAGTCATCCCACC |
| o4421_ShP | CCGTCTCAGAGGTATGACTAGCTGAATC |
| o4422_ShP | GATTCAGCTAGTCATACCTCTGAGACGG |
| o4511_ShP | CCTGTAATGAGGATCCGCGGCCGCGGACGGATCGGGACGTTAC |
| o4595_ShP | ATTGTCGACTCAGCTAGTCATACCTCTGAGACGG |
| o4651_ShP | CGTGCGGCCGCTTACTTGTCATCGTCATCCTTGTAATCGATGTCATGATCTTTATAATCACCGTCA TGGTCTTTGTAGTCGTTAATTACAAATGTGTTGTAAGGCCC |
| o4652_ShP | TATGGCGCGCCCAAAGTCAG |
| o4653_ShP | GTACTATAATGGATGTAAGTTTGCCAGAAGC |
| o4654_ShP | GCTTCTGGCAAACTTACATCCATTATAGTAC |
| o4657_ShP | GAGATTTGGGACCTTGCAACTTTTCATCTTTTG |
| o4658_ShP | CAAAAGATGAAAAGTTGCAAGGTCCCAAATCTC |
| o4946_SuB | GCCTCTAGACAGGCTCAGGTC |
| o4947_SuB | ATCGGTAGGAATTCGCGGCCG |
| o4948_SuB | GAGGAGGACCTGAGCCTGTCTAGAGGCATGCAGATCTTCGTGAGGACC |
| o4949_SuB | CGGCCGCGAATTCCTACCGATTCACCCACCTCTGAGACGGAG |
| o5226_SuB | GATCTTTGCTGGGAAACAGCTGGAAG |
| o5227_SuB | CTTCCAGCTGTTTCCCAGCAAAGATC |
| o5228_SuB | GATCTTTGCTGGGTAGCAGCTGGAAG |

| | |
|---|---|
| o5229_SuB | CTTCCAGCTGCTACCCAGCAAAGATC |
| o5274_SuB | GACTACAACATCCAGAAAGAGTCCACCCTG |
| o5275_SuB | CAGGGTGGACTCTTTCTGGATGTTGTAGTC |
| o5276_SuB | GACTACAACATCCAGTAGGAGTCCACCCTG |
| o5277_SuB | CAGGGTGGACTCCTACTGGATGTTGTAGTC |
| o5278_SuB | GGACCCTGACTGGTAAGACCATCACTCTCG |
| o5279_SuB | CGAGAGTGATGGTCTTACCAGTCAGGGTCC |
| o5280_SuB | GGACCCTGACTGGTTAGACCATCACTCTCG |
| o5281_SuB | CGAGAGTGATGGTCTAACCAGTCAGGGTCC |

**Table S2: Protein coding sequences used in this study.**

| Protein | Sequence |
|---|---|
| Myc-Ub wt | EQKLISEEDLSLSRGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLE DGRTLSDYNIQKESTLHLVLRLRGGMTS |
| Myc-Ub nc | EQKLISEEDLSLSRGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLE DGRTLSDYNIQKESTLHLVLRLRG-MTS |
| Myc-Ub K0 | EQKLISEEDLSLSRGMQIFVRTLTGRTITLEVEPSDTIENVRARIQDREGIPPDQQRLIFAGRQLE DGRTLSDYNIQRESTLHLVLRLRGG |
| Myc-Ub K11 | EQKLISEEDLSLSRGMQIFVRTLTGKTITLEVEPSDTIENVRARIQDREGIPPDQQRLIFAGRQLE DGRTLSDYNIQRESTLHLVLRLRGG |
| Myc-Ub pcK11 | EQKLISEEDLSLSRGMQIFVRTLTGpcKTITLEVEPSDTIENVRARIQDREGIPPDQQRLIFAGRQ LEDGRTLSDYNIQRESTLHLVLRLRGG |
| Myc-Ub K48 | EQKLISEEDLSLSRGMQIFVRTLTGRTITLEVEPSDTIENVRARIQDREGIPPDQQRLIFAGKQLE DGRTLSDYNIQRESTLHLVLRLRGG |
| Myc-Ub pcK48 | EQKLISEEDLSLSRGMQIFVRTLTGRTITLEVEPSDTIENVRARIQDREGIPPDQQRLIFAGpcKQ LEDGRTLSDYNIQRESTLHLVLRLRGG |
| Myc-Ub K63 | EQKLISEEDLSLSRGMQIFVRTLTGRTITLEVEPSDTIENVRARIQDREGIPPDQQRLIFAGRQLE DGRTLSDYNIQKESTLHLVLRLRGG |
| Myc-Ub pcK63 | EQKLISEEDLSLSRGMQIFVRTLTGRTITLEVEPSDTIENVRARIQDREGIPPDQQRLIFAGRQLE DGRTLSDYNIQpcKESTLHLVLRLRGG |

| | |
|---|---|
| mTET2CD -HA | QSQNGKCEGCNPDKDEAPYYTHLGAGPDVAAIRTLMEERYGEKGKAIRIEKVIYTGKEGKSSQGCP<br>IAKWVYRRSSEEEKLLCLVRVRPNHTCETAVMVIAIMLWDGIPKLLASELYSELTDILGKCGICTN<br>RRCSQNETKKKQSPPRNCCCQGENPETCGASFSFGCSWSMYYNGCKFARSKKPRKFRLHGAEPKEE<br>ERLGSHLQNLATVIAPIYKKLAPDAYNNQVEFEHQAPDCCLGLKEGRPFSGVTACLDFSAHSHRDQ<br>QNMPNGSTVVVTLNREDNREVGAKPEDEQFHVLPMYIIAPEDEFGSTEGQEKKIRMGSIEVLQSFR<br>RRRVIRIGELPKSCKKKAEPKKAKTKKAARKHSSLENCSSRTEKGKSSSHTKLMENASHMKQMTAQ<br>PQLSGPVIRQPPTLQRHLQQGQRPQQPQPPQPQPQTTPQPQPQPQHIMPGNSQSVGSHCSGSTSVY<br>TRQPTPHSPYPSSAHTSDIYGDTNHVNFYPTSSHASGSYLNPSNYMNPYLGLLNQNNQYAPFPYNG<br>SVPVDNGSPFLGSYSPQAQSRDLHRYPNQDHLTNQNLPPIHTLHQQTFGDSPSKYLSYGNQNMQRD<br>AFTTNSTLKPNVHHLATFSPYPTPKMDSHFMGAASRSPYSHPHTDYKTSEHHLPSHTVYSYTAAAS<br>GSSSSHAFHNKENDNIANGLSRVLPGFNHDRTASAQELLYSLTGSSQEKQPEVSGQDAAAVQEIEY<br>WSDSEHNFQDPCIGGVAIAPTHGSILIECAKCEVHATTKVNDPDRNHPTRISLVLYRHKNLFLPKH<br>CLALWEAKMAEKARKEEECGKNGSDHVSQKNHGKQEKREPTGPQEPSYLRFIQSLAENTGSVTTDS<br>TVTTSPYAFTQVTGPYNTFVINYPYDVPDYA |
| mTET2CD (K1212E)-HA | QSQNGKCEGCNPDKDEAPYYTHLGAGPDVAAIRTLMEERYGEKGKAIRIEKVIYTGKEGKSSQGCP<br>IAKWVYRRSSEEEKLLCLVRVRPNHTCETAVMVIAIMLWDGIPKLLASELYSELTDILGKCGICTN<br>RRCSQNETKKKQSPPRNCCCQGENPETCGASFSFGCSWSMYYNGCEFARSKKPRKFRLHGAEPKEE<br>ERLGSHLQNLATVIAPIYKKLAPDAYNNQVEFEHQAPDCCLGLKEGRPFSGVTACLDFSAHSHRDQ<br>QNMPNGSTVVVTLNREDNREVGAKPEDEQFHVLPMYIIAPEDEFGSTEGQEKKIRMGSIEVLQSFR<br>RRRVIRIGELPKSCKKKAEPKKAKTKKAARKHSSLENCSSRTEKGKSSSHTKLMENASHMKQMTAQ<br>PQLSGPVIRQPPTLQRHLQQGQRPQQPQPPQPQPQTTPQPQPQPQHIMPGNSQSVGSHCSGSTSVY<br>TRQPTPHSPYPSSAHTSDIYGDTNHVNFYPTSSHASGSYLNPSNYMNPYLGLLNQNNQYAPFPYNG<br>SVPVDNGSPFLGSYSPQAQSRDLHRYPNQDHLTNQNLPPIHTLHQQTFGDSPSKYLSYGNQNMQRD<br>AFTTNSTLKPNVHHLATFSPYPTPKMDSHFMGAASRSPYSHPHTDYKTSEHHLPSHTVYSYTAAAS<br>GSSSSHAFHNKENDNIANGLSRVLPGFNHDRTASAQELLYSLTGSSQEKQPEVSGQDAAAVQEIEY<br>WSDSEHNFQDPCIGGVAIAPTHGSILIECAKCEVHATTKVNDPDRNHPTRISLVLYRHKNLFLPKH<br>CLALWEAKMAEKARKEEECGKNGSDHVSQKNHGKQEKREPTGPQEPSYLRFIQSLAENTGSVTTDS<br>TVTTSPYAFTQVTGPYNTFVINYPYDVPDYA |
| mTET2CD (K1212N)-HA | QSQNGKCEGCNPDKDEAPYYTHLGAGPDVAAIRTLMEERYGEKGKAIRIEKVIYTGKEGKSSQGCP<br>IAKWVYRRSSEEEKLLCLVRVRPNHTCETAVMVIAIMLWDGIPKLLASELYSELTDILGKCGICTN<br>RRCSQNETKKKQSPPRNCCCQGENPETCGASFSFGCSWSMYYNGCNFARSKKPRKFRLHGAEPKEE<br>ERLGSHLQNLATVIAPIYKKLAPDAYNNQVEFEHQAPDCCLGLKEGRPFSGVTACLDFSAHSHRDQ<br>QNMPNGSTVVVTLNREDNREVGAKPEDEQFHVLPMYIIAPEDEFGSTEGQEKKIRMGSIEVLQSFR<br>RRRVIRIGELPKSCKKKAEPKKAKTKKAARKHSSLENCSSRTEKGKSSSHTKLMENASHMKQMTAQ<br>PQLSGPVIRQPPTLQRHLQQGQRPQQPQPPQPQPQTTPQPQPQPQHIMPGNSQSVGSHCSGSTSVY<br>TRQPTPHSPYPSSAHTSDIYGDTNHVNFYPTSSHASGSYLNPSNYMNPYLGLLNQNNQYAPFPYNG<br>SVPVDNGSPFLGSYSPQAQSRDLHRYPNQDHLTNQNLPPIHTLHQQTFGDSPSKYLSYGNQNMQRD<br>AFTTNSTLKPNVHHLATFSPYPTPKMDSHFMGAASRSPYSHPHTDYKTSEHHLPSHTVYSYTAAAS<br>GSSSSHAFHNKENDNIANGLSRVLPGFNHDRTASAQELLYSLTGSSQEKQPEVSGQDAAAVQEIEY<br>WSDSEHNFQDPCIGGVAIAPTHGSILIECAKCEVHATTKVNDPDRNHPTRISLVLYRHKNLFLPKH<br>CLALWEAKMAEKARKEEECGKNGSDHVSQKNHGKQEKREPTGPQEPSYLRFIQSLAENTGSVTTDS<br>TVTTSPYAFTQVTGPYNTFVINYPYDVPDYA |
| mTET2CD -3X Flag | QSQNGKCEGCNPDKDEAPYYTHLGAGPDVAAIRTLMEERYGEKGKAIRIEKVIYTGKEGKSSQGCP<br>IAKWVYRRSSEEEKLLCLVRVRPNHTCETAVMVIAIMLWDGIPKLLASELYSELTDILGKCGICTN<br>RRCSQNETKKKQSPPRNCCCQGENPETCGASFSFGCSWSMYYNGCKFARSKKPRKFRLHGAEPKEE<br>ERLGSHLQNLATVIAPIYKKLAPDAYNNQVEFEHQAPDCCLGLKEGRPFSGVTACLDFSAHSHRDQ<br>QNMPNGSTVVVTLNREDNREVGAKPEDEQFHVLPMYIIAPEDEFGSTEGQEKKIRMGSIEVLQSFR<br>RRRVIRIGELPKSCKKKAEPKKAKTKKAARKHSSLENCSSRTEKGKSSSHTKLMENASHMKQMTAQ<br>PQLSGPVIRQPPTLQRHLQQGQRPQQPQPPQPQPQTTPQPQPQPQHIMPGNSQSVGSHCSGSTSVY<br>TRQPTPHSPYPSSAHTSDIYGDTNHVNFYPTSSHASGSYLNPSNYMNPYLGLLNQNNQYAPFPYNG<br>SVPVDNGSPFLGSYSPQAQSRDLHRYPNQDHLTNQNLPPIHTLHQQTFGDSPSKYLSYGNQNMQRD<br>AFTTNSTLKPNVHHLATFSPYPTPKMDSHFMGAASRSPYSHPHTDYKTSEHHLPSHTVYSYTAAAS<br>GSSSSHAFHNKENDNIANGLSRVLPGFNHDRTASAQELLYSLTGSSQEKQPEVSGQDAAAVQEIEY<br>WSDSEHNFQDPCIGGVAIAPTHGSILIECAKCEVHATTKVNDPDRNHPTRISLVLYRHKNLFLPKH<br>CLALWEAKMAEKARKEEECGKNGSDHVSQKNHGKQEKREPTGPQEPSYLRFIQSLAENTGSVTTDS<br>TVTTSPYAFTQVTGPYNTFVINDYKDHDGDYKDHDIDYKDDDDK |
| mTET2CD (K1212#)-3X Flag | QSQNGKCEGCNPDKDEAPYYTHLGAGPDVAAIRTLMEERYGEKGKAIRIEKVIYTGKEGKSSQGCP<br>IAKWVYRRSSEEEKLLCLVRVRPNHTCETAVMVIAIMLWDGIPKLLASELYSELTDILGKCGICTN<br>RRCSQNETKKKQSPPRNCCCQGENPETCGASFSFGCSWSMYYNGCpcKFARSKKPRKFRLHGAEPK<br>EEERLGSHLQNLATVIAPIYKKLAPDAYNNQVEFEHQAPDCCLGLKEGRPFSGVTACLDFSAHSHR<br>DQQNMPNGSTVVVTLNREDNREVGAKPEDEQFHVLPMYIIAPEDEFGSTEGQEKKIRMGSIEVLQS |

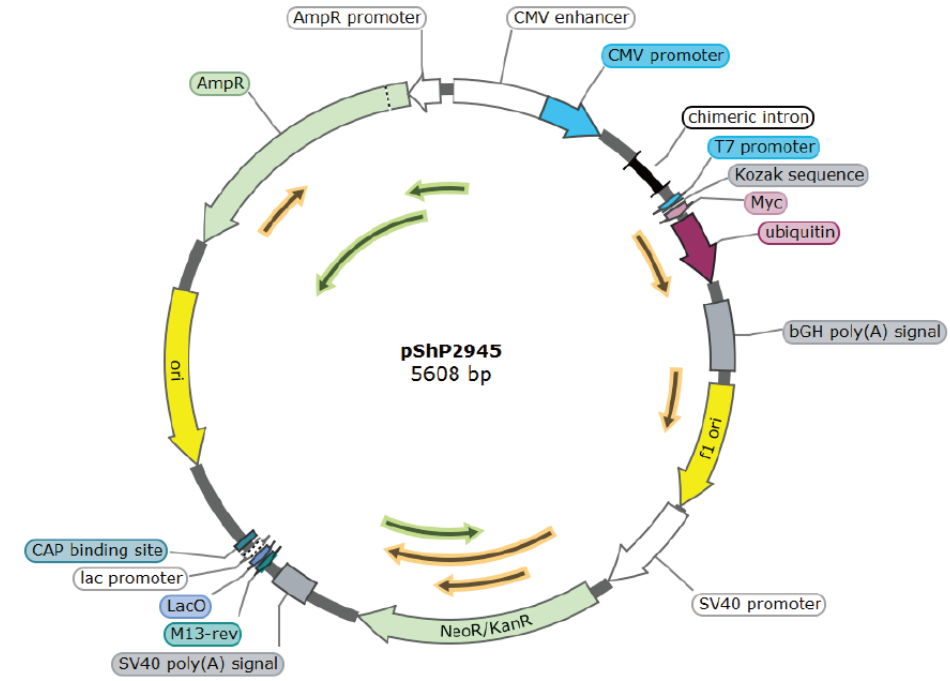
| | |
|---|---|
| | FRRRRVIRIGELPKSCKKKAEPKKAKTKKAARKHSSLENCSSRTEKGKSSSHTKLMENASHMKQMT<br>AQPQLSGPVIRQPPTLQRHLQQGQRPQQPQPPQPQPQTTPQPQPQPQHIMPGNSQSVGSHCSGSTS<br>VYTRQPTPHSPYPSSAHTSDIYGDTNHVNFYPTSSHASGSYLNPSNYMNPYLGLLNQNNQYAPFPY<br>NGSVPVDNGSPFLGSYSPQAQSRDLHRYPNQDHLTNQNLPPIHTLHQQTFGDSPSKYLSYGNQNMQ<br>RDAFTTNSTLKPNVHHLATFSPYPTPKMDSHFMGAASRSPYSHPHTDYKTSEHHLPSHTVYSYTAA<br>ASGSSSSHAFHNKENDNIANGLSRVLPGFNHDRTASAQELLYSLTGSSQEKQPEVSGQDAAAVQEI<br>EYWSDSEHNFQDPCIGGVAIAPTHGSILIECAKCEVHATTKVNDPDRNHPTRISLVLYRHKNLFLP<br>KHCLALWEAKMAEKARKEEECGKNGSDHVSQKNHGKQEKREPTGPQEPSYLRFIQSLAENTGSVTT<br>DSTVTTSPYAFTQVTGPYNTFVIN**DYKDHDGDYKDHDIDYKDDDDK** |
| HA-VprBP | **YPYDVPDYA**EFDITSLYKKAGSLKEPIQSTGSGTEFTGKAMTTVVVHVDSKAELTTLLEQWEKEHG<br>SGQDMVPILTRMSQLIEKETEEYRKGDPDPFDDRHPGRADPECMLGHLLRILFKNDDFMNALVNAY<br>VMTSREPPLNTAACRLLLDIMPGLETAVVFQEKEGIVENLFKWAREADQPLRTYSTGLLGGAMENQ<br>DIAANYRDENSQLVAIVLRRLRELQLQEVALRQENKRPSPRKLSSEPLLPLDEEAVDMDYGDMAVD<br>VVDGDQEEASGDMEISFHLDSGHKTSSRVNSTTKPEDGGLKKNKSAKQGDRENFRKAKQKLGFSSS<br>DPDRMFVELSNSSWSEMSPWVIGTNYTLYPMTPAIEQRLILQYLTPLGEYQELLPIFMQLGSRELM<br>MFYIDLKQTNDVLLTFEALKHLASLLLHNKFATEFVAHGGVQKLLEIPRPSMAATGVSMCLYYLSY<br>NQDAMERVCMHPHNVLSDVVNYTLWLMECSHASGCCHATMFFSICFSFRAVLELFDRYDGLRRLVN<br>LISTLEILNLEDQGALLSDDEIFASRQTGKHTCMALRKYFEAHLAIKLEQVKQSLQRTEGGILVHP<br>QPPYKACSYTHEQIVEMMEFLIEYGPAQLYWEPAEVFLKLSCVQLLLQLISIACNWKTYYARNDTV<br>RFALDVLAILTVVPKIQLQLAESVDVLDEAGSTVSTVGISIILGVAEGEFFIHDAEIQKSALQIII<br>NCVCGPDNRISSIGKFISGTPRRKLPQNPKSSEHTLAKMWNVVQSNNGIKVLLSLLSIKMPITDAD<br>QIRALACKALVGLSRSSTVRQIISKLPLFSSCQIQQLMKEPVLQDKRSDHVKFCKYAAELIERVSG<br>KPLLIGTDVSLARLQKADVVAQSRISFPEKELLLLIRNHLISKGLGETATVLTKEADLPMTAASHS<br>SAFTPVTAAASPVSLPRTPRIANGIATRLGSHAAVGASAPSAPTAHPQPRPPQGPLALPGPSYAGN<br>SPLIGRISFIRERPSPCNGRKIRVLRQKSDHGAYSQSPAIKKQLDRHLPSPPTLDSIITEYLREQH<br>ARCKNPVATCPPFSLFTPHQCPEPKQRRQAPINFTSRLNRRASFPKYGGVDGGCFDRHLIFSRFRP<br>ISVFREANEDESGFTCCAFSARERFLMLGTCTGQLKLYNVFSGQEEASYNCHNSAITHLEPSRDGS<br>LLLTSATWSQPLSALWGMKSVFDMKHSFTEDHYVEFSKHSQDRVIGTKGDIAHIYDIQTGNKLLTL<br>FNPDLANNYKRNCATFNPTDDLVLNDGVLWDVRSAQAIHKFDKFNMNISGVFHPNGLEVIINTEIW<br>DLRTFHLLHTVPALDQCRVVFNHTGTVMYGAMLQADDEDDLMEERMKSPFGSSFRTFNATDYKPIA<br>TIDVKRNIFDLCTDTKDCYLAVIENQGSMDALNMDTVCRLYEVGRQRLAEDEDEEEDQEEEEQEEE<br>DDDEDDDDTDDLDELDTDQLLEAELEEDDNNENAGEDGDNDFSPSDEELANLLEEGEDGEDEDSDA<br>DEEVELILGDTDSSDNSDLEDDIILSLNEPGLEISRPSFLVQSGDILEHASRGPYSIVSPKC |
| HA-VprBP (RARA) | **YPYDVPDYA**EFDITSLYKKAGSLKEPIQSTGSGTEFTGKAMTTVVVHVDSKAELTTLLEQWEKEHG<br>SGQDMVPILTRMSQLIEKETEEYRKGDPDPFDDRHPGRADPECMLGHLLRILFKNDDFMNALVNAY<br>VMTSREPPLNTAACRLLLDIMPGLETAVVFQEKEGIVENLFKWAREADQPLRTYSTGLLGGAMENQ<br>DIAANYRDENSQLVAIVLRRLRELQLQEVALRQENKRPSPRKLSSEPLLPLDEEAVDMDYGDMAVD<br>VVDGDQEEASGDMEISFHLDSGHKTSSRVNSTTKPEDGGLKKNKSAKQGDRENFRKAKQKLGFSSS<br>DPDRMFVELSNSSWSEMSPWVIGTNYTLYPMTPAIEQRLILQYLTPLGEYQELLPIFMQLGSRELM<br>MFYIDLKQTNDVLLTFEALKHLASLLLHNKFATEFVAHGGVQKLLEIPRPSMAATGVSMCLYYLSY<br>NQDAMERVCMHPHNVLSDVVNYTLWLMECSHASGCCHATMFFSICFSFRAVLELFDRYDGLRRLVN<br>LISTLEILNLEDQGALLSDDEIFASRQTGKHTCMALRKYFEAHLAIKLEQVKQSLQRTEGGILVHP<br>QPPYKACSYTHEQIVEMMEFLIEYGPAQLYWEPAEVFLKLSCVQLLLQLISIACNWKTYYARNDTV<br>RFALDVLAILTVVPKIQLQLAESVDVLDEAGSTVSTVGISIILGVAEGEFFIHDAEIQKSALQIII<br>NCVCGPDNRISSIGKFISGTPRRKLPQNPKSSEHTLAKMWNVVQSNNGIKVLLSLLSIKMPITDAD<br>QIRALACKALVGLSRSSTVRQIISKLPLFSSCQIQQLMKEPVLQDKRSDHVKFCKYAAELIERVSG<br>KPLLIGTDVSLARLQKADVVAQSRISFPEKELLLLIRNHLISKGLGETATVLTKEADLPMTAASHS<br>SAFTPVTAAASPVSLPRTPRIANGIATRLGSHAAVGASAPSAPTAHPQPRPPQGPLALPGPSYAGN<br>SPLIGRISFIRERPSPCNGRKIRVLRQKSDHGAYSQSPAIKKQLDRHLPSPPTLDSIITEYLREQH<br>ARCKNPVATCPPFSLFTPHQCPEPKQRRQAPINFTSRLNRRASFPKYGGVDGGCFDRHLIFSRFRP<br>ISVFREANEDESGFTCCAFSARERFLMLGTCTGQLKLYNVFSGQEEASYNCHNSAITHLEPSRDGS<br>LLLTSATWSQPLSALWGMKSVFDMKHSFTEDHYVEFSKHSQDRVIGTKGDIAHIYDIQTGNKLLTL<br>FNPDLANNYKRNCATFNPTDDLVLNDGVLWDV**A**SAQAIHKFDKFNMNISGVFHPNGLEVIINTEIW<br>DL**A**TFHLLHTVPALDQCRVVFNHTGTVMYGAMLQADDEDDLMEERMKSPFGSSFRTFNATDYKPIA<br>TIDVKRNIFDLCTDTKDCYLAVIENQGSMDALNMDTVCRLYEVGRQRLAEDEDEEEDQEEEEQEEE<br>DDDEDDDDTDDLDELDTDQLLEAELEEDDNNENAGEDGDNDFSPSDEELANLLEEGEDGEDEDSDA<br>DEEVELILGDTDSSDNSDLEDDIILSLNEPGLEISRPSFLVQSGDILEHASRGPYSIVSPKC |

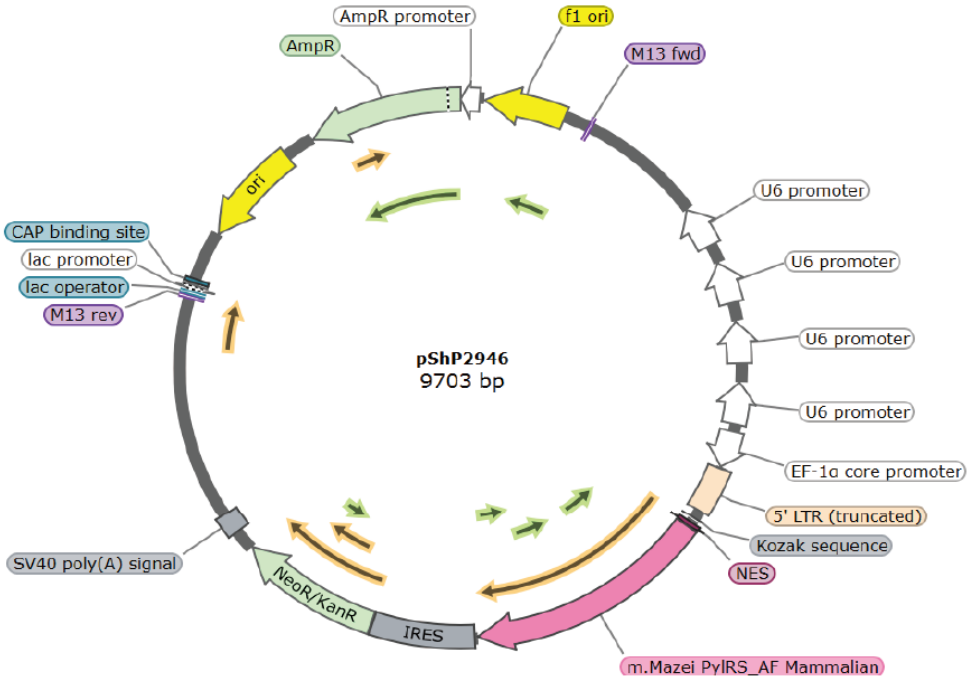
**Table S3: Plasmids**

| Plasmid Name | Gene 1 | Gene 2 | Resistance |
|---|---|---|---|
| pMaS787 | M. mazei PylRS/tRNA$^{Pyl}$ | - | Cam |
| pShP2392 | mTET2CD-HA | mCherry | Amp |
| pShP2927 | mTET2CD K1212E -HA | | Amp |
| pShP2928 | mTET2CD K1212N -HA | | Amp |
| pShP2932 | mTET2CD K1212$^{TAG}$ -HA | | Amp |
| pShP2945 | Myc-Ub | - | Amp |
| pShP2946 | MmPylRS (M276F, A302S, Y306C, L309M)/ 4X tRNA$^{Pyl}$ | - | Amp |
| pShP2947 | mTET2CD K1212$^{TAG}$ -HA/ 4X tRNA$^{Pyl}$ | - | Amp |
| pShP2950 | Myc-Ub (delG76) | - | Amp |
| pShP2956 | MmPylRS (M276F, A302S, Y306C, L309M)/ 4X tRNA$^{Pyl}$ | Myc-Ub | Amp |
| pShP2958 | MmPylRS (M276F, A302S, Y306C, L309M)/ 4X tRNA$^{Pyl}$ | Myc-Ub (delG76) | Amp |
| pShP2962 | mTET2CD K1212$^{TAG}$ -3X FLAG/ 4X tRNA$^{Pyl}$ | - | Amp |
| pShP2963 | mTET2CD -3X FLAG/ 4X tRNA$^{Pyl}$ | - | Amp |
| pShP2965 | HA-VprBP (RARA) | | Amp |
| pShP2970 | MmPylRS (wt)/ 4X tRNA$^{Pyl}$ | NeoR/KanR | Amp |
| pSuB3041 | Myc-Ub (K0) | - | Amp |
| pSuB3439 | Myc-Ub (K48) | - | Amp |
| pSuB3440 | Myc-Ub (K48$^{TAG}$) | - | Amp |
| pSuB3455 | Myc-Ub (K11) | - | Amp |
| pSuB3456 | Myc-Ub (K11$^{TAG}$) | - | Amp |
| pSuB3457 | Myc-Ub (K63) | - | Amp |
| pSuB3458 | Myc-Ub (K63$^{TAG}$) | - | Amp |

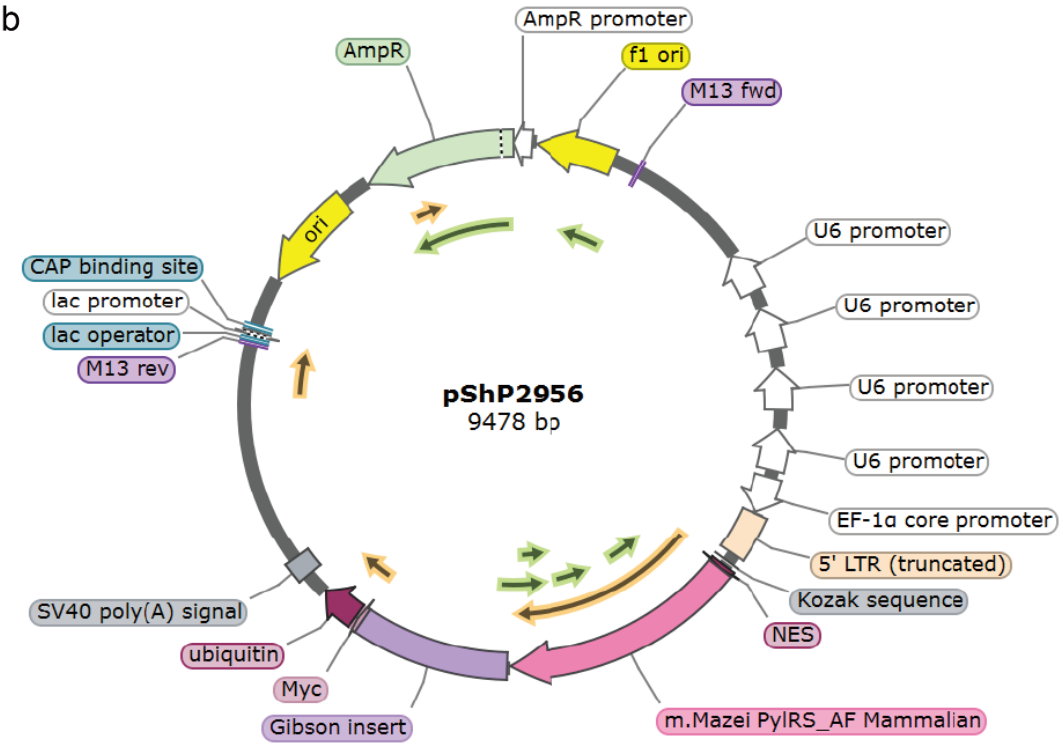
## A.3 Plasmid Maps

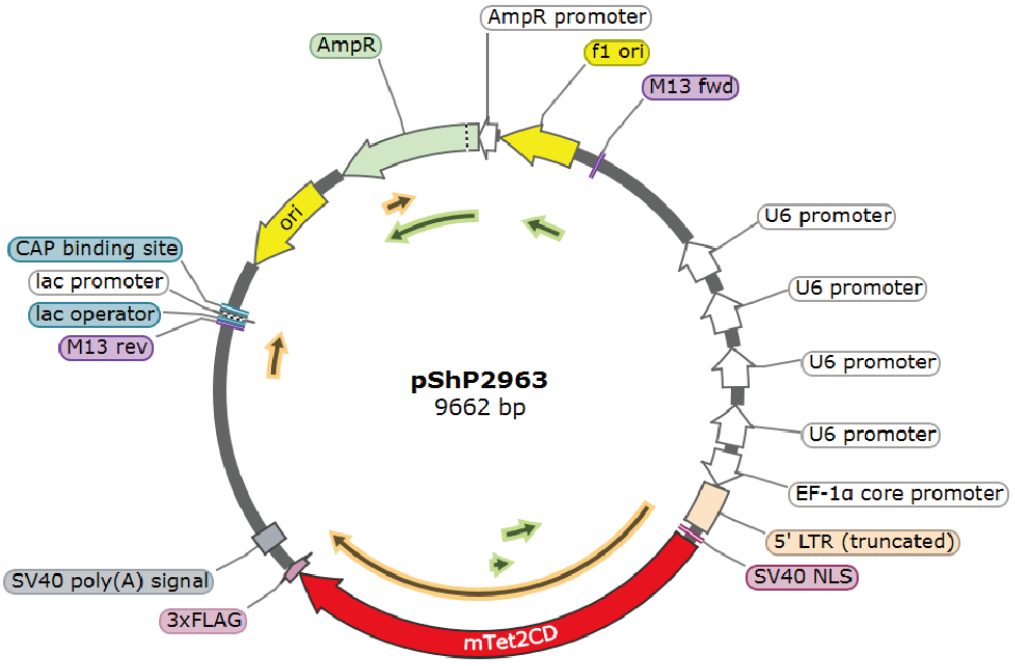
a



b


AmpR promoter
AmpR
f1 ori
M13 fwd
ori
U6 promoter
CAP binding site
U6 promoter
lac promoter
lac operator
M13 rev
U6 promoter
pShP2956
9478 bp
U6 promoter
EF-1α core promoter
5' LTR (truncated)
Kozak sequence
SV40 poly(A) signal
NES
ubiquitin
Myc
Gibson insert
m.Mazei PylRS_AF Mammalian


AmpR promoter
AmpR
f1 ori
M13 fwd
ori
U6 promoter
CAP binding site
U6 promoter
lac promoter
lac operator
M13 rev
U6 promoter
pShP2963
9662 bp
U6 promoter
EF-1α core promoter
5' LTR (truncated)
SV40 NLS
SV40 poly(A) signal
3xFLAG
mTet2CD

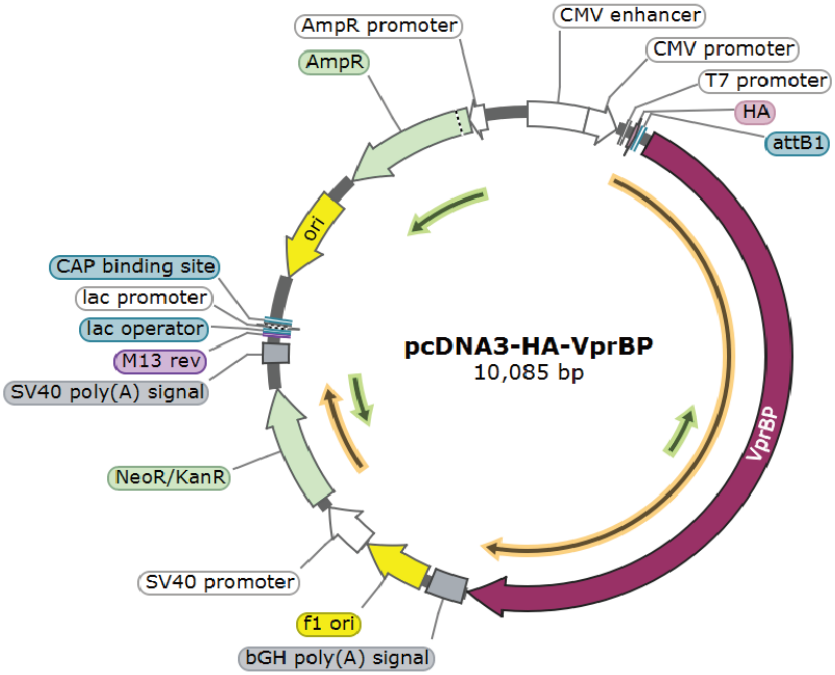



**Figure S10. Plasmid maps of expression vectors used in this study.** Maps of the expression vector for: **a.** N-terminally myc-tagged Ub WT (pShP2945; upper panel). The myc-tagged mutants of Ub, including Ub K0, Ub K11, Ub K48, Ub K63, Ub K11TAG, Ub K48TAG, and Ub K63TAG, share the same vector features except for the indicated mutations; Modified pcKRS (pShP2946; lower pannel), used for light activated Ub studies. **b.** Modified pcKRS with fused Ub containing N-terminal myc tag (pShP2956; upper pannel). Contruct containing the inactive Ub share the same vector features except for the indicated mutation (G76 deletion); mTET2CD (wt) with C-terminal 3X-FLAG tag (pShP2963; middle pannel). Contruct containing the TET mutants including K1212N, K1212E or K1212TAG, share the same vector features except for the indicated mutations. Also, construct containing mTET2CD with C-terminal HA tag share the same vector features except for the tag; VprBP (wt) containing a N-terminal HA tag (Addgene #133586). VprBP variant RARA share the same vector features except for the mutations.

## A.4 Copyright License

[7] Figure 6 is adapted with permission from (Qi et al., 2017). Copyright ©2016 by Elsevier Ltd.

[8] Figure 7 is adapted and modified with permission. Fig. I8 is reprinted from (Nedelsky et al., 2008). Copyright © 2008 Elsevier B.V.

[9] Figure 8 is adapted and modified with permission from (Kulathu & Komander, 2012). Copyright © 2012, Springer Nature Limited.

[10] Figure 9 is adapted and modified with permission from (Komander et al., 2009). Copyright © 1969, Springer Nature Limited.

[11] Figure 10 c is adapted and modified with permission from (Békés et al., 2013). Copyright © 2013, The Authors.

[12] Figure 11 b is adapted from (Laplante & Zhang, 2021). Copyright © 2021 by the Authors. The original figure is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit: https://creativecommons.org/licenses/by/4.0/

[13] Figure 12 a is adapted from (Ross et al., 2021). Copyright © 2020, The Author(s). The original figure is licensed under Creative Commons CC BY. To view a copy of this license, visit https://creativecommons.org/licenses/

[14] Figure 12 b is adapted and modified with permission from (Kudriaeva et al., 2021). Copyright © 2021, Elsevier Ltd.

[15] Figure 12 c is adapted from (Ifrim et al., 2022a). Copyright © 2022, The Authors. The original figure is licensed under the Creative Commons CC-BY-NC-ND.

[16] Figure 13 a, Figure 15 and Figure 16 b are adapted from (Kneuttinger, 2022). Copyright © 2022, The Author(s). The original figure is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit: https://creativecommons.org/licenses/by/4.0/

[17] Figure 13 b and Figure 14 are adapted from (Banerjee & Mitra, 2020). Copyright © 2019, Elsevier Ltd.

[18] Figure 16 a is created based on the work of (Dumas et al., 2015). The original figure is licensed under CC BY 3.0. To view a copy of this license, visit https://creativecommons.org/licenses/by/3.0/.

[19] Figure 17 a and Figure 17 b are adapted with permission from (Chin, 2017). Copyright © Macmillan Publishers Limited, part of Springer Nature.

[20] Figure 17 c is adapted from (Lang & Chin, 2014b). Copyright © 2014, American Chemical Society.

[21] Figure 19 b is adapted from (Zhu et al., 2023). Copyright © 2023, The Author(s). The original figure is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit: https://creativecommons.org/licenses/by/4.0/

[22] Figure 19 c is adapted with permission from (Muñoz-López & Summerer, 2018). Copyright © 2018 The Chemical Society of Japan & Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

[23] Figure 20 b is adapted with permission from (Hu et al., 2013). Copyright © 2013 Elsevier Inc.

[24] Figure 20 c is adapted with permission from (Hu et al., 2015). Copyright © 2015, Springer Nature Limited.

[25] Figure B1 a, Figure B1 b, Figure B1 d (right panel) and Figure B1 e are reprinted with permission from (Tomko & Hochstrasser, 2013). Copyright © 2013 by Annual Reviews.

[26] Figure B1 c and Figure B1 d (left panel) are reprinted with permission from (Ciechanover & Stanhill, 2014). Copyright © 2013 Elsevier B.V.

# A. Bibliography


Abdul Rehman, S. A., Kristariyanto, Y. A., Choi, S. Y., Nkosi, P. J., Weidlich, S., Labib, K., Hofmann, K., & Kulathu, Y. (2016). MINDY-1 Is a Member of an Evolutionarily Conserved and Structurally Distinct New Family of Deubiquitinating Enzymes. *Molecular Cell*, *63*(1), 146–155. https://doi.org/10.1016/j.molcel.2016.05.009

Adams, J., et al. (1998). Potent and selective inhibitors of the proteasome: Dipeptidyl boronic acids. *Bioorg. Med. Chem. Lett.* 8, 333. https://doi.org/10.1016/S0960-894X(98)00029-8

Aghajan, M. et al. (2010). Chemical genetics screen for enhancers of rapamycin identifies a specific inhibitor of an SCF family E3 ubiquitin ligase. *Nat. Biotechnol.* 28, 738 – 742. https://doi.org/10.1038/nbt.1645

Aleo, E., Henderson, C.J., Fontanini, A., Solazzo, B., & Brancolini, C. (2006). Identification of new compounds that trigger apoptosome-independent caspase activation and apoptosis. *Cancer Res.* 66, 9235–9244. https://doi.org/10.1158/0008-5472.CAN-06-0702

Altun, M., et al. (2011). Activity-Based Chemical Proteomics Accelerates Inhibitor Development for Deubiquitylating Enzymes. *Chem. Biol.* 18, 1401–1412. https://doi.org/10.1016/j.chembiol.2011.08.018

Ambrogelly, A., Palioura, S., & So ll, D. (2007). Natural expansion of the genetic code. *Nature Chemical Biology*, *3*(1), 29–35. https://doi.org/10.1038/nchembio847

Ankenbruck, N., Courtney, T., Naro, Y., & Deiters, A. (2018). Optochemische Steuerung biologischer Vorgänge in Zellen und Tieren. *Angewandte Chemie*, *130*(11), 2816–2848. https://doi.org/10.1002/ange.201700171

Aoki, K., Kumagai, Y., Sakurai, A., Komatsu, N., Fujita, Y., Shionyu, C., & Matsuda, M. (2013). Stochastic ERK activation induced by noise and cell-to-cell propagation regulates cell density-dependent proliferation. *Molecular Cell*, *52*(4), 529–540. https://doi.org/10.1016/j.molcel.2013.09.015

Arbely, E., Torres-Kolbus, J., Deiters, A., & Chin, J. W. (2012). Photocontrol of tyrosine phosphorylation in mammalian cells via genetic encoding of photocaged tyrosine. *Journal of the American Chemical Society*, *134*(29), 11912–11915. https://doi.org/10.1021/ja3046958

Arroyo, M., Hastert, F. D., Zhadan, A., Schelter, F., Zimbelmann, S., Rausch, C., Ludwig, A. K., Carell, T., & Cardoso, M. C. (2022). Isoform-specific and ubiquitination dependent recruitment of Tet1 to replicating heterochromatin modulates methylcytosine oxidation. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-32799-8

Baker, R. T. & Board, P. G. The human ubiquitin-52 amino acid fusion protein gene shares several structural features with mammalian ribosomal protein genes. *Nucleic Acids Res.* **19,** 1035–1040 (1991). https://doi.org/10.1093/nar/19.5.1035

Baker, A. S., & Deiters, A. (2014). Optical control of protein function through unnatural amino acid mutagenesis and other optogenetic approaches. In *ACS Chemical Biology* (Vol. 9, Issue 7, pp. 1398–1407). American Chemical Society. https://doi.org/10.1021/cb500176x

Banerjee, S., Cakil, Z. V., Gallant, K., Boom, J. van den, Palei, S., Meyer, H., Gersch, M., & Summerer, D. (2024). Light-Activatable Ubiquitin for Studying Linkage-Specific Ubiquitin Chain Formation Kinetics. *Advanced Science*. https://doi.org/10.1002/advs.202406570

Banerjee, S., & Mitra, D. (2020). Structural Basis of Design and Engineering for Advanced Plant Optogenetics. In *Trends in Plant Science* (Vol. 25, Issue 1, pp. 35–65). Elsevier Ltd. https://doi.org/10.1016/j.tplants.2019.10.002

Banghart, M., Borges, K., Isacoff, E., Trauner, D., & Kramer, R. H. (2004). Light-activated ion channels for remote control of neuronal firing. *Nature Neuroscience*, *7*(12), 1381–1386. https://doi.org/10.1038/nn1356

Bardhan, A., & Deiters, A. (2019). Development of photolabile protecting groups and their application to the optochemical control of cell signaling. *Current Opinion in Structural Biology*, *57*, 164–175. https://doi.org/10.1016/j.sbi.2019.03.028

Bauer, C., et al. (2015). Phosphorylation of TET Proteins Is Regulated via O-GlcNAcylation by the O-Linked N-Acetylglucosamine Transferase (OGT). *J. Biol. Chem.* 290, 8, 4801–4812. https://doi.org/10.1074/jbc.M114.605881

Békés, M., Langley, D. R., & Crews, C. M. (2022). PROTAC targeted protein degraders: the past is prologue. In *Nature Reviews Drug Discovery* (Vol. 21, Issue 3, pp. 181–200). Nature Research. https://doi.org/10.1038/s41573-021-00371-6

Békés, M., Okamoto, K., Crist, S. B., Jones, M. J., Chapman, J. R., Brasher, B. B., Melandri, F. D., Ueberheide, B. M., LazzeriniDenchi, E., & Huang, T. T. (2013). DUB-Resistant Ubiquitin to Survey Ubiquitination Switches in Mammalian Cells. *Cell Reports*, *5*(3), 826–838. https://doi.org/10.1016/j.celrep.2013.10.008

Bentley, M.L., Corn, J.E., Dong, K.C., Phung, Q., Cheung, T.K., & Cochran, A.G. (2011). Recognition of UbcH5c and the nucleosome by the Bmi1/Ring1b ubiquitin ligase complex. *EMBO J.* 30:3285–97. https://doi.org/10.1038/emboj.2011.243

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes and Development*, *16*(1), 6–21. https://doi.org/10.1101/gad.947102

Blythe, E. E., Olson, K. C., Chau, V., & Deshaies, R. J. (2017). Ubiquitin- And ATP-dependent unfoldase activity of P97/VCP•NPLOC4•UFD1L is enhanced by a mutation that causes multisystem proteinopathy. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(22), E4380–E4388. https://doi.org/10.1073/pnas.1706205114

Bose, M., Groff, D., Xie, J., Brustad, E., & Schultz, P. G. (2006). The incorporation of a photoisomerizable amino acid into proteins in E. coli. *Journal of the American Chemical Society*, *128*(2), 388–389. https://doi.org/10.1021/ja055467u

Boulegue, C., Loweneck, M., Renner, C., & Moroder, L. (2007). Redox potential of azobenzene as an amino acid residue in peptides. *ChemBioChem*, *8*(6), 591–594. https://doi.org/10.1002/cbic.200600495

Boulina, M., Samarajeewa, H., Baker, J. D., Kim, M. D., & Chiba, A. (2013). Live imaging of multicolor-labeled cells in Drosophila. *Development (Cambridge)*, *140*(7), 1605–1613. https://doi.org/10.1242/dev.088930

Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G., & Deisseroth, K. (2005). Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neuroscience*, *8*(9), 1263–1268. https://doi.org/10.1038/nn1525

Bremm, A., Freund, S.M., & Komander, D. (2010). Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne. *Nat. Struct. Mol. Biol.* 17:939–47. https://doi.org/10.1038/nsmb.1873

Buckley, D. L., & Crews, C. M. (2014). Small-molecule control of intracellular protein levels through modulation of the ubiquitin proteasome system. In *Angewandte Chemie - International Edition* (Vol. 53, Issue 9, pp. 2312–2330). https://doi.org/10.1002/anie.201307761

Bulatov, E., Zagidullin, A., Valiullina, A., Sayarova, R., & Rizvanov, A. (2018). Small molecule modulators of RING-type E3 ligases: MDM and cullin families as targets. *Frontiers in Pharmacology* (Vol. 9, Issue MAY). Frontiers Media S.A. https://doi.org/10.3389/fphar.2018.00450

Callis, J. (2014). The Ubiquitination Machinery of the Ubiquitin System. *The Arabidopsis Book*, *12*, e0174. https://doi.org/10.1199/tab.0174

Callis, J. and Vierstra, R.D. (1989). Ubiquitin and ubiquitin genes in higher plants. Oxford Surveys in Plant Mol. Biol. 6: 1-30.

Ceccarelli, D. F., Tang, X., Pelletier, B., Orlicky, S., Xie, W., Plantevin, V., Neculai, D., Chou, Y. C., Ogunjimi, A., Al-Hakim, A., Varelas, X., Koszela, J., Wasney, G. A., Vedadi, M., Dhe-Paganon, S., Cox, S., Xu, S., Lopez-Girona, A., Mercurio, F., … Sicheri, F. (2011). An allosteric inhibitor of the human Cdc34 ubiquitin-conjugating enzyme. *Cell*, *145*(7), 1075–1087. https://doi.org/10.1016/j.cell.2011.05.039

Chan, C.H. et al. (2013). Pharmacological Inactivation of Skp2 SCF Ubiquitin Ligase Restricts Cancer Stem Cell Traits and Cancer Progression. *Cell.* 154, 556 – 568. https://doi.org/10.1016/j.cell.2013.06.048

Chau, V., et al. (1989). A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* 243:1576–83. https://doi.org/10.1126/science.2538923

Chen, L., Smith, M. D., Lv, L., Nakagawa, T., Li, Z., Sun, S.-C., Brown, N. G., Xiong, Y., & Xu, Y. (2020). C E L L B I O L O G Y USP15 suppresses tumor immunity via deubiquitylation and inactivation of TET2. In *Sci. Adv* (Vol. 6). http://advances.sciencemag.org/

Chen, P. R., Groff, D., Guo, J., Ou, W., Cellitti, S., Geierstanger, B. H., & Schultz, P. G. (2009). A facile system for encoding unnatural amino acids in mammalian cells. *Angewandte Chemie -International Edition*, *48*(22), 4052–4055. https://doi.org/10.1002/anie.200900683

Chin, J. W. (2017). Expanding and reprogramming the genetic code. In *Nature* (Vol. 550, Issue 7674, pp. 53–60). Nature Publishing Group. https://doi.org/10.1038/nature24031

Chou, C., & Deiters, A. (2011). Light-activated gene editing with a photocaged zinc-finger nuclease. *Angewandte Chemie - International Edition*, *50*(30), 6839–6842. https://doi.org/10.1002/anie.201101157

Choudhury, S. R., Cui, Y., Narayanan, A., Gilley, D. P., Huda, N., Lo, C.-L., Zhou, F. C., Yernool, D., & Irudayaraj, J. (n.d.). *Oncotarget 50380 www.impactjournals.com/oncotarget Optogenetic regulation of site-specific subtelomeric DNA methylation*. www.impactjournals.com/oncotarget/

Christensen, D.E., Brzovic, P.S., & Klevit, R.E. (2007). E2-BRCA1 RING interactions dictate synthesis of mono or specific polyubiquitin chain linkages. *Nat. Struct. Mol. Biol.* 14:941–48. https://doi.org/10.1038/nsmb1295

Ciechanover, A., & Stanhill, A. (2014). The complexity of recognition of ubiquitinated substrates by the 26S proteasome. In *Biochimica et Biophysica Acta - Molecular Cell Research* (Vol. 1843, Issue 1, pp. 86–96). https://doi.org/10.1016/j.bbamcr.2013.07.007

Clague, M. J., Heride, C. & Urbe, S. (2015). The demographics of the ubiquitin system. *Trends Cell Biol.* 25, 417–426. https://doi.org/10.1016/j.tcb.2015.03.002

Clague, M. J., Urbé, S., & Komander, D. (2019). Breaking the chains: deubiquitylating enzyme specificity begets function. In *Nature Reviews Molecular Cell Biology* (Vol. 20, Issue 6, pp. 338–352). Nature Publishing Group. https://doi.org/10.1038/s41580-019-0099-1

Cohen, J.D., Thompson, S., & Ting, A.Y. (2011). Structure-guided engineering of a Pacific Blue fluorophore ligase for specific protein imaging in living cells. *Biochemistry.* 50, 8221–8225. https://doi.org/10.1021/bi201037r

Cohen, P. & Tcherpakov, M. (2010). Will the Ubiquitin System Furnish as Many Drug Targets as Protein Kinases? *Cell.* 143, 686 – 693. https://doi.org/10.1016/j.cell.2010.11.016

Cohn, M.A., Kowal, P., Yang, K., Haas, W., Huang, T. T., Gygi, S.P., & D'Andrea, A.D. (2007). A UAF1-containing multisubunit protein complex regulates the fanconi anemia pathway. *Mol. Cell*. 28, 786–797. https://doi.org/10.1016/j.molcel.2007.09.031

Corson, T.W., Aberle, N.S., & Crews, C.M. (2008). Design and applications of bifunctional small molecules: why two heads are better than one. *ACS Chem. Biol.* 3, 677 – 692. https://doi.org/10.1021/cb8001792

Crick, F. (1970). Central Dogma of Molecular Biology. *Nature 1970 227:5258*, *227*(5258), 561–563. https://doi.org/10.1038/227561a0

Crick, F. H. C. (1968). The origin of the genetic code. *Journal of Molecular Biology*, *38*(3), 367–379. https://doi.org/10.1016/0022-2836(68)90392-6

Crosas, B., Hanna, J., Kirkpatrick, D.S., Zhang, D.P., Tone, Y., et al. (2006). Ubiquitin chains are remodeled at the proteasome by opposing ubiquitin ligase and deubiquitinating activities. *Cell*. 127:1401–1413. https://doi.org/10.1016/j.cell.2006.09.051

Dagliyan, O., Tarnawski, M., Chu, P.-H., Shirvanyants, D., Schlichting, I., Dokholyan, N. V, & Hahn, K. M. (n.d.). *Engineering extrinsic disorder to control protein activity in living cells*. http://science.sciencemag.org/

Datta, A.B., Hura, G.L., Wolberger, C. (2009). The structure and conformation of Lys63-linked tetraubiquitin. *J. Mol. Biol.* 392:1117–24. https://doi.org/10.1016/j.jmb.2009.07.090

de la Torre, D., & Chin, J. W. (2021). Reprogramming the genetic code. In *Nature Reviews Genetics* (Vol. 22, Issue 3, pp. 169–184). Nature Research. https://doi.org/10.1038/s41576-020-00307-7

Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes and Development*, *25*(10), 1010–1022. https://doi.org/10.1101/gad.2037511

Demo, S.D., et al. (2007). Antitumor Activity of PR-171, a Novel Irreversible Inhibitor of the Proteasome. *Cancer Res.* 67, 6383 – 6391. https://doi.org/10.1158/0008-5472.CAN-06-4086

Deng, L., et al. (2000). Activation of the IkappaB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. *Cell.* 103:351–61. https://doi.org/10.1016/s0092-8674(00)00126-4

Deiters, A., Groff, D., Ryu, Y., Xie, J., & Schultz, P. G. (2006). A genetically encoded photocaged tyrosine. *Angewandte Chemie - International Edition*, *45*(17), 2728–2731. https://doi.org/10.1002/anie.200600264

Dieterich, D. C., Hodas, J. J. L., Gouzer, G., Shadrin, I. Y., Ngo, J. T., Triller, A., Tirrell, D. A., & Schuman, E. M. (2010). In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. *Nat. Neurosci.* 13, 897–905. https://doi.org/10.1038/nn.2580

Dieterich, D. C., Link, A. J., Graumann, J., Tirrell, D. A., & Schuman, E. M. (2006). Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proc. Natl. Acad. Sci. U. S. A.* 103, 9482–9487. https://doi.org/10.1073/pnas.0601637103

Dikic, I., & Schulman, B. A. (2022). An expanded lexicon for the ubiquitin code. Nat. Rev. Mol. Cell Biol. 24 (4): 273–287. https://doi.org/10.1038/s41580-022-00543-1

Dikic, I., Wakatsuki, S., & Walters, K.J. (2009). Ubiquitin-binding domains—from structures to functions. *Nat. Rev. Mol. Cell Biol.* 10:659–71. https://doi.org/10.1038/nrm2767

Drag, M., Mikolajczyk, J., Bekes, M., Reyes-Turcu, F. E., Ellman, J. A., Wilkinson, K. D., & Salvesen, G. S. (2008). Positional-scanning fluorigenic substrate libraries reveal unexpected specificity determinants of DUBs (deubiquitinating enzymes). *Biochemical Journal*, *415*(3), 367–375. https://doi.org/10.1042/BJ20080779

Du, Q., Wang, Z., & Schramm, V. L. (2016). Human DNMT1 transition state structure. *Proceedings of the National Academy of Sciences*, *113*(11), 2916–2921. https://doi.org/10.1073/pnas.1522491113

Duda, D. M., Borg, L. A., Scott, D. C., Hunt, H. W., Hammel, M., & Schulman, B. A. (2008). Structural insights into NEDD8 activation of cullin-RING ligases: conformational control of conjugation. *Cell*. 134, 995–1006. https://doi.org/10.1016/j.cell. 2008.07.022

Duda, D. M., Scott, D. C., Calabrese, M. F., Zimmerman, E. S., Zheng, N., & Schulman, B. A. (2011). Structural regulation of cullin-RING ubiquitin ligase complexes. *Curr. Opin. Struct. Biol.* 21, 257–264. https://doi.org/10.1016/j.sbi.2011.01.003

Dumas, A., Lercher, L., Spicer, C. D., & Davis, B. G. (2015). Designing logical codon reassignment – Expanding the chemistry in biology. *Chemical Science*, *6*(1), 50–69. https://doi.org/10.1039/C4SC01534G

Dynek, J.N., et al. (2010). c-IAP1 and UbcH5 promote K11-linked polyubiquitination of RIP1 in TNF signalling. *EMBO J.* 29:4198–209. https://doi.org/10.1038/emboj.2010.300

Eddins, M.J., Carlile, C.M., Gomez, K.M., Pickart, C.M., & Wolberger, C. (2006). Mms2-Ubc13 covalently bound to ubiquitin reveals the structural basis of linkage-specific polyubiquitin chain formation. *Nat. Struct. Mol. Biol.* 13:915–20. https://doi.org/10.1038/nsmb1148

Eddins, M.J., Varadan, R., Fushman, D., Pickart, C.M., & Wolberger, C. (2007). Crystal structure and solution NMR studies of Lys48-linked tetraubiquitin at neutral pH. *J. Mol. Biol.* 367:204–11. https://doi.org/10.1016/j.jmb.2006.12.065

Edwards, W. F., Young, D. D., & Deiters, A. (2009). Light-activated Cre recombinase as a tool for the spatial and temporal control of gene function in mammalian cells. *ACS Chemical Biology*, *4*(6), 441–445. https://doi.org/10.1021/cb900041s

Elia, A. E. et al. (2015). Quantitative proteomic atlas of ubiquitination and acetylation in the DNA damage response. *Mol. Cell.* 59, 867–881. https://doi.org/10.1016/j.molcel.2015.05.006

Faesen, A. C. et al. (2011). The differential modulation of USP activity by internal regulatory domains, interactors and eight ubiquitin chain types. *Chem. Biol.* 18, 1550–1561. https://doi.org/10.1016/j.chembiol.2011.10.017

Falquet, L. *et al.* (1995) A human de-ubiquitinating enzyme with both isopeptidase and peptidase activities *in vitro*. *FEBS Lett.* **359,** 73–77. https://doi.org/10.1016/0014-5793(94)01451-6

Fang, D., & Kerppola, T.K. (2004). Ubiquitin-mediated fluorescence complementation reveals that Jun ubiquitinated by Itch/AIP4 is localized to lysosomes. *Proc. Natl. Acad. Sci. USA.* 101, 14782–14787. https://doi.org/10.1073/pnas.0404445101

Fehrentz, T., Scho nberger, M., & Trauner, D. (2011). Optochemical Genetics. *Angewandte Chemie International Edition*, *50*(51), 12156–12182. https://doi.org/10.1002/anie.201103236

Fiil, B. K. et al. (2013). OTULIN restricts Met1-linked ubiquitination to control innate immune signaling. *Mol. Cell.* 50, 818–830. https://doi.org/10.1016/j.molcel.2013.06.004

Fisk, A. S., Tam, S. K. E., Brown, L. A., Vyazovskiy, V. V., Bannerman, D. M., & Peirson, S. N. (2018). Light and cognition: Roles for circadian rhythms, sleep, and arousal. In *Frontiers in Neurology* (Vol. 9, Issue FEB). Frontiers Media S.A. https://doi.org/10.3389/fneur.2018.00056

Foster, G. L., Schoenheimer, R., & Rittenberg, D. (1939). Studies in protein metabolism V. The utilization of ammonia for amino acid and creatine formation in animals. *J. Biol. Chem.* 127, 319–327. https://doi.org/10.1016/S0021-9258(18)73844-1

French, M. E., Koehler, C. F., & Hunter, T. (2021). Emerging functions of branched ubiquitin chains. In *Cell Discovery* (Vol. 7, Issue 1). Springer Nature. https://doi.org/10.1038/s41421-020-00237-y

Gautier, A., Deiters, A., & Chin, J. W. (2011). Light-activated kinases enable temporal dissection of signaling networks in living cells. *Journal of the American Chemical Society*, *133*(7), 2124–2127. https://doi.org/10.1021/ja1109979

Gautier, A., Nguyen, D. P., Lusic, H., An, W., Deiters, A., & Chin, J. W. (2010). Genetically encoded photocontrol of protein localization in mammalian cells. *Journal of the American Chemical Society*, *132*(12), 4086–4088. https://doi.org/10.1021/ja910688s

Gavory, G., et al. (2018). Discovery and characterization of highly potent and selective allosteric USP7 inhibitors. *Nat. Chem. Biol.* 14, 118–125. https://doi.org/10.1038/nchembio.2528

Gilardini, A., Marmiroli, P., & Cavaletti, G. (2008). Proteasome inhibition: a promising strategy for treating cancer, but what about neurotoxicity? *Curr. Med. Chem.* 15, 3025 – 3035. https://doi.org/10.2174/092986708786848622

Gonzalez-Lopez de Turiso, F., et al. (2013). Rational Design and Binding Mode Duality of MDM2–p53 Inhibitors. *Med. Chem.* 56, 4053 – 4070. https://doi.org/10.1021/jm400293z

Grice, G. L., & Nathan, J. A. (2016). The recognition of ubiquitinated proteins by the proteasome. In *Cellular and Molecular Life Sciences* (Vol. 73, Issue 18, pp. 3497–3506). Birkhauser Verlag AG. https://doi.org/10.1007/s00018-016-2255-5

Groff, D., Wang, F., Jockusch, S., Turro, N. J., & Schultz, P. G. (2010). A new strategy to photoactivate green fluorescent protein. *Angewandte Chemie - International Edition*, *49*(42), 7677–7679. https://doi.org/10.1002/anie.201003797

Groll, M., Kim, K.B., Kairies, N., Huber, R, & Crews, C.M. (2000). *J. Am. Chem. Soc.* 122, 1237 – 1238. https://doi.org/10.1021/ja993588m

Grou, C. P., Pinto, M. P., Mendes, A. V., Domingues, P., & Azevedo, J. E. (2015). The de novo synthesis of ubiquitin: Identification of deubiquitinases acting on ubiquitin precursors. *Scientific Reports*, *5*. https://doi.org/10.1038/srep12836

Haahr, P. et al. (2018). ZUFSP deubiquitylates K63-linked polyubiquitin chains to promote genome stability. *Mol. Cell* 70, 165–174. https://doi.org/10.1016/j.molcel.2018.02.024

Haas, A.L., & Rose, I.A. (1982). The mechanisms of ubiquitin activation enzyme: A kinetic and equilibrium analysis. J. Biol. Chem. 257: 10329-10337. https://doi.org/10.1016/S0021-9258(18)34024-9

Haas, A.L., Warms, J.B.V., & Rose, I.A. (1982a). Ubiquitin adenylate: Structure and role in ubiquitin activation. Biochemistry 22: 4388-4394. https://doi.org/10.1021/bi00288a007

Haas, A.L., Warms, J.V.B., Hershko, A., & Rose, I.A. (1982b). Ubiquitin activating enzyme: Mechanism and role in protein-ubiquitin conjugation. J. Biol. Chem. 257: 2543-2548.

Haldeman, M.T., Xia, G., Kasperek, E.M., & Pickart, C.M. (1997). Structure and function of ubiquitin conjugating enzyme E2-25K: The tail is a core-dependent activity element. *Biochemistry* 36:10526–37. https://doi.org/10.1021/bi970750u

Hanada, M., et al. (1992). Epoxomicin, a new antitumor agent of microbial origin. *J. Antibiot.* 45, 1746 – 1752. https://doi.org/10.7164/antibiotics.45.1746

Hao, B., Gong, W., Ferguson, T. K., James, C. M., Krzycki, J. A., & Chan, M. K. (2002). A New UAG Encoded Residue in the Structure of a Methanogen Methyltransferase. *Science*, *296*(5572), 1462–1466. https://doi.org/10.1126/science.1069556

He, Y.-F., et al. (2011). Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA. *Science*, *333*(6047), 1303–1307. https://doi.org/10.1126/science.1210944

Hemphill, J., Borchardt, E. K., Brown, K., Asokan, A., & Deiters, A. (2015). Optical control of CRISPR/Cas9 gene editing. *Journal of the American Chemical Society*, *137*(17), 5642–5645. https://doi.org/10.1021/ja512664v

Hershko, A., & Ciechanover, A. (1998). THE UBIQUITIN SYSTEM. In *Annu. Rev. Biochem* (Vol. 67). www.annualreviews.org

Hibbert, R.G., Huang, A., Boelens, R., & Sixma, T.K. (2011). E3 ligase Rad18 promotes monoubiquitination rather than ubiquitin chain formation by E2 enzyme Rad6. *Proc. Natl. Acad. Sci. USA* 108:5590–95. https://doi.org/10.1073/pnas.1017516108

Hines, J., Groll, M., Fahnestock, M., & Crews, C.M. (2008). Proteasome Inhibition by Fellutamide B Induces Nerve Growth Factor Synthesis. *Chem. Biol.* 15, 501 – 512. https://doi.org/10.1016/j.chembiol.2008.03.020

Hoppmann, C., Lacey, V. K., Louie, G. v., Wei, J., Noel, J. P., & Wang, L. (2014). Genetically encoding photoswitchable click amino acids in Escherichia coli and mammalian cells. *Angewandte Chemie - International Edition*, *53*(15), 3932–3936. https://doi.org/10.1002/anie.201400001

Hu, C.D., Chinenov, Y., & Kerppola, T.K. (2002). Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. *Mol. Cell*. 9, 789–798. https://doi.org/10.1016/S1097-2765(02)00496-3

Hu, L., Li, Z., Cheng, J., Rao, Q., Gong, W., Liu, M., Shi, Y. G., Zhu, J., Wang, P., & Xu, Y. (2013). Crystal Structure of TET2-DNA Complex: Insight into TET-Mediated 5mC Oxidation. *Cell*, *155*(7), 1545–1555. https://doi.org/10.1016/j.cell.2013.11.020

Hu, L., Lu, J., Cheng, J., Rao, Q., Li, Z., Hou, H., Lou, Z., Zhang, L., Li, W., Gong, W., Liu, M., Sun, C., Yin, X., Li, J., Tan, X., Wang, P., Wang, Y., Fang, D., Cui, Q., … Xu, Y. (2015). Structural insight into substrate preference for TET-mediated oxidation. *Nature*, *527*(7576), 118–122. https://doi.org/10.1038/nature15713.

Hu, M., Li, P., Li, M., Li, W., Yao, T., et al. (2002). Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. *Cell* 111:1041–54. https://doi.org/10.1016/s0092-8674(02)01199-6

Hu, X., Chen, Y., & Zhao, Z.J. (2016). Structure, regulation, and function of TET family proteins. *Epigenetic Gene Expression and Regulation.* 379-395. http://dx.doi.org/10.1016/B978-0-12-799958-6.00017-2

Huang, T. T. et al. (2006). Regulation of monoubiquitinated PCNA by DUB autocleavage. *Nat. Cell Biol.* 8, 339–347. https://doi.org/10.1038/ncb1378

Huber, E.M., & Groll, M. (2012). Inhibitors for the Immuno- and Constitutive Proteasome: Current and Future Trends in Drug Development. *Angew. Chem. Int. Ed.* 51, 8708 – 8720. https://doi.org/10.1002/anie.201201616

Huibregtse, J.M., Scheffner, M., Beaudenon, S., & Howley, P.M. (1995). A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. *Proc. Natl. Acad. Sci. USA* 92:2563–67. https://doi.org/10.1073/pnas.92.7.2563

Hwang, J., & Qi, L. (2018). Quality Control in the Endoplasmic Reticulum: Crosstalk between ERAD and UPR pathways. In *Trends in Biochemical Sciences* (Vol. 43, Issue 8, pp. 593–605). Elsevier Ltd. https://doi.org/10.1016/j.tibs.2018.06.005

Hyer, M. L., et al. (2018). A small-molecule inhibitor of the ubiquitin activating enzyme for cancer treatment. *Nature Medicine*, *24*(2), 186–193. https://doi.org/10.1038/nm.4474

Ibba, M., & Söll, D. (2004). Aminoacyl-tRNAs: Setting the limits of the genetic code. In *Genes and Development* (Vol. 18, Issue 7, pp. 731–738). https://doi.org/10.1101/gad.1187404

Idevall-Hagren, O., Dickson, E. J., Hille, B., Toomre, D. K., & De Camilli, P. (2012). Optogenetic control of phosphoinositide metabolism. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(35). https://doi.org/10.1073/pnas.1211305109

Ifrim, M. F., Janusz-Kaminska, A., & Bassell, G. J. (2022a). Development of single-molecule ubiquitination mediated fluorescence complementation to visualize protein ubiquitination dynamics in dendrites. *Cell Reports*, *41*(7). https://doi.org/10.1016/j.celrep.2022.111658

Ifrim, M. F., Janusz-Kaminska, A., & Bassell, G. J. (2022b). Development of single-molecule ubiquitination mediated fluorescence complementation to visualize protein ubiquitination dynamics in dendrites. *Cell Reports*, *41*(7). https://doi.org/10.1016/j.celrep.2022.111658

Il'ichev, Y. v., & Wirz, J. (2000). Rearrangements of 2-Nitrobenzyl Compounds. 1. Potential Energy Surface of 2-Nitrotoluene and Its Isomers Explored with ab Initio and Density Functional Theory Methods. *The Journal of Physical Chemistry A*, *104*(33), 7856–7870. https://doi.org/10.1021/jp000261v

*INSIGHTS OF THE DECADE*. (2010). https://www.science.org

Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., He, C., & Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, *333*(6047), 1300–1303. https://doi.org/10.1126/science.1210597

Ito, T., Ando, H., Suzuki, T., Ogura, T., Hotta, K., Imamura, Y., Yamaguchi, Y., & Handa, H. (2010). Identification of a Primary Target of Thalidomide Teratogenicity. *Science*. 327, 1345 – 1350. https://doi.org/10.1126/science.1177319

Johansson, L., Gafvelin, G., & Arne r, E. S. J. (2005). Selenocysteine in proteins—properties and biotechnological use. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1726*(1), 1–13. https://doi.org/10.1016/j.bbagen.2005.05.010

Jin, B., Li, Y., & Robertson, K. D. (2011). DNA methylation: Superior or subordinate in the epigenetic hierarchy? In *Genes and Cancer* (Vol. 2, Issue 6, pp. 607–617). https://doi.org/10.1177/1947601910393957

Jin, L., Williamson, A., Banerjee, S., Philipp, I., & Rape, M. (2008). Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex. *Cell* 133:653–65. https://doi.org/10.1016/j.cell.2008.04.012

Kaiser, S. E., Riley, B. E., Shaler, T. A., Trevino, R. S., Becker, C. H., Schulman, H., & Kopito, R. R. (2011). Protein standard absolute quantification (PSAQ) method for the measurement of cellular ubiquitin pools. *Nature Methods*, *8*(8), 691–696. https://doi.org/10.1038/nmeth.1649

Kamadurai, H.B., Souphron, J., Scott, D.C., Duda, D.M., Miller, D.J., *et al*. (2009). Insights into ubiquitin transfer cascades from a structure of a UbcH5B~ubiquitin-HECTNEDD4L complex. *Mol. Cell* 36:1095–102. https://doi.org/10.1016/j.molcel.2009.11.010

Kang, J. Y., Kawaguchi, D., Coin, I., Xiang, Z., O'Leary, D. D. M., Slesinger, P. A., & Wang, L. (2013). In vivo expression of a light-activatable potassium channel using unnatural amino acids. *Neuron*, *80*(2), 358–370. https://doi.org/10.1016/j.neuron.2013.08.016

Kennedy, M. J., Hughes, R. M., Peteya, L. A., Schwartz, J. W., Ehlers, M. D., & Tucker, C. L. (2010). Rapid blue-light-mediated induction of protein interactions in living cells. *Nature Methods*, *7*(12), 973–975. https://doi.org/10.1038/nmeth.1524

Kerscher, O., Felberbaum, R., & Hochstrasser, M. (2006). Modification of Proteins by Ubiquitin and Ubiquitin-Like Proteins. *Annual Review of Biochemistry,* 22:159–80. https://doi.org/10.1146/annurev.cellbio.22.010605.093503

Keusekotten, K. *et al.* (2013). OTULIN antagonizes LUBAC signaling by specifically hydrolyzing Met1-linked polyubiquitin. *Cell.* 153, 1312–1326. https://doi.org/10.1016/j.cell.2013.05.014

Kim, H.C., & Huibregtse, J.M. (2009). Polyubiquitination by HECT E3s and the determinants of chain type specificity. *Mol. Cell. Biol.* 29:3307–18. https://doi.org/10.1128/mcb.00240-09

Kim, H.T., et al. (2007). Certain pairs of ubiquitin conjugating enzymes (E2s) and ubiquitin-protein ligases (E3s) synthesize nondegradable forked ubiquitin chains containing all possible isopeptide linkages. *J. Biol. Chem.* 282:17375–86. https://doi.org/10.1074/jbc.m609659200

Kim, J., et al. 2009. RAD6-mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. *Cell* 137:459–71. https://doi.org/10.1016/j.cell.2009.02.027

Kirkpatrick, D.S., et al. (2006). Quantitative analysis of in vitro ubiquitinated cyclin B1 reveals complex chain topology. *Nat. Cell Biol.* 8:700–10. https://doi.org/10.1038/ncb1436

Kisselev, A.F., & Goldberg, A.L. (2001). Proteasome inhibitors: from research tools to drug candidates. *Chem. Biol.* 8, 739 – 758. https://doi.org/10.1016/S1074-5521(01)00056-4

Klan, P., Solomek, T., Bochet, C. G., Blanc, A., Givens, R., Rubina, M., Popik, V., Kostikov, A., & Wirz, J. (2013). Photoremovable Protecting Groups in Chemistry and Biology: Reaction Mechanisms and Efficacy. *Chemical Reviews*, *113*(1), 119–191. https://doi.org/10.1021/cr300177k

Kneuttinger, A. C. (2022). A guide to designing photocontrol in proteins: methods, strategies and applications. In *Biological Chemistry* (Vol. 403, Issues 5–6, pp. 573–613). De Gruyter Open Ltd. https://doi.org/10.1515/hsz-2021-0417

Knight, R. D., Freeland, S. J., & Landweber, L. F. (2001). Rewiring the keyboard: evolvability of the genetic code. *Nature Reviews Genetics*, *2*(1), 49–58. https://doi.org/10.1038/35047500

Koegl, M., Hoppe, T., Schlenker, S., Ulrich, H.D., Mayer, T.U., & Jentsch, S. (1999). A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. *Cell* 96:635–44. https://doi.org/10.1016/s0092-8674(00)80574-7

Kohli, R. M., & Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. In *Nature* (Vol. 502, Issue 7472, pp. 472–479). https://doi.org/10.1038/nature12750

Komander, D. et al. (2008). The structure of the CYLD USP domain explains its specificity for Lys63-linked polyubiquitin and reveals a B box module. *Mol. Cell.* 29, 451–464. https://doi.org/10.1016/j.molcel.2007.12.018

Komander, D., Clague, M. J., & Urbé, S. (2009). Breaking the chains: Structure and function of the deubiquitinases. In *Nature Reviews Molecular Cell Biology* (Vol. 10, Issue 8, pp. 550–563). https://doi.org/10.1038/nrm2731

Komander, D., & Rape, M. (2012). The ubiquitin code. *Annual Review of Biochemistry*, *81*, 203–229. https://doi.org/10.1146/annurev-biochem-060310-170328

Komander, D. & Barford, D. (2008). Structure of the A20 OUT domain and mechanistic insights into deubiquitination. *Biochem. J.* 409, 77–85. https://doi.org/10.1042/bj20071399

Komander, D., Reyes-Turcu, F., Licchesi, J.D., Odenwaelder, P., Wilkinson, K.D., Barford, D. (2009). Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains. *EMBO Rep.* 10:466–73. https://doi.org/10.1038/embor.2009.55

Konermann, S., Brigham, M. D., Trevino, A. E., Hsu, P. D., Heidenreich, M., Cong, L., Platt, R. J., Scott, D. A., Church, G. M., & Zhang, F. (2013). Optical control of mammalian endogenous transcription and epigenetic states. *Nature*, *500*(7463), 472–476. https://doi.org/10.1038/nature12466

Kriaucionis, S., & Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in purkinje neurons and the brain. *Science*, *324*(5929), 929–930. https://doi.org/10.1126/science.1169786

Kristariyanto, Y. A., Abdul Rehman, S. A., Weidlich, S., Knebel, A. & Kulathu, Y. (2017). A single MIU motif of MINDY-1 recognizes K48-linked polyubiquitin chains. *EMBO Rep.* 18, 392–402. https://doi.org/10.15252/embr.201643205

Kudriaeva, A. A., Livneh, I., Baranov, M. S., Ziganshin, R. H., Tupikin, A. E., Zaitseva, S. O., Kabilov, M. R., Ciechanover, A., & Belogurov, A. A. (2021). In-depth characterization of ubiquitin turnover in mammalian cells by fluorescence tracking. *Cell Chemical Biology*, *28*(8), 1192-1205.e9. https://doi.org/10.1016/j.chembiol.2021.02.009

Kulathu, Y., & Komander, D. (2012). Atypical ubiquitylation-the unexplored world of polyubiquitin beyond Lys48 and Lys63 linkages. In *Nature Reviews Molecular Cell Biology* (Vol. 13, Issue 8, pp. 508–523). https://doi.org/10.1038/nrm3394

Kupka, S. et al. (2016). SPATA2-mediated binding of CYLD to HOIP enables CYLD recruitment to signaling complexes. *Cell Rep.* 16, 2271–2280. https://doi.org/10.1016/j.celrep.2016.07.086

Kuznetsov, G., Goodman, D.B., Filsinger, G.T., Landon, M., Rohland, N., Aach, J., Lajoie, M.J., & Church, G.M. (2017). Optimizing complex phenotypes through model-guided multiplex genome engineering. *Genome Biol*. 18: 100. https://doi.org/10.1186/s13059-017-1217-z

Kwasna, D., Abdul Rehman, S. A., Natarajan, J., Matthews, S., Madden, R., De Cesare, V., Weidlich, S., Virdee, S., Ahel, I., Gibbs-Seymour, I., & Kulathu, Y. (2018). Discovery and Characterization of ZUFSP/ZUP1, a Distinct Deubiquitinase Class Important for Genome Stability. *Molecular Cell*, *70*(1), 150-164.e6. https://doi.org/10.1016/j.molcel.2018.02.023

Lang, K., & Chin, J. W. (2014). Cellular incorporation of unnatural amino acids and bioorthogonal Labeling of Proteins. In *Chemical Reviews* (Vol. 114, Issue 9, pp. 4764–4806). American Chemical Society. https://doi.org/10.1021/cr400355w

Lange, O.F., et al. (2008). Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science* 320:1471–75. https://doi.org/10.1126/science.1157092

Laplante, G., & Zhang, W. (2021). Targeting the ubiquitin-proteasome system for cancer therapeutics by small-molecule inhibitors. In *Cancers* (Vol. 13, Issue 12). MDPI AG. https://doi.org/10.3390/cancers13123079

Laplantine, E., et al. (2009). NEMO specifically recognizes K63-linked poly-ubiquitin chains through a new bipartite ubiquitin-binding domain. *EMBO J.* 28:2885–95. https://doi.org/10.1038/emboj.2009.241

Larance, M., Ahmad, Y., Kirkwood, K. J., Ly, T., & Lamond, A. I. (2013) Global subcellular characterization of protein degradation using quantitative proteomics. *Mol. Cell Proteomics*. 12, 638–650. https://doi.org/10.1074/mcp.M112.024547

Larsen, C. N., Krantz, B. A. & Wilkinson, K. D. (1998). Substrate specificity of deubiquitinating enzymes: ubiquitin C-terminal hydrolases. *Biochemistry* **37,** 3358–3368. https://doi.org/10.1021/bi972274d

Lawrence, D. S. (2005). The preparation and in vivo applications of caged peptides and proteins. In *Current Opinion in Chemical Biology* (Vol. 9, Issue 6, pp. 570–575). https://doi.org/10.1016/j.cbpa.2005.09.002

Leinfelder, W., Zehelein, E., MandrandBerthelot, M., & Bock, A. (1988). Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. *Nature*, *331*(6158), 723–725. https://doi.org/10.1038/331723a0

Lemke, E. A., Summerer, D., Geierstanger, B. H., Brittain, S. M., & Schultz, P. G. (2007). Control of protein phosphorylation with a genetically encoded photocaged amino acid. *Nature Chemical Biology*, *3*(12), 769–772. https://doi.org/10.1038/nchembio.2007.44

Li, M., Brooks, C.L., Kon, N., & Gu, W. (2004). A Dynamic Role of HAUSP in the p53-Mdm2 Pathway. *Mol. Cell.* 13, 879–886. https://doi.org/10.1016/S1097-2765(04)00157-1

Li, M., Chen, D., Shiloh, A., Luo, J., Nikolaev, A.Y., Qin, J., & Gu, W. (2002). Deubiquitination of p53 by HAUSP is an important pathway for p53 stabilization. *Nat. Cell Biol.* 416, 648–653. https://doi.org/10.1038/nature737

Li, Z. Q., Chen, X., & Wang, Y. (2021). Small molecules targeting ubiquitination to control inflammatory diseases. In *Drug Discovery Today* (Vol. 26, Issue 10, pp. 2414–2422). Elsevier Ltd. https://doi.org/10.1016/j.drudis.2021.04.029

Li, W., Tu, D., Brunger, A.T., & Ye, Y. (2007). A ubiquitin ligase transfers preformed polyubiquitin chains from a conjugating enzyme to a substrate. *Nature*. 446:333–37. https://doi.org/10.1038/nature05542

Liaunardy-Jopeace, A., Murton, B. L., Mahesh, M., Chin, J. W., & James, J. R. (2017). Encoding optical control in LCK kinase to quantitatively investigate its activity in live cells. *Nature Structural and Molecular Biology*, *24*(12), 1155–1163. https://doi.org/10.1038/nsmb.3492

Lin, T. C., Engelhard, L., Söldner, B., Linser, R., & Summerer, D. (2024). Light-Activatable MBD-Readers of 5-Methylcytosine Reveal Domain-Dependent Chromatin Association Kinetics In Vivo. *Advanced Science*, *11*(11). https://doi.org/10.1002/advs.202307930

Liu, C. C., & Schultz, P. G. (2010). Adding new chemistries to the genetic code. In *Annual Review of Biochemistry* (Vol. 79, pp. 413–444). https://doi.org/10.1146/annurev.biochem.052308.105824

Liu, C., Liu, W., Ye, Y. & Li, W. (2017) Ufd2p synthesizes branched ubiquitin chains to promote the degradation of substrates modified with atypical chains. *Nat.Commun*. 8, 14274. https://doi.org/10.1038/ncomms14274

Liu, D.S., et al. (2014). Computational design of a red fluorophore ligase for site-specific protein labelling in living cells. *Proc. Natl. Acad. Sci. U S A.* 111, E4551–E4559. https://doi.org/10.1073/pnas.1404736111

Liu, P., Gan, W., Su, S., Hauenstein, A. V, Fu, T.-M., Brasher, B., Schwerdtfeger, C., Liang, A. C., Xu, M., & Wei, W. (2018). D N A R E P A I R K63-linked polyubiquitin chains bind to DNA to facilitate DNA damage repair. In *Sci. Signal* (Vol. 11). https://www.science.org

Lo, C. L., Choudhury, S. R., Irudayaraj, J., & Zhou, F. C. (2017). Epigenetic editing of Ascl1 gene in neural stem cells by optogenetics. *Scientific Reports*, *7*. https://doi.org/10.1038/srep42047

Locke, M., Toth, J. I., & Petroski, M. D. (2014). Lys11- and Lys48-linked ubiquitin chains interact with p97 during endoplasmic-reticulum-associated degradation. *Biochemical Journal*, *459*(1), 205–216. https://doi.org/10.1042/BJ20120662

Lopez-Mosqueda, J. & Dikic, I. (2014) Deciphering functions of branched ubiquitin chains. *Cell*. 157, 767–769. https://doi.org/10.1016/j.cell.2014.04.026

Lu, H., Bhoopatiraju, S., Wang, H., Schmitz, N. P., Wang, X., Freeman, M. J., Forster, C. L., Verneris, M.R., Linden, M. A., & Hallstrom, T. C. (2016). *Loss of UHRF2 expression is associated with human neoplasia, promoter hypermethylation, decreased 5-hydroxymethylcytosine, and high proliferative activity* (Vol. 7, Issue 46). www.impactjournals.com/oncotarget

Lu, X., Zhao, B. S., & He, C. (2015). TET family proteins: Oxidation activity, interacting molecules, and functions in diseases. In *Chemical Reviews* (Vol. 115, Issue 6, pp. 2225–2239). American Chemical Society. https://doi.org/10.1021/cr500470n

Luo, J., Arbely, E., Zhang, J., Chou, C., Uprety, R., Chin, J. W., & Deiters, A. (2016). Genetically encoded optical activation of DNA recombination in human cells. *Chemical Communications*, *52*(55), 8529–8532. https://doi.org/10.1039/c6cc03934k

Luo, J., Uprety, R., Naro, Y., Chou, C., Nguyen, D. P., Chin, J. W., & Deiters, A. (2014). Genetically encoded optochemical probes for simultaneous fluorescence reporting and light activation of protein function with two-photon excitation. *Journal of the American Chemical Society*, *136*(44), 15551–15558. https://doi.org/10.1021/ja5055862

Lutz, J., Höllmüller, E., Scheffner, M., Marx, A., & Stengel, F. (2020). The Length of a Ubiquitin Chain: A General Factor for Selective Recognition by Ubiquitin-Binding Proteins. *Angewandte Chemie - International Edition*, *59*(30), 12371–12375. https://doi.org/10.1002/anie.202003058

Magnaghi, P., D'Alessio, R., Valsasina, B., Avanzi, N., Rizzi, S., Asa, D., Gasparri, F., Cozzi, L., Cucchi, U., Orrenius, C., Polucci, P., Ballinari, D., Perrera, C., Leone, A., Cervi, G., Casale, E., Xiao, Y., Wong, C., Anderson, D. J., … Isacchi, A. (2013). Covalent and allosteric inhibitors of the ATPase VCP/p97 induce cancer cell death. *Nature Chemical Biology*, *9*(9), 548–559. https://doi.org/10.1038/nchembio.1313

Mainz, E. R., Wang, Q., Lawrence, D. S., & Allbritton, N. L. (2016). An Integrated Chemical Cytometry Method: Shining a Light on Akt Activity in Single Cells. *Angewandte Chemie*, *128*(42), 13289–13292. https://doi.org/10.1002/ange.201606914

Maiti, A., & Drohat, A. C. (2011). Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: Potential implications for active demethylation of CpG sites. *Journal of Biological Chemistry*, *286*(41), 35334–35338. https://doi.org/10.1074/jbc.C111.284620

Manoilov, K. Y., Verkhusha, V. V., & Shcherbakova, D. M. (2021). A guide to the optogenetic regulation of endogenous molecules. In *Nature Methods* (Vol. 18, Issue 9, pp. 1027–1037). Nature Research. https://doi.org/10.1038/s41592-021-01240-1

Maytal-Kivity, V., Reis, N., Hofmann, K. & Glickman, M. H. (2002). MPN+, a putative catalytic motif found in a subset of MPN domain proteins from eukaryotes and prokaryotes, is critical for Rpn11 function. *BMC Biochem.* 3, 28. https://doi.org/10.1186/1471-2091-3-28

Metzger, M.B., Pruneda, J.N., Klevit, R.E., & Weissman, A.M. (2014) RING-type E3 ligases: master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. Biochim Biophys Acta. 1843(1):47–60. https://doi.org/10.1016/j.bbamcr.2013.05.026.

Mevissen, T. E. T., Hospenthal, M. K., Geurink, P. P., Elliott, P. R., Akutsu, M., Arnaudo, N., Ekkebus, R., Kulathu, Y., Wauer, T., El Oualid, F., Freund, S. M. V., Ovaa, H., & Komander, D. (2013). XOTU deubiquitinases reveal mechanisms of linkage specificity and enable ubiquitin chain restriction analysis. *Cell*, *154*(1), 169. https://doi.org/10.1016/j.cell.2013.05.046

Mevissen, T. E. T., & Komander, D. (2024). *Mechanisms of Deubiquitinase Specificity and Regulation*. https://doi.org/10.1146/annurev-biochem

Meyers, R. M. et al. (2017). Computational correction of copy number effect improves specificity of CRISPR- Cas9 essentiality screens in cancer cells. *Nat. Genet.* 49, 1779–1784. https://doi.org/10.1038/ng.3984

Mizushima, N. (2024). Ubiquitin in autophagy and non-protein ubiquitination. In *Nature Structural and Molecular Biology* (Vol. 31, Issue 2, pp. 208–209). Nature Research. https://doi.org/10.1038/s41594-024-01217-6

Möglich, A., Yang, X., Ayers, R. A., & Moffat, K. (2010). Structure and function of plant photoreceptors. *Annual Review of Plant Biology*, *61*, 21–47. https://doi.org/10.1146/annurev-arplant-042809-112259

Muñoz-López, Á., & Summerer, D. (2018). Recognition of Oxidized 5-Methylcytosine Derivatives in DNA by Natural and Engineered Protein Scaffolds. In *Chemical Record* (Vol. 18, Issue 1, pp. 105–116). John Wiley and Sons Inc. https://doi.org/10.1002/tcr.201700088

Münzel, M., Lischke, U., Stathis, D., Pfaffeneder, T., Gnerlich, F. A., Deiml, C. A., Koch, S. C., Karaghiosoff, K., & Carell, T. (2011). Improved synthesis and mutagenicity of oligonucleotides containing 5-hydroxymethylcytosine, 5-formylcytosine and 5-carboxylcytosine. *Chemistry - A European Journal*, *17*(49), 13782–13788. https://doi.org/10.1002/chem.201102782

Nagel, G., Szellas, T., Huhn, W., Kateriya, S., Adeishvili, N., Berthold, P., Ollig, D., Hegemann, P., & Bamberg, E. (2003). *Channelrhodopsin-2, a directly light-gated cation-selective membrane channel*. www.pnas.orgcgidoi10.1073pnas.1936192100

Nakagawa, T., Lv, L., Nakagawa, M., Yu, Y., Yu, C., D'Alessio, A. C., Nakayama, K., Fan, H. Y., Chen, X., & Xiong, Y. (2015). CRL4VprBP E3 ligase promotes monoubiquitylation and chromatin binding of TET dioxygenases. *Molecular Cell*, *57*(2), 247–260. https://doi.org/10.1016/j.molcel.2014.12.002

Nandi, D., Woodward, E., Ginsburg, D.B., Monaco, J.J. (1997). Intermediates in the formation of mouse 20S proteasomes: implications for the assembly of precursor beta subunits. *Embo J*. 16:5363–5375. https://doi.org/10.1093/emboj/16.17.5363

Nedelsky, N. B., Todd, P. K., & Taylor, J. P. (2008). Autophagy and the ubiquitin-proteasome system: Collaborators in neuroprotection. In *Biochimica et Biophysica Acta - Molecular Basis of Disease* (Vol. 1782, Issue 12, pp. 691–699). https://doi.org/10.1016/j.bbadis.2008.10.002

Neumann, H., Wang, K., Davis, L., Garcia-Alai, M., & Chin, J. W. (2010). Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature*, *464*(7287), 441–444. https://doi.org/10.1038/nature08817

Nguyen, D. P., Mahesh, M., Elsa sser, S. J., Hancock, S. M., Uttamapinant, C., & Chin, J. W. (2014). Genetic encoding of photocaged cysteine allows photoactivation of TEV protease in live mammalian cells. *Journal of the American Chemical Society*, *136*(6), 2240–2243. https://doi.org/10.1021/ja412191m

Nguyen, K.T., Mun, S., Lee, C., & Hwang, C. (2018). Control of protein degradation by N-terminal acetylation and the N-end rule pathway. *Experimental & Molecular Medicine*. 50:91. https://doi.org/10.1038/s12276-018-0097-y

Nihongaki, Y., Kawano, F., Nakajima, T., & Sato, M. (2015). Photoactivatable CRISPR-Cas9 for optogenetic genome editing. *Nature Biotechnology*, *33*(7), 755–760. https://doi.org/10.1038/nbt.3245

Nishi, R. et al. (2014). Systematic characterization of deubiquitylating enzymes for roles in maintaining genome integrity. *Nat. Cell Biol.* 16, 1016–1026. https://doi.org/10.1038/ncb3028

Ohtake, F., Tsuchiya, H., Saeki, Y. & Tanaka, K. (2018). K63 ubiquitylation triggers proteasomal degradation by seeding branched ubiquitin chains. *Proc. Natl Acad. Sci. USA.* 115, E1401–E1408. https://doi.org/10.1073/pnas.1716673115

Osei-Amponsa, V., & Walters, K. J. (2022). Proteasome substrate receptors and their therapeutic potential. In *Trends in Biochemical Sciences* (Vol. 47, Issue 11, pp. 950–964). Elsevier Ltd. https://doi.org/10.1016/j.tibs.2022.06.006

Oshikawa, K., Matsumoto, M., Oyamada, K., & Nakayama, K. I. (2012). Proteome-wide identification of ubiquitylation sites by conjugation of engineered lysine-less ubiquitin. *Journal of Proteome Research*, *11*(2), 796–807. https://doi.org/10.1021/pr200668y

Palei, S., Buchmuller, B., Wolffgramm, J., Muñoz-Lopez, Á., Jung, S., Czodrowski, P., & Summerer, D. (2020). Light-Activatable TET-Dioxygenases Reveal Dynamics of 5-Methylcytosine Oxidation and Transcriptome Reorganization. *Journal of the American Chemical Society*, *142*(16), 7289–7294. https://doi.org/10.1021/jacs.0c01193

Palombella, V.J., Rando, O.J., Goldberg, A.L., & Maniatis, T. (1994). The ubiquitinproteasome pathway is required for processing the NF-κB1 precursor protein and the activation of NF-κB. *Cell.* 78, 773 – 785. https://doi.org/10.1016/S0092-8674(94)90482-0

Partch, C. L., Green, C. B., & Takahashi, J. S. (2014). Molecular architecture of the mammalian circadian clock. In *Trends in Cell Biology* (Vol. 24, Issue 2, pp. 90–99). https://doi.org/10.1016/j.tcb.2013.07.002

Patchornik, A., Amit, B., & Woodward, R. B. (1970). Photosensitive protecting groups. *Journal of the American Chemical Society*, *92*(21), 6333–6335. https://doi.org/10.1021/ja00724a041

Peterson, L.F., et al. (2015). Targeting deubiquitinase activity with a novel small-molecule inhibitor as therapy for B-cell malignancies. *Blood*. 125, 3588–3597. https://doi.org/10.1182/blood-2014-10-605584

Petroski, M.D., & Deshaies, R.J. (2005). Mechanism of lysine 48-linked ubiquitin-chain synthesis by the cullin-RING ubiquitin-ligase complex SCF-Cdc34. *Cell* 123:1107–20. https://doi.org/10.1016/j.cell.2005.09.033

Pfaffeneder, T., Hackner, B., Truß, M., Mu nzel, M., Mu ller, M., Deiml, C. A., Hagemeier, C., & Carell, T. (2011). The discovery of 5-formylcytosine in embryonic stem cell DNA. *Angewandte Chemie - International Edition*, *50*(31), 7008–7012. https://doi.org/10.1002/anie.201103899

Pohl, C., & Dikic, I. (n.d.). *Cellular quality control by the ubiquitin-proteasome system and autophagy*. http://science.sciencemag.org/

Polycarpo, C., Ambrogelly, A., Be rube , A., Winbush, S. M., McCloskey, J. A., Crain, P. F., Wood, J. L., & Soll, D. (2004). An aminoacyl-tRNA synthetase that specifically activates pyrrolysine. *Proceedings of the National Academy of Sciences*, *101*(34), 12450–12454. https://doi.org/10.1073/pnas.0405362101

Prus, G., Satpathy, S., Weinert, B. T., Narita, T., & Choudhary, C. (2024). Global, site-resolved analysis of ubiquitylation occupancy and turnover rate reveals systems properties. *Cell*, *187*(11), 2875-2892.e21. https://doi.org/10.1016/j.cell.2024.03.024

Pudasaini, A., El-Arab, K. K., & Zoltowski, B. D. (2015). LOV-based optogenetic devices: light driven modules to impart photoregulated control of cellular signaling. *Frontiers in Molecular Biosciences*, *2*(MAY). https://doi.org/10.3389/fmolb.2015.00018

Pulvino, M., Liang, Y., Oleksyn, D., DeRan, M., Van Pelt, E., Shapiro, J., Sanz, I., Chen, L., & Zhao, J. (2012). Inhibition of proliferation and survival of diffuse large B-cell lymphoma cells by a small-molecule inhibitor of the ubiquitin-conjugating enzyme Ubc13-Uev1A. *Blood*, *120*(8), 1668–1677. https://doi.org/10.1182/blood-2012-02-406074

Qi, L., Tsai, B., & Arvan, P. (2017). New Insights into the Physiological Role of Endoplasmic Reticulum-Associated Degradation. In *Trends in Cell Biology* (Vol. 27, Issue 6, pp. 430–440). Elsevier Ltd. https://doi.org/10.1016/j.tcb.2016.12.002

Rackham, O., & Chin, J. W. (2005). A network of orthogonal ribosome·mrna pairs. *Nature Chemical Biology*, *1*(3), 159–166. https://doi.org/10.1038/nchembio719

Rahighi, S., Ikeda, F., Kawasaki, M., Akutsu, M., Suzuki, N., et al. 2009. Specific recognition of linear ubiquitin chains by NEMO is important for NF-kappaB activation. Cell 136:1098–109. https://doi.org/10.1016/j.cell.2009.03.007

Rakauskaite, R., Urbanavic iute, G., Ruks enaite, A., Liutkevic iute, Z., Jus kenas, R., Masevic ius, V., & Klimas auskas, S. (2015). Biosynthetic selenoproteins with genetically-encoded photocaged selenocysteines. *Chemical Communications*, *51*(39), 8245–8248. https://doi.org/10.1039/c4cc07910h

Rao, V.K., et al. (2020). Phosphorylation of Tet3 by cdk5 is critical for robust activation of BRN2 during neuronal differentiation. *Nucleic Acids Research*. 48 (3), 1225-1238. https://doi.org/10.1093/nar/gkz1144

Razin, A., & Riggs, A. D. (1980). DNA Methylation and Gene Function. *Science*, *210*(4470), 604–610. https://www.science.org

Redman, K. L. & Rechsteiner, M. Identification of the long ubiquitin extension as ribosomal protein S27a. *Nature* 338**,** 438–440 (1989). https://doi.org/10.1038/338438a0

Rentsch, A., Landsberg, D., Brodmann, T., low, L.B., Girbig, A.K., & Kalesse, M. (2013). Synthesis and Pharmacology of Proteasome Inhibitors. *Angew. Chem. Int. Ed.* 52, 5450 – 5488. https://doi.org/10.1002/anie.201207900

Reiner, T., Parrondo, R., Pozas, A.D.L., Palenzuela, D., & Perez-Stable, C. (2013) Betulinic acid selectively increases protein degradation and enhances prostate cancer-specific apoptosis: possible role for inhibition of deubiquitinase activity. *PLoS ONE.* 8, e56234. https://doi.org/10.1371/journal.pone.0056234

Resnick, E., et al. (2019). Rapid covalent-probe discovery by electrophile-fragment screening. *J. Am. Chem. Soc*. 141, 8951–8968. https://doi.org/10.1021/jacs.9b02822

Reyes-Turcu, F.E., Horton, J.R., Mullally, J.E., Heroux, A., Cheng, X., & Wilkinson, K.D. (2006). The ubiquitin binding domain ZnF UBP recognizes the C-terminal diglycine motif of unanchored ubiquitin. *Cell* 124:1197–208. https://doi.org/10.1016/j.cell.2006.02.038

Richardson, L. A. et al. (2012). A conserved deubiquitinating enzyme controls cell growth by regulating RNA polymerase I stability. *Cell Rep.* 2, 372–385. https://doi.org/10.1016/j.celrep.2012.07.009

Rodrigo-Brenni, M.C., Foster, S.A., & Morgan, D.O. (2010). Catalysis of lysine 48-specific ubiquitin chain assembly by residues in E2 and ubiquitin. *Mol. Cell.* 39:548–59. https://doi.org/10.1016/j.molcel.2010.07.027

Ross, A. B., Langer, J. D., & Jovanovic, M. (2021). Proteome turnover in the spotlight: Approaches, applications, and perspectives. In *Molecular and Cellular Proteomics* (Vol. 20). American Society for Biochemistry and Molecular Biology Inc. https://doi.org/10.1074/MCP.R120.002190

Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T. & Eliceiri, K. W. (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics. 18*, 529. https://doi.org/10.1186/s12859-017-1934-z

Rusk, N. (2011). Programming molecular instruments. In *Nature Methods* (Vol. 8, Issue 1, p. 11). https://doi.org/10.1038/nmeth0111-11

Ryu, M.Y., Cho, S.K., Hong, Y., Kim, J., Kim, J.H., Kim, G.M., *et al*. (2019). Classification of barley U-box E3 ligases and their expression patterns in response to drought and pathogen stresses. BMC Genomics. 20(1):326. https://doi.org/10.1186/s12864-019-5696-z

Sakamoto, K., Kim, K.B., Kumagai, A., Mercurio, F., Crews, C.M., & Deshaies, R.J. (2001). Protacs: Chimeric molecules that target proteins to the Skp1–Cullin–F box complex for ubiquitination and degradation. *Proc. Natl. Acad. Sci. USA*. 98, 8554 – 8559. https://doi.org/10.1073/pnas.141230798

Sanders, M. A., Brahemi, G., Nangia-Makker, P., Balan, V., Morelli, M., Kothayer, H., Westwell, A. D., & Shekhar, M. P. V. (2013). Novel inhibitors of Rad6 ubiquitin conjugating enzyme: Design, synthesis, identification, and functional characterization. *Molecular Cancer Therapeutics*, *12*(4), 373–383. https://doi.org/10.1158/1535-7163.MCT-12-0793

*science.346.6209.568*. (n.d.).

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat Methods.* 9, 671-675. https://doi.org/10.1038/nmeth.2089

Schoenheimer, R. (1946). The Dynamic State of Body Constituents. Harvard University Press, Cambridge, MA.

Seeger, M., et al. (2003). Interaction of the anaphase-promoting complex/cyclosome and proteasome protein complexes with multiubiquitin chain-binding proteins. *J. Biol. Chem.* 278:16791–96. https://doi.org/10.1074/jbc.M208281200

Shi, F. T., et al. (2013). Ten-Eleven Translocation 1 (Tet1) Is Regulated by *O*-Linked *N*-acetylglucosamine Transferase (OGT) for Target Gene Repression in Mouse Embryonic Stem Cells. *J. Biol. Chem*. 288 (29), 20776-20784. https://doi.org/10.1074/jbc.M113.460386

Shimizu-Sato, S., Huq, E., Tepperman, J. M., & Quail, P. H. (2002). A light-switchable gene promoter system. *Nature Biotechnology*, *20*(10), 1041–1044. https://doi.org/10.1038/nbt734

Sloper-Mould, K.E., Jemc, J.C., Pickart, C.M., & Hicke, L. 2001. Distinct functional surface regions on ubiquitin. *J. Biol. Chem.* 276:30483–89. https://doi.org/10.1074/jbc.M103248200

Smith, Z. D., & Meissner, A. (2013). DNA methylation: Roles in mammalian development. *Nature Reviews Genetics*, *14*(3), 204–220. https://doi.org/10.1038/nrg3354

Soucy, T. A., et al. (2009). An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature*, *458*(7239), 732–736. https://doi.org/10.1038/nature07884

Srinivasan, G., James, C. M., & Krzycki, J. A. (2002). Pyrrolysine Encoded by UAG in Archaea: Charging of a UAG-Decoding Specialized tRNA. *Science*, *296*(5572), 1459–1462. https://doi.org/10.1126/science.1069588

Strickland, D., Moffat, K., & Sosnick, T. R. (2008). *Light-activated DNA binding in a designed allosteric protein*. www.pymol.org

Swatek, K. N., & Komander, D. (2016). Ubiquitin modifications. In *Cell Research* (Vol. 26, Issue 4, pp. 399–422). Nature Publishing Group. https://doi.org/10.1038/cr.2016.39

Swords, R. T., et al. (2015). Pevonedistat (MLN4924), a First-in-Class NEDD8-activating enzyme inhibitor, in patients with acute myeloid leukaemia and myelodysplastic syndromes: a phase 1 study. *Br. J. Haematol.* 169, 534–543. https://doi.org/10.1111/bjh.13323

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L., & Rao, A. (2009). Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science*, *324*(5929), 930–935. https://doi.org/10.1126/science.1169786

Tarhonskaya, H., Rydzik, A. M., Leung, I. K. H., Loik, N. D., Chan, M. C., Kawamura, A., McCullagh, J. S. O., Claridge, T. D. W., Flashman, E., & Schofield, C. J. (2014). Non-enzymatic chemistry enables 2-hydroxyglutarate-mediated activation of 2-oxoglutarate oxygenases. *Nature Communications*, *5*. https://doi.org/10.1038/ncomms4423

Taslimi, A., Vrana, J. D., Chen, D., Borinskaya, S., Mayer, B. J., Kennedy, M. J., & Tucker, C. L. (2014). An optimized optogenetic clustering tool for probing protein interaction and function. *Nature Communications*, *5*. https://doi.org/10.1038/ncomms5925

Taslimi, A., Zoltowski, B., Miranda, J. G., Pathak, G. P., Hughes, R. M., & Tucker, C. L. (2016). Optimized second-generation CRY2-CIB dimerizers and photoactivatable Cre recombinase. *Nature Chemical Biology*, *12*(6), 425–430. https://doi.org/10.1038/nchembio.2063

Tatavarty, V., Kim, E.J., Rodionov, V., & Yu, J. (2009). Investigating subspine actin dynamics in rat hippocampal neurons with super-resolution optical imaging. *PLoS One*. 4, e7724. https://doi.org/10.1371/journal.pone.0007724

Tenno, T., Fujiwara, K., Tochio, H., Iwai, K., Morita, E.H., *et al.* 2004. Structural basis for distinct roles of Lys63- and Lys48-linked polyubiquitin chains. *Genes Cells* 9:865–75. https://doi.org/10.1111/j.1365-2443.2004.00780.x

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

Toettcher, J. E., Weiner, O. D., & Lim, W. A. (2013). XUsing optogenetics to interrogate the dynamic control of signal transmission by the Ras/Erk module. *Cell*, *155*(6). https://doi.org/10.1016/j.cell.2013.11.004

Tokunaga, F., et al. (2009). Involvement of linear polyubiquitylation of NEMO in NF-kappaB activation. *Nat. Cell Biol.* 11:123–32. https://doi.org/10.1038/ncb1821

tom Dieck, S., Kochen, L., Hanus, C., Heumüller, M., Bartnik, I., Nassim-Assir, B., Merk, K., Mosler, T., Garg, S., Bunse, S., Tirrell, D. A., & Schuman, E. M. (2015) Direct visualization of newly synthesized target proteins in situ. *Nat. Methods.* 12, 411–414. https://doi.org/10.1038/nmeth.3319

Tomko, R. J., & Hochstrasser, M. (2013). Molecular architecture and assembly of the eukaryotic proteasome. In *Annual Review of Biochemistry* (Vol. 82, pp. 415–445). Annual Reviews Inc. https://doi.org/10.1146/annurev-biochem-060410-150257

Tran, H. J., Allen, M. D., Lowe, J. & Bycroft, M. (2003). Structure of the Jab1/MPN domain and its implications for proteasome function. *Biochemistry.* 42, 11460–11465. https://doi.org/10.1021/bi035033g

Uttamapinant, C., White, K.A., Baruah, H., Thompson, S., Fernandez-Suarez, M., Puthenveetil, S., & Ting, A.Y. (2010). A fluorophore ligase for site-specific protein labeling inside living cells. *Proc. Natl. Acad. Sci. U S A.* 107, 10914–10919. https://doi.org/10.1073/pnas.091406710

van den Boom, J. & Meyer, H. (2018). VCP/p97-mediated unfolding as a principle in protein homeostasis and signaling. Mol. *Cell.* 69, 182–194. https://doi.org/10.1016/j.molcel.2017.10.028

Vijay-Kumar3, S., Bugg, C. E., & Cook, W. J. (1987). Structure of Ubiquitin Refined at 143 A Resolution. In *I. Mol. Bid* (Vol. 194).

Walden, M., Masandi, S. K., Pawlowski, K. & Zeqiraj, E. (2018). Pseudo- DUBs as allosteric activators and molecular scaffolds of protein complexes. *Biochem. Soc. Trans.* 46, 453–466. https://doi.org/10.1042/bst20160268

Walker, J. W., Martin, H., Schmitt, F. R., & Barsotti, R. J. (1993). Rapid release of an .alpha.- adrenergic receptor ligand from photolabile analogs. *Biochemistry*, *32*(5), 1338–1345. https://doi.org/10.1021/bi00056a020

Wang, C., Deng, L., Hong, M., Akkaraju, G.R., Inoue, J., & Chen, Z.J. (2001). TAK1 is a ubiquitin-dependent kinase of MKK and IKK. *Nature* 412:346–51. https://doi.org/10.1038/35085597

Wang, J., et al. (2019). Time-resolved protein activation by proximal decaging in living systems. Nature. 569: 509–513. https://doi.org/10.1038/s41586-019-1188-1

Wang, K., Neumann, H., Peak-Chew, S. Y., & Chin, J. W. (2007). Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. *Nature Biotechnology*, *25*(7), 770–777. https://doi.org/10.1038/nbt1314

Wang, M., & Pickart, C.M. (2005). Different HECT domain ubiquitin ligases employ distinct mechanisms of polyubiquitin chain synthesis. *EMBO J.* 24:4324–33. https://doi.org/10.1038/sj.emboj.7600895

Wang, Y., & Zhang, Y. (2014). Regulation of TET Protein Stability by Calpains. *Cell Reports.* 6, 278–284. http://dx.doi.org/10.1016/j.celrep.2013.12.031

Weber, A. R., Krawczyk, C., Robertson, A. B., Kusnierczyk, A., Va gbø, C. B., Schuermann, D., Klungland, A., & Schar, P. (2016). Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. *Nature Communications*, *7*. https://doi.org/10.1038/ncomms10806

Wenzel, D.M., Lissounov, A., Brzovic, P.S., & Klevit, R.E. (2011). UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids. *Nature* 474:105–8. https://doi.org/10.1038/nature09966

Wilkins, B. J., Marionni, S., Young, D. D., Liu, J., Wang, Y., di Salvo, M. L., Deiters, A., & Cropp, T. A. (2010). Site-specific incorporation of fluorotyrosines into proteins in *Escherichia coli* by photochemical disguise. *Biochemistry*, *49*(8), 1557–1559. https://doi.org/10.1021/bi100013s

Wilkinson, D. J. (2018). Historical and contemporary stable isotope tracer approaches to studying mammalian protein metabolism. *Mass Spectrom. Rev.* 37, 57–80. https://doi.org/10.1002/mas.21507

Williamson, A. (2011). Regulation of ubiquitin chain initiation to determine the timing of substrate degradation. *Mol. Cell* 42:744–57. https://doi.org/10.1016/j.molcel.2011.04.022

Williamson, A., Wickliffe, K.E., Mellone, B.G., Song, L., Karpen, G.H., & Rape, M. (2009). Identification of a physiological E2 module for the human anaphase-promoting complex. *Proc. Natl. Acad. Sci. USA* 106:18213–18. https://doi.org/10.1073/pnas.0907887106

Winborn, B. J. et al. (2008) The deubiquitinating enzyme ataxin-3, a polyglutamine disease protein, edits Lys63 linkages in mixed linkage ubiquitin chains. *J. Biol. Chem.* 283, 26436–26443. https://doi.org/10.1074/jbc.M803692200

Wolffgramm, J., Buchmuller, B., Palei, S., Muñoz-López, Á., Kanne, J., Janning, P., Schweiger, M. R., & Summerer, D. (2021). Light-Activation of DNA-Methyltransferases. *Angewandte Chemie - International Edition*, *60*(24), 13507–13512. https://doi.org/10.1002/anie.202103945

Wu, D., et al. (2018). Glucose-regulated phosphorylation of TET2 by AMPK reveals a pathway linking diabetes to cancer. *Nature*. 559, 637- 658. https://doi.org/10.1038/s41586-018-0350-5

Wu, K., Kovacev, J., & Pan, Z.Q. (2010). Priming and extending: a UbcH5/Cdc34 E2 handoff mechanism for polyubiquitination on a SCF substrate. *Mol. Cell* 37:784–96. https://doi.org/10.1016/j.molcel.2010.02.025

Wu, N., Deiters, A., Cropp, T. A., King, D., & Schultz, P. G. (2004). A genetically encoded photocaged amino acid. Journal of the American Chemical Society, 126(44), 14306–14307. https://doi.org/10.1021/ja040175z

Xie, Y., & Varshavsky, A. (2002). UFD4 lacking the proteasome-binding region catalyses ubiquitination but is impaired in proteolysis. *Nat. Cell Biol.* 4:1003–7. https://doi.org/10.1038/ncb889

Xue, Q., Tang, D., Chen, X., & Liu, J. (2023). Targeting deubiquitinases for cancer therapy. In *Clinical and Translational Discovery* (Vol. 3, Issue 5). Blackwell Publishing. https://doi.org/10.1002/ctd2.242

Yang, Q., Zhao, J., Chen, D., & Wang, Y. (2021). E3 ubiquitin ligases: styles, structures and functions. In *Molecular Biomedicine* (Vol. 2, Issue 1). Springer. https://doi.org/10.1186/s43556-021-00043-2

Yao, Y., Toth, C.R., Huang, L., Wong, M.L., Dias, P., et al. (1999). alpha5 subunit in Trypanosoma brucei proteasome can self-assemble to form a cylinder of four stacked heptamer rings. *Biochem J*. 344(Pt 2):349–358. https://doi.org/10.1042/0264-6021:3440349

Yau, R. G., Doerner, K., Castellanos, E.R., Haakonsen, D.L., *et al*. (2017). Assembly and function of heterotypic ubiquitin chains in cell cycle and protein quality control. *Cell*. 171, 918–933.e920. https://doi.org/10.1016/j.cell.2017.09.040

Ye, Y., Akutsu, M., Reyes-Turcu, F., Enchev, R.I., Wilkinson, K.D., & Komander, D. 2011. Polyubiquitin binding and cross-reactivity in the USP domain deubiquitinase USP21. *EMBO Rep.* 12:350–57. https://doi.org/10.1038/embor.2011.17

Yee, E.F. et al. (2015). Signal transduction in light-oxygen-voltage receptors lacking the adduct forming cysteine residue. *Nat. Commun*. 6, 10079. https://doi.org/10.1038/ncomms10079

Yin, X., & Xu, Y. (2016). DNA Methyltransferases - Role and Function. *Advances in Experimental Medicine and Biology*. 945, 275-302. https://doi.org/10.1007/978-3-319-43624-1_12

Young, T. S., Ahmad, I., Yin, J. A., & Schultz, P.G. (2010). An enhanced system for unnatural amino acid mutagenesis in E. coli. *J Mol Biol.* 395, 361-374. https://doi.org/10.1016/j.jmb.2009.10.030

Zangi, R., Arrieta, A., & Cossí o, F. P. (2010). Mechanism of DNA Methylation: The Double Role of DNA as a Substrate and as a Cofactor. *Journal of Molecular Biology*, *400*(3), 632–644. https://doi.org/10.1016/J.JMB.2010.05.021

Zeng, W., et al. (2010). Reconstitution of the RIG-I pathway reveals a signaling role of unanchored polyubiquitin chains in innate immunity. *Cell* 141:315–30. https://doi.org/10.1016/j.cell.2010.03.029

Zhang, Q., et al. (2014). Differential Regulation of the Ten-Eleven Translocation (TET) Family of Dioxygenases by *O*-Linked_-*N*-Acetylglucosamine Transferase (OGT). J. Biol. Chem. 289 (9), 5986-5996. https://doi.org/10.1074/jbc.M113.524140

Zhang, T., et al. (2019). Phosphorylation of TET2 by AMPK is indispensable in myogenic differentiation. *Epigenetics & Chromatin*. 12, 32. https://doi.org/10.1186/s13072-019-0281-x

Zhang, Y. W., et al. (2017). Acetylation Enhances TET2 Function in Protecting against Abnormal DNA Methylation during Oxidative Stress. *Mol. Cell*. 65, 323–335. http://dx.doi.org/10.1016/j.molcel.2016.12.013

Zhao, J., Zhao, J., Lin, S., Huang, Y., & Chen, P. R. (2013). Mechanism-based design of a photoactivatable firefly luciferase. *Journal of the American Chemical Society*, *135*(20), 7410–7413. https://doi.org/10.1021/ja4013535

Zhu, H., Ding, G., Liu, X., & Huang, H. (2023). Developmental origins of diabetes mellitus: Environmental epigenomics and emerging patterns. In *Journal of Diabetes* (Vol. 15, Issue 7, pp. 569–582). John Wiley and Sons Inc. https://doi.org/10.1111/1753-0407.13403

Zimmerman, G., Chow, L.-Y., & Paik, U.-J. (1958). The Photochemical Isomerization of Azobenzene[1]. *Journal of the American Chemical Society*, *80*(14), 3528–3531. https://doi.org/10.1021/ja01547a010