Proteomics and Functional Investigation of SUMO and Ubiquitin E3 ligases

Salas Lloret, D.

Proteomics and functional investigation of SUMO and ubiquitin E3 ligases

Daniel Salas Lloret

Proteomics and functional investigation of SUMO and ubiquitin E3 ligases

Proefschrift

ter verkrijging van

de graad van doctor aan de Universiteit Leiden,

op gezag van rector magnificus prof.dr.ir. H. Bijl,

volgens besluit van het college voor promoties

te verdedigen op dinsdag 10 oktober 2023

klokke 11.15 uur

door

Daniel Salas Lloret

geboren te Alicante, Spanje

in 1994

Promotor

Prof. Dr. A.C.O. Vertegaal

Co-promotor

Dr. R. González Prieto

Members of the promotion committee

Prof. Dr. P. ten Dijke	LUMC
Prof. Dr. M. Tijsterman	LUMC
Dr. S.M. Noordermeer	LUMC
Prof. Dr. M.A.T.M. van Vugt	UMCG
Dr. J.J.L. Jacobs	NKI

PhD Thesis, Leiden University, 2023.

The research described in this thesis was mostly performed at Leiden University Medical Center (LUMC), Department of Cell and Chemical Biology (CCB), Leiden, The Netherlands. This research was primarily supported by the Dutch Cancer Society (KWF; grant 11367), the European Research Council (ERC; grant 310913) and the Dutch Research Council (NWO; grant 724.016.003).

Cover design and layout by: Daniel Salas Lloret

Printed by: Ridderprint

ISBN: 978-94-6483-416-1

Table of Contents

Summary

In this thesis, we focus on the post-translational modification of proteins (PTM) with special interest in Ubiquitination and SUMOylation. We describe both the Ubiquitin and SUMO E3 ligase families, showing not only their different mechanisms to conjugate Ubiquitin or SUMO respectively to a substrate protein, but also the signaling pathways they are involved in. E3s can conjugate one or several moieties to a substrate protein giving rise to polymers. Depending on the chain type and polymer morphology, the substrate protein will lead to different cellular outcomes. At the end of chapter 1, we analyze current Mass-Spectrometry (MS) approaches to identify both Ubiquitin and SUMO conjugates with their advantages and limitations.

In chapter 2, we review recent literature about the DNA damage repair (DDR) and how it is orchestrated by different E3 ligases. Among these E3s, we focus on BRCA1-BARD1 as one of the main coordinators of the DDR and its implications in breast and ovarian cancers. The BRCA1-BARD1 heterodimer is formed by the interaction of BRCA1 and BARD1 through their Really Interesting New Gene (RING) domains. The formation of this heterodimer confers its only known intrinsic enzymatic activity, with only one well-defined ubiquitination target, the histone H2A. After more than 30 years studying BRCA1 protein, its E3 ligase activity remains unclear and its role in homologous recombination (HR) repair pathway and its tumor suppression role is still unresolved. We discuss the controversy about this heterodimer and the current directions in the field. To finish this chapter, we disclose several ways to target the ubiquitination machinery and control the DDR in order to tackle cancer.

In chapter 3, we developed a Mass-Spectrometry methodology that we termed Targets for Ubiquitin Ligases Identified by Proteomics 2 (TULIP2). There are many diseases including neurodegenerative disorders and cancers that come from dysregulation of E3 ligases. TULIP2 allows the identification of specific targets for these E3 ligases using MS and offers new possibilities to target a particular disease. Compared to previous methodologies, TULIP2 increases more than 50-fold the purification yield of ubiquitination conjugates and 8-fold the signal of the E3 ligase under study. Overall, this chapter offers a clear and detailed methodology that can be implemented in any laboratory interested in the identification of targets for an E3 ligase of interest.

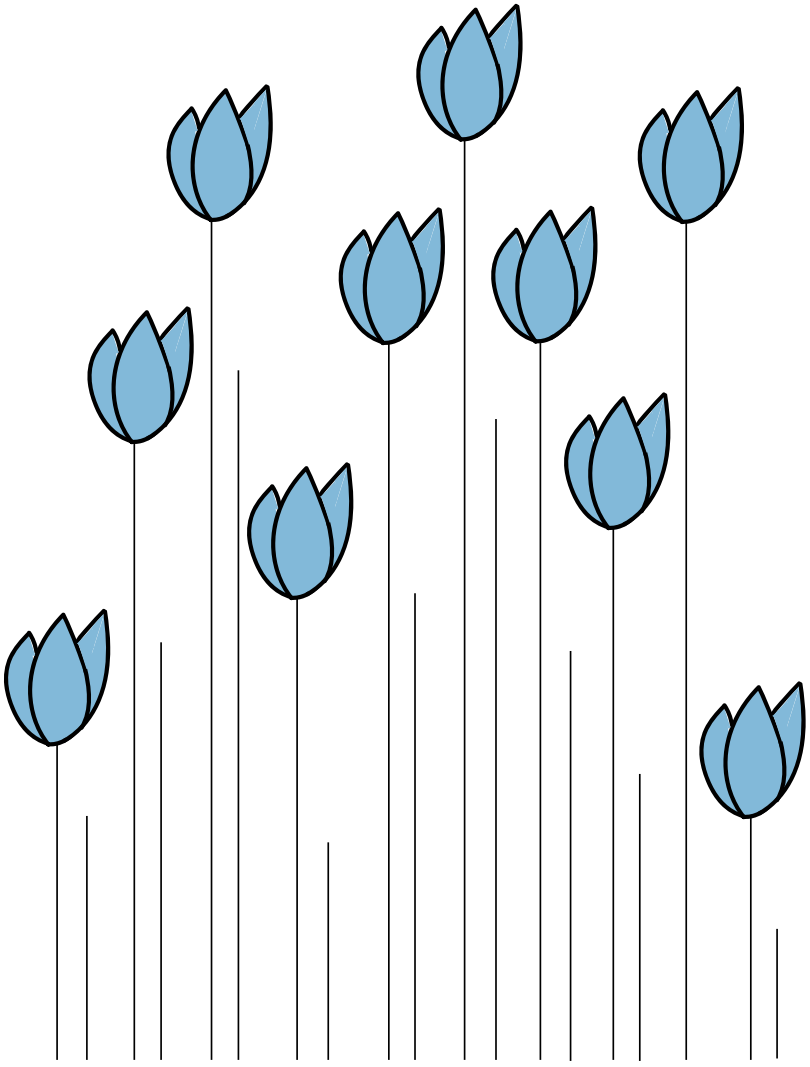
TULIP2 methodology can be modified to trap SUMO conjugates by changing the fused ubiquitin by a SUMO moiety. In chapter 4, we developed SUMO Activated Target Traps (SATTs), which allowed the identification of an E3-specific SUMO proteome. Employing eight different SUMO E3 ligases (PIAS1, PIAS2, PIAS3, PIAS4, NSMCE2, ZNF451, LAZSUL (ZNF451-3) and ZMIZ2) we identified 427 SUMO1 and 961 SUMO2/3 potential targets in an E3-specific manner. Although we found redundancy between E3 ligases, we also found high specificity, even at the substrate isoform level. To browse the dataset in an user-

friendly way, we developed the Polar Volcano plots online web app, which is freely accessible (https://amsterdamstudygroup.shinyapps.io/polarVolcaNoseR_revised/).

Next, we employed TULIP2 technology to study the BRCA1-BARD1 E3 ligase. In chapter 5, we generated both BRCA1-TULIP2 and BARD1-TULIP2 to search BRCA1-BARD1-specific ubiquitination targets. BRCA1 and BARD1 germline mutations are associated with high risk of developing breast and ovarian cancer. Finding BRCA1-BARD1 ubiquitination targets is crucial to understand its E3 activity and therefore, new possibilities for cancer treatment. Performing TULIP2, we found that BARD1 preferentially directs the ubiquitination of targets differing from BRCA1. In addition to its well-known target H2A, we found that the macro H2A (mH2A) variant is a specific target, and PCNA seems to be the main ubiquitination target during normal grow conditions. We report that BARD1-mediated PCNA ubiquitination is involved in the prevention of single stranded DNA (ssDNA) accumulation in contrast to RAD18-mediated PCNA ubiquitination, which is involved in UV-induced damage repair. Additionally, we found that the BRCA1 catalytic dead mutant is not phosphorylated at S114, leading to a compromised replication fork protection. Overall, we conclude that BRCA1-BARD1 E3 activity is important to keep genetic stability by controlling replication fork homeostasis instead of being involved in HR pathway, which opens new therapeutic opportunities.

In chapter 6, we aimed to find specific targets for an E2 conjugating enzyme, UBE2D3. This E2 has been reported to interact with BRCA1-BARD1 at DNA damage sites and with RNF8 for PCNA ubiquitination. Commonly, it has been used as promiscuous E2 for *in vitro* ubiquitination. However, the *in vivo* role has not been defined yet. In chapter 6, we performed SILAC-based and label-free quantitative ubiquitin diGly proteomics, together with TULIP2 technology, to study global proteome and ubiquitinome changes associated with UBE2D3 depletion. The ubiquitin proteome completely changed after UBE2D3 depletion, having a prominent effect on molecular pathways related to mRNA translation. With UBE2D3-TULIP2 methodology we were able to detect, in addition to PCNA as already known target, two ribosomal proteins (RPS10 and RPS20) as specific targets, which are critical for ribosome-associated protein quality control (RQC). Ubiquitination of RPS10 and RPS20 was also UBE2D3-dependent in SILAC and diGly proteomics, revealing a new *in vivo* role for UBE2D3.

In the last chapter, we discuss about different MS approaches in shotgun or bottom-up proteomics regarding sample complexity and challenges in finding E3 specific targets. Then, we explain the potential and limitations of the TULIP2 technology developed in this thesis and discuss how all our findings can be of potential use in the clinic. Finally, we contrast our findings on the SUMO field with published literature to show how a previously believed small post-translational modification network could become very complex and sophisticated.



1

Insights in Post-translational modifications: Ubiquitin and SUMO

Daniel Salas-Lloret[1] and Román González-Prieto[1,2]

[1] Cell and Chemical Biology, Leiden University Medical Centre, Einthovenweg 20, 2333ZC Leiden, The Netherlands

[2] Genome Proteomics Group, Department of Genome Biology, Andalusian Centre for Regenerative Medicine and Molecular Biology (CABIMER), Américo Vespucio 24, 41092 Seville, Spain

This chapter has been published in International Journal of Molecular Sciences. 23, 3281 (2022).

Abstract

Both ubiquitination and SUMOylation are dynamic post-translational modifications that regulate thousands of target proteins to control virtually every cellular process. Unfortunately, the detailed mechanisms of how all these cellular processes are regulated by both modifications remain unclear. Target proteins can be modified by one or several moieties, giving rise to polymers of different morphology. The conjugation cascades of both modifications comprise a few activating and conjugating enzymes but close to thousands of ligating enzymes (E3s) in the case of ubiquitination. As a result, these E3s give substrate specificity and can form polymers on a target protein. Polymers can be quickly modified forming branches or cleaving chains leading the target protein to its cellular fate. The recent development of mass spectrometry(MS) -based approaches has increased the understanding of ubiquitination and SUMOylation by finding essential modified targets in particular signaling pathways. Here, we perform a concise overview comprising from the basic mechanisms of both ubiquitination and SUMOylation to recent MS-based approaches aimed to find specific targets for particular E3 enzymes.

Keywords: Ubiquitin; SUMO; E3 enzymes; Proteomics

1. INTRODUCTION

The environment is constantly changing. Consequently, cells need to adapt and give a quick and appropriate response against different stimuli or stresses. Proteins control the vast majority of cellular processes, and their function is regulated by a variety of posttranslational modifications (PTMs). Commonly, PTMs are covalent enzymatic modifications at the amino acid chain or termini of proteins that may happen after translation and can affect the function of a target protein, including the activity, folding, localization, interaction partners, and protein homeostasis. PTMs can be completed in seconds [1] and are generally reversible by the action of specific deconjugating enzymes, which enable the cells to provide a rapid and precise response to environmental changes. More than 200 different PTMs have been described, which can either consist of the addition of a small chemical group (e.g., acetylation, methylation, phosphorylation, etc.), the modification by complex molecules (i.e., glycosylation, AMPylation, ADPribosylation, prenylation, etc.) or the addition of long polypeptide chains such as ubiquitin or ubiquitin-like modifiers (UBLs). Additionally, irreversible PTMs exist (i.e., proteolysis, deamidation, etc.) and recent research has reported the role of deamidation in the regulation of ubiquitin E3 enzymes [2] (**Figure 1**).

In this review, we provide a general overview of the current tendencies in the ubiquitination and SUMOylation fields, discussing the basic principles of these modifications and their importance in complex signaling pathways. We compare the enzymes involved in the catalytic cycle of both PTMs and analyze different approaches to identify target proteins of these modifications. Finally, we discuss the future directions in the field with the use of new inhibitors, genome-wide screenings, and MS-based proteomics approaches.

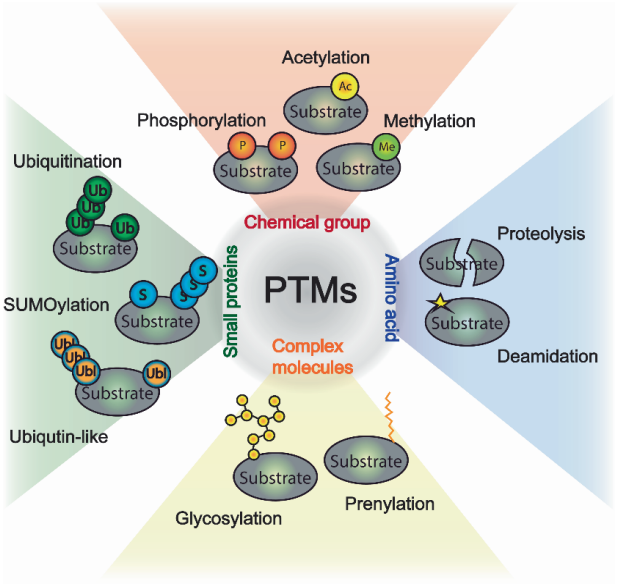


Figure 1. Post-translational modifications of proteins. Chemical group modifications are shown in red, amino acid modifications in blue, complex molecules are in yellow, and additions of small proteins are displayed in green.

1

2. UBIQUITINATION

Ubiquitination, namely, the covalent attachment of ubiquitin (Ub) to acceptor residues in target proteins is, after phosphorylation, the second PTM in abundance [3] and regulates virtually all events in cells. Ubiquitin received its name from being ubiquitous in all cell types among eukaryotes. It is a highly conserved 76 amino acids protein which differs by only 2 amino acids from yeast to humans and by 3 amino acids from plants to humans. However, Ub has not been identified in procaryotes. Although, its origins may reside in bacteria protein molybdopterin converting factor subunit 1(MoaD) and thiamine biosynthesis protein S (ThiS). These proteins share the same structure with Ub but are mainly involved in biosynthetic pathways [4,5]. Years later, other bacterial proteins such as the small prokaryotic ubiquitinlike protein (Pup) from Mycobacterium tuberculosis have also been described. Similar to Ub, Pup becomes conjugated to 26S proteasome targets, sharing comparable function to ubiquitin in eukaryotes [6]. The prokaryotic ancestor may not only create Ub, but also ubiquitin-like proteins including SUMO, NEDD8, and ISG15. In total, 20 human UBLs have been reported to be conjugated to other molecules. They share a common overall fold and a 3-step conjugation cascade consisting of their own E1, E2, and E3 enzymes. Despite their similarity, they impart distinct functions to their substrate proteins [7] and participate together coordinating different cellular functions [8–11]. Most eukaryotic genomes contain multiple Ub genes. In humans, Ub is encoded by four genes: UBB, UBC, UBA52, and RPS27A. UBA52 and RPS27A encode single copies of Ub, while UBB and UBC encode polyubiquitin chains that are cleaved by the Ubiquitin Specific Protease USP5 to produce monomeric active Ub molecules. These monomeric active Ub molecules fold into a compact and globular structure with a terminal diglycine (diGly) sequence essential for its conjugation [12]. During the ubiquitination cascade, the activating enzyme (E1) hydrolyzes ATP to form a thioester bond with the C-terminal of ubiquitin. Then, ubiquitin is transferred via the thioester-like complex to the ubiquitin-conjugating enzyme (E2). Finally, the ubiquitin ligase enzyme (E3) mediates the conjugation of the ubiquitin moiety to either a lysine residue or the extreme amino terminus of the targeted protein in a highly controlled manner (**Figure 2**) [13–17]. All research focused on the ubiquitination cascade and the ubiquitination field was recognized with the Nobel Prize in Chemistry in 2004 awarded to Aaron Ciechanover, Avram Hershko, and Irwin Rose by the Royal Swedish Academy of Sciences.

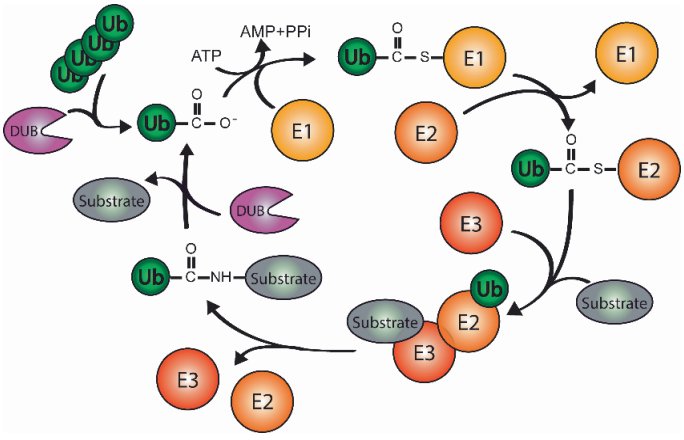


Figure 2. Ubiquitination cascade. Free and active ubiquitin (Ub) is conjugated to the activating enzyme (E1) in an ATP-dependent manner. Then, Ub is transferred to the conjugating enzyme (E2) to be finally covalently attached to the substrate protein assisted by the ligating enzyme (E3) which provides substrate specificity. Subsequently, Ub can be deconjugated from substrates by DeUBiqutinating enzymes (DUBs).

2.1 Number of Ubiquitination Enzymes

Humans have eight E1 enzymes which share the adenylation domain necessary for UBL recognition. These E1 enzymes are found as monomers, heterodimers, and homodimers (Table 1). The human genome encodes two ubiquitin E1s, Uba1 and Uba6. Uba6 conjugates less than 1% of Ub while Uba1 is responsible for the 99% conjugation of cellular Ub [18–20]. Uba1 has two isoforms, Uba1A, which is predominantly nuclear, and Uba1B, which is cytosolic [21,22]. The complexity of the conjugation cascade increases for every ubiquitin-like modifier while progressing through the cascade. However, the complexity in ubiquitination can be observed in early stages of the cycle, as the Ub-charged E1 can interact with 38 Ub-E2s differing from other UBLs, where only one E2 has been found [23] (**Table 1**). Reporting the complexity of ubiquitination, possibilities for substrate conjugation escalate even more when Ub-charged E2 interacts with the E3 enzyme, considering that there are more than 600 different E3 ubiquitin ligases providing substrate specificity [24].

Table 1. Summary of E1 and E2 enzymes for ubiquitin-like modifiers.

E1	Complex Formation	UBL	Cellular Process	E2	Reference
Uba1 Uba6	Monomer Monomer	Ub	Ubiquitination	38	[23]
Uba7	Monomer	ISG15	ISGylation	Ubch8	[25]
Sae1/Uba2	Heterodimer	SUMO	SUMOylation	Ubc9	[26]
Nae1/Uba3	Heterodimer	NEDD8	NEDDylation	Uba12	[27]
Uba4/Mocs3	Homodimer	URM1	URMylation	Unknown	[28]
Uba5	Homodimer	UFM1	UFMylation	Ufc1	[29]
Atg7	Homodimer	ATG8/ATG12	Autophagy	ATG3/ATG10	[30] [31]

The final conjugation of Ub to the substrate depends on the catalytic structure of the E3 enzyme, which can be classified into three major families: the Really Interesting New Gene (RING) family, the Homologous to E6-associated protein C-terminus (HECT) family, and the RING-in-between-RING (RBR) family (**Figure 3**).

RING E3 enzymes. This is the largest family of E3 enzimes and it is well characterized by its zinc-binding RING domain or by an U-box domain, which adopts the same RING fold but does not contain zinc [32]. This domain facilitates direct Ub transfer from the E2 enzyme to the substrate, functioning as a scaffold to orient the ubiquitin-charged E2 to the substrate protein [33]. RING E3 enzymes can function as monomers (c-CBL), homodimers (RNF4), or heterodimers (BRCA1/BARD1). A noteworthy difference between homodimers and heterodimers is the capacity of homodimers to bind two E2s (one each monomer), while active heterodimers bind only one E2 enzyme [34,35]. RING E3 dimers also differ depending on how they are formed. Dimerization can occur through sequences flanking the RING domain (i.e., BRCA1/BARD1) [36] or sequences within

the RING per se (i.e., RNF4) [37]. In the first case, dimers are usually formed via α-helix and in the second case the dimer is formed via interleaved C-termini. In both cases, the two RINGs are positioned allowing the E2-binding surfaces to face away from each other to enable the interaction with the E2 enzyme [38]. Finally, there are some RING E3s that exist as multi-subunit assemblies, such as the Cullin-RING E3 ubiquitin ligases (CRLs). CRLs are assembled on a cullin (CUL) scaffold, containing a globular domain with an embedded RING finger protein (RBX1/RBX2/HRT1) in the C-terminus. This embedded RING finger protein serves as the site for E2 binding and the ubiquitin transfer activity. CRLs also contain an adaptor protein in the CUL N-terminus that binds to distinct sets of substrate receptors (SR), which ensures substrate specificity [39] (**Figure 3**). A notable example of a large multi-subunit complex E3 is APC/C, which is an assembly of 19 subunits whose cullin-like subunit is Apc2 [40].

HECT E3 enzymes. This family is characterized by a conserved HECT domain (350 amino acids), which is located at the C-terminus of the E3 ligase. On the other side, the N-terminus domain is very diverse and mediates substrate targeting. The HECT domain has two well-characterized lobes; the N-terminal lobe interacts with the ubiquitin-charged E2, whereas the C-terminal lobe contains the catalytic cysteine that catalyzes the ubiquitin transfer to the substrate protein in a two-step reaction [41]. First, Ub is transferred from the E2 to the catalytic cysteine of the E3. Secondly, the Ub is conjugated to the target protein from the catalytic cysteine of the E3. In order to facilitate the Ub transfer between the E2 and the catalytic cysteine of the E3, the N- and C-terminal lobes are connected through a flexible hinge that allows them to come together [42]. Based on the N-terminal extensions of these enzymes, HECT E3 enzymes can be classified into three subfamilies: the Nedd4 family, which contains tryptophan–tryptophan (WW) motifs, the HERC family, which possesses one or more (regulators of chromosome condensation 1) RCC1-like domains (RLDs), and the remaining HECT E3 enzymes that contain various domains. The HECT Int. J. Mol. Sci. 2022, 23, 3281 5 of 23 E3s family is often regulated by intramolecular interactions that keep the protein in an autoinhibited state, that is released in response to various signals. An example is Smurf1, a NEDD4 HECT E3 ligase which autoinhibits itself through its C2 domain [43].

RBR E3 enzymes. Similar to the HECT E3 family, RBR E3s catalyze ubiquitin transfer in a two-step reaction where ubiquitin is first transferred from the E2 to the catalytic cysteine on the E3 and then to the substrate protein [44]. The RBR family is the smallest E3 family and is characterized by the presence of two RING domains (RING1 and RING2) separated by an in-between-RING domain (IBR). RING1 is required for the recruitment of the ubiquitin-charged E2, and RING2 possesses the catalytic cysteine. The IBR adopts the same fold as RING2 but lacks the catalytic cysteine residue. RBR E3s contain additional specific domains that are involved in intramolecular interactions that keep the protein inactive. This inactivation state can be modified by PTMs such as phosphorylation or by protein–protein interactions [45]. Within this family, we can find PARKIN and HOIP as notorious RBR E3 members, respectively, involved in Parkinson disease and being the central catalytic factor of the LUBAC (linear ubiquitin chain assembly complex) [46,47]. The main mechanistical difference is that RING E3s facilitate the direct transfer of ubiquitin from the E2 to the target protein, whilst HECT and RBR E3s contain an active site cysteine that forms a thioester bond with ubiquitin before transferring it to the substrate [7,48,49] (**Figure 3**).

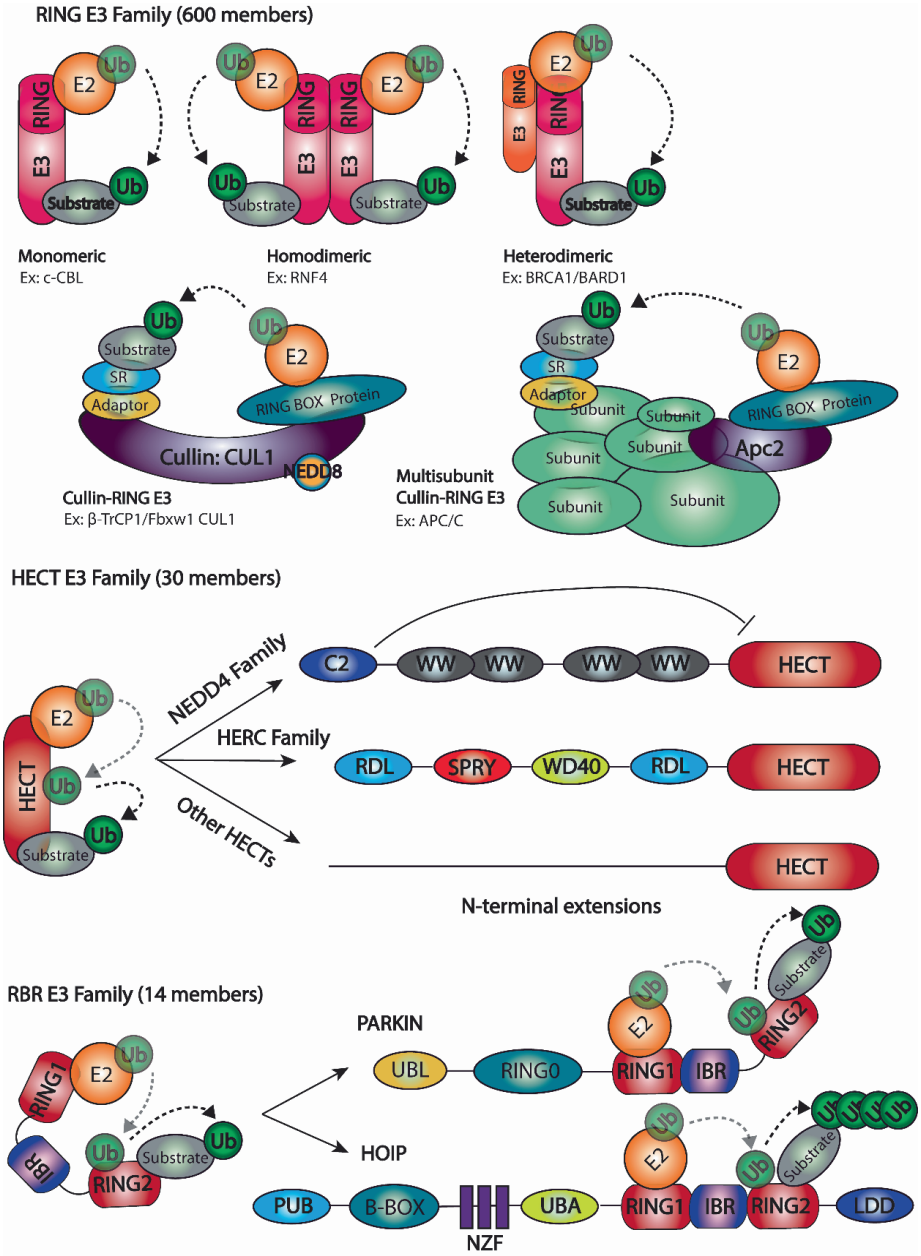


Figure 3. The E3 enzyme families and subclasses. RING E3 enzymes are shown as monomeric, homodimeric, heterodimeric, and forming multi-subunit complexes (CULLINs) E3 enzymes. HECT E3 enzymes are exhibited in three sub-families: NEDD4 Family, HERC Family, and Other HECT. RBR E3 Family is a 14 members family where two members are shown (PARKIN and HOIP), and the ubiquitination mechanism is displayed. The ubiquitination process is depicted. While the Ub transfer from the E2 to the substrate occur in a 2-step reaction for HECT and RBR families, there is a direct ubiquitin transfer from the E2 to the substrate in the RING E3 family

1

2.2 Linkages and cellular functions

After a first ubiquitination cascade, the tethered ubiquitin can become a target for additional ubiquitination, giving rise to ubiquitin polymerization and the formation of polyubiquitination chains and branches. Ubiquitin contains seven acceptor lysines and the amino terminus where chains can be formed (M1, K6, K11, K27, K29, K33, K48, and K63). This offers multiple possibilities to assemble a specific polymer, from monoubiquitination and multi-monoubiquitination to diverse polyubiquitination chains. Polyubiquitination chains can have different types, topologies, and configurations. Ubiquitin chains that comprise only a single linkage type are called homotypic chains (for example a K48 linkage). In contrast, chains that contain mixed linkages are called heterotypic chains. Heterotypic chains can be more complex if one ubiquitin molecule is ubiquitinated at two or more sites creating branches which are known as branched chains (for example K11/K48 linkages) [50]. Depending on the acceptor lysine and the configuration of the linkages, the ubiquitination signal can drive different cellular outcomes (**Figure 4**).

The conjugation of a single ubiquitin molecule to one (monoubiquitination) or several lysines (multi-monoubiquitination) is the major ubiquitination event. In yeast, this accounts for over half of all conjugated ubiquitin [51]. Monoubiquitination has a special role in DNA damage repair, signal transduction, and endocytosis [52]. Recently, novel roles in DNA damage repair have been described, being required in DNA crosslink repair by FANCD2 monoubiquitination to promote the closure of FANCD2/FANCI heterodime [53] and DNA trans-lesion synthesis where PCNA monoubiquitination is thighly regulated [54,55]. While the role of monoubiquitination in the DNA damage response has been predominantly studied in histones and associated to malignancies and neurodevelopmental disorders [56–58], multi-monoubiquitination has been mainly reported in endocytosis [59]. After the addition of one ubiquitin moiety, monoubiquitination can lead to the formation of ubiquitin chains. Early research connected the formation of K48 homotypic ubiquitin linkages to the delivery of a target protein to proteasomal degradation [60]. Later, K11 homotypic chains were discovered to target proteins for proteasomal degradation similarly as K48 chains [61]. In the last decade, K11/K48 heterotypic chains were also reported to not only deliver a substrate protein to proteolytic degradation, but also to be a stronger proteolytic signal [50]. However, non-proteolytic functions have also been elucidated for ubiquitin linkages. For example, polyubiquitination linkages in K63 have been reported to have a key role in several cellular processes including signal transduction, growth response, transcriptional regulation, protein kinase activation, viral protein activation, DNA replication, and DNA repair [62–65]. K11 linkages have been involved not only in proteasome degradation, but also in several cellular functions with non-proteolytic commitments [66]. Together with K48 linkages, K11 and K63 linkages are known as the most abundant ubiquitin chains in cells, being K48 the canonical ubiquitin chains for protein degradation by the proteasome [67,68]. However, the non-conventional or atypical ubiquitin linkages have retained attention in recent years. One of the most studied is the essential role of the PARKIN E3 ligase and its function in regulating mitophagy by atypical K6 linkages [69,70]. In contrast, other E3 enzymes, such as BRCA1/BARD1 and HUWE1 have been linked to K6 linkages, with a role in DNA replication and repair, although the exact function of these K6 linkages remains unclear [71–73]. Another atypical linkage involved in DNA repair is K27. RNF168 E3 ligase seems to form K27 polyubiquitin chains to signal DNA damage [74]. The development of chemical probes and affimers against specific linkages in combination with proteomics have brought light to our understanding of cellular function of both K6 and K27 linkages [73,75]. Other atypical ubiquitin linkages occur at K29 and K33. The K29 linkages seem to have a role in embryogenesis and

tumorigenesis [76], while K33 linkages have been associated with the regulation of the immune response [77]. Finally, linear ubiquitination chains at M1 are formed by the linear ubiquitin chain assembly complex (LUBAC) and have been reported to be associated to pathologies and the immune response [78,79].

New cellular functions of heterotypic and branched chains are also being uncovered. Besides the abovementioned K11/K48 heterotypic chains, which have a role in cell cycle regulation [50], other branched chains such as K29/K48, K48/K63 and M1/K63 are associated with endothelial reticulum associated degradation (ERAD) and control of membrane fluidity (OLE pathway), NF-κB signaling, apoptosis, and immune response, respectively [80–83]. Moreover, proteomic analyses have revealed the existence of K6/K48, K11/K33, K27/K29, and K29/K33 branched chains [73,84–86]. Recently, a novel study using *Saccharomyces cerevisiae* as a model organism presented a strategy enabling the formation of tailored linkage-specific ubiquitin chains on targeted substrates [87]. Specifically, they studied the consequences of modifying PCNA with different ubiquitin chains. While K63 or M1 ubiquitin chains on PCNA contributed to error-free template switching (TS), K48 polyubiquitination linkages on PCNA induced its degradation by the proteasome. These approaches open new opportunities for the study of specific chain-on-substrate consequences for protein and pathway fate.

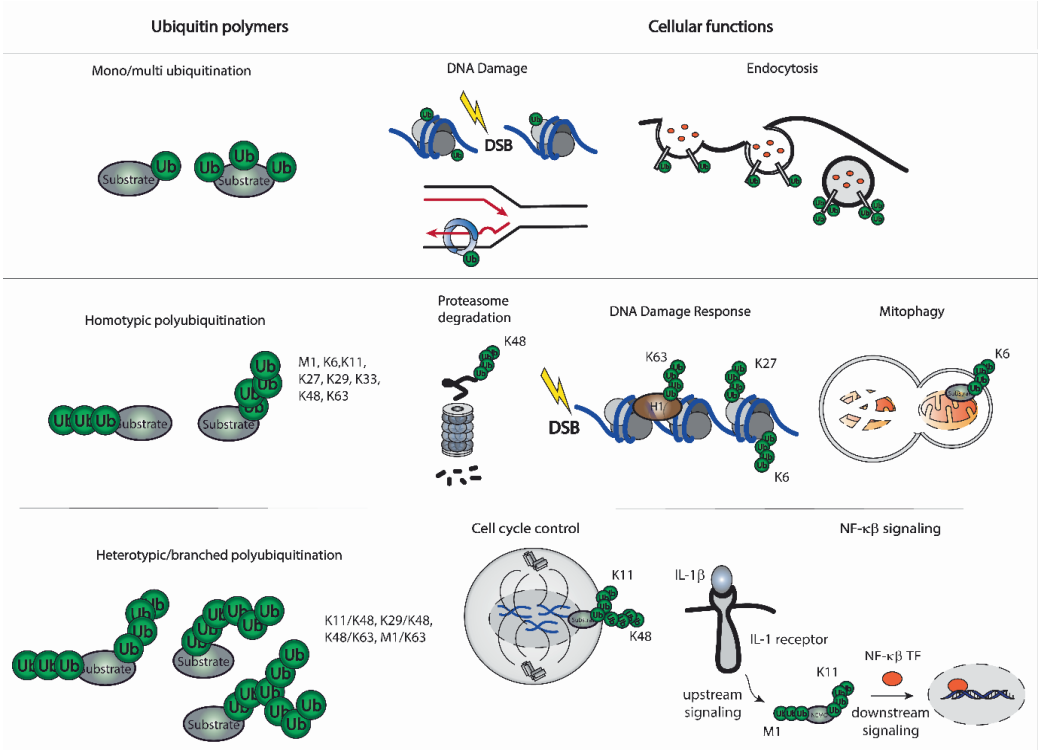


Figure 4. Ubiquitination polymers. Ub moieties can modify proteins at one (mono ubiquitination) or several (multiple mono ubiquitination) Lys residues. Ub can form eight distinctive homotypic linkages, either through M1 (linear Ub chain) or 7 internal Lys residues (K6, K11, K27, K29, K33, K48, and K63 Ub chains). Additional complexity is achieved through the formation of heterotypic Ub chains, which contain multiple Ub linkages and adopt mixed or branched topology. Cellular functions associated to these ubiquitin polymers are displayed.

2.3 Crosstalk with Other Ubls

This complex ubiquitin signaling system controls virtually all cellular functions and acts as "the ubiquitin code". It was previously described by Komander and Rape, and it is composed by "writers", the E1-E2-E3 enzymes, "readers", the proteins that recognize ubiquitinated proteins by their ubiquitin-binder domains (UBDs), and "erasers", the deubiquitinating enzymes (DUBs) that can disassemble ubiquitin chains [88]. The complexity of the code increases even more when different PTMs are involved in the same process. Ubiquitin can also be modified by acetylation, phosphorylation, ADP-ribosylation, phosphoribosylation, deamidation, succinylation, and SUMOylation (Kliza 2020). A suitable example of Ub modification is the S65 ubiquitin phosphorylation, which is mediated by PTEN-induce putative kinase 1 (PINK1). This Ub modification mediates substrate specificity and unlocks the autoinhibition of PARKIN E3 ligase [89]. Besides free Ub modification, the crosstalk between ubiquitination and other PTMs during cellular processes boosts the spread, subtlety, and complexity of the ubiquitin code. In fact, neddylation is involved in the ubiquitination process itself. As was shown by cryoEM, Nedd8 coordinates ubiquitination in Cullin-RING E3 enzymes by binding CUL1 and forming a globular activation module which helps the recruitment of the E2 enzyme for subsequent Ub transfer to the substrate. The presence of Nedd8 stimulates the reaction 2000-fold comparing to the absence of Nedd8, indicating the need of this PTM for the resulting ubiquitination [10]. Similar to Nedd8, other ubiquitin-like proteins like SUMO are able to coordinate ubiquitination responses [9,90].

3. SUMOylation

Small Ubiquitin-like Modifiers (SUMOs) share a similar three-dimensional structure with other UBLs. However, SUMOs differ due to their flexible N-terminus, which also contains the site for SUMO chain formation. All eukaryotes express at least one SUMO paralogue. Five SUMO family members have been identified in humans (SUMO1, SUMO2, SUMO3, SUMO4, and SUMO5) [91]. However, SUMO1, SUMO2, and SUMO3 are the main family members where they are commonly classified as SUMO1 and SUMO2/3 because of the high similarity between mature SUMO2 and SUMO3. All SUMO paralogues are similar in structure but differ in expression levels. SUMO2 is the most abundant family member in mammalian cells. Studies in mice show that the knockout of SUMO2 is embryonic lethal, while SUMO1 and SUMO3 knockout mice were associated to mild phenotypes, possibly because SUMO2 might compensate the loss of either SUMO1 or SUMO3 [92,93]. In contrast to ubiquitination, SUMOylation occurs predominantly in the nucleus and is involved in all nuclear processes. Mis-regulation of these UBLs (Ubiquitin and SUMOs) are connected to diseases including cancer [94,95]. The conjugation of any SUMO member to a substrate is termed SUMOylation. Similarly to ubiquitin, SUMO is conjugated in a 3-step enzymatic cascade that involves a dimeric E1 activating enzyme (SAE1 and SAE2), an E2 conjugating enzyme (Ubc9), and several SUMO E3 enzymes (**Figure 5**). Unlike Ub, each SUMO is encoded by one functional gene, which is translated into a premature SUMO moiety (pre-SUMO) [96]. The pre-SUMO matures with the action of SUMO specific proteases known as SENtrin-specific Proteases (SENPs) [97]. The SENPs remove C-terminal amino acids in order to expose the C-terminal diGly motif, which is needed for the conjugation to the specific lysine of a substrate protein. In vitro experiments have shown that high concentration of Ubc9 and the E1 enzyme is often sufficient to SUMOylate a substrate protein [98,99]. However, in vivo, most substrates require the presence of a SUMO E3 enzyme, which facilitates the transfer of SUMO to the acceptor K of a substrate by enhancing the interaction between the SUMO-charged Ubc9 and the substrate protein [100].

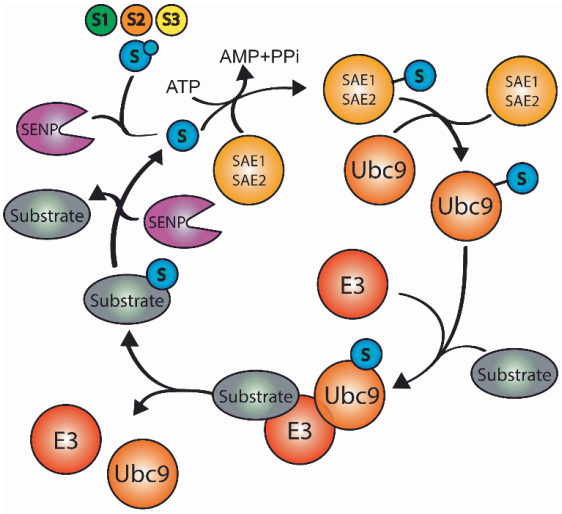


Figure 5. SUMOylation cascade. SUMO precursor matures by the action of a SENP that cleavages the SUMO C-terminal, leaving a diGly motive that forms a thioester bond with the activating enzyme E1 in an ATP dependent manner. Then, activated SUMO is transferred to the conjugating enzyme E2. Finally, the E2 conjugates SUMO to the acceptor lysine (usually in the consensus motive ψKxE) with or without the ligation enzyme E3 which confers substrate specificity. Additional rounds of this cascade form SUMO polymers that can be cleaved by specific SENPs.

In contrast to the ubiquitin system where more than 600 ubiquitin E3 enzymes have been elucidated, only a few SUMO E3 enzymes have been discovered. SUMO E3 activity in vivo and in vitro was observed for first time in 2001, where SIZ1 and SIZ2 SUMO E3s were required for most SUMO conjugation in *yeast* [101] and PIAS1 acted as a SUMO E3 towards p53 in human cells [102]. In mammals, unrelated classes of proteins appear to act as SUMO E3 enzymes, including the protein inhibitor of activated STAT (PIAS) family, the nuclear pore complex (NPC) component RanBP2, the zinc finger 451 (ZNF451) class, the SLX4 complex and other possible SUMO E3 enzymes that enhance SUMOylation of one or more substrates such as the human polycomb protein Pc2/CBX4, topoisomerase I binding protein (TOPORS), the transcription factor Krox20, the tumor suppressor p14/Arf, the histone deacetylase HDAC4, and the Ras homologue enriched in striatum (Rhes) which must be further evaluated (**Table 2**).

3.1 *SUMO E3 enzymes*

The major class of SUMO E3 enzymes is the PIAS family composed of five members (Table 2). This family is characterized by its Siz/Pias Really Interesting New Gene (SPRING) domain that binds to the SUMO E2 enzyme Ubc9. The five members share structural similarity and act in a similar manner as RING E3 ubiquitin ligases [103]. However, in contrast to Ub RING E3 enzymes, knockout studies in mice showed that PIAS SUMO E3 enzymes seem to be redundant and not essential. Mice displayed mild phenotypes, indicating that the lack of one member of the PIAS family is either dispensable or compensated for other members of the PIAS family [121,122]. The substrate specificity that Ub RING E3 enzymes exhibit is yet to be questioned in the PIAS family.

Table 2. SUMO E3 Ligases.

Class	E3 ligase	SUMO	Substrates	Chains	Functions	Reference
PIAS	PIAS1	SUMO1, SUMO2/3	p53, PCNA, Vimentin…	K?	Check points regulation, DNA damage, Cell migration	[94,95,152]
	PIAS2	SUMO1	p53	K7?	Check points	[177]
	PIAS3	SUMO2/3	ATR	K11?	DNA damage	[178]
	PIAS4	SUMO1, SUMO2/3	LEF1	K11?	Wnt-signaling, DNA damage	[179]
	NSMCE2/NSE2	SUMO1	Rad18, TRAX, MMS21…	K7?	DNA damage	[180,181]
NCP	RanBP2/Nup358	SUMO 1	Sp100, Top2, Borealin, DNApolλ	K7?	Nuclear import, Mitosis, DNA repair	[91,101,182]
ZNF451	ZNF451-1/2	SUMO2/3	MCM4, PML	K11	PML stability	[103,183]
	ZNF451-3	SUMO2/3	MCM4	K11		[103]
	KIAA 1586	SUMO1, SUMO2/3	MCM4	K7? K11		[103]
SLX4	SLX4 Complex	SUMO1, SUMO2/3	SLX4, XPF	K?	Genome maintenance, cell division	[107]
Other SUMO E3 ligases	Pc2/CBX4	SUMO1	CtBP	-	Polycomb bodies	[184]
	TOPORS	SUMO1	p53	-	Check points regulation	[185]
	Krox20	SUMO1	Nab	-	Krox20 autoregulation	[186]
	p14/Arf	SUMO1, SUMO2/3	MDM2, NPM/B23	K?	Tumor suppresion	[187]
	HDAC4	SUMO1	MEF2	Mono	Muscle cell diferentiation	[188]
	Rhes	SUMO1, SUMO2/3	Ubc9	Multimono?	SUMOylation	[189]

On the other side, the nucleoporin RanBP2/Nup358 does not contain the SP-RING domain, but instead can form a complex with RanGAP1, Ubc9, and SUMO1 to enable E3 ligase activity [123]. This component of the NPC appears to have different roles during the cell cycle. RanBP2 enriches at kinetochores and the mitotic spindle having essential functions in nucleoplasmic transport during interphase [124] and chromosome segregation mitosis [125]. In addition, new functions are emerging in DNA damage, as recent research has reported the E3 activity of RanBP2 in DNA polymerase lambda SUMOylation [110]. All these functions match with knockout mice studies showing that RanBP2−/− mice were unviable, highlighting the essential role of RanBP2 [126].

Another class of SUMO E3 enzymes was discovered recently. The zinc finger ZNF451 family is composed of ZNF451 isoform 1 (ZNF451-1), isoform 2 (ZNF451-2), isoform 3 (ZNF451-3), and the primate-specific KIAA1586. ZNF451-1 and ZNF451-2 are very similar in contrast to the distant members ZNF451 isoform 3 (ZNF451-3) and KIAA1586. All members share a practically identical N-terminus that includes catalytic tandem SUMO interactive motifs (SIMs) which are necessary for SUMO conjugation. Both SIMs work together, the first SIM places the donor SUMO, while a second SIM binds SUMO on the back side of the E2 enzyme for subsequent SUMO conjugation to the substrate protein [112,127]. In contrast to the N-terminus, the C-terminus differs between family members. ZNF451- 1 contains C_2H_2-Zinc finger domains, whereas ZNF451-2 lacks residues 870-917 due to alternative splicing. ZNF451-3 encodes a C-terminal deletion of 933-1061 and holds a LAP2alpha (LAP2α) domain which is not present in the other family members [128]. The ZNF451 class seems to be inefficient in the initial conjugation of SUMO, although this tandem-SIMs region is sufficient to extend a SUMO chain and form a SUMO polymer. This activity is referred to as E4 elongase. In addition, the ZNF451 class has been implicated in SUMO2/3 polymers formation during both proteasome inhibition and DNA damage stresses [112]. Years later, due to its role in DNA-Protein crosslink repair by stalled TOP2 SUMOylation, it was suggested to re-name this ZNF451 class “ZATT” (zinc finger protein associated with TDP2 and TOP2) although these are probably not the only substrates [129].

SLX4 contains a BTB domain and three putative SIMs essential for SUMO binding and SUMOylation. The BTB domain seems to be important for protein–protein interaction and oligomerization necessary for the formation of the SLX4 complex. Pull-down experiments employing SLX4 SIMs mutants show the capacity of SLX4 to SUMOylate xeroderma pigmentosum group-F (XPF). Interestingly, SLX4 can SUMOylate itself with both SUMO1 and SUMO3. However, SLX4 seems to preferentially use SUMO3 for XPF SUMOylation. SLX4 SUMO E3 activity plays a role in response to global and local replication stress [114].

Additional E3 enzymes have been identified in other organisms, such as herpesvirus where they have a possible role during infection. Examples are SM, UL69, and UL54. Interestingly, the SM and UL69 show preference for SUMO1 and UL54 for SUMO2 [130]. There is still a lot to discover and research to be done, but a complex network could be emerging where SUMO E3 enzymes use different SUMO modifiers to form different chains, in order to lead a substrate protein to a particular function in the cell, which seems to be tightly regulated.

3.2 SUMO Polymers

The discovery of the ZNF451 class and its E4 elongase activity gave rise to the study of a possible physiological role of poly-SUMOylation, although the knowledge about SUMO chains signaling remains limited compared to the ubiquitin chain field. A notable difference between SUMO1 and SUMO2/3 polymers resides in the ability to form SUMO chains.

In contrast to SUMO1, SUMO2 and SUMO3 possess a K11 in their flexible N-terminus, which is located in a sequence motif, ψKXE, where ψ represents a large hydrophobic amino acid and X any amino acid. This sequence is referred to as SUMO consensus motif. The consensus motif is preferentially targeted for SUMOylation and it is also present in other potential SUMOylation target proteins [131]. This K11 in the SUMO consensus motif allows SUMO2/3 to form K11 SUMO chains [132]. Although this K11 seems to be the main site for SUMO chains, site-specific mass spectrometry approaches have revealed several other SUMO acceptor lysines within SUMO1, SUMO2, and SUMO3 [133]. For example, SUMO1 contains an inverted SUMO consensus site, ExK [134], and harbors an N-terminal K7, which tolerates low efficient SUMO1 chains formation as demonstrated in vitro and in vivo by site-specific mass spectrometry [133,135]. However, it seems that SUMO1 works as a capping factor, terminating SUMO2/3 chain formation [132,136].

The main consequence of SUMO conjugation seems to be the alteration of binding surfaces of the substrate protein, which can either hinder or promote intra- or intermolecular interactions. Additionally, SUMO is able to promote molecular interaction due to its affinity to SIMs. SIMs are short peptide sequences mostly located in unstructured regions of the modified protein or interacting proteins [137]. These SIMs allow non-covalent interaction between SUMOylated proteins and effector proteins which contain SIMs [138]. Given the fact that SUMOylation occurs predominantly in the nucleus and nuclear bodies, its role varies from transcription regulation and chromatin remodeling to DNA repair and cell cycle progression [139]. Although SUMO polymers have been previously reviewed [140,141], to date there is no indication that different SUMO chain linkages fulfil distinct roles within human cells. However, work in S. pombe revealed the possible role of two different SUMO chain linkages (K14 and K30) in response to replication arrest [142].

3.3 SUMO and Ubiquitin crosstalk

SUMO chains on target substrates can be a signal for the recruitment of SUMO-targeted ubiquitin ligases (STUbLs), leading to a crosstalk between SUMOylation and Ubiquitination. STUbLs, such as the RING-finger protein 4 (RNF4), contain a RING domain that binds to an E2 ubiquitin enzyme, and SIMs that allow the interaction with SUMOylated substrates and increase the preference for SUMO modified targets. RNF4 possesses at least three closely spaced SIMs and shows a clear preference for substrates that are modified by a SUMO chain or at least three SUMO moieties [143]. RNF4-mediated ubiquitination results in either K48 or K63 ubiquitin linkages, which, respectively, label the substrate protein for proteasomal degradation or for the recruitment of ubiquitin-binding motif containing proteins. This mechanism has been implicated in a variety of cellular processes, including promyelocytic leukemia (PML) nuclear body (NB) integrity, mitosis, and DNA Damage Response (DDR) [143–146]. RNF4 regulates DNA damage signaling by controlling the lifetime of proteins involved in DNA repair such as the check point mediator MDC1. It also regulates the whole SUMOylation machinery, E1, E2, and E3s, by labeling the members for proteasomal degradation [9,145,147,148]. However, it is yet to be discovered how only some poly-SUMOylated proteins are targeted by STUbLs and how different STUbLs can bind selective SUMOylated targets.

The SUMO and Ubiquitin interplay can also be modified by SUMO-targeted ubiquitin proteases (STUPs). STUPs can recognize poly-SUMOylated proteins and are able to modify the ubiquitin chains on SUMOylated targets by their ability to remove Ub. To date, three STUPs have been identified. USP7 seems to remove ubiquitin from SUMO targets with a role in DNA replication [149]. USP11 might have a role regulating nuclear bodies by limiting RNF4 activity [150]. The last

STUP is Ataxin-3 (ATX3), which seems to participate in the regulation of MDC1 counteracting the RNF4 activity [151]. Finally, there is not only crosstalk with ubiquitination but also with other PTMs. This was observed in a very deep profiling of the SUMO proteome [152].

4. PROTEOMICS FOR SUBSTRATE IDENTIFICATION

The major challenge to fully understand the role of a specific E3 ligase is to determine its target proteins. Due to the hierarchical structure of the ubiquitination cascade, one E1 enzyme (Uba1/Uba6) can bind dozens of E2, these E2s can bind hundreds of E3 which are responsible for determining the substrate specificity for ubiquitination. Thus, mapping specific targets for a particular E3 ligase became challenging (**Figure 6a**). In addition, the transient and weak interaction between the E3 ligase and the target, the dynamic and reversible nature of these modifications, the relative low abundance and expression of substrates, the fact that several substrates are labeled for proteasome degradation and rapidly degraded, increase the difficulty of the identification. Moreover, the ubiquitination is often dependent on physiological conditions and spatiotemporal organization. Therefore, many substrate proteins can only be identified upon different cell treatments. Additionally, under a stabilized physiological condition, several E3 enzymes can target individual substrate proteins at different residues, making the identification even more complex.

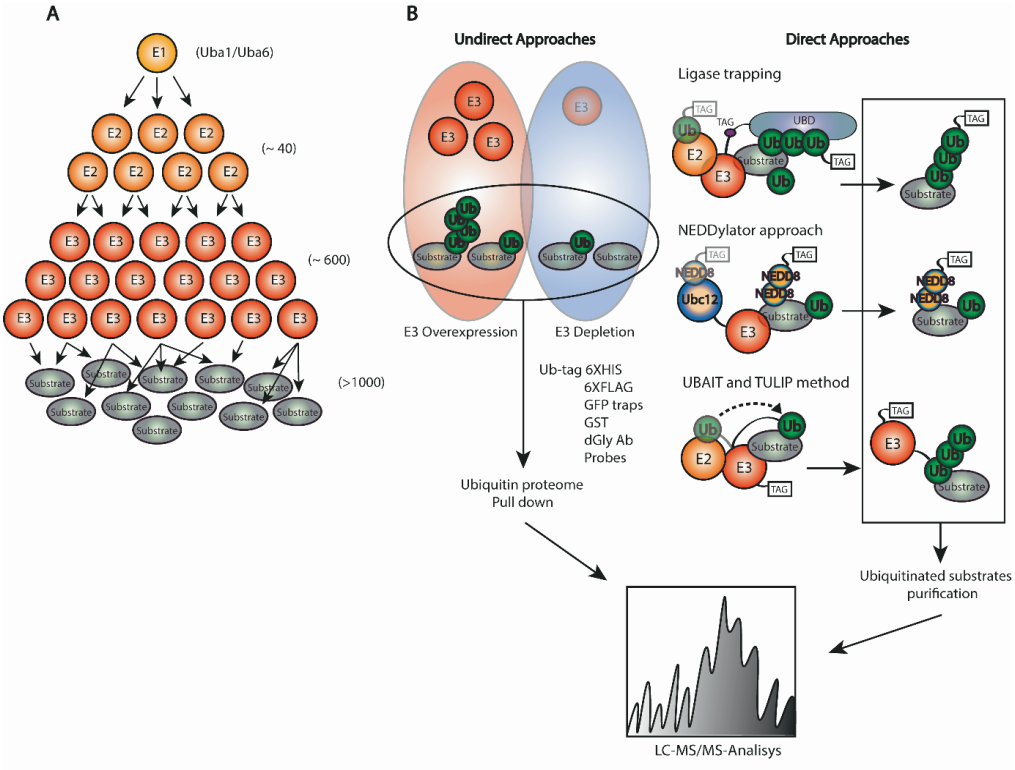


Figure 6. Proteomics for E3 ligase target identification. (a) Hierarchical organization of the ubiquitination cascade. Emphasizing the difficulty of mapping substrate proteins for specific E3 enzymes. (b) Strategies for

E3 ligase substrate identification divided into undirect and direct approaches. Within undirect methods, in red, the overexpression of an E3 ligase results in the increase in ubiquitination levels for putative substrates. Opposing, in blue, the depletion of an E3 ligase displays a decrease in ubiquitination levels of putative substrates. The direct approaches allow identification of specific-E3 ligase substrates where ligase trapping, NEDDylation approach, and UBAIT/TULIP methodology are shown.

The vast majority of the E3 enzymes identified in the human genome remain relatively uncharacterized. Generally, the identification and validation of substrates rely on relatively slow, low-throughput biochemical methods reviewed in [153]. Currently, the increasing advances in proteomics and mass spectrometry (MS) provide new approaches to not only identify the ubiquitin and SUMO modified proteins, but also decipher ubiquitination and SUMOylation sites in target proteins [139,154].

4.1 *Mapping Sites*

In order to map ubiquitination sites, many strategies reside in the diGly residue that is covalently attached to the modified lysine of a substrate protein. The serine protease trypsin is commonly used to generate peptides from a protein sample for consequent MS identification. Trypsin cleaves Ub after R74, leaving this remnant diGly at the ubiquitinated residue of a substrate protein. Immunopurification methods using monoclonal antibodies against selective diGly residues have been exploited for the MS identification of ubiquitination sites [155,156]. However, in addition to ubiquitin, other UBL-modifiers (ISG15 and NEDD8) also leave the diGly residue after trypsin digestion, which is one of the major limitations of these approaches.

In 2018, Blagoy Blagoev's group developed the UbiSite antibody which recognizes the 13 amino acids of Ub that remain covalently attached to the modified residue of the substrate proteins after digestion with endo-protease LysC. Another advantage of this antibody is that it allows detection of N-terminal ubiquitination but not linear Ub signatures. This method enabled the identification of 63.000 unique ubiquitination sites [154]. More strategies have been recently developed regarding the diGly dipeptide. David Komander's group established the Ub-clipping approach which utilizes an engineered viral protease (Lbpro) to remove the Ub from substrate proteins but leaving the C-terminal diGly conjugated to the modification site. Unlike trypsin digestion, the formed ubiquitin branches can be detected by MS analysis [157].

Immunopurification using antibodies also enables the determination of ubiquitination chains topology. There are Ub linkage-specific antibodies which specifically recognize M1, K11, K27, K48, and K63 linkages [158,159]. More recently, chemical synthesis has enabled the engineering of Ub-binding modules or probes. These probes can be designed as traps for binding specific chains, branches, and even hybrid chains that allow their enrichment for subsequent MS analysis [160,161].

In a similar strategy to the UbiSite, 14869 SUMO2ylation sites were identified at the endogenous level under heat stress and proteasomal inhibition in human cells using the epitope for the commercially available SUMO2 antibody (8A2) [139].

4.2 Identification of novel substrates

Likewise, MS-based approaches have been performed to identify novel ubiquitination substrates. We could classify these into undirect and direct MS-based approaches (**Figure 6b**). Within the undirect group, finding differences in ubiquitinated proteins upon either overexpression or depletion of an E3 ligase of interest has been commonly used [162–165].

Substrate proteins that are enriched or depleted in their ubiquitination levels are considered as putative ubiquitination substrates for the E3 ligase under investigation. However, these indirect approaches are based on full ubiquitin proteomes once the E3 ligase of interest is either overexpressed or depleted. Due to the complexity of the ubiquitin proteome and the low abundance ubiquitination targets, some results might be obtained because of overexpressed artifacts in the case of overexpression of an E3 ligase. Conversely, in knockdown-based screenings, E3s can be redundant on their targets, and some targets might be missed because their ubiquitination is still performed by another E3 enzyme resulting from an epistatic effect. As a consequence, every potential target must be carefully validated.

Regarding direct approaches, several methods have been employed including Ub ligase substrate trapping [166,167], Michaelis intermediates traps [168], NEDDylator approaches [169], Ub activated interaction traps, UBAITs [170], and optimized methods such as Targets for Ubiquitin Ligases Identified by Proteomics, TULIP and TULIP2 [9,148] which allow the identification of E3 ligase-specific ubiquitination targets by mass spectrometry analysis.

Ligase trapping is based on E3 enzymes fused to Ubiquitin Binding Domains (UBDs) that are used for isolation of ubiquitinated substrates by affinity purification. The rationale resides in the fact that UBDs increase the binding affinity of the E3 ligase of interest toward its targets and enabled the identification of potential substrates by MS analysis. The selection of proper UBD, as well as the fusion point, are the major limitations of this approach, as effective enrichment of substrates and potential disruption of the substrate recruitment are essential for the proper functionality of the ligase trap [166,167].

Michaelis intermediate strategies are based on the generation of a E2-SUMO thioester mimetic than can be crosslinked to the substrate protein. The resulting E2-SUMO-Substrate complex can be purified and the structure can be determined by crystallization [168].

The NEDDylator approach relies on the fusion of the NEDD8 E2 enzyme (Ubc12) to the substrate-binding region of the E3 ligase of interest. Such configuration allows artificial NEDDylation of endogenous E3 ligase substrates. The enrichment and subsequent MS analysis relies on the NEDDylation proteome which is far less complex than the ubiquitination. As NEDDylation does not occur at a high level in a cell, it is possible to distinguish between endogenous and NEDDylator-induced modifications to identify E3 ligase substrates [169].

UBAIT methodology allows identification of substrates for both RING and HECT E3 enzymes. The UBAIT tool consists of the utilization of E3 enzyme–ubiquitin fusions. The fusion of Ub to the E3 ligase abolishes the transient interaction between the E3 ligase and the substrate, since the E3 ligase remains fused to the Ub which is covalently bound to the substrate protein after its putative ubiquitination. The presence of Ub enables E1- and E2-mediated activation of UBAIT and subsequent covalent capture of E3 ligase substrates. This allows the later purification of the E3 together with its ubiquitination target for subsequent identification by MS analysis. The UBAIT strategy has been later optimized to enable the systematic identification of Ubiquitin E3 substrates, in the TULIP and TULIP2 methodologies [9,148].

4.3 Other Ub/SUMO-Related MS-Based Proteomics Approaches

We can also find methods using proximity-dependent biotin identification (BioID), which can be used to identify weak/transient protein interactions of proximal/surrounding proteins of a particular E3 ligase. It is based on the fusion of the E3 ligase with the mutated form of biotin ligase

(BirA), which biotinylates all proteins in the close vicinity, if biotin is available. These biotinylated proteins can be enriched by affinity purification and identified by MS analysis, where the labeled proteins can be potential E3 substrate targets [171]. Recently, this approach has been updated for the identification of SUMOdependent interactors [172]. In this strategy, TurboID (upgraded version of BioID) was split in two, one fragment was fused to SUMO and the complementary fragment to a protein of interest. When the protein of interest and SUMO interact, the TurboID enzyme can be reconstituted and able to label proximal complexes, that can be later purified and identified by MS.

The identification of a substrate can be even more complex when two different E3 enzymes work together for the ubiquitination of a particular substrate. It is known that many neddylated culling-RING E3 enzymes (CRLs) and RBR enzymes in the ARIH family form E3-E3 super-assemblies [173], hence, new activity-based chemical probes that enable cryo-electron microscopy visualization of E3-E3 ubiquitination have been developed, facilitating the visualization of the ubiquitination intermediates [174].

Some of these strategies can be modified for SUMO targets. In a recent study, PIAS1 was overexpressed to produce profile changes in protein SUMOylation. This profiles can be determined by MS [104]. This is an example of an indirect approach for SUMOylation targets.

It has been shown that defects in ubiquitin and SUMO E3 enzymes may be contributing factors in cancer [175–177] and human neurodegenerative diseases such as Alzheimer's, Huntington's, amyotrophic lateral sclerosis, and Parkinson's [178]. Therefore, it becomes crucial to identify substrates that are ubiquitylated/SUMOylated by specific E3 enzymes, in order to obtain novel insights in particular diseases.

5. CONCLUSIONS AND FUTURE PERSPECTIVES

During recent years, new tools and techniques have allowed us to exploit the potential of already known features or to discover new ones. A good example of this was the observation of degradation signals in short-lived proteins in 1986 [179]. Nowadays, the use of MS approaches makes it possible to characterize a phosphodegron site, the E3 ligase responsible for the ubiquitination and, with biochemical validation, to know the biological processes this degradation signal is involved in [180].

New tools to explore ubiquitin and SUMO are ubiquitination and SUMOylation inhibitors. Recently, a new SUMO inhibitor has been discovered [181] with optimistic results in the treatment of cancer [182,183]. The employment of inhibitors combined with MS have revealed a large set of SUMOylation and ubiquitination targets. The inhibition of either ubiquitination or SUMOylation in cells stably expressing tag-ubiquitin or tagSUMO, facilitates the identification of these targets when comparing the treated cells with non-treated ones [184,185].

Another key strategy for the decade ahead is the use of CRISPR genome-wide screening combined with either ubiquitination or SUMOylation inhibitors to study up- and downregulated genes upon a desired condition [186,187].

To get deeper insight into the ubiquitin code and decipher the complexity of the ubiquitination network, the use of chemical synthetic probes combined with high resolution MS equipment would be a key topic for the next decade. Using these probes, it would be possible to find particular ubiquitin or SUMO linkages/branches, or even hybrid chains between SUMO and ubiquitin [188–

190]. Some strategies have been already reviewed in the ubiquitin filed [191], nonetheless, they can be extended for the SUMO field.

Author Contributions: Writing—original draft preparation, D.S.-L.; writing—review and editing, D.S.-L. and R.G.-P.; supervision and funding acquisition R.G.-P. All authors have read and agreed to the published version of the manuscript**.**

Funding: The authors are supported by the Dutch Cancer Foundation (KFW-YIG 11367/2017-2).

REFERENCES

1. Blazek, M.; Santisteban, T.S.; Zengerle, R.; Meier, M. Analysis of fast protein phosphorylation kinetics in single cells on a microfluidic chip. Lab Chip 2015, 15, 726–734.
2. Mohanty, P.; Chatterjee, K.S.; Das, R. NEDD8 Deamidation Inhibits Cullin RING Ligase Dynamics. Front. Immunol. 2021, 12, 695331.
3. Duan, G.; Walther, D. The roles of post-translational modifications in the context of protein interaction networks. PLoS Comput. Biol. 2015, 11, e1004049.
4. Lake, M.W.; Wuebbens, M.M.; Rajagopalan, K.V.; Schindelin, H. Mechanism of ubiquitin activation revealed by the structure of a bacterial MoeB-MoaD complex. Nature 2001, 414, 325–329.
5. Wang, C.; Xi, J.; Begley, T.P.; Nicholson, L.K. Solution structure of ThiS and implications for the evolutionary roots of ubiquitin. Nat. Struct. Biol. 2001, 8, 47–51.
6. Pearce, M.J.; Mintseris, J.; Ferreyra, J.; Gygi, S.P.; Darwin, K.H. Ubiquitin-like protein involved in the proteasome pathway of Mycobacterium tuberculosis. Science 2008, 322, 1104–1107.
7. Schulman, B.A.; Harper, J.W. Ubiquitin-like protein activation by E1 enzymes: The apex for downstream signalling pathways. Nat. Rev. Ther. Mol. Cell Biol. 2009, 10, 319–331.
8. Aichem, A.; Sailer, C.; Ryu, S.; Catone, N.; Stankovic-Valentin, N.; Schmidtke, G.; Melchior, F.; Stengel, F.; Groettrup, M. The ubiquitin-like modifier FAT10 interferes with SUMO activation. Nat. Commun. 2019, 10, 4452.
9. Kumar, R.; Gonzalez-Prieto, R.; Xiao, Z.; Verlaan-de Vries, M.; Vertegaal, A.C.O. The STUbL RNF4 regulates protein group SUMOylation by targeting the SUMO conjugation machinery. Nat. Commun. 2017, 8, 1809.
10. Baek, K.; Krist, D.T.; Prabu, J.R.; Hill, S.; Klugel, M.; Neumaier, L.M.; von Gronau, S.; Kleiger, G.; Schulman, B.A. NEDD8 nucleates a multivalent cullin-RING-UBE2D ubiquitin ligation assembly. Nature 2020, 578, 461–466.
11. Fan, J.B.; Arimoto, K.; Motamedchaboki, K.; Yan, M.; Wolf, D.A.; Zhang, D.E. Identification and characterization of a novel ISG15-ubiquitin mixed chain and its role in regulating protein homeostasis. Sci. Rep. 2015, 5, 12704.
12. Grou, C.P.; Pinto, M.P.; Mendes, A.V.; Domingues, P.; Azevedo, J.E. The de novo synthesis of ubiquitin: Identification of deubiquitinases acting on ubiquitin precursors. Sci. Rep. 2015, 5, 12836.
13. Schlesinger, D.H.; Goldstein, G.; Niall, H.D. The complete amino acid sequence of ubiquitin, an adenylate cyclase stimulating polypeptide probably universal in living cells. Biochemistry 1975, 14, 2214–2218.
14. Scheffner, M.; Nuber, U.; Huibregtse, J.M. Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade. Nature 1995, 373, 81–83.
15. Hershko, A.; Ciechanover, A. The ubiquitin system. Annu. Rev. Biochem. 1998, 67, 425–479.
16. Kirisako, T.; Kamei, K.; Murata, S.; Kato, M.; Fukumoto, H.; Kanie, M.; Sano, S.; Tokunaga, F.; Tanaka, K.; Iwai, K. A ubiquitin ligase complex assembles linear polyubiquitin chains. EMBO J. 2006, 25, 4877–4887.
17. Hershko, A.; Heller, H.; Elias, S.; Ciechanover, A. Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. J. Biol. Chem. 1983, 258, 8206–8214.
18. Jin, J.; Li, X.; Gygi, S.P.; Harper, J.W. Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging. Nature 2007, 447, 1135–1138.
19. Pelzer, C.; Kassner, I.; Matentzoglu, K.; Singh, R.K.; Wollscheid, H.P.; Scheffner, M.; Schmidtke, G.; Groettrup, M. UBE1L2, a novel E1 enzyme specific for ubiquitin. J. Biol. Chem. 2007, 282, 23010–23014.

20. Chiu, Y.H.; Sun, Q.; Chen, Z.J. E1-L2 activates both ubiquitin and FAT10. Mol. Cell 2007, 27, 1014–1023.
21. Handley-Gearhart, P.M.; Stephen, A.G.; Trausch-Azar, J.S.; Ciechanover, A.; Schwartz, A.L. Human ubiquitin-activating enzyme, E1. Indication of potential nuclear and cytoplasmic subpopulations using epitope-tagged cDNA constructs. J. Biol. Chem. 1994, 269, 33171–33178.
22. Sugaya, K.; Ishihara, Y.; Inoue, S.; Tsuji, H. Characterization of ubiquitin-activating enzyme Uba1 in the nucleus by its mammalian temperature-sensitive mutant. PLoS ONE 2014, 9, e96666.
23. Stewart, M.D.; Ritterhoff, T.; Klevit, R.E.; Brzovic, P.S. E2 enzymes: More than just middle men. Cell Res. 2016, 26, 423–440.
24. Li, W.; Bengtson, M.H.; Ulbrich, A.; Matsuda, A.; Reddy, V.A.; Orth, A.; Chanda, S.K.; Batalov, S.; Joazeiro, C.A. Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling. PLoS ONE 2008, 3, e1487.
25. Zhao, C.; Beaudenon, S.L.; Kelley, M.L.; Waddell, M.B.; Yuan, W.; Schulman, B.A.; Huibregtse, J.M.; Krug, R.M. The UbcH8 ubiquitin E2 enzyme is also the E2 enzyme for ISG15, an IFN-alpha/beta-induced ubiquitin-like protein. Proc. Natl. Acad. Sci. USA 2004, 101, 7578–7582.
26. Lee, G.W.; Melchior, F.; Matunis, M.J.; Mahajan, R.; Tian, Q.; Anderson, P. Modification of Ran GTPase-activating protein by the small ubiquitin-related modifier SUMO-1 requires Ubc9, an E2-type ubiquitin-conjugating enzyme homologue. J. Biol. Chem. 1998, 273, 6503–6507.
27. Huang, D.T.; Ayrault, O.; Hunt, H.W.; Taherbhoy, A.M.; Duda, D.M.; Scott, D.C.; Borg, L.A.; Neale, G.; Murray, P.J.; Roussel, M.F.; et al. E2-RING expansion of the NEDD8 cascade confers specificity to cullin modification. Mol. Cell 2009, 33, 483–495.
28. Judes, A.; Bruch, A.; Klassen, R.; Helm, M.; Schaffrath, R. Sulfur transfer and activation by ubiquitin-like modifier system Uba4*Urm1 link protein urmylation and tRNA thiolation in yeast. Microb. Cell 2016, 3, 554–564.
29. Tatsumi, K.; Sou, Y.S.; Tada, N.; Nakamura, E.; Iemura, S.; Natsume, T.; Kang, S.H.; Chung, C.H.; Kasahara, M.; Kominami, E.; et al. A novel type of E3 ligase for the Ufm1 conjugation system. J. Biol. Chem. 2010, 285, 5417–5427.
30. Shintani, T.; Mizushima, N.; Ogawa, Y.; Matsuura, A.; Noda, T.; Ohsumi, Y. Apg10p, a novel protein-conjugating enzyme essential for autophagy in yeast. EMBO J. 1999, 18, 5234–5241.
31. Ichimura, Y.; Kirisako, T.; Takao, T.; Satomi, Y.; Shimonishi, Y.; Ishihara, N.; Mizushima, N.; Tanida, I.; Kominami, E.; Ohsumi, M.; et al. A ubiquitin-like system mediates protein lipidation. Nature 2000, 408, 488–492.
32. Ohi, M.D.; Vander Kooi, C.W.; Rosenberg, J.A.; Chazin, W.J.; Gould, K.L. Structural insights into the U-box, a domain associated with multi-ubiquitination. Nat. Struct. Biol. 2003, 10, 250–255.
33. Lorick, K.L.; Jensen, J.P.; Fang, S.; Ong, A.M.; Hatakeyama, S.; Weissman, A.M. RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. Proc. Natl. Acad. Sci. USA 1999, 96, 11364–11369.
34. Mace, P.D.; Linke, K.; Feltham, R.; Schumacher, F.R.; Smith, C.A.; Vaux, D.L.; Silke, J.; Day, C.L. Structures of the cIAP2 RING domain reveal conformational changes associated with ubiquitin-conjugating enzyme (E2) recruitment. J. Biol. Chem. 2008, 283, 31633–31640.
35. Bentley, M.L.; Corn, J.E.; Dong, K.C.; Phung, Q.; Cheung, T.K.; Cochran, A.G. Recognition of UbcH5c and the nucleosome by the Bmi1/Ring1b ubiquitin ligase complex. EMBO J. 2011, 30, 3285–3297.
36. Brzovic, P.S.; Rajagopal, P.; Hoyt, D.W.; King, M.C.; Klevit, R.E. Structure of a BRCA1-BARD1 heterodimeric RING-RING complex. Nat. Struct. Biol. 2001, 8, 833–837.
37. Liew, C.W.; Sun, H.; Hunter, T.; Day, C.L. RING domain dimerization is essential for RNF4 function. Biochem. J. 2010, 431, 23–29.
38. Metzger, M.B.; Pruneda, J.N.; Klevit, R.E.; Weissman, A.M. RING-type E3 ligases: Master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. Biochim. Biophys. Acta 2014, 1843, 47–60.
39. Lydeard, J.R.; Schulman, B.A.; Harper, J.W. Building and remodelling Cullin-RING E3 ubiquitin ligases. EMBO Rep. 2013, 14, 1050–1061.

40. Schreiber, A.; Stengel, F.; Zhang, Z.; Enchev, R.I.; Kong, E.H.; Morris, E.P.; Robinson, C.V.; da Fonseca, P.C.; Barford, D. Structural basis for the subunit assembly of the anaphase-promoting complex. Nature 2011, 470, 227–232.
41. Huang, L.; Kinnucan, E.; Wang, G.; Beaudenon, S.; Howley, P.M.; Huibregtse, J.M.; Pavletich, N.P. Structure of an E6AP-UbcH7 complex: Insights into ubiquitination by the E2-E3 enzyme cascade. Science 1999, 286, 1321–1326.
42. Verdecia, M.A.; Joazeiro, C.A.; Wells, N.J.; Ferrer, J.L.; Bowman, M.E.; Hunter, T.; Noel, J.P. Conformational flexibility underlies ubiquitin ligation mediated by the WWP1 HECT domain E3 ligase. Mol. Cell 2003, 11, 249–259.
43. Wiesner, S.; Ogunjimi, A.A.; Wang, H.R.; Rotin, D.; Sicheri, F.; Wrana, J.L.; Forman-Kay, J.D. Autoinhibition of the HECT-type ubiquitin ligase Smurf2 through its C2 domain. Cell 2007, 130, 651–662.
44. Wenzel, D.M.; Lissounov, A.; Brzovic, P.S.; Klevit, R.E. UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids. Nature 2011, 474, 105–108.
45. Reiter, K.H.; Klevit, R.E. Characterization of RING-Between-RING E3 Ubiquitin Transfer Mechanisms. Methods Mol. Biol. 2018, 1844, 3–17.
46. Lechtenberg, B.C.; Rajput, A.; Sanishvili, R.; Dobaczewska, M.K.; Ware, C.F.; Mace, P.D.; Riedl, S.J. Structure of a HOIP/E2~ubiquitin complex reveals RBR E3 ligase mechanism and regulation. Nature 2016, 529, 546–550.
47. Kitada, T.; Asakawa, S.; Hattori, N.; Matsumine, H.; Yamamura, Y.; Minoshima, S.; Yokochi, M.; Mizuno, Y.; Shimizu, N. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. Nature 1998, 392, 605–608.
48. Metzger, M.B.; Hristova, V.A.; Weissman, A.M. HECT and RING finger families of E3 ubiquitin ligases at a glance. J. Cell Sci. 2012, 125, 531–537.
49. Sluimer, J.; Distel, B. Regulating the human HECT E3 ligases. Cell Mol. Life Sci. 2018, 75, 3121–3141.
50. Meyer, H.J.; Rape, M. Enhanced protein degradation by branched ubiquitin chains. Cell 2014, 157, 910–921.
51. Ziv, I.; Matiuhin, Y.; Kirkpatrick, D.S.; Erpapazoglou, Z.; Leon, S.; Pantazopoulou, M.; Kim, W.; Gygi, S.P.; Haguenauer-Tsapis, R.; Reis, N.; et al. A perturbed ubiquitin landscape distinguishes between ubiquitin in trafficking and in proteolysis. Mol. Cell Proteom. 2011, 10, M111.009753.
52. Passmore, L.A.; Barford, D. Getting into position: The catalytic mechanisms of protein ubiquitylation. Biochem. J. 2004, 379, 513–525.
53. Alcon, P.; Shakeel, S.; Chen, Z.A.; Rappsilber, J.; Patel, K.J.; Passmore, L.A. FANCD2-FANCI is a clamp stabilized on DNA by monoubiquitination of FANCD2 during DNA repair. Nat. Struct. Mol. Biol. 2020, 27, 240–248.
54. Hedglin, M.; Aitha, M.; Pedley, A.; Benkovic, S.J. Replication protein A dynamically regulates monoubiquitination of proliferating cell nuclear antigen. J. Biol. Chem. 2019, 294, 5157–5168.
55. Thakar, T.; Leung, W.; Nicolae, C.M.; Clements, K.E.; Shen, B.; Bielinsky, A.K.; Moldovan, G.L. Ubiquitinated-PCNA protects replication forks from DNA2-mediated degradation by regulating Okazaki fragment maturation and chromatin assembly. Nat. Commun. 2020, 11, 2147.
56. Marsh, D.J.; Dickson, K.A. Writing Histone Monoubiquitination in Human Malignancy-The Role of RING Finger E3 Ubiquitin Ligases. Genes 2019, 10, 67.
57. Srivastava, A.; McGrath, B.; Bielas, S.L. Histone H2A Monoubiquitination in Neurodevelopmental Disorders. Trends Genet. 2017, 33, 566–578.
58. Becker, J.R.; Clifford, G.; Bonnet, C.; Groth, A.; Wilson, M.D.; Chapman, J.R. BARD1 reads H2A lysine 15 ubiquitination to direct homologous recombination. Nature 2021, 596, 433–437.
59. Haglund, K.; Sigismund, S.; Polo, S.; Szymkiewicz, I.; Di Fiore, P.P.; Dikic, I. Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation. Nat. Cell Biol. 2003, 5, 461–466.
60. Chau, V.; Tobias, J.W.; Bachmair, A.; Marriott, D.; Ecker, D.J.; Gonda, D.K.; Varshavsky, A. A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. Science 1989, 243, 1576–1583.

61. Jin, L.; Williamson, A.; Banerjee, S.; Philipp, I.; Rape, M. Mechanism of ubiquitin-chain formation by the human anaphasepromoting complex. Cell 2008, 133, 653–665.
62. Deng, L.; Wang, C.; Spencer, E.; Yang, L.; Braun, A.; You, J.; Slaughter, C.; Pickart, C.; Chen, Z.J. Activation of the IkappaB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. Cell 2000, 103, 351–361.
63. Plans, V.; Scheper, J.; Soler, M.; Loukili, N.; Okano, Y.; Thomson, T.M. The RING finger protein RNF8 recruits UBC13 for lysine 63-based self polyubiquitylation. J. Cell Biochem. 2006, 97, 572–582.
64. Shembade, N.; Harhaj, N.S.; Yamamoto, M.; Akira, S.; Harhaj, E.W. The human T-cell leukemia virus type 1 Tax oncoprotein requires the ubiquitin-conjugating enzyme Ubc13 for NF-kappaB activation. J. Virol. 2007, 81, 13735–13742.
65. Ulrich, H.D.; Walden, H. Ubiquitin signalling in DNA replication and repair. Nat. Rev. Ther. Mol. Cell Biol. 2010, 11, 479–489.
66. Bremm, A.; Komander, D. Emerging roles for Lys11-linked polyubiquitin in cellular regulation. Trends Biochem. Sci. 2011, 36, 355–363.
67. Xu, P.; Duong, D.M.; Seyfried, N.T.; Cheng, D.; Xie, Y.; Robert, J.; Rush, J.; Hochstrasser, M.; Finley, D.; Peng, J. Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation. Cell 2009, 137, 133–145.
68. Phu, L.; Izrael-Tomasevic, A.; Matsumoto, M.L.; Bustos, D.; Dynek, J.N.; Fedorova, A.V.; Bakalarski, C.E.; Arnott, D.; Deshayes, K.; Dixit, V.M.; et al. Improved quantitative mass spectrometry methods for characterizing complex ubiquitin signals. Mol. Cell Proteom. 2011, 10, M110.003756.
69. Durcan, T.M.; Tang, M.Y.; Perusse, J.R.; Dashti, E.A.; Aguileta, M.A.; McLelland, G.L.; Gros, P.; Shaler, T.A.; Faubert, D.; Coulombe, B.; et al. USP8 regulates mitophagy by removing K6-linked ubiquitin conjugates from parkin. EMBO J. 2014, 33, 2473–2491.
70. Ordureau, A.; Sarraf, S.A.; Duda, D.M.; Heo, J.M.; Jedrychowski, M.P.; Sviderskiy, V.O.; Olszewski, J.L.; Koerber, J.T.; Xie, T.; Beausoleil, S.A.; et al. Quantitative proteomics reveal a feedforward mechanism for mitochondrial PARKIN translocation and ubiquitin chain synthesis. Mol. Cell 2014, 56, 360–375.
71. Wu-Baer, F.; Lagrazon, K.; Yuan, W.; Baer, R. The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin. J. Biol. Chem. 2003, 278, 34743–34746.
72. Morris, J.R.; Solomon, E. BRCA1: BARD1 induces the formation of conjugated ubiquitin structures, dependent on K6 of ubiquitin, in cells during DNA replication and repair. Hum. Mol. Genet. 2004, 13, 807–817.
73. Michel, M.A.; Swatek, K.N.; Hospenthal, M.K.; Komander, D. Ubiquitin Linkage-Specific Affimers Reveal Insights into K6-Linked Ubiquitin Signaling. Mol. Cell 2017, 68, 233–246.e235.
74. Gatti, M.; Pinato, S.; Maiolica, A.; Rocchio, F.; Prato, M.G.; Aebersold, R.; Penengo, L. RNF168 promotes noncanonical K27 ubiquitination to signal DNA damage. Cell Rep. 2015, 10, 226–238.
75. Zhang, X.; Smits, A.H.; van Tilburg, G.B.; Jansen, P.W.; Makowski, M.M.; Ovaa, H.; Vermeulen, M. An Interaction Landscape of Ubiquitin Signaling. Mol. Cell 2017, 65, 941–955.e948.
76. Fei, C.; Li, Z.; Li, C.; Chen, Y.; Chen, Z.; He, X.; Mao, L.; Wang, X.; Zeng, R.; Li, L. Smurf1-mediated Lys29-linked nonproteolytic polyubiquitination of axin negatively regulates Wnt/beta-catenin signaling. Mol. Cell Biol. 2013, 33, 4095–4105.
77. Huang, H.; Jeon, M.S.; Liao, L.; Yang, C.; Elly, C.; Yates, J.R., 3rd; Liu, Y.C. K33-linked polyubiquitination of T cell receptor-zeta regulates proteolysis-independent T cell signaling. Immunity 2010, 33, 60–70.
78. Dittmar, G.; Winklhofer, K.F. Linear Ubiquitin Chains: Cellular Functions and Strategies for Detection and Quantification. Front. Chem. 2019, 7, 915.
79. Spit, M.; Rieser, E.; Walczak, H. Linear ubiquitination at a glance. J. Cell Sci. 2019, 132, jcs208512.
80. Liu, C.; Liu, W.; Ye, Y.; Li, W. Ufd2p synthesizes branched ubiquitin chains to promote the degradation of substrates modified with atypical chains. Nat. Commun. 2017, 8, 14274.
81. Ohtake, F.; Saeki, Y.; Ishido, S.; Kanno, J.; Tanaka, K. The K48-K63 Branched Ubiquitin Chain Regulates NF-kappaB Signaling. Mol. Cell 2016, 64, 251–266.

82. Ohtake, F.; Tsuchiya, H.; Saeki, Y.; Tanaka, K. K63 ubiquitylation triggers proteasomal degradation by seeding branched ubiquitin chains. Proc. Natl. Acad. Sci. USA 2018, 115, E1401–E1408.
83. Wertz, I.E.; Newton, K.; Seshasayee, D.; Kusam, S.; Lam, C.; Zhang, J.; Popovych, N.; Helgason, E.; Schoeffler, A.; Jeet, S.; et al. Erratum: Phosphorylation and linear ubiquitin direct A20 inhibition of inflammation. Nature 2016, 532, 402.
84. Kim, H.T.; Kim, K.P.; Lledias, F.; Kisselev, A.F.; Scaglione, K.M.; Skowyra, D.; Gygi, S.P.; Goldberg, A.L. Certain pairs of ubiquitinconjugating enzymes (E2s) and ubiquitin-protein ligases (E3s) synthesize nondegradable forked ubiquitin chains containing all possible isopeptide linkages. J. Biol. Chem. 2007, 282, 17375–17386.
85. Michel, M.A.; Elliott, P.R.; Swatek, K.N.; Simicek, M.; Pruneda, J.N.; Wagstaff, J.L.; Freund, S.M.; Komander, D. Assembly and specific recognition of k29- and k33-linked polyubiquitin. Mol. Cell 2015, 58, 95–109.
86. Peng, J.; Schwartz, D.; Elias, J.E.; Thoreen, C.C.; Cheng, D.; Marsischky, G.; Roelofs, J.; Finley, D.; Gygi, S.P. A proteomics approach to understanding protein ubiquitination. Nat. Biotechnol. 2003, 21, 921–926.
87. Wegmann, S.; Meister, C.; Renz, C.; Yakoub, G.; Wollscheid, H.P.; Takahashi, D.T.; Mikicic, I.; Beli, P.; Ulrich, H.D. Linkage reprogramming by tailor-made E3s reveals polyubiquitin chain requirements in DNA-damage bypass. Mol. Cell 2022.
88. Komander, D.; Rape, M. The ubiquitin code. Annu. Rev. Biochem. 2012, 81, 203–229.
89. Koyano, F.; Okatsu, K.; Kosako, H.; Tamura, Y.; Go, E.; Kimura, M.; Kimura, Y.; Tsuchiya, H.; Yoshihara, H.; Hirokawa, T.; et al. Ubiquitin is phosphorylated by PINK1 to activate parkin. Nature 2014, 510, 162–166.
90. Lamoliatte, F.; McManus, F.P.; Maarifi, G.; Chelbi-Alix, M.K.; Thibault, P. Uncovering the SUMOylation and ubiquitylation crosstalk in human cells using sequential peptide immunopurification. Nat. Commun. 2017, 8, 14109.
91. Liang, Y.C.; Lee, C.C.; Yao, Y.L.; Lai, C.C.; Schmitz, M.L.; Yang, W.M. SUMO5, a Novel Poly-SUMO Isoform, Regulates PML Nuclear Bodies. Sci. Rep. 2016, 6, 26509.
92. Evdokimov, E.; Sharma, P.; Lockett, S.J.; Lualdi, M.; Kuehn, M.R. Loss of SUMO1 in mice affects RanGAP1 localization and formation of PML nuclear bodies, but is not lethal as it can be compensated by SUMO2 or SUMO3. J. Cell Sci. 2008, 121, 4106–4113.
93. Wang, L.; Wansleeben, C.; Zhao, S.; Miao, P.; Paschen, W.; Yang, W. SUMO2 is essential while SUMO3 is dispensable for mouse embryonic development. EMBO Rep. 2014, 15, 878–885.
94. Khoo, K.H.; Verma, C.S.; Lane, D.P. Drugging the p53 pathway: Understanding the route to clinical efficacy. Nat. Rev. Drug Discov. 2014, 13, 217–236.
95. Kessler, J.D.; Kahle, K.T.; Sun, T.; Meerbrey, K.L.; Schlabach, M.R.; Schmitt, E.M.; Skinner, S.O.; Xu, Q.; Li, M.Z.; Hartman, Z.C.; et al. A SUMOylation-dependent transcriptional subprogram is required for Myc-driven tumorigenesis. Science 2012, 335, 348–353.
96. Su, H.L.; Li, S.S. Molecular features of human ubiquitin-like SUMO genes and their encoded proteins. Gene 2002, 296, 65–73.
97. Nayak, A.; Muller, S. SUMO-specific proteases/isopeptidases: SENPs and beyond. Genome Biol. 2014, 15, 422.
98. Bernier-Villamor, V.; Sampson, D.A.; Matunis, M.J.; Lima, C.D. Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. Cell 2002, 108, 345–356.
99. Pichler, A.; Gast, A.; Seeler, J.S.; Dejean, A.; Melchior, F. The nucleoporin RanBP2 has SUMO1 E3 ligase activity. Cell 2002, 108, 109–120.
100. Reverter, D.; Lima, C.D. Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. Nature 2005, 435, 687–692.
101. Johnson, E.S.; Gupta, A.A. An E3-like factor that promotes SUMO conjugation to the yeast septins. Cell 2001, 106, 735–744.

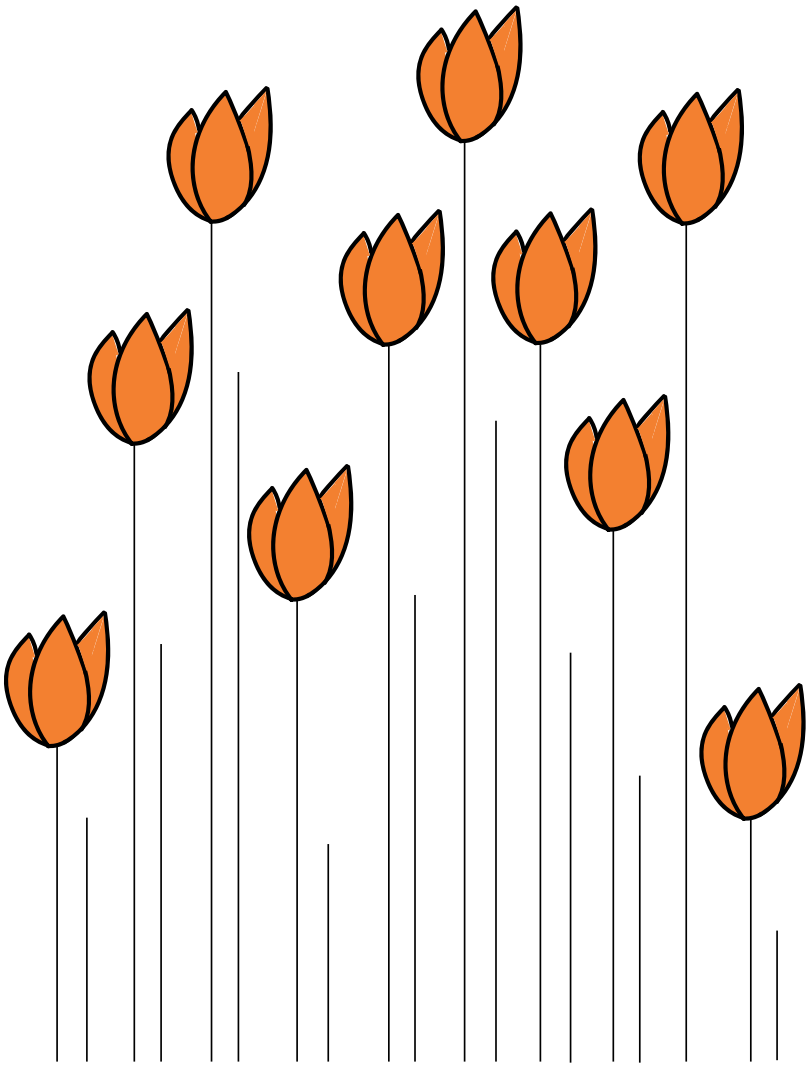
102. Kahyo, T.; Nishida, T.; Yasuda, H. Involvement of PIAS1 in the sumoylation of tumor suppressor p53. Mol. Cell 2001, 8, 713–718.
103. Yunus, A.A.; Lima, C.D. Structure of the Siz/PIAS SUMO E3 ligase Siz1 and determinants required for SUMO modification of PCNA. Mol. Cell 2009, 35, 669–682.
104. Li, C.; McManus, F.P.; Plutoni, C.; Pascariu, C.M.; Nelson, T.; Alberici Delsin, L.E.; Emery, G.; Thibault, P. Quantitative SUMO proteomics identifies PIAS1 substrates involved in cell migration and motility. Nat. Commun. 2020, 11, 834.
105. Schmidt, D.; Muller, S. Members of the PIAS family act as SUMO ligases for c-Jun and p53 and repress p53 activity. Proc. Natl. Acad. Sci. USA 2002, 99, 2872–2877.
106. Wu, C.S.; Zou, L. The SUMO (Small Ubiquitin-like Modifier) Ligase PIAS3 Primes ATR for Checkpoint Activation. J. Biol. Chem. 2016, 291, 279–290.
107. Sachdev, S.; Bruhn, L.; Sieber, H.; Pichler, A.; Melchior, F.; Grosschedl, R. PIASy, a nuclear matrix-associated SUMO E3 ligase, represses LEF1 activity by sequestration into nuclear bodies. Genes Dev. 2001, 15, 3088–3103.
108. Potts, P.R.; Yu, H. Human MMS21/NSE2 is a SUMO ligase required for DNA repair. Mol. Cell Biol. 2005, 25, 7021–7032.
109. Andrews, E.A.; Palecek, J.; Sergeant, J.; Taylor, E.; Lehmann, A.R.; Watts, F.Z. Nse2, a component of the Smc5-6 complex, is a SUMO ligase required for the response to DNA damage. Mol. Cell Biol. 2005, 25, 185–196.
110. Moreno-Onate, M.; Herrero-Ruiz, A.M.; Garcia-Dominguez, M.; Cortes-Ledesma, F.; Ruiz, J.F. RanBP2-Mediated SUMOylation Promotes Human DNA Polymerase Lambda Nuclear Localization and DNA Repair. J. Mol. Biol. 2020, 423, 3965–3979.
111. Dawlaty, M.M.; Malureanu, L.; Jeganathan, K.B.; Kao, E.; Sustmann, C.; Tahk, S.; Shuai, K.; Grosschedl, R.; van Deursen, J.M. Resolution of sister centromeres requires RanBP2-mediated SUMOylation of topoisomerase IIalpha. Cell 2008, 133, 103–115.
112. Eisenhardt, N.; Chaugule, V.K.; Koidl, S.; Droescher, M.; Dogan, E.; Rettich, J.; Sutinen, P.; Imanishi, S.Y.; Hofmann, K.; Palvimo, J.J.; et al. A new vertebrate SUMO enzyme family reveals insights into SUMO-chain assembly. Nat. Struct. Mol. Biol. 2015, 22, 959–967.
113. Koidl, S.; Eisenhardt, N.; Fatouros, C.; Droescher, M.; Chaugule, V.K.; Pichler, A. The SUMO2/3 specific E3 ligase ZNF451-1 regulates PML stability. Int. J. Biochem. Cell Biol. 2016, 79, 478–487.
114. Guervilly, J.H.; Takedachi, A.; Naim, V.; Scaglione, S.; Chawhan, C.; Lovera, Y.; Despras, E.; Kuraoka, I.; Kannouche, P.; Rosselli, F.; et al. The SLX4 complex is a SUMO E3 ligase that impacts on replication stress outcome and genome stability. Mol. Cell 2015, 57, 123–137.
115. Kagey, M.H.; Melhuish, T.A.; Wotton, D. The polycomb protein Pc2 is a SUMO E3. Cell 2003, 113, 127–137.
116. Weger, S.; Hammer, E.; Heilbronn, R. Topors acts as a SUMO-1 E3 ligase for p53 in vitro and in vivo. FEBS Lett. 2005, 579, 5007–5012.
117. Garcia-Gutierrez, P.; Juarez-Vicente, F.; Gallardo-Chamizo, F.; Charnay, P.; Garcia-Dominguez, M. The transcription factor Krox20 is an E3 ligase that sumoylates its Nab coregulators. EMBO Rep. 2011, 12, 1018–1023.
118. Tago, K.; Chiocca, S.; Sherr, C.J. Sumoylation induced by the Arf tumor suppressor: A p53-independent function. Proc. Natl. Acad. Sci. USA 2005, 102, 7689–7694.
119. Zhao, X.; Sternsdorf, T.; Bolger, T.A.; Evans, R.M.; Yao, T.P. Regulation of MEF2 by histone deacetylase 4- and SIRT1 deacetylasemediated lysine modifications. Mol. Cell Biol. 2005, 25, 8456–8464.
120. Subramaniam, S.; Mealer, R.G.; Sixt, K.M.; Barrow, R.K.; Usiello, A.; Snyder, S.H. Rhes, a physiologic regulator of sumoylation, enhances cross-sumoylation between the basic sumoylation enzymes E1 and Ubc9. J. Biol. Chem. 2010, 285, 20428–20432.
121. Roth, W.; Sustmann, C.; Kieslinger, M.; Gilmozzi, A.; Irmer, D.; Kremmer, E.; Turck, C.; Grosschedl, R. PIASy-deficient mice display modest defects in IFN and Wnt signaling. J. Immunol. 2004, 173, 6189–6199.

122. Wong, K.A.; Kim, R.; Christofk, H.; Gao, J.; Lawson, G.; Wu, H. Protein inhibitor of activated STAT Y (PIASy) and a splice variant lacking exon 6 enhance sumoylation but are not essential for embryogenesis and adult life. Mol. Cell Biol. 2004, 24, 5577–5586.
123. Werner, A.; Flotho, A.; Melchior, F. The RanBP2/RanGAP1*SUMO1/Ubc9 complex is a multisubunit SUMO E3 ligase. Mol. Cell 2012, 46, 287–298.
124. Joseph, J.; Tan, S.H.; Karpova, T.S.; McNally, J.G.; Dasso, M. SUMO-1 targets RanGAP1 to kinetochores and mitotic spindles. J. Cell Biol. 2002, 156, 595–602.
125. Hamada, M.; Haeger, A.; Jeganathan, K.B.; van Ree, J.H.; Malureanu, L.; Walde, S.; Joseph, J.; Kehlenbach, R.H.; van Deursen, J.M. Ran-dependent docking of importin-beta to RanBP2/Nup358 filaments is essential for protein import and cell viability. J. Cell Biol. 2011, 194, 597–612.
126. Aslanukov, A.; Bhowmick, R.; Guruju, M.; Oswald, J.; Raz, D.; Bush, R.A.; Sieving, P.A.; Lu, X.; Bock, C.B.; Ferreira, P.A. RanBP2 modulates Cox11 and hexokinase I activities and haploinsufficiency of RanBP2 causes deficits in glucose metabolism. PLoS Genet. 2006, 2, e177.
127. Abascal, F.; Tress, M.L.; Valencia, A. Alternative splicing and co-option of transposable elements: The case of TMPO/LAP2alpha and ZNF451 in mammals. Bioinformatics 2015, 31, 2257–2261.
128. Karvonen, U.; Jaaskelainen, T.; Rytinki, M.; Kaikkonen, S.; Palvimo, J.J. ZNF451 is a novel PML body- and SUMO-associated transcriptional coregulator. J. Mol. Biol. 2008, 382, 585–600.
129. Schellenberg, M.J.; Lieberman, J.A.; Herrero-Ruiz, A.; Butler, L.R.; Williams, J.G.; Munoz-Cabello, A.M.; Mueller, G.A.; London, R.E.; Cortes-Ledesma, F.; Williams, R.S. ZATT (ZNF451)-mediated resolution of topoisomerase 2 DNA-protein cross-links. Science 2017, 357, 1412–1416.
130. De La Cruz-Herrera, C.F.; Shire, K.; Siddiqi, U.Z.; Frappier, L. A genome-wide screen of Epstein-Barr virus proteins that modulate host SUMOylation identifies a SUMO E3 ligase conserved in herpesviruses. PLoS Pathog. 2018, 14, e1007176.
131. Rodriguez, M.S.; Dargemont, C.; Hay, R.T. SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. J. Biol. Chem. 2001, 276, 12654–12659.
132. Tatham, M.H.; Jaffray, E.; Vaughan, O.A.; Desterro, J.M.; Botting, C.H.; Naismith, J.H.; Hay, R.T. Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. J. Biol. Chem. 2001, 276, 35368–35374.
133. Hendriks, I.A.; Vertegaal, A.C. A comprehensive compilation of SUMO proteomics. Nat. Rev. Mol. Cell Biol. 2016, 17, 581–595.
134. Matic, I.; Schimmel, J.; Hendriks, I.A.; van Santen, M.A.; van de Rijke, F.; van Dam, H.; Gnad, F.; Mann, M.; Vertegaal, A.C. Sitespecific identification of SUMO-2 targets in cells reveals an inverted SUMOylation motif and a hydrophobic cluster SUMOylation motif. Mol. Cell 2010, 39, 641–652.
135. Yang, M.; Hsu, C.T.; Ting, C.Y.; Liu, L.F.; Hwang, J. Assembly of a polymeric chain of SUMO1 on human topoisomerase I in vitro. J. Biol. Chem. 2006, 281, 8264–8274.
136. Vertegaal, A.C.; Ogg, S.C.; Jaffray, E.; Rodriguez, M.S.; Hay, R.T.; Andersen, J.S.; Mann, M.; Lamond, A.I. A proteomic study of SUMO-2 target proteins. J. Biol. Chem. 2004, 279, 33791–33798.
137. Vogt, B.; Hofmann, K. Bioinformatical detection of recognition factors for ubiquitin and SUMO. Methods Mol. Biol. 2012, 832, 249–261.
138. González-Prieto, R.; Eifler-Olivi, K.; Claessens, L.A.; Willemstein, E.; Xiao, Z.; Ormeno, C.M.T.; Ovaa, H.; Ulrich, H.D.; Vertegaal, A.C.O. Global non-covalent SUMO interaction networks reveal SUMO-dependent stabilization of the non-homologous end joining complex. Cell Rep. 2021, 34, 108691.
139. Hendriks, I.A.; Lyon, D.; Su, D.; Skotte, N.H.; Daniel, J.A.; Jensen, L.J.; Nielsen, M.L. Site-specific characterization of endogenous SUMOylation across species and organs. Nat. Commun. 2018, 9, 2456.
140. Jansen, N.S.; Vertegaal, A.C.O. A Chain of Events: Regulating Target Proteins by SUMO Polymers. Trends Biochem. Sci. 2021, 46, 113–123.
141. Vertegaal, A.C. SUMO chains: Polymeric signals. Biochem. Soc. Trans. 2010, 38, 46–49.
142. Skilton, A.; Ho, J.C.; Mercer, B.; Outwin, E.; Watts, F.Z. SUMO chain formation is required for response to replication arrest in S. pombe. PLoS ONE 2009, 4, e6750.

143. Tatham, M.H.; Geoffroy, M.C.; Shen, L.; Plechanovova, A.; Hattersley, N.; Jaffray, E.G.; Palvimo, J.J.; Hay, R.T. RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. Nat. Cell Biol. 2008, 10, 538–546.
144. Lallemand-Breitenbach, V.; Jeanne, M.; Benhenda, S.; Nasr, R.; Lei, M.; Peres, L.; Zhou, J.; Zhu, J.; Raught, B.; de The, H. Arsenic degrades PML or PML-RARalpha through a SUMO-triggered RNF4/ubiquitin-mediated pathway. Nat. Cell Biol. 2008, 10, 547–555.
145. Yin, Y.; Seifert, A.; Chua, J.S.; Maure, J.F.; Golebiowski, F.; Hay, R.T. SUMO-targeted ubiquitin E3 ligase RNF4 is required for the response of human cells to DNA damage. Genes Dev. 2012, 26, 1196–1208.
146. Mukhopadhyay, D.; Arnaoutov, A.; Dasso, M. The SUMO protease SENP6 is essential for inner kinetochore assembly. J. Cell Biol. 2010, 188, 681–692.
147. Galanty, Y.; Belotserkovskaya, R.; Coates, J.; Jackson, S.P. RNF4, a SUMO-targeted ubiquitin E3 ligase, promotes DNA doublestrand break repair. Genes Dev. 2012, 26, 1179–1195.
148. Salas-Lloret, D.; Agabitini, G.; Gonzalez-Prieto, R. TULIP2: An Improved Method for the Identification of Ubiquitin E3-Specific Targets. Front. Chem. 2019, 7, 802.
149. Lecona, E.; Rodriguez-Acebes, S.; Specks, J.; Lopez-Contreras, A.J.; Ruppen, I.; Murga, M.; Munoz, J.; Mendez, J.; FernandezCapetillo, O. USP7 is a SUMO deubiquitinase essential for DNA replication. Nat. Struct. Mol. Biol. 2016, 23, 270–277.
150. Hendriks, I.A.; Schimmel, J.; Eifler, K.; Olsen, J.V.; Vertegaal, A.C.O. Ubiquitin-specific Protease 11 (USP11) Deubiquitinates Hybrid Small Ubiquitin-like Modifier (SUMO)-Ubiquitin Chains to Counteract RING Finger Protein 4 (RNF4). J. Biol. Chem. 2015, 290, 15526–15537.
151. Pfeiffer, A.; Luijsterburg, M.S.; Acs, K.; Wiegant, W.W.; Helfricht, A.; Herzog, L.K.; Minoia, M.; Bottcher, C.; Salomons, F.A.; van Attikum, H.; et al. Ataxin-3 consolidates the MDC1-dependent DNA double-strand break response by counteracting the SUMO-targeted ubiquitin ligase RNF4. EMBO J. 2017, 36, 1066–1083.
152. Hendriks, I.A.; Lyon, D.; Young, C.; Jensen, L.J.; Vertegaal, A.C.; Nielsen, M.L. Site-specific mapping of the human SUMO proteome reveals co-modification with phosphorylation. Nat. Struct. Mol. Biol. 2017, 24, 325–336.
153. Iconomou, M.; Saunders, D.N. Systematic approaches to identify E3 ligase substrates. Biochem. J. 2016, 473, 4083–4101.
154. Akimov, V.; Barrio-Hernandez, I.; Hansen, S.V.F.; Hallenborg, P.; Pedersen, A.K.; Bekker-Jensen, D.B.; Puglia, M.; Christensen, S.D.K.; Vanselow, J.T.; Nielsen, M.M.; et al. UbiSite approach for comprehensive mapping of lysine and N-terminal ubiquitination sites. Nat. Struct. Mol. Biol. 2018, 25, 631–640.
155. Kim, W.; Bennett, E.J.; Huttlin, E.L.; Guo, A.; Li, J.; Possemato, A.; Sowa, M.E.; Rad, R.; Rush, J.; Comb, M.J.; et al. Systematic and quantitative assessment of the ubiquitin-modified proteome. Mol. Cell 2011, 44, 325–340.
156. Wagner, S.A.; Beli, P.; Weinert, B.T.; Nielsen, M.L.; Cox, J.; Mann, M.; Choudhary, C. A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. Mol. Cell Proteom. 2011, 10, M111.013284.
157. Swatek, K.N.; Usher, J.L.; Kueck, A.F.; Gladkova, C.; Mevissen, T.E.T.; Pruneda, J.N.; Skern, T.; Komander, D. Insights into ubiquitin chain architecture using Ub-clipping. Nature 2019, 572, 533–537.
158. Matsumoto, M.L.; Dong, K.C.; Yu, C.; Phu, L.; Gao, X.; Hannoush, R.N.; Hymowitz, S.G.; Kirkpatrick, D.S.; Dixit, V.M.; Kelley, R.F. Engineering and structural characterization of a linear polyubiquitin-specific antibody. J. Mol. Biol. 2012, 418, 134–144.
159. Newton, K.; Matsumoto, M.L.; Ferrando, R.E.; Wickliffe, K.E.; Rape, M.; Kelley, R.F.; Dixit, V.M. Using linkage-specific monoclonal antibodies to analyze cellular ubiquitylation. Methods Mol. Biol. 2012, 832, 185–196.
160. Mattern, M.; Sutherland, J.; Kadimisetty, K.; Barrio, R.; Rodriguez, M.S. Using Ubiquitin Binders to Decipher the Ubiquitin Code. Trends Biochem. Sci. 2019, 44, 599–615.
161. Kliza, K.; Husnjak, K. Resolving the Complexity of Ubiquitin Networks. Front. Mol. Biosci. 2020, 7, 21.
162. Song, M.; Hakala, K.; Weintraub, S.T.; Shiio, Y. Quantitative proteomic identification of the BRCA1 ubiquitination substrates. J. Proteome Res. 2011, 10, 5191–5198.

163. Sarraf, S.A.; Raman, M.; Guarani-Pereira, V.; Sowa, M.E.; Huttlin, E.L.; Gygi, S.P.; Harper, J.W. Landscape of the PARKINdependent ubiquitylome in response to mitochondrial depolarization. Nature 2013, 496, 372–376.
164. Thompson, J.W.; Nagel, J.; Hoving, S.; Gerrits, B.; Bauer, A.; Thomas, J.R.; Kirschner, M.W.; Schirle, M.; Luchansky, S.J. Quantitative Lys–Gly-Gly (diGly) proteomics coupled with inducible RNAi reveals ubiquitin-mediated proteolysis of DNA damage-inducible transcript 4 (DDIT4) by the E3 ligase HUWE1. J. Biol. Chem. 2014, 289, 28942–28955.
165. Kim, B.J.; Chan, D.W.; Jung, S.Y.; Chen, Y.; Qin, J.; Wang, Y. The Histone Variant MacroH2A1 Is a BRCA1 Ubiquitin Ligase Substrate. Cell Rep. 2017, 19, 1758–1766.
166. Mark, K.G.; Simonetta, M.; Maiolica, A.; Seller, C.A.; Toczyski, D.P. Ubiquitin ligase trapping identifies an SCF(Saf1) pathway targeting unprocessed vacuolar/lysosomal proteins. Mol. Cell 2014, 53, 148–161.
167. Loveless, T.B.; Topacio, B.R.; Vashisht, A.A.; Galaang, S.; Ulrich, K.M.; Young, B.D.; Wohlschlegel, J.A.; Toczyski, D.P. DNA Damage Regulates Translation through beta-TRCP Targeting of CReP. PLoS Genet. 2015, 11, e1005292.
168. Streich, F.C., Jr.; Lima, C.D. Strategies to Trap Enzyme-Substrate Complexes that Mimic Michaelis Intermediates During E3- Mediated Ubiquitin-Like Protein Ligation. Methods Mol. Biol. 2018, 1844, 169–196.
169. Zhuang, M.; Guan, S.; Wang, H.; Burlingame, A.L.; Wells, J.A. Substrates of IAP ubiquitin ligases identified with a designed orthogonal E3 ligase, the NEDDylator. Mol. Cell 2013, 49, 273–282.
170. O'Connor, H.F.; Lyon, N.; Leung, J.W.; Agarwal, P.; Swaim, C.D.; Miller, K.M.; Huibregtse, J.M. Ubiquitin-Activated Interaction Traps (UBAITs) identify E3 ligase binding partners. EMBO Rep. 2015, 16, 1699–1712.
171. Coyaud, E.; Mis, M.; Laurent, E.M.; Dunham, W.H.; Couzens, A.L.; Robitaille, M.; Gingras, A.C.; Angers, S.; Raught, B. BioID-based Identification of Skp Cullin F-box (SCF)beta-TrCP1/2 E3 Ligase Substrates. Mol. Cell Proteom. 2015, 14, 1781–1795.
172. Barroso-Gomila, O.; Trulsson, F.; Muratore, V.; Canosa, I.; Merino-Cacho, L.; Cortazar, A.R.; Perez, C.; Azkargorta, M.; Iloro, I.; Carracedo, A.; et al. Identification of proximal SUMO-dependent interactors using SUMO-ID. Nat. Commun. 2021, 12, 6671.
173. Scott, D.C.; Rhee, D.Y.; Duda, D.M.; Kelsall, I.R.; Olszewski, J.L.; Paulo, J.A.; de Jong, A.; Ovaa, H.; Alpi, A.F.; Harper, J.W.; et al. Two Distinct Types of E3 Ligases Work in Unison to Regulate Substrate Ubiquitylation. Cell 2016, 166, 1198–1214.e1124.
174. Horn-Ghetko, D.; Krist, D.T.; Prabu, J.R.; Baek, K.; Mulder, M.P.C.; Klugel, M.; Scott, D.C.; Ovaa, H.; Kleiger, G.; Schulman, B.A. Ubiquitin ligation to F-box protein targets by SCF-RBR E3-E3 super-assembly. Nature 2021, 590, 671–676.
175. Senft, D.; Qi, J.; Ronai, Z.A. Ubiquitin ligases in oncogenic transformation and cancer therapy. Nat. Rev. Ther. Cancer 2018, 18, 69–88.
176. Deng, L.; Meng, T.; Chen, L.; Wei, W.; Wang, P. The role of ubiquitination in tumorigenesis and targeted drug discovery. Signal. Transduct. Target 2020, 5, 11.
177. Rabellino, A.; Melegari, M.; Tompkins, V.S.; Chen, W.; Van Ness, B.G.; Teruya-Feldstein, J.; Conacci-Sorrell, M.; Janz, S.; Scaglioni, P.P. PIAS1 Promotes Lymphomagenesis through MYC Upregulation. Cell Rep. 2016, 15, 2266–2278.
178. Rayner, S.L.; Morsch, M.; Molloy, M.P.; Shi, B.; Chung, R.; Lee, A. Using proteomics to identify ubiquitin ligase-substrate pairs: How novel methods may unveil therapeutic targets for neurodegenerative diseases. Cell Mol. Life Sci. 2019, 76, 2499–2510.
179. Bachmair, A.; Finley, D.; Varshavsky, A. In vivo half-life of a protein is a function of its amino-terminal residue. Science 1986, 234, 179–186.
180. Liu, P.; Cong, X.; Liao, S.; Jia, X.; Wang, X.; Dai, W.; Zhai, L.; Zhao, L.; Ji, J.; Ni, D.; et al. Global identification of phospho-dependent SCF substrates reveals a FBXO22 phosphodegron and an ERK-FBXO22-BAG3 axis in tumorigenesis. Cell Death Differ. 2022, 29, 1–13.

181. Langston, S.P.; Grossman, S.; England, D.; Afroze, R.; Bence, N.; Bowman, D.; Bump, N.; Chau, R.; Chuang, B.C.; Claiborne, C.; et al. Discovery of TAK-981, a First-in-Class Inhibitor of SUMO-Activating Enzyme for the Treatment of Cancer. J. Med. Chem. 2021, 64, 2501–2520.
182. Kroonen, J.S.; Vertegaal, A.C.O. Targeting SUMO Signaling to Wrestle Cancer. Trends Cancer 2021, 7, 496–510.
183. Kumar, S.; Schoonderwoerd, M.J.A.; Kroonen, J.S.; de Graaf, I.J.; Sluijter, M.; Ruano, D.; Gonzalez-Prieto, R.; Verlaan-de Vries, M.; Rip, J.; Arens, R.; et al. Targeting pancreatic cancer by TAK-981: A SUMOylation inhibitor that activates the immune system and blocks cancer cell cycle progression in a preclinical model. Gut 2022, 1–18.
184. Sha, Z.; Blyszcz, T.; Gonzalez-Prieto, R.; Vertegaal, A.C.O.; Goldberg, A.L. Inhibiting ubiquitination causes an accumulation of SUMOylated newly synthesized nuclear proteins at PML bodies. J. Biol. Chem. 2019, 294, 15218–15234.
185. Mojsa, B.; Tatham, M.H.; Davidson, L.; Liczmanska, M.; Branigan, E.; Hay, R.T. Identification of SUMO Targets Associated With the Pluripotent State in Human Stem Cells. Mol. Cell Proteom. 2021, 20, 100164.
186. Joung, J.; Konermann, S.; Gootenberg, J.S.; Abudayyeh, O.O.; Platt, R.J.; Brigham, M.D.; Sanjana, N.E.; Zhang, F. Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. Nat. Protoc. 2017, 12, 828–863.
187. Tang, C.P.; Clark, O.; Ferrarone, J.R.; Campos, C.; Lalani, A.S.; Chodera, J.D.; Intlekofer, A.M.; Elemento, O.; Mellinghoff, I.K. GCN2 kinase activation by ATP-competitive kinase inhibitors. Nat. Chem. Biol. 2022, 18, 207–215.
188. Weller, C.E.; Dhall, A.; Ding, F.; Linares, E.; Whedon, S.D.; Senger, N.A.; Tyson, E.L.; Bagert, J.D.; Li, X.; Augusto, O.; et al. Aromatic thiol-mediated cleavage of N-O bonds enables chemical ubiquitylation of folded proteins. Nat. Commun. 2016, 7, 12979.
189. Bondalapati, S.; Eid, E.; Mali, S.M.; Wolberger, C.; Brik, A. Total chemical synthesis of SUMO-2-Lys63-linked diubiquitin hybrid chains assisted by removable solubilizing tags. Chem. Sci. 2017, 8, 4027–4034.
190. Mulder, M.P.C.; Merkx, R.; Witting, K.F.; Hameed, D.S.; El Atmioui, D.; Lelieveld, L.; Liebelt, F.; Neefjes, J.; Berlin, I.; Vertegaal, A.C.O.; et al. Total Chemical Synthesis of SUMO and SUMO-Based Probes for Profiling the Activity of SUMO-Specific Proteases. Angew. Chem. Int. Ed. Engl. 2018, 57, 8958–8962.
191. Wang, Y.S.; Wu, K.P.; Jiang, H.K.; Kurkute, P.; Chen, R.H. Branched Ubiquitination: Detection Methods, Biological Functions and Chemical Synthesis. Molecules 2020, 25, 5200



2

Unveiling BRCA1-BARD1 ubiquitin ligase heterodimer. DNA repair, Ubiquitin and Cancer

Daniel Salas-Lloret[1] and Román González-Prieto[1,2]

[1] Cell and Chemical Biology, Leiden University Medical Centre, Einthovenweg 20, 2333ZC Leiden, The Netherlands

[2] Genome Proteomics Group, Department of Genome Biology, Andalusian Centre for Regenerative Medicine and Molecular Biology (CABIMER), Américo Vespucio 24, 41092 Seville, Spain

This chapter has been reviewed by DNA Repair (2022).

Abstract

Double strands breaks (DSBs) are the major source of genetic instability. Cells are able to sense the damage and develop a quick and accurate DNA damage response (DDR), which normally ends fixing the DSB. One of the fastest ways to give a response is by post-translational modification (PTM) of proteins. Ubiquitination is a PTM that governs the DDR, from the beginning to the end. E3 ubiquitin ligases are key players during the DDR and their dysregulation is associated to cancer. How different E3 ligases work together to regulate the spread, the repair pathway choice and the termination of the DDR is graphically depicted in this review. We focus on BRCA1/BARD1 as a multifunctional E3 ligase with crucial roles in DNA damage repair and tumor suppression. Here, we decipher the, yet controversial, role of BRCA1/BARD1 E3 ligase activity in homologous recombination and the possible roles on DNA replication and anti-tumorigenesis. Finally, we discuss novel strategies to target the ubiquitination machinery during the DDR and the future directions in the field.

Keywords: DNA damage; Ubiquitin; E3 ligases; BRCA1/BARD1 heterodimer

1. INTRODUCTION

The maintenance of genome integrity is crucial to avoid genetic alterations, from chromosome rearrangements to point mutations, that are associated with pathological disorders, premature aging, inherited diseases and cancer (1). DNA is constantly threatened by exogenous agents such as ionizing radiation (IR), mutagenic chemicals or UV light, as well as endogenous agents such as free radicals, single stranded DNA (ssDNA) and DNA replication problems, resulting in more than 70,000 lesions per cell per day (2). DNA replication is frequently challenged with obstructions of the DNA polymerase (DNA secondary structures, R-loops, etc) and transcription-replication conflicts that could cause replication fork stalling or collapse. Failure to restart stalled replication forks and difficulties resolving single strand breaks (SSB), can lead to double strand breaks (DSBs), which are the most harmful DNA lesions.

Fortunately, cells are able to sense the DNA damage and elaborate a fast and accurate response to solve the lesion. During this process ubiquitination is crucial not only to recruit downstream proteins to DNA damage sites, but also to repress transcription, choose the repair pathway and control the lifespan of proteins during the repair. As key player with E3 ligase activity in the repair of DSB, BRCA1/BARD1 heterodimer is known as tumor suppressor and germline mutations in either BRCA1 or BARD1 lead to development of breast and ovarian cancer (3). However, its E3 activity is not well-defined neither during DSB repair nor as tumor suppressor.

In this review, we will focus on ubiquitination as the key PTM that governs the DDR; the interplay of E3 ligases to achieve proper DNA repair; and the role of BRCA1/BARD1 E3 activity in homologous recombination, replication and tumor suppression.

2. UBIQUITINATION REGULATING THE DDR

In order to repair DSBs, cells are equipped with a plethora of conserved DNA damage sensing and repair mechanisms that combined are known as the DNA damage response (DDR) (13). DDR is tightly regulated by PTMs, including phosphorylation, SUMOylation, methylation and ubiquitination (14, 15). The cross-talk between PTMs can manage the whole DDR. The damage is sensed by kinases which produce phosphorylation cascades in seconds. The continuity of the DDR relies on the recruitment of proteins to DNA damage sites through polyubiquitination chains and scaffold proteins. The choice of repair pathway is influenced by histone modifications, as one modification not only leads to a specific repair pathway but inhibits another one. Finally, the shutdown of the response is also managed by PTMs, where in all mentioned steps, ubiquitination and E3 ligases play a crucial role to successfully repair the DNA damage.

2.1 EARLY STAGE

The first step to sense DNA damage and initiate the DDR is the recruitment of proteins to DNA damage sites, such as DSBs. ATM kinase is a main player in sensing the damage. It is responsible for downstream substrates activation, such as p53, BRCA1 and 53BP1 which will lead to DNA repair, cell cycle progression or apoptosis (16). ATM phosphorylates p53 for its recruitment to chromatin and also the E3 ligase that regulates p53 stability, mouse double minute 2 (MDM2). MDM2 controls p53 levels by proteasomal degradation and its overexpression is related to several cancers (17, 18). ATM-mediated phosphorylation of MDM2 upon DNA damage, results in the

inhibition of its E3 activity against p53. Thus MDM2 phosphorylation allows p53 stabilization at DNA damage sites (19).

Another player recognizing DSBs is the MRN complex (Mre11, Rad50 and NBS1) which is responsible for the initial DNA end resection at DSBs. ATM associates with the MRN complex to promote Histone H2AX phosphorylation (γH2AX) that will be propagated along the DSB. This histone phosphorylation functions as DNA damage mark, and several proteins are recruited to damage sites using γH2AX as a platform (20). One of the proteins that is recruited to DNA damage sites through γH2AX is Mediator of DNA damage Checkpoint protein 1 (MDC1), that interacts with γH2AX by its BRCT domain. Once it is recruited, ATM phosphorylates MDC1, which is necessary for the recruitment of one of the first ubiquitin E3 ligases, RNF8 (21).

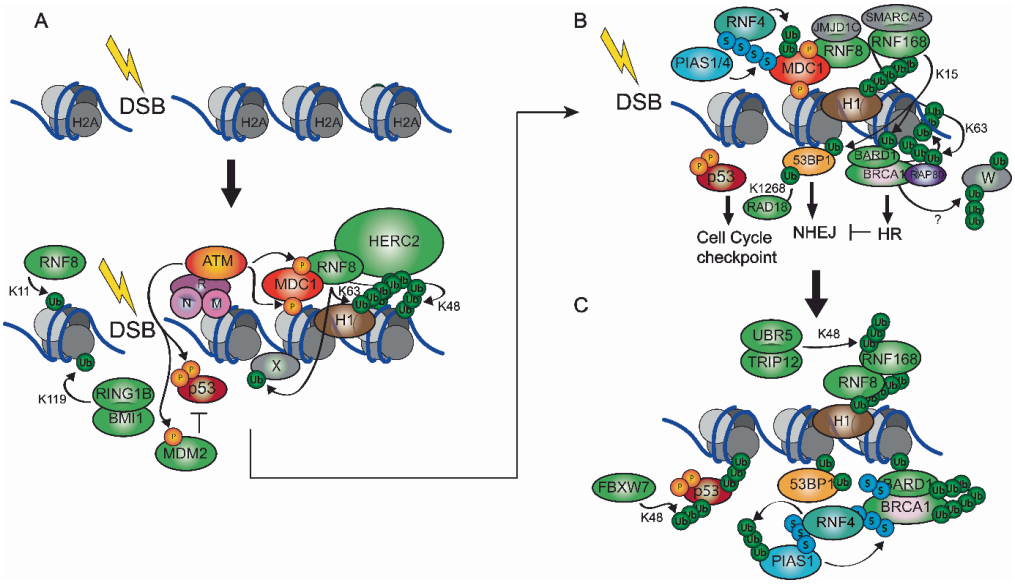


Figure 1. E3 ligases in DDR. DDR after DSBs displaying the role of several E3 ligases (green) with kinases (red) and SUMO E3 ligases (blue). **A.** Early stage where ATM and MRN complex recognize the DBS, ATM phosphorylates H2A and MDM2, it also promotes p53 and MDC1 recruitment to DSBs. MDM2 phosphorylation inhibits p53 degradation, leading to p53 stabilization at DSBs. MDC1 gets phosphorylated by ATM and promotes RNF8 recruitment, which promotes histone ubiquitination leading to RNF168 recruitment. **B.** RNF8-RNF168 stimulates additional histone ubiquitination necessary for the recruitment of downstream proteins such as 53BP1 and BRCA1/BARD1, which will lead to repair pathway choice and cell fate. **C.** Ending of the response, where UBR5 and TIP12 E3 ligases target RNF168 for degradation. STUbL RNF4 labels MDC1 and the SUMO machinery for its proteasomal degradation and E3 ligase FBXW7 targets p53 for degradation after DSBs repair.

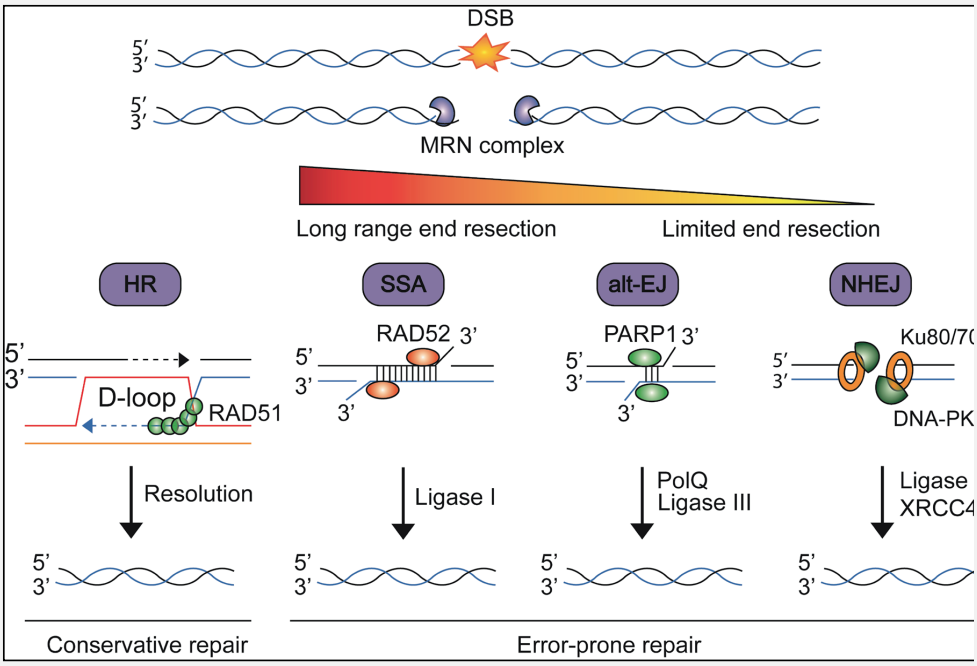
To assure proper DNA repair, RNF8 regulates DNA damage-induced transcription inhibition by K11 linkages on damaged chromatin, protecting damage sites from transcription-repair conflicts (22). Another E3 ligase with a possible role also in genome silencing after DNA damage is E3 BMI1-

RING1B, which is a subunit of the Polycomb Repressive Complex 1 (PRC1), and accumulates at DSB sites to locally increase H2A K119 monoubiquitination. It is recruited after H2AX phosphorylation and might also have a role during the DDR, as it remains at DNA damage sites for 8h post-damage. (23, 24). On the other hand, upon MDC1 interaction, RNF8 poly-K63 ubiquitinates H1 type linker histones, which is responsible for RNF168 recruitment through its motif interacting with ubiquitin (MIU) to amplify the signal (**Figure 1A**). However, spatiotemporal recruitment experiments and the identification of new ubiquitination substrates would be key to spot every E3 ligase and better understand how this ubiquitination network is formed around DNA damage sites.

Once RNF168 is recruited, together with RNF8, it stimulates additional histone ubiquitination and promote the accumulation of downstream proteins such as 53BP1 and BRCA1/BARD1 at DNA damage sites, which will lead to the repair pathway choice (**Figure 1B**) (**BOX 1. DNA repair pathways**) (25).

BOX 1. DNA repair pathways.

Once DNA damage is sensed and the DDR is active and propagated, the DSB must be repaired. Cells have developed several conserved but mechanistically different DSBs repair pathways, including Homologous Recombination (HR), non-homologous end joining (NHEJ), alternative end joining (alt-EJ) and single-strand annealing (SSA).



HR is an error-free repair pathway as it can faithfully restore the original configuration of the broken DNA molecule using the intact sister chromatid as template for repair. Therefore, HR is restricted to S and G2 phase of the cell cycle and requires DNA resection. In contrast, NHEJ, alt-EJ and SSA are error-prone repair pathways. Classical-NHEJ (cNHEJ) is considered the default mechanism for DSB repair in a cell cycle independent manner. Generally, cNHEJ does not require end-trimming and the resolution of the DNA damage typically comprises the deletion or insertion of a few nucleotides. Alt-EJ and SSA require DNA end resection and, in the case of SSA, no limit of end-resection has been determined yet. However, these mechanisms are prone to generate deletions and chromosomes translocations (8). The nucleotide depletion/insertion of the repair pathway can be monitored by employing CRISPR technologies with consequent sequencing analysis, giving rise to different mutational signatures (9).

2.2 MID-LATE STAGE

Before going through mitosis, DNA damage should be resolved and the DDR must come to an end. An efficient way to close a signaling pathway is by targeting its key regulators. One key player in the response, is RNF168. Two E3 ligases, UBR6 and TRIP12, cooperate together to efficiently control the downstream RNF168 events by targeting RNF168 to proteasomal degradation.

Depletion of these two E3 ligases results in supra-physiological accumulation of RNF168 and downstream DDR factors, compromising proper DNA repair (26).

The Sumo Targeting Ubiquitin Ligase (STUbL) RNF4 also plays an important role in this tightly regulated DDR, controlling the residence time of MDC1, which gets SUMOylated by PIAS1/4 and is subsequently labelled for proteasomal degradation by RNF4, controlling the response in a middle stage (27). Persistence of MDC1 prevents downstream signaling through the HR pathway. Therefore, MDC1 removal is required prior to CtIP and RAD51 recruitment (**Figure 3**) (28). In previous work, RNF168 was also found as RNF4 target for proteasomal degradation (29, 30) involving RNF4 in the regulation of RNF168. Together with the fact that BRCA1/BARD1 is SUMOylated upon DNA damage by PIAS1/4 and that RNF4 targets the SUMOylation machinery for proteasomal degradation (29, 30), RNF4 might be also a key factor in the closure of the DDR.

Finally, other proteins like ubiquitin-selective chaperone/ segregase (VCP/p97) have a role regulating the spatiotemporal localization of the recruited chromatin-associated proteins (31) (**Figure 1C**). The DDR is orchestrated by crosstalk between PTMs and it is tightly regulated in space and time. We display the role of main E3 ligases in DDR (**Figure 1**) (**Table 1**). However, more and more E3 ligases are emerging in the DDR field, such as Pellino 1, which appears to interact with phosphorylated p53 (32) (**BOX 2. Other E3 ligases in DDR**). Additionally, new roles are being contributed to the E3 ligases we mentioned above. For example, RNF168 has been recently involved in the recruitment of SLX4 for inter-strand crosslink (ICL) damage repair (33). This denotes that E3 ligases are key regulators in the DDR and their dysregulation may lead to cancer and other diseases.

3. BRCA1/BARD1 AS UBIQUITIN E3 LIGASE

The breast cancer susceptibility gene 1 (BRCA1) and its obligated partner BRCA1-associated RING domain protein 1 (BARD1) form a heterodimer through a four-helix bundle flanking their Really Interesting New Gene (RING) domains (34). Despite BRCA1 is a large protein that participates in multiple cellular activities through forming different complexes with other proteins (35), the E3 ligase activity coming from BARD1 interaction is the only known intrinsic enzymatic activity of this heterodimer. In vitro experiments might show independent E3 ligase activity for each partner, although the formation of the heterodimer exhibits substantially greater E3 ligase activity (36, 37). However, it is generally accepted that BARD1 does not possess inherent E3 ligase activity as it lacks the alpha-helix required for binding to the E2 (38). The residual BARD1-E3 activity might come from the presence of either endogenous BRCA1 or another E3 after immunoprecipitation from cell lysates (37). Later, cryo-EM structures showed that BRCA1 RING domain preferentially binds the E2 while BARD1 RING domain shows higher affinity to the substrate during H2A ubiquitination (**Figure 2**) (38, 39).

Table 1. E3 ligases during the DDR.

E3 Ligase	DDR Stage	Target	Linkage	Function	Ref.
RNF8	Early	H2A/H2AX H1	K11 K63	Transcription inhibition Recruitment	(22) (21)
BMI1/PRC1	Early	H2A	Mono	Genome silencing	(24)
MDM2	Early	P53	?	Cell cycle	(2)
HERC2	Early	H2A/H2AX P53	K63	Recruitment Cell cycle	(1) (2)
RNF168	Medium	H2A/H2AX	K63 Mono	Recruitment DNA signalling BRCA1/BARD1	(21) (43) (47)
BRCA1/BARD1	Medium	H2A/H2AX mH2A	Mono	HR Cellular senescence	(39) (105)
RAD18	Medium	53BP1	Mono	NHEJ	(4)
UBR6	Late	RNF168	K48?	DDR regulation	(26)
TRIP12	Late	RNF168	K48?	DDR regulation	(26)
FBXW7	Late	P53	K48	Cell cycle and DDR	(5)
PELLINO1	Late	?	?	Cell cycle and DDR	(32)
RNF4	Medium Late	MDC1 RNF168	K48? K48?	DDR regulation DDR regulation	(28) (29)

3.1 BRCA1/BARD1 IN REPAIR PATHWAY CHOICE

BRCA1 and 53BP1 are essentially engaged in a tug of war that determines commitments to two different DSBs repair pathways, HR or NHEJ, where the cell cycle plays a crucial role (40) (**BOX 1. DNA Repair Pathways**). BRCA1-RAP80 interacts through its UIMs with K63-linked ubiquitin facilitating the recruitment to DNA damage sites.

Furthermore, knocking down either RAP80, RNF8 or RNF168 prevents BRCA1/BARD1 recruitment to DSBs (41-43). A recent study, using RAP80 CRISPR knock outs (RAP80-KO), found that RAP80-BRCA1 complex is dispensable for the initial recruitment of BRCA1 to DSB and the BRCA1 RING domain is critical for the recruitment in RAP80 deficient cells (44). IR-induced foci localization showed that the BRCA1 recruitment to DSB was completely dependent on RNF8 and partially dependent on RNF168, but displayed a near-normal recruitment in 7 independent RAP80-KO cell lines. Previous data also corroborate the importance of BRCA1 RING domain as mutation of K70/K71 in BRCA1 specifically disrupts the ability of the RING domain to bind to the nucleosome acidic path (45). Likewise, this result is supported by a recent study where RNF168 seems responsible for localizing the BRCA1-PALB2 complex in DSBs (46). These findings propose two possible ways of BRCA1 recruitment, being RNF168-mediated monoubiquitination on H2AK13/15 the predominant pathway, while BRCA1-RAP80 complex contributes as backup pathway.

BOX 2. Other E3 ligases in DDR

Another E3 ligase that aids RNF8 for initiating the recruitment of downstream proteins, is the large HECT-type ligase (HERC2), which interacts with RNF8 to promote downstream ubiquitination at DSB sites. Most of the ubiquitination chains formed by RNF8 and HERC2 are on K63 (1). Recent studies propose HERC2 E3 as regulator of the p53-MDM2 pathway. HERC2 might form a complex with MDM2 that ubiquitinates p53 for proteasomal degradation. Upon DNA damage and consequent MDM2 phosphorylation, the complex is dissociated and p53 is active at DSBs (2).

RNF8-mediated K63 polyubiquitylation also leads to the recruitment of another ubiquitin E3 ligase, RAD18, in a ubiquitin binding domain (UBD) dependent manner. RAD18 mediates RAD9 recruitment through its RING and Zinc finger domains, which might play a role in DSB repair and downstream activation of checkpoints (3). Another possible role of RAD18 in DSB is the monoubiquitination of 53BP1 at K1268, which may retain 53BP1 at DNA damage sites during G1-phase promoting NHEJ repair pathway (4) (**Figure 1B**).

Recently, the role of FBXW7 E3 ligase has been reported to ensure proper DNA repair. ATM is responsible for p53 phosphorylation at S33 and S37 in response to DSBs, as mentioned above. This phosphorylation is required for the interaction with substrate recognition component of FBXW7 that mediates p53 ubiquitination for proteasome degradation once it has been recruited to DNA damaged sites (5) (**Figure 1C**). For regulating cell cycle checkpoints and inducing apoptosis when the DNA damage cannot be solved, p53 is known as the “guardian of the genome”, thus it is not surprising that p53 levels must be precisely regulated under DSBs, letting DNA be repaired before continuing cell cycle progression. This mechanism is poorly understood compared with the widely investigated role of MDM2 targeting p53 for proteasomal degradation (6,7).

Recently, it has been shown how both BARD1 and 53BP1 can bind the same mono H2AK15Ub (**Figure 1B**), but subsequent post-replication histone modifications might govern the DSB repair choice (47). As shown before, the DDR is strongly regulated, and it is not surprising that other mechanisms collaborate to regulate BRCA1 and 53BP1 recruitment. It was reported that the TIP60 complex competes with RNF168 for H2AK15 modification. Acetylation by TIP60 inhibits RNF168 H2AK15 ubiquitination, disrupting the recruitment of 53BP1 and BRCA1/BARD1 to DSBs (48).

Additionally, not only H2A modification is involved in the recruitment of these proteins, but histone 4 (H4) methylation at lysine 20 (H4K20m) has been revealed as a key factor for BRCA1/BARD1 recruitment (47, 49). During G1 phase of the cell cycle, H4K20 is methylated and BRCA1/BARD1 cannot be recruited to DSBs. However, H4K20 is unmethylated right after replication allowing BRCA1/BARD1 recruitment and HR performance during S phase. In concordance with this findings, cryo-EM experiments showed how the Ankyrin and tandem BRCT (BUDR) domains of BARD1 can adopt a compact fold and bind nucleasomal histones, DNA and the monoubiquitin attached to H2A amino-terminal K13/15, to promote ubiquitination of the flexible carboxy-terminal tails of H2A and H2AX (**Figure 2**) (50). BRCA1/BARD1 not only binds and ubiquitinates H2A at K125/127/129 (50-52), but also blocks other ubiquitination events on H2A as K63 linkages, which are responsible for 53BP1 recruitment (50, 53). Deubiquitinating enzymes are also involved in this recruitment. BRCA1/BARD1-mediated H2A ubiquitination seems to be regulated by the deubiquitinating enzyme USP48, which acts as modulator antagonizing BRCA1/BARD1 E3 activity. Loss of USP48 resulted in further 53BP1 repositioning from the break site and extending resection lengths (51). The possible role of BRCA1/BARD1-mediated H2A ubiquitylation may reside in the eviction of 53BP1, which is an agonist of HR and promotes NHEJ repair, and continuation of the HR repair pathway (51, 54).

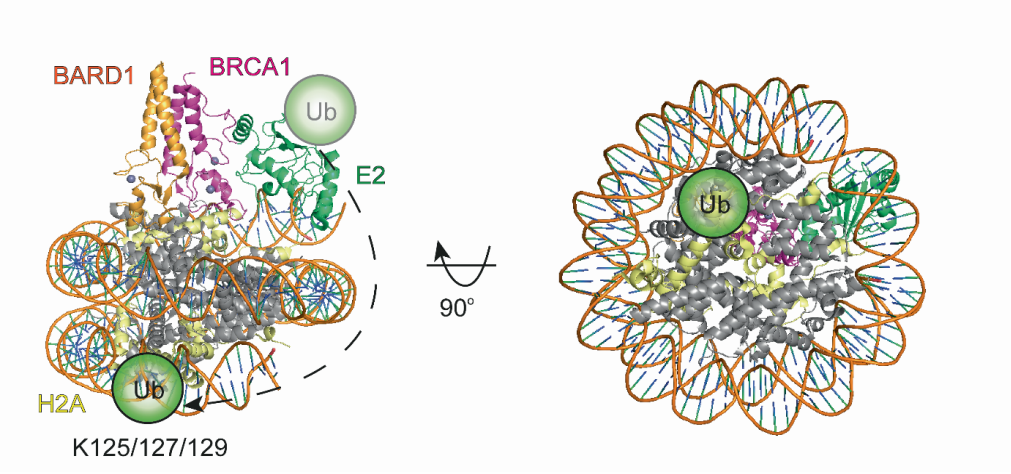


Figure 2. BRCA1-BARD1 RING domains interacting with the nucleosome. Structure obtained from PDB:7JZV. BRCA1 interacts with the E2 enzyme UbcH5c and BARD1 directs H2A ubiquitination at K125/127/129.

3.2 BRCA1/BARD1 IN HR

The first role of BRCA1/BARD1 in HR was related to the interaction between BRCA1 and the recombinase RAD51, which colocalized after ionizing radiation (IR) (55). Following upstream events of the HR pathway (**Figure 3**), BRCA1 has been related to DNA end resection. BRCA1 can form a complex with CtIP and MRN in a cell cycle dependent manner. CtIP associates with BRCA1 through the BRCT domains whereas MRN interacts with the N terminus of BRCA1. This complex promotes the essential steps of DNA resection by opposing the block on resection by 53BP1 and its effector proteins (56). The major evidence of E3 ligase activity involved in this process was observed in BARD1 knock down cells complemented with either a BARD1-WT or BARD1-R99E mutant version. Cell complemented with the BARD1-R99E showed decreased numbers, size and intensity of RPA foci compared with cells complemented with BARD1-WT (54). Additionally, it was reported that BRCA1 is required for CtIP ubiquitination upon DNA damage for G2/M checkpoint control (57). However, rescue experiments including a BRCA1 catalytic dead mutant could indicate whether CtIP ubiquitination depends on BRCA1 E3 activity.

In early years, most investigators assumed that the E3 activity of BRCA1/BARD1 would also be essential for BRCA1/BARD1-mediated HR. However, in 2008 it was reported that the HR function of BRCA1-I26A cells was indistinguishable from isogenic BRCA1-WT cells, as measured by RAD51 foci formation and recombination of a chromosomally-integrated HR reporter constructs (58). In 2016, these conclusions were challenged by the Morris laboratory, who reported that cells expressing the BARD1-R99E mutation, which also disrupts the E3 activity of BRCA1/BARD1 without impairing the heterodimer formation, are both HR deficient and PARPi hypersensitive (54).

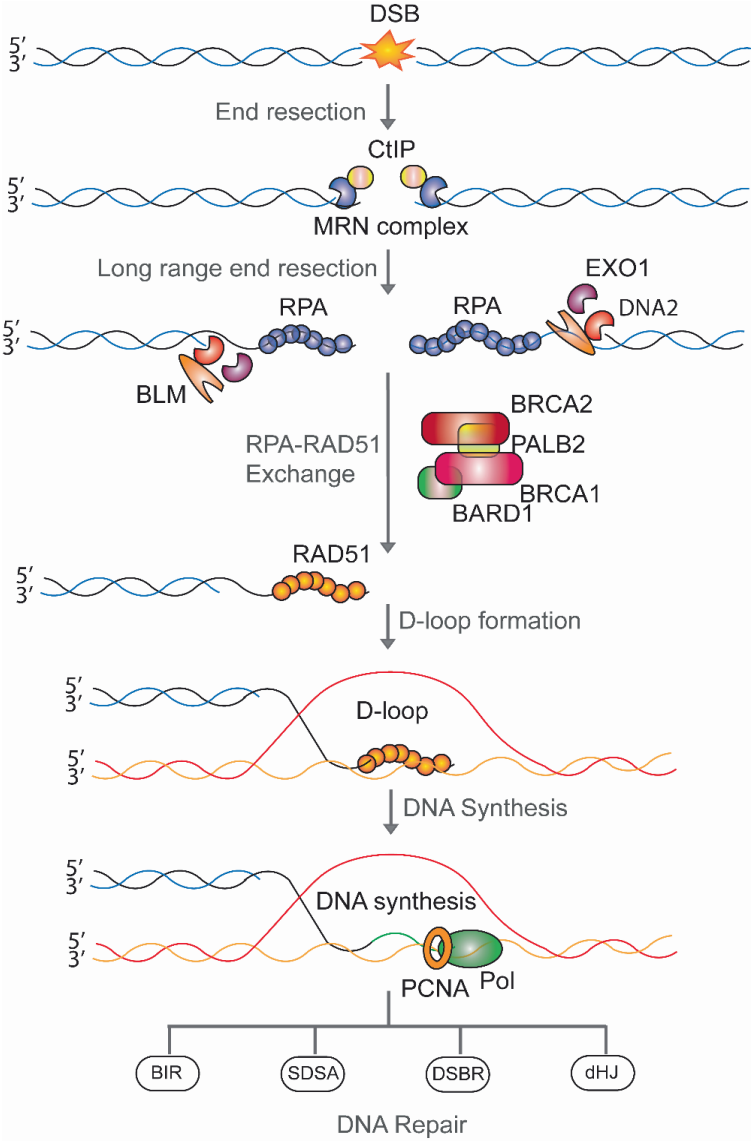


Figure 3. Homologous Recombination Pathway. Once the double strand break (DSB) is produced, BRCA1 interacts with CtIP and the MRN complex to initiate DNA end resection. Exonuclease 1 (EXO1), endonuclease DNA2 and the helicase BLM promote long range end resection. This process is also controlled by the BRCA1-ABRAXAS complex. Replication protein A (RPA) coats ssDNA for subsequent exchange with the recombinase RAD51. This process is mediated by the BRCA1-PALB2-BRCA2 complex. RAD51 will procced to DNA invasion and the formation of the D-loop. Then, DNA polymerase together with PCNA will synthesize new DNA for final resolution by synthesis-dependent single-strand annealing (SDSA), canonical double-strand break repair (DSBR), double Holliday junction (dHJ) resolution or dissolution and break-induced DNA replication (BIR).

Additionally, BRCA1-BARD1 E3 activity was associated to HR by the link between ubiquitination of histone H2A and IR-induced RAD51 foci formation. In the absence of BARD1, there was no presence of RAD51 foci formation, but the foci accumulation was rescued with a H2A-Ub variant and not with a H2A-WT or a Ub-H2A form, suggesting that BRCA1/BARD1-mediated H2A ubiquitination was essential for the progression of the HR pathway (54). In support, a more detailed mechanism was published using cryo-EM (50). BRCA1/BARD1 binds the N-terminus of H2A to produce C-terminus H2A ubiquitination, which seems to be important for the pathway choice and HR continuation (**Figure 4A**). However, HR deficiencies and PARPi hypersensitivity do not appear to be reproducible. Other laboratories observed no effect of the BARD1-R99E mutation on PARPi sensitivity (46, 49). Likewise, cells expressing another E3-impaired mutant (BRCA1-3A) were found to be fully competent for PARPi resistance and IR-induced RAD51 foci formation (44).

In their recent study, Sherker et al. (44), also reported that HR was fully proficient in cells expressing the BRCA1-3A mutant, while HR is abrogated in RAP80-null cells that express BRCA1-3A. Although the interaction between RAP80 and BRCA1/BARD1-mediated ubiquitination needs further validation and the relevant substrates remain unclear, this result might explain why one laboratory observed a requirement for the E3 activity in HR (54) while most laboratories do not (44, 46, 49, 58). Nonetheless, since the RAP80 complex is broadly expressed across mammalian tissues, loss of RAP80 functions seems to be either an artificial or very limited condition. Thus, it remains unclear whether BRCA1/BARD1-mediated ubiquitination is relevant to HR during either normal or malignant development. It may be possible to identify a physiological role by further analysis of BRCA1-I26A mice. Unlike BRCA1-C61G mice, which die early during embryogenesis, BRCA1-I26A mice only exhibit modest developmental defects, such as reduced body weight and male sterility due to a late block in spermatogenesis. These defects may point to the missing physiological functions of BRCA1/BARD1-mediated ubiquitination. Therefore, identifying BRCA1/BARD1-mediated ubiquitination targets would be a crucial step forward to resolving the enigmatic role of this E3 ligase.

3.3 BRCA1/BARD1 E3 LIGASE IN TUMOR SUPRESSION

BRCA1 is a well-known tumor suppressor protein (59) and since the RING domain is required for both its interaction with BARD1 and its E3 activity, the function of the enzymatic activity has been studied using carefully-designed separation-of-function mutations that specifically ablate the ligase activity without impairing heterodimer formation (**Figure 4A**).

Mutations affecting the RING domain of the heterodimer can be found in both partners and divided in two groups. On one hand, BARD1 L44R and BRCA1 C61G mutations disrupt the formation of the heterodimer leading not only to E3 ligase activity depletion but also to BRCA1 instability, resulting in its proteasomal degradation (54, 60) (**Figure 4A**). These mutations have been found in tumors (61, 62) and showed a tumorigenic profile in mice studies. BRCA1-C61G mice showed increased breast tumor formation similar to BRCA1 deficient mice (63). In human cells, BARD1-L44R and BRCA1-C61G shared similar HR and genotoxic stress phenotypes compared to BARD1 and BRCA1 deficient cells (54).

On the other hand, BARD1-R99E and BRCA1-I26A RING domain mutations disrupt the interaction with the E2 and abrogate the E3 ligase activity while keeping the heterodimer assembly

(38, 54) (**Figure 4A**). These mutations have not been found in tumors yet, and mice studies showed that the E3 ligase activity is dispensable for the suppression of tumorigenesis. In contrast to BRCA1-C61G mice that failed to suppress tumor development, BRCA1-I26A mice were able to suppress tumor formation (64). This conclusion is supported by previous and recent work in mouse embryonic stem cells and human cells, where the BRCA1-I26A mutation presented similar HR and genotoxic stress phenotypes compared to BRCA1-WT (46, 49, 58). Together, these studies argued that heterodimer formation, but not BRCA1/BARD1-mediated ubiquitination, is essential for tumor suppression.

Nevertheless, mutations in BARD1 (C53W/C71Y/C83R) that impair H2A ubiquitination have been identified in families afflicted with breast cancer (65). These mutations alter three zinc-binding residues in the BARD1 RING domain but allow the formation of the BRCA1/BARD1 heterodimer. Previous work suggested that BRCA1/BARD1-mediated H2A ubiquitination was involved in tumor suppression by the maintenance of the global heterochromatin integrity (66). By chromatin immunoprecipitation (ChIP), BRCA1 was observed to bind satellite DNA regions, and its deficiency was accompanied by de-repression of normally silenced genes and loss of H2A monoubiquitination in a murine model. Satellite DNA regions are normally transcriptionally repressed and α-satellite RNAs overexpression has been related to genomic instability and breast cancer development in mice (67). BRCA1/BARD1 E3 activity on H2A was shown by rescuing BRCA1 deficient cells with ectopic expression of H2A-Ub, which not only restored silencing but also proliferative and HR defects. Additionally, BRCA1 deficient cells were also reconstituted with either a BRCA1-WT or with a BRCA1-I26A mutant version. While BRCA1-WT reconstituted cells were able to rescue the ubiquitinated H2A accumulation, the BRCA1-I26A mutant failed to significantly enrich ubiquitinated H2A at satellite repeats, suggesting that H2A ubiquitination is specific for BRCA1/BARD1 and has an essential role in heterochromatin integrity and tumor suppression.

The main controversy comes from the non-tumorigenic role of BRCA1/BARD1 E3 activity and the tumorigenic role of BRCA1/BARD1-mediated H2A ubiquitination (**Figure 4A**). It would be possible that other E3 ligases could overcome histone H2A ubiquitination after BRCA1/BARD1 depletion or malfunctioning. Although there has been much speculation about the BRCA1/BARD1 E3 activity role in tumor suppression, to date only the BRCA1-I26A mutant has been examined in animals, which is the most appropriate setting to experimentally assess tumor suppression. Surprisingly, homozygous BRCA1-I26A mice displayed a very mild phenotype, unlike homozygous null or tumor-associated alleles mice, most of which suffer embryonic lethality. Likewise, a conditional mouse model of human triple-negative breast cancer carrying BRCA1-I26A retained the ability to suppress mammary tumor formation, in contrast to BRCA1-S1598F homozygous mice (63, 64). Since 2011, this result has not been challenged experimentally despite the data in human and mice cells proposing an antitumoral role of BRCA1/BARD1 E3 activity. Therefore, more studies employing mice models or organoids with well-defined BRCA1/BARD1 mutants are required to fully understand the potential role as tumor suppressor.

3.4 BRCA1-BARD1 E3 ACTIVITY IN REPLICATION

BRCA1/BARD1 has crucial roles in the repair and restart of stalled and damaged DNA replication forks and in their protection against nucleolytic attack. The first time BRCA1/BARD1 was associated with replication was in 1997 when breast cancer cells were subjected to hydroxyurea (HU) treatment, which inhibits ribonucleotide reductase, leading to depletion of nucleotides causing replication stress and fork stalling. In this study, BRCA1 and BARD1 colocalized with RAD51 and PCNA upon HU treatment, suggesting that BRCA1/BARD1 are present in stalled replication forks (68). Later, mass spectrometry studies employing isolation of proteins on nascent DNA (iPOND), placed BRCA1/BARD1 at ongoing and stalled replication forks upon HU treatment (69, 70). DNA fiber experiments showed a role of BRCA1/BARD1 preventing replication fork degradation also after HU treatment (71). Whether the E3 ligase activity of BRCA1/BARD1 has a role during fork protection is still uncertain. BRCA1 contains a S114-P115 regulatory region, and BARD1 contains a RAD51 binding region. Disruption of these regions leads to defective fork protection and have been reported in cancer patients. Cells supplemented with a BARD1 F133A/D135A/A136E mutant, which abrogates the RAD51 binding site, exhibited defective fork protection phenotypes upon HU treatment. However, cells complemented with a BARD1-I26A mutant, which is deficient in E3 ligase activity, did not present replication fork protection problems, similarly to cells complemented with BARD1-WT (72).

However, not only a decrease of the nucleotide pool affects DNA replication. DNA polymerase can encounter difficulties during DNA replication such as replication barriers (**Figure 4B**). A well-known barrier is a RNA-DNA hybrid with an appended displaced ssDNA known as R-loop. BRCA1/BARD1 interacts with senataxin (SETX) and participates in the resolution of R-loops structures (73). The role of BRCA1 in R-loop resolution and prevention has been reviewed and more recent studies arise in this field (74, 75), but there is yet no relationship between the BRCA1/BARD1 E3 activity with R-loops. Another replication barrier is the G-quadruplex DNA structure that is formed at G-rich sites. G-quadruplex structures present strong impediments to replication fork progression. BRCA1 deficient cells have shown hypersensitivity to compounds stabilizing these DNA structures, suggesting that BRCA1/BARD1 promotes restart of stalled replication forks at G-quadruplex DNA structures (76). It could also be possible that the stabilization of the G-quadruplex leads to an accumulation of ssDNA gaps during replication. The excess of ssDNA might be toxic for BRCA1 deficient cells independently of its role in HR or replication fork restart, similarly to PARPi treatments (77) (**Figure 4B**). In line, mass-spectrometry data revealed a possible role of BRCA1 in replication fork protection by neutralizing RNA satellite overexpression, which leads to replication problems and chromosome breaks (67). Unfortunately, the E3 activity of BRCA1/BARD1 has not been interrogated during these replication difficulties, leaving unclear the relationship between BRCA1/BARD1 E3 activity and its role in compromised DNA replication. Addressing this topic would be key for understanding BRCA1/BARD1 E3 activity in the future.

Nevertheless, as mentioned above, BRCA1 has been placed on ongoing replication forks, meaning that it could have a role during this process. It is possible that BRCA1/BARD1 E3 activity could be involved in replication fork homeostasis by ubiquitination of key players. Recently, PCNA was found to be ubiquitinated during normal S-phase progression. Human cells carrying PCNA

K164R (PCNA-KR) mutation, being K164 the major ubiquitination site on PCNA, showed an increased ssDNA gap formation and degradation of stalled replication forks. Additionally, BRCA1 knock down in PCNAKR cell lines resulted in hypersensitivity to PARP inhibitors (78). These results suggest that lack of ubiquitination on replication forks, which includes BRCA1/BARD1 deficient and PCNA-KR mutant cells, is associated with ssDNA gap formation and replication fork protection (**Figure 4B**).

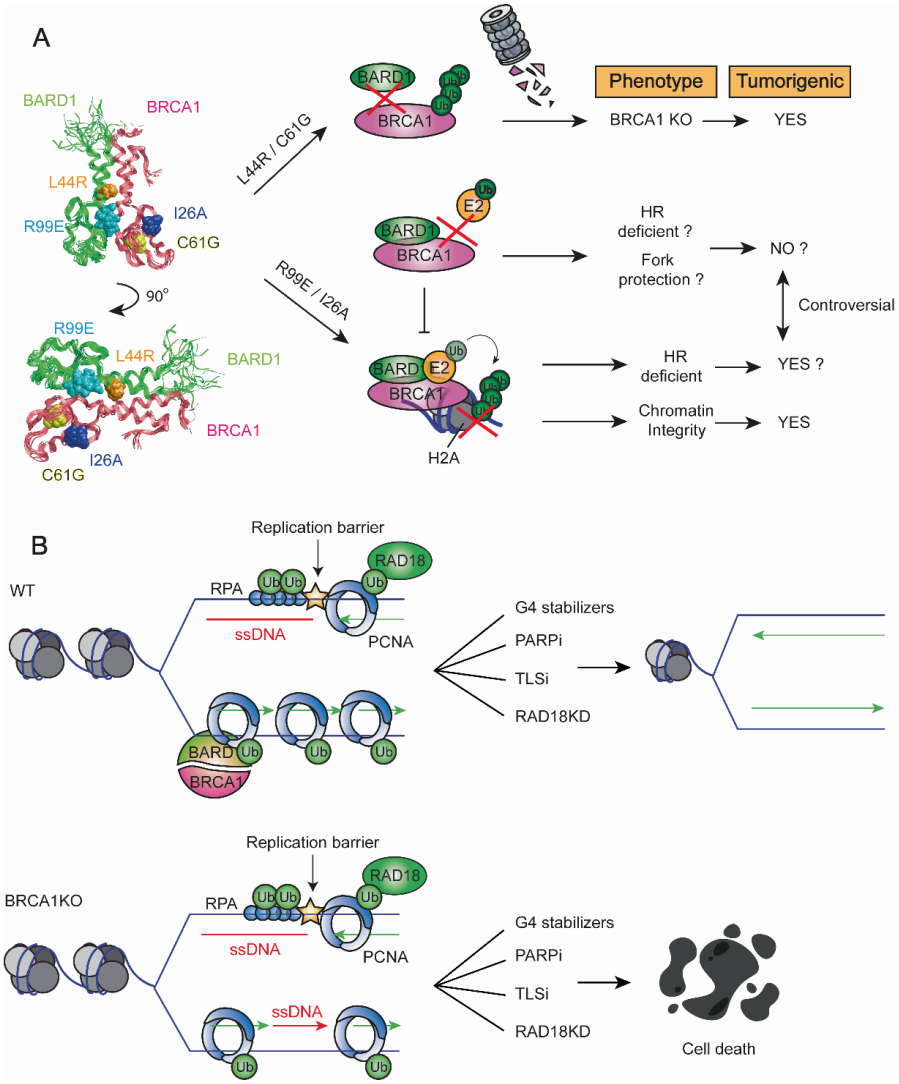


Figure 4. BRCA1/BARD1 in homologous recombination and replication. A. Solution structure of the BRCA1/BARD1 RING domain heterodimer was obtained from PDB (1JM7). Both L44R and C61G mutations in BARD1 and BRCA1 respectively disrupt the formation of the heterodimer resulting in BRCA1 instability and its subsequent proteasomal degradation. These mutations lead to a BRCA1KO phenotype being tumorigenic. BARD1 R99E and BRCA1 I26A mutations abrogate the E3 ligase activity but allow the formation of the heterodimer. These mutations have not been found in cancer patients and mice carrying these

mutations suppressed the formation of tumors. Therefore, they are no tumorigenic. However, R99E and I26A mutations in BARD1 and BRCA1 respectively, influence histone H2A ubiquitination, which impairment has been found in cancer patients and it is therefore tumorigenic. BRCA1/BARD1-mediated histone H2A ubiquitination has also been related to functional HR and chromatin integrity. **B.** Replication forks encounter a replication barrier. WT cells are able to survive G4 stabilizers, PARPi, TLSi and the knock down of the RAD18 E3 ligase. However, BRCA1 deficient cells accumulate ssDNA gaps under normal grow conditions and are unable to overcome the excessive replication stress upon inhibitors or disruption of PCNA ubiquitination, leading BRCA1 deficient cancer cells to death.

According to this data, a recent study revealed that BRCA1 deficient cells accumulate ssDNA gaps and this can be exploited therapeutically using PARP inhibitors (77). Authors suggest that PARPi sensitivity in BRCA1 deficient cells derives from ssDNA replication gaps and not due to BRCA1 function during HR or fork protection. How to take advantage of these ssDNA gaps to overcome BRCA1 deficient cancer cells is a current topic in the breast cancer field. Nayak and colleges found that TLS inhibition using a REV1 inhibitor caused cell death in cancer cells. This finding suggests that cancer cells rely on gap suppression during DNA replication (79). In agreement, Taglialatela et al., found that BRCA1 deficient cancer cells need TLS polymerases to maintain viability. TLS depends on PCNA ubiquitination, which is mediated by RAD18 E3 ligase. Therefore, it is not surprising that RAD18 knock down in BRCA1 deficient cells will lead to cell death in a similar way as using TLS inhibitors (80) (**Figure 4B)**.

However, the biological mechanism underlaying why BRCA1 deficient cells accumulate ssDNA gaps is still unclear. Again, the E3 activity of the heterodimer has not been studied in this context yet, and it would be key for future research. Afterall, there is still no evidence of BRCA1/BARD1 E3 activity involved in DNA replication. Therefore, experiments using BRCA1 and BARD1 mutants disrupting the E3 ligase activity will be one of the future directions in the field.

4. TARGETING THE UBIQUITINATION MACHINERY FOR CANCER TREATMENT

Virtually every cellular process is regulated by ubiquitination and its deregulation is associated to pathological disorders and cancer (81). Ubiquitination is carried out by an enzymatic cascade, in which Ub is activated by an activating enzyme (E1), which hydrolyzes ATP to form a thioester bond with ubiquitin. Then, ubiquitin is transferred via thioester-like complex to the ubiquitin-conjugating enzyme (E2), and the ubiquitin ligase enzyme (E3) mediates the conjugation of the ubiquitin moiety to either a lysine residue or the extreme amino terminus of the targeted protein (82-85). There are two known E1 enzymes, UBA1 and UBA6, nonetheless just UBA1 seems to charge the 99% of cellular ubiquitin. Therefore, targeting UBA1 will inhibit the majority of the ubiquitination events. There is already one commercial UBA1 inhibitor, TAK-243. This inhibitor forms a TAK-243-ubiquitin adduct that drastically decrease the formation of cellular ubiquitin conjugates, affecting cell cycle progression and DNA repair, leading to cancer cell death (86). This inhibitor, used at a tolerated dose, has shown promising results in mice bearing human xenograft tumors and in preclinical evaluation of acute myeloid leukemia (AML) (86, 87).

The complexity of the ubiquitination cascade scales up when we look at the numbers of the E2s and E3s enzymes, with 38 E2s and more than 600 E3s (88, 89). Potent inhibitors and new strategies have been developed targeting E2s enzymes. CC0651 is an allosteric inhibitor of CDC34,

an E2 involved in cell proliferation. CDC34 inhibition by CC0651 exhibited the accumulation of its ubiquitination targets and inhibition of cell proliferation (90). Recently, a new strategy concerning protein-based reagents, called ubiquitin variants (UbVs), can be used to develop protein-based inhibitor against E2s. The designed UbVs bind UBE2K and block both, ubiquitin charging and E3 catalyzed ubiquitin transfer to the target protein (91).

The proteome profile is constantly changing and the changes during cancer disease can be tracked. Consequently, controlling the protein turnover by inhibiting upregulated E3 ligases is a key factor to disturb cancer cells. However, one of the major limitations to properly control the protein levels of an E3 ligase substrate, is the identification of the targets for the E3 ligase under consideration. Fortunately, some targets for specific E3s have been elucidated, and the ability of these E3s to form K48 linkages on the target protein for its proteasomal degradation has identified these E3s as potential targets for cancer treatments (92-94). A major example of E3 inhibitors is *nutlin*, the first small-molecule inhibitor of the p53-MDM2 E3 interaction (95). P53 is inactivated in 50% of cancers, and in some cases this is due to MDM2 overexpression. Advanced MDM2 inhibitors are in clinical trials for solid tumours, haematological neoplasms, liposarcomas, soft tissue sarcoma and AML (96).

Recently, two new small molecules with anticancer potential have been reported as Culling-RING E3 ligases (CRLs) inhibitors (97). These inhibitors seem to impact the Culling-E2 interaction, inhibiting the ubiquitination process. Moreover, the compounds exhibited in vivo antitumoral activity in AML MV4-11 xenograft mouse models. These inhibitors could have great potential in tumors with low expression of CRLs. Another way to make use of the ubiquitin proteasome system (UPS) as cancer vulnerability, is targeting deubiquitinating enzymes (DUBs) (98). This is the case of TAK-659, a DUB inhibitor in clinical trials, which targets USP10 and induces degradation of the spleen tyrosine kinase (SYK). SYK seems to be critical for AML transformation and maintenance in AML patients, and its degradation results in death of AML cancer cells (99). Nevertheless, not only E3s involved in proteasomal degradation are associated with cancer. As it was illustrated above, cells have developed signaling mechanisms to end the DDR. However, we can try to control the response with the use of inhibitors that target upstream players, as is the case for the early recruited BMI1-RING1B ligase which ubiquitinates H2A K119. PRT4165 inhibits BMI1-RING1B-mediated H2A ubiquitination in vitro and in vivo (100). The use of this inhibitor revealed not only its potential use as cancer treatment but also supports the early role of this E3 ligase in the DDR.

The association between E3 deregulation and cancer has been previously reviewed and several E3s with non-degradative or still yet to know substrates appear as potential targets for cancer therapy (81, 101, 102). The BRCA1/BARD1 E3 role in DNA damage and cell cycle checkpoints regulation has been linked to cancer development (103). BRCA1/BARD1 mediates monoubiquitination, degradative and non-degradative polyubiquitination of its targets, meaning signaling purposes in multiple cellular processes (**Table 2**). However, very few targets have been validated as direct targets for BRCA1/BARD1. There are several considerations to keep in mind when validating the ubiquitination targets: the use of catalytically dead mutants that still can use a different E2 ligase (104). Depletion and overexpression of an E3 ligase can completely change the ubiquitinated proteome of a cell and lead to artificial protein ubiquitination. Finally, the overexpression of a substrate could lead to its ubiquitination for the mere reason of being in

excess within the cell. Up to date, there is only one well-described ubiquitination target for BRCA1/BARD1 and it is the histone H2A.

Besides the canonical H2A as well-defined BRCA1/BARD1 ubiquitination target, the histone variant mH2A has emerged as solid ubiquitination target (105). In addition to its role in senescence, the ubiquitination of mH2A could be related to the repair pathway choice and the Alt-EJ repair pathway, due to its recently described roles in these cellular processes (106). It would be beneficial to develop BRCA1/BARD1 inhibitors as they could be used not only as chemotherapy strategy but also to study the BRCA1/BARD1 function during the DDR. However, in order to develop successful inhibitors, it will be crucial to identify and carefully validate its targets. In this way, some inhibitors could work blocking the interaction between BRCA1/BARD1 and a specific substrate.

Table 2. BRCA1/BARD1 Ubiquitination targets.

Substrate	Cellular function	Linkage	Reference
Histone H2A	DNA damage repair Heterochromatin integrity	Mono	(39, 50, 52)
Macro H2A1 (H2AFY)	Cellular senescence Alt-EJ repair	Mono	(105) (106)
Cyclin B	Cell cycle regulation	Poly (K6)	(112)
CtIP	DNA damage repair	Poly	(57)
Merlin (NF2)	Hippo growth signaling	Poly (K63)	(113)
Claspin	DNA damage repair	-	(114)
RNApol II	DNA damage repair	Poly	(115)

5. CONCLUDING REMARKS AND FUTURE DIRECTIONS

The DDR is tightly regulated in time and space, and a large set of new players has arisen in a short period of time. To completely understand the whole DDR and be able to take advantage of it when there is disease as cancer, years of research will be needed. In the meantime, we dissected the DDR from an ubiquitination point of view, where E3 ligases are key regulators. Controlling the ubiquitination signaling of the DDR could influence the whole response. In the case of E3 ligases, finding their ubiquitination substrates discloses their function. Lots of effort have been done to elucidate BRCA1/BARD1 ubiquitination substrates and completely understand its E3 ligase function (**Table 2**). In the last decade, the proteomic field has evolved notoriously and mass-spectrometry (MS) approaches has been used to find BRCA1/BARD1 targets (105, 107-111). However, the validation of targets from proteomic screenings is very challenging.

Different strategies can be used to find potential substrates for BRCA1/BARD1. To investigate the role of BRCA1/BARD1 in biological functions, the employment of BRCA1 mutants combined

with MS approaches could lead to the identification of its interactors and ubiquitination substrates. The development of organoids harboring BRCA1/BARD1 null and E3 dead mutants could contribute to unveil the role of this E3 ligase. Finally, the development of inhibitors interrupting the E3 activity might be of interest to address the E3 role of BRCA1/BARD1 and finding its targets by MS strategies.

Author Contributions: Writing—original draft preparation, D.S.-L.; writing—review and editing, D.S.-L. and R.G.-P.; supervision and funding acquisition R.G.-P. All authors have read and agreed to the published version of the manuscript**.**

Funding: The authors are supported by the Dutch Cancer Foundation (KFW-YIG 11367/2017-2).

REFERENCES

1. J. R. Danielsen *et al.*, DNA damage-inducible SUMOylation of HERC2 promotes RNF8 binding via a novel SUMO-binding Zinc finger. *J Cell Biol* **197**, 179-187 (2012).
2. J. Garcia-Cano *et al.*, Regulation of the MDM2-p53 pathway by the ubiquitin ligase HERC2. *Mol Oncol* **14**, 69-86 (2020).
3. A. Inagaki *et al.*, Human RAD18 interacts with ubiquitylated chromatin components and facilitates RAD9 recruitment to DNA double strand breaks. *PLoS One* **6**, e23155 (2011).
4. K. Watanabe *et al.*, RAD18 promotes DNA double-strand break repair during G1 phase through chromatin retention of 53BP1. *Nucleic Acids Res* **37**, 2176-2193 (2009).
5. D. Cui *et al.*, FBXW7 Confers Radiation Survival by Targeting p53 for Degradation. *Cell Rep* **30**, 497-509 e494 (2020).
6. Y. Haupt, R. Maya, A. Kazaz, M. Oren, Mdm2 promotes the rapid degradation of p53. *Nature* **387**, 296-299 (1997).
7. C. M. Eischen, Role of Mdm2 and Mdmx in DNA repair. *J Mol Cell Biol* **9**, 69-73 (2017).
8. L. S. Symington, J. Gautier, Double-strand break end resection and repair pathway choice. *Annu Rev Genet* **45**, 247-271 (2011).
9. J. Schimmel, N. Munoz-Subirana, H. Kool, R. van Schendel, M. Tijsterman, Small tandem DNA duplications result from CST-guided Pol alpha-primase action at DNA break termini. *Nat Commun* **12**, 4843 (2021).
10. A. Aguilera, B. Gomez-Gonzalez, Genome instability: a mechanistic view of its causes and consequences. *Nat Rev Genet* **9**, 204-217 (2008).
11. T. Lindahl, D. E. Barnes, Repair of endogenous DNA damage. *Cold Spring Harb Symp Quant Biol* **65**, 127-133 (2000).
12. N. Weber-Lassalle *et al.*, Germline loss-of-function variants in the BARD1 gene are associated with early-onset familial breast cancer but not ovarian cancer. *Breast Cancer Res* **21**, 55 (2019).
13. A. Ciccia, S. J. Elledge, The DNA damage response: making it safe to play with knives. *Mol Cell* **40**, 179-204 (2010).
14. D. Wang, L. Ma, B. Wang, J. Liu, W. Wei, E3 ubiquitin ligases in cancer and implications for therapies. *Cancer Metastasis Rev* **36**, 683-702 (2017).
15. N. P. Dantuma, H. van Attikum, Spatiotemporal regulation of posttranslational modifications in the DNA damage response. *EMBO J* **35**, 6-23 (2016).
16. S. P. Jackson, Sensing and repairing DNA double-strand breaks. *Carcinogenesis* **23**, 687-696 (2002).
17. C. Capoulade *et al.*, Overexpression of MDM2, due to enhanced translation, results in inactivation of wild-type p53 in Burkitt's lymphoma cells. *Oncogene* **16**, 1603-1610 (1998).
18. A. Zafar *et al.*, Targeting the p53-MDM2 pathway for neuroblastoma therapy: Rays of hope. *Cancer Lett* **496**, 16-29 (2021).

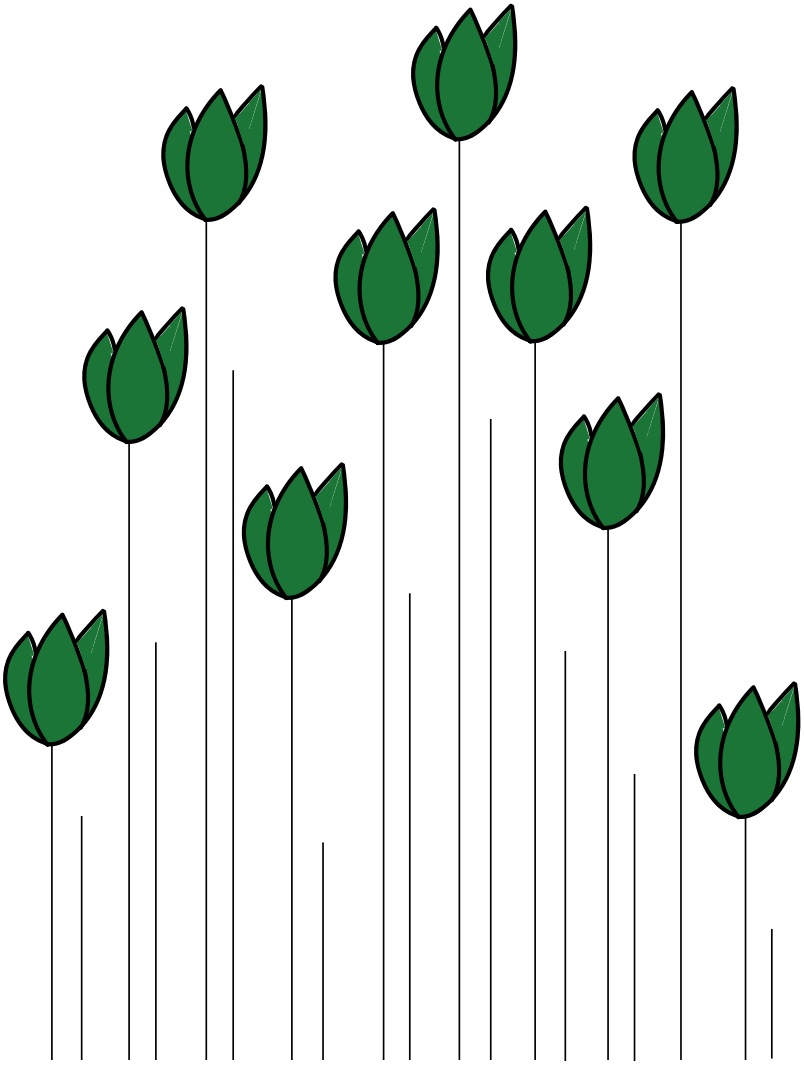
19. Q. Cheng *et al.*, Regulation of MDM2 E3 ligase activity by phosphorylation after DNA damage. *Mol Cell Biol* **31**, 4951-4963 (2011).
20. S. Bekker-Jensen *et al.*, Spatial organization of the mammalian genome surveillance machinery in response to DNA strand breaks. *J Cell Biol* **173**, 195-206 (2006).
21. N. Mailand *et al.*, RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins. *Cell* **131**, 887-900 (2007).
22. A. Paul, B. Wang, RNF8- and Ube2S-Dependent Ubiquitin Lysine 11-Linkage Modification in Response to DNA Damage. *Mol Cell* **66**, 458-472 e455 (2017).
23. R. Cao, Y. Tsukada, Y. Zhang, Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol Cell* **20**, 845-854 (2005).
24. V. Ginjala *et al.*, BMI1 is recruited to DNA breaks and contributes to DNA damage-induced H2A ubiquitination and repair. *Mol Cell Biol* **31**, 1972-1982 (2011).
25. N. K. Kolas *et al.*, Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. *Science* **318**, 1637-1640 (2007).
26. T. Gudjonsson *et al.*, TRIP12 and UBR5 suppress spreading of chromatin ubiquitylation at damaged chromosomes. *Cell* **150**, 697-709 (2012).
27. Y. Galanty *et al.*, Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks. *Nature* **462**, 935-939 (2009).
28. K. Luo, H. Zhang, L. Wang, J. Yuan, Z. Lou, Sumoylation of MDC1 is important for proper DNA damage response. *EMBO J* **31**, 3008-3019 (2012).
29. R. Kumar, R. Gonzalez-Prieto, Z. Xiao, M. Verlaan-de Vries, A. C. O. Vertegaal, The STUbL RNF4 regulates protein group SUMOylation by targeting the SUMO conjugation machinery. *Nat Commun* **8**, 1809 (2017).
30. D. Salas-Lloret, G. Agabitini, R. Gonzalez-Prieto, TULIP2: An Improved Method for the Identification of Ubiquitin E3-Specific Targets. *Front Chem* **7**, 802 (2019).
31. N. P. Dantuma, T. Hoppe, Growing sphere of influence: Cdc48/p97 orchestrates ubiquitin-dependent extraction from chromatin. *Trends Cell Biol* **22**, 483-491 (2012).
32. L. Dai *et al.*, Pellino1 specifically binds to phospho-Thr18 of p53 and is recruited to sites of DNA damage. *Biochem Biophys Res Commun* **513**, 714-720 (2019).
33. Y. Katsuki *et al.*, RNF168 E3 ligase participates in ubiquitin signaling and recruitment of SLX4 during DNA crosslink repair. *Cell Rep* **37**, 109879 (2021).
34. P. S. Brzovic, P. Rajagopal, D. W. Hoyt, M. C. King, R. E. Klevit, Structure of a BRCA1-BARD1 heterodimeric RING-RING complex. *Nat Struct Biol* **8**, 833-837 (2001).
35. C. M. Christou, K. Kyriacou, BRCA1 and Its Network of Interacting Partners. *Biology (Basel)* **2**, 40-63 (2013).
36. Y. Xia, G. M. Pao, H. W. Chen, I. M. Verma, T. Hunter, Enhancement of BRCA1 E3 ubiquitin ligase activity through direct interaction with the BARD1 protein. *J Biol Chem* **278**, 5255-5263 (2003).
37. R. Hashizume *et al.*, The RING heterodimer BRCA1-BARD1 is a ubiquitin ligase inactivated by a breast cancer-derived mutation. *J Biol Chem* **276**, 14537-14540 (2001).
38. P. S. Brzovic *et al.*, Binding and recognition in the assembly of an active BRCA1/BARD1 ubiquitin-ligase complex. *Proc Natl Acad Sci U S A* **100**, 5646-5651 (2003).
39. S. R. Witus *et al.*, BRCA1/BARD1 site-specific ubiquitylation of nucleosomal H2A is directed by BARD1. *Nat Struct Mol Biol* **28**, 268-277 (2021).
40. S. M. Noordermeer *et al.*, The shieldin complex mediates 53BP1-dependent DNA repair. *Nature* **560**, 117-121 (2018).
41. B. Sobhian *et al.*, RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science* **316**, 1198-1202 (2007).

42. M. S. Huen *et al.*, RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly. *Cell* **131**, 901-914 (2007).
43. F. Mattiroli *et al.*, RNF168 ubiquitinates K13-15 on H2A/H2AX to drive DNA damage signaling. *Cell* **150**, 1182-1195 (2012).
44. A. Sherker *et al.*, Two redundant ubiquitin-dependent pathways of BRCA1 localization to DNA damage sites. *EMBO Rep* **22**, e53679 (2021).
45. R. K. McGinty, R. C. Henrici, S. Tan, Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. *Nature* **514**, 591-596 (2014).
46. J. J. Krais *et al.*, RNF168-mediated localization of BARD1 recruits the BRCA1-PALB2 complex to DNA damage. *Nat Commun* **12**, 5016 (2021).
47. J. R. Becker *et al.*, BARD1 reads H2A lysine 15 ubiquitination to direct homologous recombination. *Nature* **596**, 433-437 (2021).
48. K. Jacquet *et al.*, The TIP60 Complex Regulates Bivalent Chromatin Recognition by 53BP1 through Direct H4K20me Binding and H2AK15 Acetylation. *Mol Cell* **62**, 409-421 (2016).
49. K. Nakamura *et al.*, H4K20me0 recognition by BRCA1-BARD1 directs homologous recombination to sister chromatids. *Nat Cell Biol* **21**, 311-318 (2019).
50. Q. Hu *et al.*, Mechanisms of BRCA1-BARD1 nucleosome recognition and ubiquitylation. *Nature* **596**, 438-443 (2021).
51. M. Uckelmann *et al.*, USP48 restrains resection by site-specific cleavage of the BRCA1 ubiquitin mark from H2A. *Nat Commun* **9**, 229 (2018).
52. R. Kalb, D. L. Mallery, C. Larkin, J. T. Huang, K. Hiom, BRCA1 is a histone-H2A-specific ubiquitin ligase. *Cell Rep* **8**, 999-1005 (2014).
53. M. D. Wilson *et al.*, The structural basis of modified nucleosome recognition by 53BP1. *Nature* **536**, 100-103 (2016).
54. R. M. Densham *et al.*, Human BRCA1-BARD1 ubiquitin ligase activity counteracts chromatin barriers to DNA resection. *Nat Struct Mol Biol* **23**, 647-655 (2016).
55. R. Scully *et al.*, Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell* **88**, 265-275 (1997).
56. L. Chen, C. J. Nievera, A. Y. Lee, X. Wu, Cell cycle-dependent complex formation of BRCA1.CtIP.MRN is important for DNA double-strand break repair. *J Biol Chem* **283**, 7713-7720 (2008).
57. X. Yu, S. Fu, M. Lai, R. Baer, J. Chen, BRCA1 ubiquitinates its phosphorylation-dependent binding partner CtIP. *Genes Dev* **20**, 1721-1726 (2006).
58. L. J. Reid *et al.*, E3 ligase activity of BRCA1 is not essential for mammalian cell viability or homology-directed repair of double-strand DNA breaks. *Proc Natl Acad Sci U S A* **105**, 20876-20881 (2008).
59. M. Tarsounas, P. Sung, The antitumorigenic roles of BRCA1-BARD1 in DNA repair and replication. *Nat Rev Mol Cell Biol* 10.1038/s41580-020-0218-z (2020).
60. P. S. Brzovic, J. Meza, M. C. King, R. E. Klevit, The cancer-predisposing mutation C61G disrupts homodimer formation in the NH2-terminal BRCA1 RING finger domain. *J Biol Chem* **273**, 7795-7799 (1998).
61. S. L. Clark, A. M. Rodriguez, R. R. Snyder, G. D. Hankins, D. Boehning, Structure-Function Of The Tumor Suppressor BRCA1. *Comput Struct Biotechnol J* **1** (2012).
62. A. Jhuraney *et al.*, BRCA1 Circos: a visualisation resource for functional analysis of missense variants. *J Med Genet* **52**, 224-230 (2015).
63. R. Drost *et al.*, BRCA1 RING function is essential for tumor suppression but dispensable for therapy resistance. *Cancer Cell* **20**, 797-809 (2011).
64. R. Shakya *et al.*, BRCA1 tumor suppression depends on BRCT phosphoprotein binding, but not its E3 ligase activity. *Science* **334**, 525-528 (2011).

65. M. D. Stewart *et al.*, BARD1 is necessary for ubiquitylation of nucleosomal histone H2A and for transcriptional regulation of estrogen metabolism genes. *Proc Natl Acad Sci U S A* **115**, 1316-1321 (2018).
66. Q. Zhu *et al.*, BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature* **477**, 179-184 (2011).
67. Q. Zhu *et al.*, Heterochromatin-Encoded Satellite RNAs Induce Breast Cancer. *Mol Cell* **70**, 842-853 e847 (2018).
68. R. Scully *et al.*, Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage. *Cell* **90**, 425-435 (1997).
69. B. M. Sirbu *et al.*, Identification of proteins at active, stalled, and collapsed replication forks using isolation of proteins on nascent DNA (iPOND) coupled with mass spectrometry. *J Biol Chem* **288**, 31458-31467 (2013).
70. H. Dungrawala *et al.*, The Replication Checkpoint Prevents Two Types of Fork Collapse without Regulating Replisome Stability. *Mol Cell* **59**, 998-1010 (2015).
71. K. Schlacher, H. Wu, M. Jasin, A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2. *Cancer Cell* **22**, 106-116 (2012).
72. M. Daza-Martin *et al.*, Isomerization of BRCA1-BARD1 promotes replication fork protection. *Nature* **571**, 521-527 (2019).
73. S. J. Hill *et al.*, Systematic screening reveals a role for BRCA1 in the response to transcription-associated DNA damage. *Genes Dev* **28**, 1957-1975 (2014).
74. M. San Martin Alonso, S. M. Noordermeer, Untangling the crosstalk between BRCA1 and R-loops during DNA repair. *Nucleic Acids Res* **49**, 4848-4863 (2021).
75. C. Racca *et al.*, BRCA1 prevents R-loop-associated centromeric instability. *Cell Death Dis* **12**, 896 (2021).
76. J. Zimmer *et al.*, Targeting BRCA1 and BRCA2 Deficiencies with G-Quadruplex-Interacting Compounds. *Mol Cell* **61**, 449-460 (2016).
77. K. Cong *et al.*, Replication gaps are a key determinant of PARP inhibitor synthetic lethality with BRCA deficiency. *Mol Cell* **81**, 3227 (2021).
78. T. Thakar *et al.*, Ubiquitinated-PCNA protects replication forks from DNA2-mediated degradation by regulating Okazaki fragment maturation and chromatin assembly. *Nat Commun* **11**, 2147 (2020).
79. S. Nayak *et al.*, Inhibition of the translesion synthesis polymerase REV1 exploits replication gaps as a cancer vulnerability. *Sci Adv* **6**, eaaz7808 (2020).
80. A. Taglialatela *et al.*, REV1-Polzeta maintains the viability of homologous recombination-deficient cancer cells through mutagenic repair of PRIMPOL-dependent ssDNA gaps. *Mol Cell* **81**, 4008-4025 e4007 (2021).
81. D. Senft, J. Qi, Z. A. Ronai, Ubiquitin ligases in oncogenic transformation and cancer therapy. *Nat Rev Cancer* **18**, 69-88 (2018).
82. D. H. Schlesinger, G. Goldstein, H. D. Niall, The complete amino acid sequence of ubiquitin, an adenylate cyclase stimulating polypeptide probably universal in living cells. *Biochemistry* **14**, 2214-2218 (1975).
83. M. Scheffner, U. Nuber, J. M. Huibregtse, Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade. *Nature* **373**, 81-83 (1995).
84. A. Hershko, A. Ciechanover, The ubiquitin system. *Annu Rev Biochem* **67**, 425-479 (1998).
85. T. Kirisako *et al.*, A ubiquitin ligase complex assembles linear polyubiquitin chains. *EMBO J* **25**, 4877-4887 (2006).
86. M. L. Hyer *et al.*, A small-molecule inhibitor of the ubiquitin activating enzyme for cancer treatment. *Nat Med* **24**, 186-193 (2018).

87. S. H. Barghout *et al.*, Preclinical evaluation of the selective small-molecule UBA1 inhibitor, TAK-243, in acute myeloid leukemia. *Leukemia* **33**, 37-51 (2019).
88. M. D. Stewart, T. Ritterhoff, R. E. Klevit, P. S. Brzovic, E2 enzymes: more than just middle men. *Cell Res* **26**, 423-440 (2016).
89. D. Salas-Lloret, R. Gonzalez-Prieto, Insights in Post-Translational Modifications: Ubiquitin and SUMO. *Int J Mol Sci* **23** (2022).
90. D. F. Ceccarelli *et al.*, An allosteric inhibitor of the human Cdc34 ubiquitin-conjugating enzyme. *Cell* **145**, 1075-1087 (2011).
91. A. J. Middleton, J. Teyra, J. Zhu, S. S. Sidhu, C. L. Day, Identification of Ubiquitin Variants That Inhibit the E2 Ubiquitin Conjugating Enzyme, Ube2k. *ACS Chem Biol* **16**, 1745-1756 (2021).
92. X. Huang, V. M. Dixit, Drugging the undruggables: exploring the ubiquitin system for drug development. *Cell Res* **26**, 484-498 (2016).
93. E. Bulatov, A. Zagidullin, A. Valiullina, R. Sayarova, A. Rizvanov, Small Molecule Modulators of RING-Type E3 Ligases: MDM and Cullin Families as Targets. *Front Pharmacol* **9**, 450 (2018).
94. V. Chau *et al.*, A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* **243**, 1576-1583 (1989).
95. L. T. Vassilev *et al.*, In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **303**, 844-848 (2004).
96. K. H. Khoo, C. S. Verma, D. P. Lane, Drugging the p53 pathway: understanding the route to clinical efficacy. *Nat Rev Drug Discov* **13**, 217-236 (2014).
97. K. Wu *et al.*, Inhibitors of cullin-RING E3 ubiquitin ligase 4 with antitumor potential. *Proc Natl Acad Sci U S A* **118** (2021).
98. A. Cremer, K. Stegmaier, Targeting DUBs to degrade oncogenic proteins. *Br J Cancer* **122**, 1121-1123 (2020).
99. J. Yang *et al.*, Inhibition of the deubiquitinase USP10 induces degradation of SYK. *Br J Cancer* **122**, 1175-1184 (2020).
100. I. H. Ismail, D. McDonald, H. Strickfaden, Z. Xu, M. J. Hendzel, A small molecule inhibitor of polycomb repressive complex 1 inhibits ubiquitin signaling at DNA double-strand breaks. *J Biol Chem* **288**, 26944-26954 (2013).
101. F. Bernassola, G. Chillemi, G. Melino, HECT-Type E3 Ubiquitin Ligases in Cancer. *Trends Biochem Sci* **44**, 1057-1075 (2019).
102. A. Sinha, P. V. Iyengar, P. Ten Dijke, E3 Ubiquitin Ligases: Key Regulators of TGFbeta Signaling in Cancer Progression. *Int J Mol Sci* **22** (2021).
103. K. I. Savage, D. P. Harkin, BRCA1, a 'complex' protein involved in the maintenance of genomic stability. *FEBS J* **282**, 630-646 (2015).
104. S. R. Witus, M. D. Stewart, R. E. Klevit, The BRCA1/BARD1 ubiquitin ligase and its substrates. *Biochem J* **478**, 3467-3483 (2021).
105. B. J. Kim *et al.*, The Histone Variant MacroH2A1 Is a BRCA1 Ubiquitin Ligase Substrate. *Cell Rep* **19**, 1758-1766 (2017).
106. R. Sebastian *et al.*, Epigenetic Regulation of DNA Repair Pathway Choice by MacroH2A1 Splice Variants Ensures Genome Stability. *Mol Cell* **79**, 836-845 e837 (2020).
107. M. Song, K. Hakala, S. T. Weintraub, Y. Shiio, Quantitative proteomic identification of the BRCA1 ubiquitination substrates. *J Proteome Res* **10**, 5191-5198 (2011).
108. A. Matsuzawa *et al.*, The BRCA1/BARD1-interacting protein OLA1 functions in centrosome regulation. *Mol Cell* **53**, 101-114 (2014).
109. J. B. Heidelberger, S. A. Wagner, P. Beli, Mass Spectrometry-Based Proteomics for Investigating DNA Damage-Associated Protein Ubiquitylation. *Front Genet* **7**, 109 (2016).

110. C. M. Eakin, M. J. Maccoss, G. L. Finney, R. E. Klevit, Estrogen receptor alpha is a putative substrate for the BRCA1 ubiquitin ligase. *Proc Natl Acad Sci U S A* **104**, 5794-5799 (2007).
111. J. K. Barrows, G. Fullbright, D. T. Long, BRCA1-BARD1 regulates transcription through BRD4 in Xenopus nucleoplasmic extract. *Nucleic Acids Res* **49**, 3263-3273 (2021).
112. S. Shabbeer *et al.*, BRCA1 targets G2/M cell cycle proteins for ubiquitination and proteasomal degradation. *Oncogene* **32**, 5005-5016 (2013).
113. S. Verma *et al.*, BRCA1/BARD1-dependent ubiquitination of NF2 regulates Hippo-YAP1 signaling. *Proc Natl Acad Sci U S A* **116**, 7363-7370 (2019).
114. K. Sato *et al.*, A DNA-damage selective role for BRCA1 E3 ligase in claspin ubiquitylation, CHK1 activation, and DNA repair. *Curr Biol* **22**, 1659-1666 (2012).
115. L. M. Starita *et al.*, BRCA1/BARD1 ubiquitinate phosphorylated RNA polymerase II. *J Biol Chem* **280**, 24498-24505 (2005).



3

TULIP2: An Improved Method for the Identification of Ubiquitin E3-Specific Targets

Daniel Salas-Lloret[1], Giulia Agabitini[1] and Román González-Prieto[1,2]

[1] Cell and Chemical Biology, Leiden University Medical Centre, Einthovenweg 20, 2333ZC Leiden, The Netherlands

[2] Genome Proteomics Group, Department of Genome Biology, Andalusian Centre for Regenerative Medicine and Molecular Biology (CABIMER), Américo Vespucio 24, 41092 Seville, Spain

This chapter has been published in Frontiers in Chemistry. 7, 802 (2019)

Abstract

Protein modification by Ubiquitin or Ubiquitin-like modifiers is mediated by an enzyme cascade composed of E1, E2, and E3 enzymes. E1s, or ubiquitin-activating enzymes, perform ubiquitin activation. Next, ubiquitin is transferred to ubiquitin-conjugating enzymes or E2s. Finally, ubiquitin ligases or E3s catalyze the transfer of ubiquitin to the acceptor proteins. E3 enzymes are responsible for determining the substrate specificity. Determining which E3 enzyme maps to which substrate is a major challenge that is greatly facilitated by the TULIP2 methodology. TULIP2 methodology is fast, precise, and cost-effective. Compared to the previous TULIP methodology protocol, TULIP2 methodology achieves a more than 50-fold improvement in the purification yield and two orders of magnitude improvement in the signal-to-background ratio after label free quantification by mass spectrometry analysis. The method includes the generation of TULIP2 cell lines, subsequent purification of TULIP2 conjugates, preparation, and analysis of samples by mass spectrometry.

Keywords: Mass-Spectrometry; Ubiquitin; RNF4; TULIP2

INTRODUCTION

The development of liquid chromatography tandem mass spectrometry (LC-MS/MS)-based proteomics technology has boomed in the past years, and, recently, a new strategy termed UbiSite, enabled the identification of around 63,000 unique sites for ubiquitination at endogenous levels of more than 10,000 proteins, including N-terminal ubiquitination (1). The identification of additional ubiquitination sites seems to be a matter of repeating the UbiSite strategy with samples from different sources.

Determining which E3 enzyme is responsible for modifying which substrate is challenging. Different strategies have been proposed for identification of specific E3 substrates. Many of these strategies are based on indirect evidence. For example, investigating differences in the ubiquitin proteome upon overexpression or depletion of a specific E3 (2-4). Proteins that are enriched or depleted, respectively, in their ubiquitination levels are considered putative ubiquitination substrates for the specific E3 under investigation. However, the complexity of full ubiquitin proteomes is high (1), and low abundant ubiquitination targets might be missed. Furthermore, results obtained from overexpression-based screens might be due to overexpression artifacts. In the case of the knock down-based screens, E3 ligases can be redundant on their targets, and some targets might be missed because their ubiquitination is still performed by another E3 enzyme. E3 enzyme cascades exist, and the absence of a specific ubiquitinated protein might be a result of an epistatic effect. Thus, every target has to be very carefully verified. As a consequence, indirect approaches are unable to find E3-specific substrates in a reliable manner.

A proposed direct approach is the employment of ubiquitin activated interaction traps, UBAITs (5), which work both for Really Interesting New Gene (RING) and Homologous to E6AP C-Terminus (HECT)-type E3 enzymes. The UBAIT approach is based on the utilization of E3 enzyme ubiquitin fusions. The rationale behind this technique is that, if a linear fusion between a specific E3 and ubiquitin is made, the E3 will be prone to use this ubiquitin to conjugate it to its ubiquitination target. Therefore, the E3 will remain covalently bound to its target after ubiquitination, which allows the later purification of the E3 together with its ubiquitination target. Enabling subsequent identification by LC-MS/MS analysis (**Figure 1**). The main pitfall of the UBait approach is that the purification of the conjugates is based on epitope-antibody interaction, which excludes the possibility of using denaturing buffers. This disadvantage makes it difficult to distinguish between ubiquitination targets and other potential strong interactors of the E3s. Additionally, it is based on overexpression of the constructs, so the occurrence of overexpression-derived artifacts is a possibility.

Nevertheless, using the UBAIT as a base, we optimized and designed a systematic methodology which we termed Targets of Ubiquitin Ligases Identified by Proteomics (TULIP) (6). TULIP methodology employs 10xHIS nickel-based purification, which allows the use of harsh denaturing buffers, solving the drawback of being unable to distinguish between ubiquitination targets and interactors of the E3. Moreover, TULIP methodology is lentiviral based, employing an all-in-one doxycycline-ON system followed by Gateway R cloning cassette and puromycin as selection marker for infected cells. TULIP methodology enables the generation of stable-inducible cell lines where the expression levels can be titrated to near-to-endogenous levels, minimizing the probability of obtaining results due to overexpression. The C-terminal GlyGly motif of ubiquitin is required for conjugation to a target. TULIP plasmids where ubiquitin lacks the C-terminal GlyGly

motif (TULIP-1GG) are also available as negative controls. Furthermore, catalytically dead mutants of the E3 enzymes are used as an additional negative control.

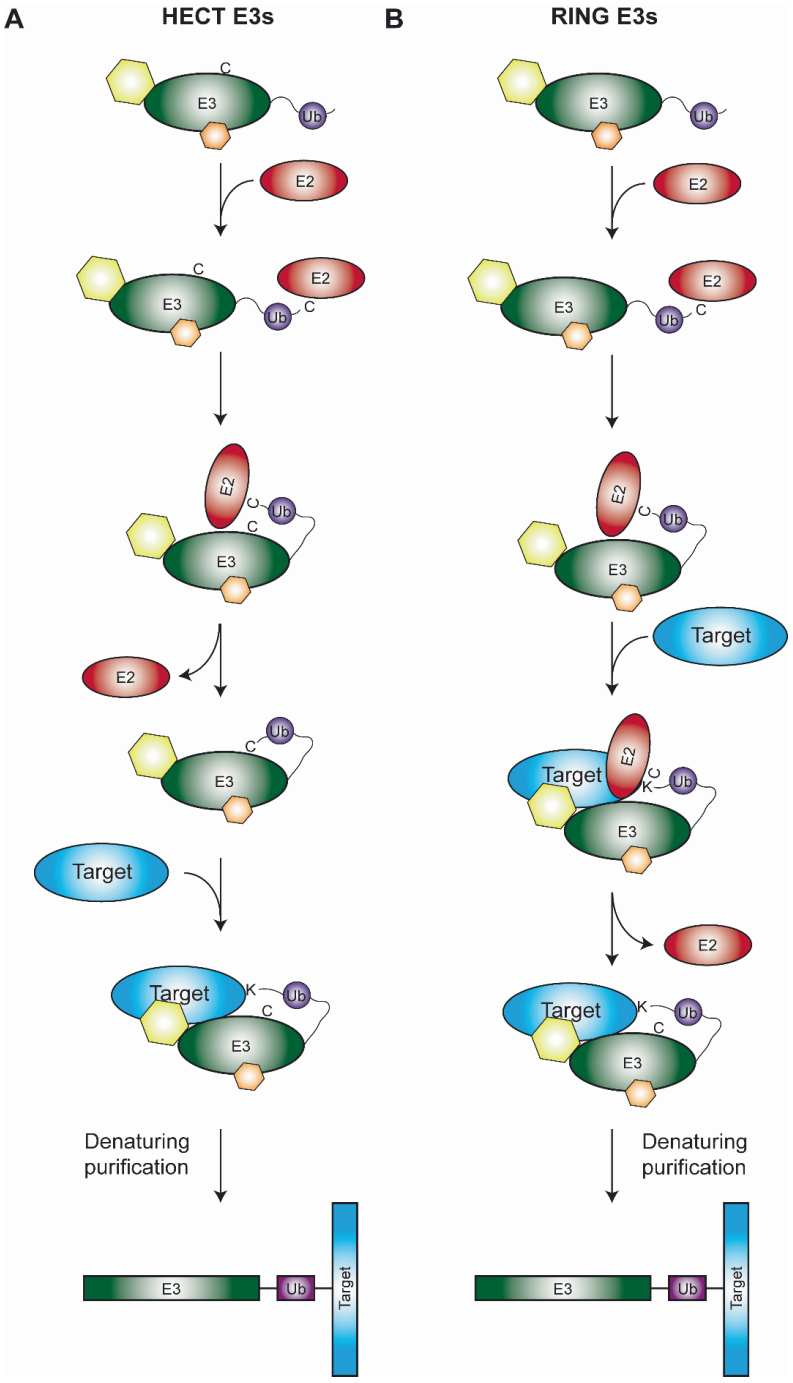


Figure 1. Rationale of the TULIP2 methodology. Rationale is depicted for both HECT (A) and RING (B) E3 enzymes. **A**. Activated ubiquitin linearly fused to a HECT E3 of interest will be conjugated to its respective E2 and transferred from the catalytic cysteine of the E2 to the catalytic cysteine of the HECT E3. Next ubiquitin will be transferred from the catalytic cysteine of the E3 to the acceptor lysine of the E3-target protein. Ubiquitination target will remain covalently bound to the E3, enabling the purification of the E3 together with the target protein. **B**. Similar to A, but in this case the RING E3 catalyzes the transfer of its attached ubiquitin directly from the catalytic cysteine of its respective E2 to the ubiquitination target. Hexagons represent non-covalent interactors of the E3s.

In this article, we describe an improved version of the TULIP methodology (6), which we have termed TULIP2. TULIP2 introduces an extra 10xHIS N-terminal tag preceding the Gateway R cloning cassette. The addition of the extra 10xHIS tag results in an average improvement of more than 50 times in terms of purification efficiency of the TULIP conjugates and an improvement of two orders of magnitude in the signal-to-background ratio after mass spectrometry and Label Free Quantification (LFQ) analysis for the SUMO-Targeted Ubiquitin Ligase (STUbL) RNF4.

METHODS

Materials, Reagents and Antibodies

Dulbecco's modified Eagle's medium, penicillin/streptomycin solution, trypsin-EDTA solution were acquired from Life Technologies (Carlsbad, CA, USA). Fetal bovine serum was from Biowest (Nuaillé, France). Di-sodium hydrogen phosphate dihydrate (Na2HPO4•2H2O) was from VWR chemicals (Radnor, PA, USA). Sodium dihydrogen phosphate monohydrate (NaH2PO4•H2O), sodium chloride, trifluoroacetic acid, tween20, puromycin dihydrochloride and imidazole were acquired from Merck (Darmstadt, Germany). Sodium dodecyl sulfate (SDS), MOPS running buffer and Guanidine hydrochloride 99.5+% were acquired from Thermo Fisher Scientific (Waltham, MA, USA). Nonidet P-40, formic acid (LC-MS grade), methanol (chromasol HPLC), acetonitrile (HPLC grade), MG132 (Z-leu-leu-leu-al) ≥90% HPLC, doxycycline, ponceau-S, polyethylenimine (PEI), urea, ammonium bicarbonate, polybrene, β-mercaptoethanol, and Triton X-100 were from Sigma Aldrich (St. Louis, MO, USA). C18 (Octadecyl) matrix for STAGE-tips was from Bioanalytical Technologies 3M Company (St. Paul, MN, USA). Phosphate-Buffered Saline (PBS) was from Fresenius Kabi (Bad Homburg, Germany). TRIS-Base was from Roche (Basel, Switzerland). Velocity DNA polymerase was from Bioline (London, UK). Elk milk powder was from Campina (Zaltbommel, The Netherlands). Rabbit-anti-RNF4 (Eurogentec, custom made (7)), HRP-conjugated Donkey-anti-Rabbit secondary antibody was from Thermo Fisher Scientific. Western Bright Quantum Western blotting detection kit was from Advansta (Menlo Park, CA, USA).

Generation of the TULIP2 Toolbox

For the construction of the TULIP2 plasmids, using the previous TULIP plasmid (Kumar et al., 2017), a 1.7 Kbp fragment was amplified by PCR with Velocity DNA polymerase using either FW-NheI-H-TULIP2: AGCTAGCATGCATCACCATCATCACCACCACCACCATCACCAATCAACAAGTTTGTACAAAAAAGCTGAACG or FW-NheI-HF-TULIP2: AGCTAGCATGCATCACCATCATCACCACCACCACCATCACGATTACAAGGATGACGACGATAAGCAATCAACAAGTTTGTACAAAAAAGCTGAACG as forward primer for H-TULIP2 and HF-TULIP2, respectively. RV-TULIP2: AGAATTCCGGATGAGCATTCATCAGG as reverse. PCR fragment was digested with NheI and AgeI restriction enzymes and cloned between the NheI and AgeI sites within the TULIP plasmids.

Generation of TULIP2 Lentiviral Plasmids

TULIP2 plasmids are generated by Gateway R cloning (Thermo Fisher Scientific) according to vendor instructions. LR reactions are performed using a donor plasmid containing an E3 enzyme cDNA without stop codon and a TULIP2 plasmid (**Figure 2**) as destination vector. cDNAs from several E3 enzymes without stop codon can be obtained from repositories such as DNASU (8) or the CCSB Human ORFeome Project (9). Additionally, cDNAs can also be subcloned into donor vectors by Gateway R cloning BP reactions (Thermo Fisher Scientific). In this article, we use pDONR207-RNF4, which was previously described (6).

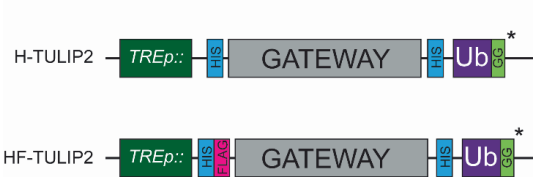


Figure 2. TULIP2 Constructs. Schematic representation of the TULIP2 cloning cassette including the TRE promoter, 10xHIS and tandem 10xHIS-FLAG tag, Gateway cloning cassette, linker containing 10xHIS and active ubiquitin. *ΔGG constructs lack the C-terminal GG motif.

Cell Culture

293T and U2OS were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS) and 100 U/mL penicillin/100µg/mL streptomycin at 37◦C and 5% CO2 unless specifically specified. The cells were regularly tested for mycoplasma contamination.

TULIP2 Lentivirus Production

293T cells were seeded at 30% confluency in a T175 flask containing 16 mL of DMEM + 10% FBS and allowed to attach overnight. Next, a 2 mL transfection mixture was prepared in 150 mM NaCl containing 7.5 µg pMD2.G (#12259, Addgene), 11.4 µg pMDLg-RRE (#12251, Addgene), 5.4 µg pRSV-REV (#12253, Addgene), 13.7 µg TULIP2 plasmid and 114 µL of 1 mg/mL Polyethylenimine (PEI) solution. All the components were mixed by vortexing and incubated 10 min at room temperature. Subsequently, the transfection mix was added to the cells. The day after transfection, culture medium was replaced by fresh DMEM/FBS/Pen/Strep. Three days after transfection, lentiviral suspension was filtered by passing through a 0.45µm syringe filter (PN4184, Pall Corporation). Lentiviral particle concentration was determined using the HIV Type 1 p24 antigen ELISA Kit (ZeptoMetrix Corporation).

TULIP2 Cell Lines

U2OS cells were seeded in 15 cm diameter plates at 10% confluency (2 × 106 cells) and allowed to attach overnight. Next day, cell culture medium was replaced with cell culture medium containing 3.2 µg of lentiviral particles and polybrene 8µg/mL final concentration. Twenty-four hours later, medium was replaced with fresh medium. Three days after lentiviral transduction, TULIP2 construct-positive clones were selected by adding puromycin 3µg/mL to the culture medium.

Purification of TULIP2 Conjugates

A method overview of TULIP2 methodology is provided in Figure 3. Five 15 cm diameter plates of U2OS cells were grown up to 60–80% confluence and the expression of TULIP2 construct was induced with 1µg/mL doxycycline for 24 h. Next, cells were treated for 5 h with proteasome inhibitor MG132 (Sigma Aldrich) at 10µM. Subsequently, cells were washed twice with ice-cold PBS, scraped and transferred to a 50 mL tube. Cells were spun down 5 min at 500 × g, supernatant was discarded and cells were transferred to a 15 mL tube with 5 mL PBS. At this point, a 100 µL aliquot was taken to serve as input sample. After spinning down 1 min at 500 x g and discarding supernatant, input sample cells were lysed in 100 µL SNTBS buffer (2% SDS, 1% NP-40, 50 mM TRIS pH 7.5, 150 mM NaCl). Rest of the sample was centrifuged 3 min at 500 × g and the supernatant discarded. Cell pellet was lysed in 10 mL Guanidinium buffer (6M guanidine-HCl, 0.1M Sodium Phosphate, 10 mM TRIS, pH 7.8). Samples were homogenized at room temperature by sonication using a tip sonicator (Q125 Sonicator, QSonica, Newtown, USA). Sonication was performed at 80% amplitude during 5 s. Subsequently, protein concentration was determined by BiCinchoninic Acid (BCA) Protein Assay Reagent (Thermo Scientific) and sample total protein content was equalized accordingly. Lysates were supplemented with 5 mM β-mercaptoethanol and 50 mM Imidazole pH 8.0. 100 µL of nickel-nitrilotriacetic acid-agarose (Ni-NTA) beads (QIAGEN), were equilibrated with Guanidinium buffer supplemented with 5 mM βmercaptoethanol and 50 mM Imidazole pH 8.0, added to the cell lysates and incubated overnight at 4◦C under rotation. After lysate-beads incubation, samples were centrifuged 5 min at 500 × g and the supernatant was discarded. Ni-NTA beads were transferred with 1 mL Wash buffer 1 (6 M GuanidineHCl, 0.1 M Sodium Phosphate, 10 mM TRIS, 10 mM Imidazole, 5 mM β-mercaptoethanol, 0.2% Triton X-100, pH 7.8) to an Eppendorf LoBind tube (Eppendorf). Centrifuged again, supernatant discarded, and moved to a new LoBind tube with Wash buffer 2 (8 M Urea, 0.1 M Sodium Phosphate, 10 mM TRIS, 10 mM imidazole, 5 mM β-mercaptoethanol, pH 8). Same procedure was repeated with Wash buffer 3 (8 M urea, 0.1 M Sodium Phosphate, 10 mM TRIS, 10 mM imidazole, 5 mM βmercaptoethanol, pH 6.3). Next, beads were washed twice with Wash buffer 4 (8 M urea, 0.1 M Sodium Phosphate, 10 mM TRIS, 5 mM β-mercaptoethanol, pH 6.3). In every wash step, beads were allowed to equilibrate with the buffer for 15 min under rotation. The steps for the purification of the TULIP2 conjugates are indicated in a simplified manner in (Supplementary Protocol 1).

Trypsin Digestion

After second wash with Wash buffer 4, Ni-NTA beads were separated from the buffer by passing through a 0.45µm filter Ultrafree-MC-HV spin column (Merck-Millipore) which had been previously equilibrated with 250 µL of ABC buffer (50 mM ammonium bicarbonate). Using 400 µL of ABC buffer, Ni-NTA beads were transferred to a new Eppendorf LoBind tube and 500 ng of sequencing grade modified trypsin (Promega) were added to the ABC buffer-beads suspension. Digestion was performed overnight at 37◦C while shaking at 1,400 rpm.

Electrophoresis and Immunoblotting

0.1% of the whole-cell extract (Inputs) and 5% of the HIS-purified proteins (TULIP and TULIP2 conjugates) were separated on Novex 4–12% gradient gels (Thermo Fisher Scientific) using NuPAGE R MOPS SDS running buffer (50 mM MOPS, 50 mM TRIS-base, 0.1% SDS, 1 mM EDTA pH 7.7) and transferred onto Amersham Protran Premium 0.45 NC Nitrocellulose blotting membrane (GE Healthcare) using a Bolt Mini-Gel system (Thermo Fisher Scientific), which was used for both the gel electrophoresis and the protein transfer to the membrane according to vendor

instructions. Membrane was stained with Ponceau-S (Sigma Aldrich) to determine total amount of protein loaded. Next membrane was de-stained with PBS + 0.1% Tween-20 and, subsequently, was blocked with Blocking solution (8% Elk milk, 0.1% Tween-20 in PBS) for 1 h. Next, membrane was incubated overnight with 2 ml of a 1:2500 dilution of anti-RNF4 antibody in blocking solution. Next day, membranes were washed 3 times 10 min with PBS + 0.1% Tween-20. Subsequently, membranes were incubated for 1 h with a 1:5000 dilution of HRP-conjugated Donkey-anti-rabbit secondary antibody in blocking solution and washed another 3 times 10 min with PBS+0.1% Tween 20. Chemiluminescence reaction was initiated with Western Bright Quantum Western blotting detection kit and measured in a ChemiDocTM imaging system (BIO-RAD, Hercules, CA, USA). The quantification of the signal corresponding to the TULIP and TULIP2 constructs was done using FIJI software (10).

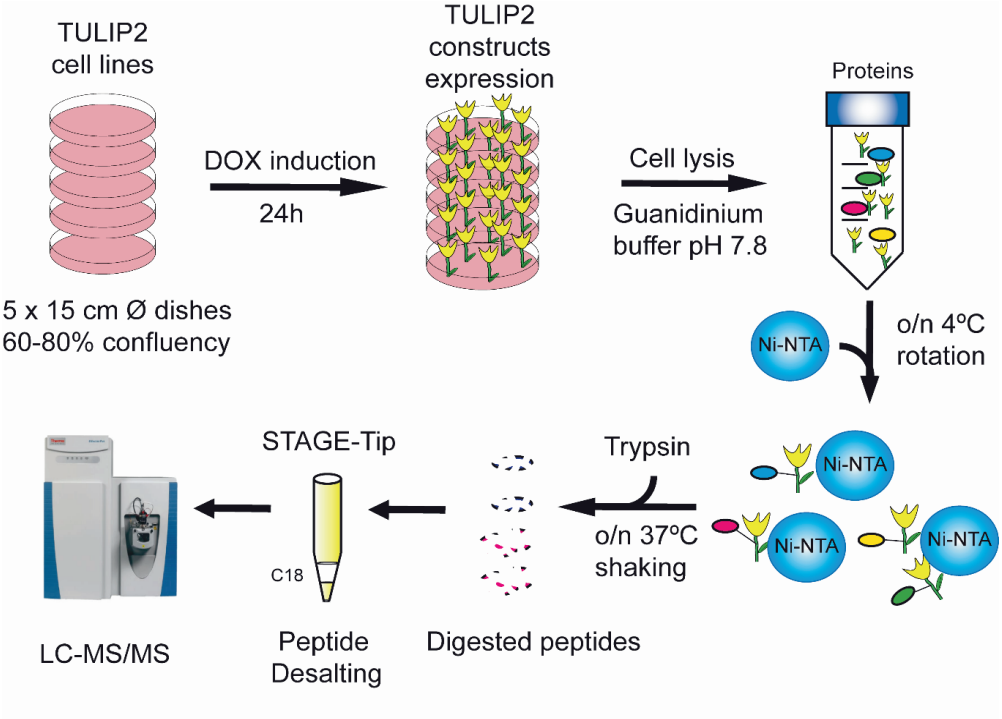


Figure 3. TULIP2 methodology overview. Cells stably containing the E3-TULIP2 expression cassettes are cultured up to 60–80% confluency. The expression of the E3-TULIP2 constructs is induced for 24 h and then they are lysed in Guanidinium buffer and incubated overnight with Ni-NTA beads. Subsequently, beads are washed with different washing buffers and on-the-beads digestion of TULIP2 conjugates with trypsin is performed overnight at 37 ◦C while shaking. Next, digested peptides are desalted by C18 STAGE-Tipping and analyzed by LC-MS/MS.

Mass Spectrometry Sample Preparation

Trypsin-digested peptides were separated from the beads by filtering through a 0.45µm filter Ultrafree-MC-HV spin column (Merck-Millipore) which had been previously equilibrated with 250 µL of ABC buffer. Flow through was collected in an Eppendorf LoBind tube and acidified by adding

2% TriFlourAcetic (TFA) acid. Subsequently, peptides were desalted and concentrated on STAGE-Tips as previously described (11). STAGE-Tips were inhouse assembled using 200 µL micro pipet tips and a C18 matrix. STAGE-Tips were activated by passing through 100 µL of methanol. Subsequently 100 µL of Buffer B (80% acetonitrile, 0.1% formic acid), 100 µL of Buffer A (0.1% formic acid), the peptide sample, and two times 100 µL Buffer A were passed through the STAGE-tip. Elution was performed in 50 µL of 50% acetonitrile, 0.1% formic acid.

Samples were vacuum dried using a SpeedVac RC10.10 (Jouan, France) and stored at −20◦C. Prior to mass spectrometry analysis, samples were reconstituted in 10 µL 0.1% Formic acid and transferred to autoload vials.

LC-MS/MS

All the experiments were performed on an EASY-nLC 1000 system (Proxeon, Odense, Denmark) connected to a QExactive Orbitrap (Thermo Fisher Scientific, Germany) through a nano-electrospray ion source. The Q-Exactive was coupled to a 25 cm silica emitter (FS360-75-15-N-5-C25, NewObjective, Woburn, MA, USA) packed in house with 1.9µm C18- AQ beads (Reprospher-DE, Pur, Dr. Manish, AmmerbuchEntringen, Germany).

Twenty percent of the sample was injected in a 100 min chromatography gradient from 0 to 30% acetonitrile and then increasing to 95% acetonitrile prior to column re-equilibration with flow rate of 200 nL/min. The mass spectrometer was operated in a Data-Dependent Acquisition (DDA) mode with a top-10 method and a scan range of 300–1,600 m/z. Full-scan MS spectra were acquired at a target value of 3 × 106 and a resolution of 70,000, and the Higher-Collisional Dissociation (HCD) tandem mass spectra (MS/MS) were recorded at a target value of 1 × 105 and with a resolution of 17,500, an isolation window of 2.2 m/z, and a normalized collision energy (NCE) of 25%. The minimum AGC target was 1 × 104 . The maximum MS1 and MS2 injection times were 250 and 60 ms, respectively.

The precursor ion masses of scanned ions were dynamically excluded (DE) from MS/MS analysis for 20 s. Ions with charge 1, and >6, were excluded from triggering MS2 analysis.

Mass Spectrometry Data Analysis

All raw data were analyzed using MaxQuant (version 1.6.7.0) as described previously (12). We performed the search against an in silico digested UniProt reference proteome for Homo sapiens including canonical and isoform sequences (27th May 2019). Database searches were performed according to standard settings with the following modifications. Digestion with Trypsin/P was used, allowing 4 missed cleavages. Oxidation (M), Acetyl (Protein N-term), and GlyGly (for ubiquitination sites) were allowed as variable modifications with a maximum number of 3. Carbamidomethyl (C) was disabled as a fixed modification. Label-Free Quantification was enabled, not allowing Fast LFQ. All peptides were used for protein quantification.

Output from MaxQuant Data were exported and processed in MS Excel for further filtering, processing of the data, and visualization.

For the statistical analysis of RNF4-TULIP2 samples, output from the analysis in MaxQuant was further processed in the Perseus computational platform (v 1.6.7.0) (13). LFQ intensity values were log2 transformed. Potential contaminants and proteins identified by site only or reverse peptide were removed. Samples were grouped in experimental categories and proteins not identified in 3

out of 3 replicates in at least one group were also removed. Missing values were imputed using normally distributed values with a 1.8 downshift (log2) and a randomized 0.3 width (log2) considering whole matrix values. Statistical analysis was performed to determine which proteins were significantly enriched in the wild type RNF4 samples compared to the 1GG samples (t-test with permutationbased False Discovery Rate (FDR) = 0.05 and S0 = 0.1).

RESULTS

TULIP vs. TULIP2

Previously, TULIP methodology was employed to identify the SUMO Targeted Ubiquitin Ligase (STUbL) RNF4 specific ubiquitination targets (6). In order to compare the new TULIP2 methodology with the previous TULIP methodology version, we cloned the RNF4 into the H-TULIP2 plasmids. Next, we generated lentiviral particles containing the RNF4-TULIP2 constructs and used them to stably introduce the RNF4-TULIP2 constructs in U2OS cells by lentiviral transduction. Positive clones were selected with puromycin.

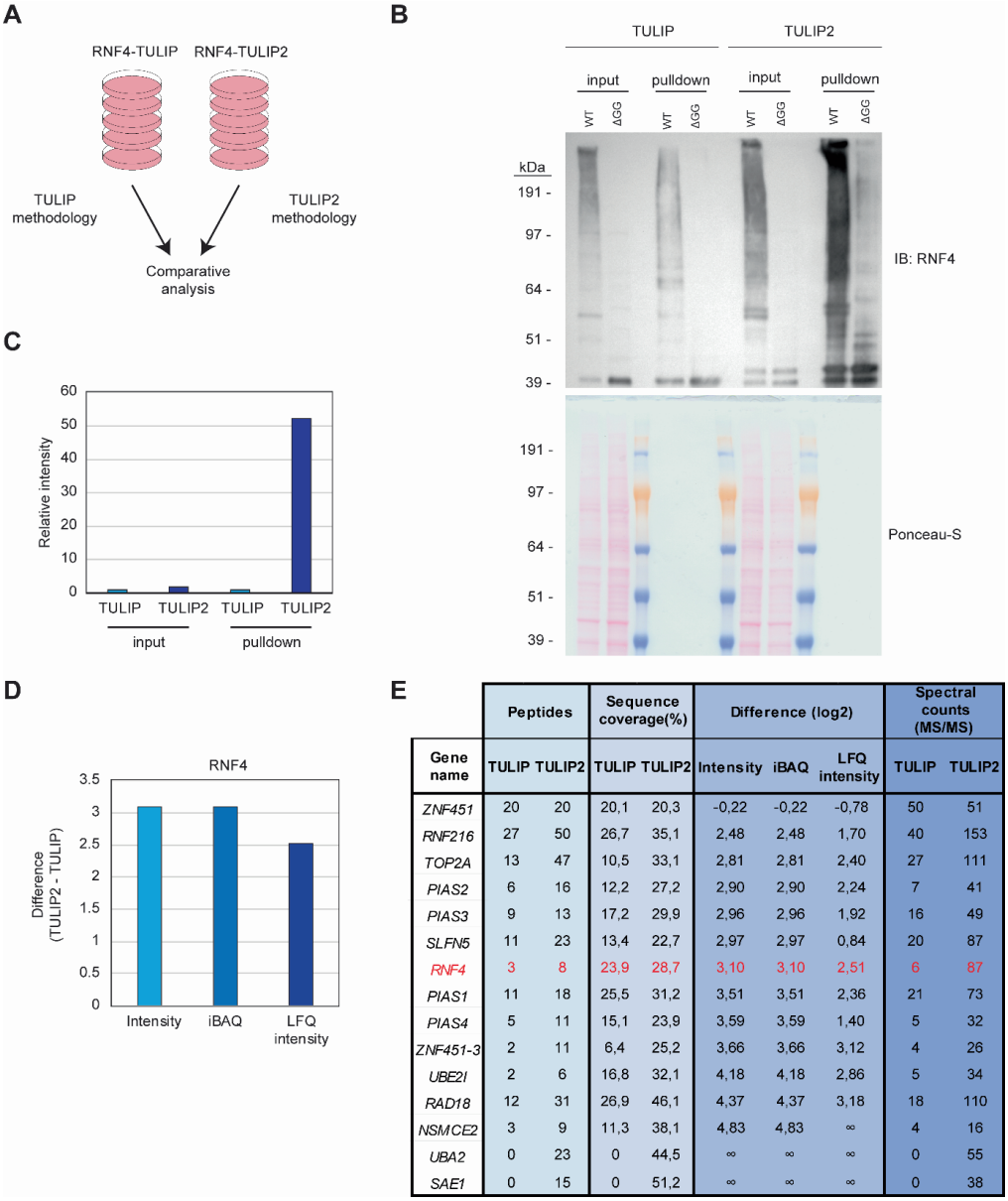
Cells expressing RNF4-TULIP and RNF4-TULIP2 constructs were grown in equal amount, induced for the same time and treated for 5 h with the proteasome inhibitor MG132. Next, cells were lysed and the RNF4-TULIP and RNF4-TULIP2 conjugates were purified in parallel following the TULIP methodology protocol (14) or the TULIP2 method introduced in this article, respectively (**Figure 4A**). Next, whole cell extracts and 5% of the HIS-pulldown samples were analyzed by immunoblotting using an anti-RNF4 antibody (**Figure 4B**). While the RNF4-TULIP2 constructs were expressed relatively higher than their RNF4-TULIP counterparts by a factor of 1.7, the amount of RN4-TULIP2 conjugates purified were 52.2 times higher compared to the amount of RNF4-TULIP conjugates while using the same amount of starting material (**Figure 4C**).

Next, we decided to perform a comparison using three biological replicates of RNF4-TULIP2 samples and the RNF4- TULIP samples from Kumar et al. 2017 (6) both generated after treating with the proteasome inhibitor MG132. In both cases, 20% of the RNF4-TULIP or RNF4-TULIP2 samples were injected in the mass spectrometer and analyzed using the same chromatography gradients. All three biological replicates of each sample set were grouped together for performing comparisons. Signal corresponding to RNF4 was more than 8 times higher in the TULIP2 samples compared to TULIP samples when looking at Intensity or iBAQ MaxQuant output values and more than 5 times in the case of the values of the Label Free Quantification intensity (**Figure 4D**).

Previously, using TULIP methodology, we identified components of the sumoylation machinery and other proteins such as TOP2A, SLFN5, RAD18, and RNF216 as the most important SIM- and MG132-dependent RNF4 targets. Using TULIP2 methodology we were able to increase the number of peptides, the percentage of sequence coverage, intensity, iBAQ, and LFQ intensity values and the number of spectral counts for all these RNF4 direct ubiquitination targets (**Figure 4E**, **Supplementary Dataset 1**).

While TULIP methodology allowed us to identify SUMO E3s and E2 as ubiquitination targets for RNF4, TULIP2 methodology also identified the SUMO E1 enzyme (SAE1/UBA2) as an RNF4 ubiquitination target, indicating that, upon SUMOylation, all the members of the SUMOylation machinery, including E1, E2, and E3 enzymes, are targeted for degradation in an RNF4- dependent manner.

Next, in order to generate a new list of RNF4 ubiquitination targets by using TULIP2 methodology, we performed a second analysis including RNF4-TULIP2 samples and RNF4-TULIP2-1GG samples as negative control. We performed 3 biological replicates of each construct in order to perform statistical comparisons. Comparison between the RNF4-TULIP2 and RNF4-TULIP2-1GG identified 409 RNF4-TULIP2 conjugated proteins (**Figure 5**, **Supplementary Dataset 2**). Moreover, mass spectrometry analysis also allowed to identify 372 specific ubiquitination sites in 209 proteins (**Supplementary Dataset 3**), including many members of the sumoylation machinery and the previously identified as main ubiquitination targets targeted for degradation by RNF4 in a SUMO-dependent manner.



E

Gene name	Peptides		Sequence coverage(%)		Difference (log2)			Spectral counts (MS/MS)	
	TULIP	TULIP2	TULIP	TULIP2	Intensity	iBAQ	LFQ intensity	TULIP	TULIP2
ZNF451	20	20	20,1	20,3	-0,22	-0,22	-0,78	50	51
RNF216	27	50	26,7	35,1	2,48	2,48	1,70	40	153
TOP2A	13	47	10,5	33,1	2,81	2,81	2,40	27	111
PIAS2	6	16	12,2	27,2	2,90	2,90	2,24	7	41
PIAS3	9	13	17,2	29,9	2,96	2,96	1,92	16	49
SLFN5	11	23	13,4	22,7	2,97	2,97	0,84	20	87
RNF4	3	8	23,9	28,7	3,10	3,10	2,51	6	87
PIAS1	11	18	25,5	31,2	3,51	3,51	2,36	21	73
PIAS4	5	11	15,1	23,9	3,59	3,59	1,40	5	32
ZNF451-3	2	11	6,4	25,2	3,66	3,66	3,12	4	26
UBE2I	2	6	16,8	32,1	4,18	4,18	2,86	5	34
RAD18	12	31	26,9	46,1	4,37	4,37	3,18	18	110
NSMCE2	3	9	11,3	38,1	4,83	4,83	∞	4	16
UBA2	0	23	0	44,5	∞	∞	∞	0	55
SAE1	0	15	0	51,2	∞	∞	∞	0	38

Figure 4. TULIP vs. TULIP2. **A**. Experimental design to compare TULIP vs. TULIP2. **B**. U2OS cells containing either RNF4-TULIP or RNF4-TULIP2 expression cassettes were induced overnight with doxycycline, lysed and TULIP/TULIP2 conjugates purified according to TULIP or TULIP2 methodology, respectively. The efficiency of the expression and the purification was analyzed by immunoblotting. Ponceau-S is provided as loading control. **C**. Quantification of the intensity from the immunoblotting analysis performed in (B). Intensity of the signal in TULIP samples is normalized as 1. **D**. Graph depicting the log2 difference between RNF4-TULIP2 and RNF4-TULIP samples for RNF4 after mass spectrometry analysis in terms of Intensity, iBAQ or LFQ

intensity. **E**. Table indicating the values for number of peptides, sequence coverage, log2 difference of intensities after LC-MS/MS analysis and spectral counts of top RNF4-specific ubiquitination targets comparing RNF4-TULIP and RNF4-TULIP2 samples.

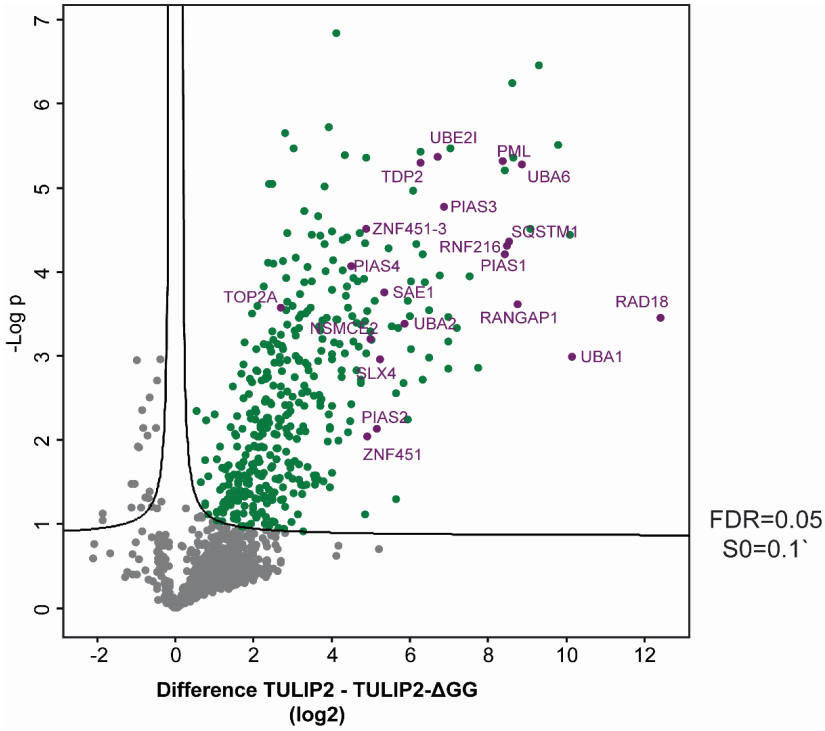


Figure 5. RNF4-TULIP2 ubiquitination targets. Volcano plot depicting RNF4-TULIP2 conjugates comparing to RNF4-TULIP2-1GG samples. Each dot represents a protein. Green dots represent proteins that are statistically enriched in the RNF4-TULIP2 samples compared to RNF4-TULIP2-1GG samples for an FDR = 0.05 and S0 = 0.1. Purple labeled dots represent proteins relates to the SUMOylation machinery or top main ubiquitination targets previously identified by TULIP methodology.

DISCUSSION, ADVANTAGES, AND PITFALLS

In this article we have performed a comparison between our previously published TULIP methodology (6) and an improved version, which we have termed TULIP2 methodology. Compared to previous version, for the STUbL RNF4, it achieves a more than 50 times improvement in terms of purification efficiency (**Figures 4B,C**). This methodology can be implemented in any laboratory interested in the identification of the ubiquitination targets of a given E3 of interest. Furthermore, the simplification of the protocol by suppressing the elution and size exclusion filter-based sample concentration results in a reduction of the execution costs of the experiments. Moreover, the introduction of the HIS-FLAG TULIP2 plasmids allow the employment of an anti-FLAG tag antibody when a good specific antibody for immunoblotting is not available for the E3 enzyme of interest or for unambiguous identification respect of the endogenous E3 enzyme.

Together, all these improvements enable the implementation of the TULIP2 methodology in any research group with access to a mass-spectrometry facility. To facilitate the implementation of the TULIP2 methodology in any laboratory we have included an annotated step-by-step protocol from the induction of the expression of the TULIP2 constructs until the isolation of the trypsin-digested peptides corresponding to the TULIP2 constructs and conjugates.

The improvement achieved by TULIP2 allowed us not only to have a better coverage of the RNF4 ubiquitination targets after mass spectrometry analysis, but also to identify new RNF4 ubiquitination substrates (**Figures 4E, 5**, **Supplementary Datasets 1, 2**). Moreover, we could determine the specific ubiquitination sites of many of the identified RNF4 targets (**Supplementary Dataset 3**). While previous TULIP methodology allowed us to identify 31 ubiquitination sites on 16 proteins (6), these numbers increased to 372 and 209, respectively, using TULIP2 methodology.

The improvement achieved by TULIP2 methodology facilitates the identification of specific substrates for other E3 enzymes which are less stable, their ubiquitination targets less abundant and/or have a lower ubiquitination activity than RNF4. The identification of the E3-specific ubiquitination substrates using TULIP methodology was still challenging and very large amounts of cells needed to be lysed to obtain the minimum amounts of TULIP conjugates to allow identification by mass spectrometry. TULIP2 methodology solves this major drawback. TULIP2 is straightforward and enables the systematic identification of the specific ubiquitination targets of virtually every HECT- and RING-type E3 enzyme. Using Gateway cloning, any E3-ligase cDNA can be cloned into the TULIP2 plasmids.

Nevertheless, the TULIP2 methodology still shares some limitations with the previous version of the method (6). Some E3-TULIP2 constructs might not be functional due to steric hindrance and the size of the E3 to be cloned into the TULIP2 plasmids is limited by the capacity of the lentiviral particles. As an indication, we have been able to clone E3 enzymes with cDNA sizes up to 6 kilobase pairs. Some E3-TULIP2 constructs might be very rapidly targeted for degradation by the proteasome via autoubiquitination given that the already present ubiquitin moiety is a signal for ubiquitin chain lengthening. Thus, inhibition of the proteasome might be required to be able to purify sufficient amount of TULIP2 conjugates to secure identification by mass spectrometry.

It is also worth noting that, although TULIP2-attached E3s represent a bulky tag that hamper the utilization of the attached ubiquitin by other E3s to ubiquitinate their targets, potentially ubiquitin moieties from the TULIP2 constructs can still be used by other E3s. Thus, including catalytically dead mutants of the E3s of interest as an additional negative control to the 1GG TULIP2 constructs might be advantageous. Finally, the probability of success in identifying the specific ubiquitination substrates for a given E3 enzyme highly depends on the sensitivity of the mass spectrometry equipment employed and the amount of sample injected. The signal corresponding to the TULIP2 conjugates is commonly below the signal corresponding to the common unspecific binders to Ni-NTA beads, making good enrichment is critical for successful identification.

DATA AVAILABILITY STATEMENT

TULIP2 construct plasmids are freely available from the González-Prieto lab upon reasonable request. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (15) partner repository with the dataset identifier PXD015437.

AUTHOR CONTRIBUTIONS

RG-P designed and constructed the TULIP2 plasmids. RG-P, DS-L, and GA performed experiments. GA was supervised by DS-L. RG-P and DS-L made the figures. RGP wrote the manuscript with input from DS-L. All authors contributed to manuscript revision, read and approved the submitted version.

FUNDING

Work in the González-Prieto lab was supported by the Dutch Cancer Society (KWF-Young Investigator Grant: 11367).

ACKNOWLEDGMENTS

Authors would like to thank Prof. Alfred Vertegaal for infrastructural support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem. 2019.00802/full#supplementary-material.

Supplementary Dataset 1 | Values from the mass spectrometry analysis comparing RNF4-TULIP and RNF4-TULIP2.

Supplementary Dataset 2 | Statistical analysis of RNF4-TULIP2 samples compared to RNF4-TULIP2-1GG samples.

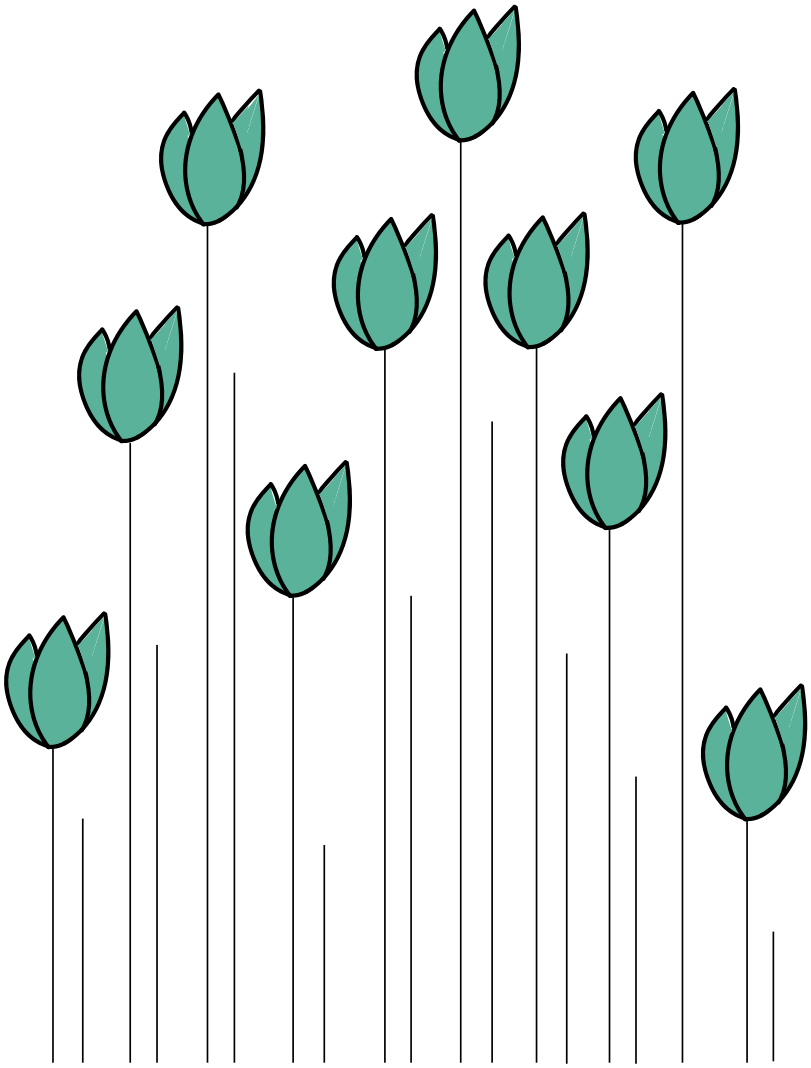
Supplementary Dataset 3 | Specific ubiquitination sites identified by mass spectrometry in RNF4-TULIP2 samples.

Supplementary Protocol 1 | TULIP2 benchtop step-by-step protocol.

REFERENCES

1. V. Akimov *et al.*, UbiSite approach for comprehensive mapping of lysine and N-terminal ubiquitination sites. *Nature structural & molecular biology* **25**, 631-640 (2018).
2. M. Song, K. Hakala, S. T. Weintraub, Y. Shiio, Quantitative proteomic identification of the BRCA1 ubiquitination substrates. *J Proteome Res* **10**, 5191-5198 (2011).
3. S. A. Sarraf *et al.*, Landscape of the PARKIN-dependent ubiquitylome in response to mitochondrial depolarization. *Nature* **496**, 372-376 (2013).
4. J. W. Thompson *et al.*, Quantitative Lys--Gly-Gly (diGly) proteomics coupled with inducible RNAi reveals ubiquitin-mediated proteolysis of DNA damage-inducible transcript 4 (DDIT4) by the E3 ligase HUWE1. *J Biol Chem* **289**, 28942-28955 (2014).
5. H. F. O'Connor *et al.*, Ubiquitin-Activated Interaction Traps (UBAITs) identify E3 ligase binding partners. *EMBO Rep* **16**, 1699-1712 (2015).
6. R. Kumar, R. Gonzalez-Prieto, Z. Xiao, M. Verlaan-de Vries, A. C. O. Vertegaal, The STUbL RNF4 regulates protein group SUMOylation by targeting the SUMO conjugation machinery. *Nat Commun* **8**, 1809 (2017).
7. R. Vyas *et al.*, RNF4 is required for DNA double-strand break repair in vivo. *Cell death and differentiation* **20**, 490-502 (2013).
8. C. Y. Seiler *et al.*, DNASU plasmid and PSI:Biology-Materials repositories: resources to accelerate biological research. *Nucleic acids research* **42**, D1253-1260 (2014).

9. P. Lamesch *et al.*, hORFeome v3.1: a resource of human open reading frames representing over 10,000 human genes. *Genomics* **89**, 307-315 (2007).
10. J. Schindelin *et al.*, Fiji: an open-source platform for biological-image analysis. *Nature methods* **9**, 676-682 (2012).
11. J. Rappsilber, M. Mann, Y. Ishihama, Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature protocols* **2**, 1896-1906 (2007).
12. S. Tyanova, T. Temu, J. Cox, The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature protocols* **11**, 2301-2319 (2016).
13. S. Tyanova *et al.*, The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature methods* **13**, 731-740 (2016).
14. R. Gonzalez-Prieto, A. C. O. Vertegaal, "TULIP: Targets of Ubiquitin Ligases Identified by Proteomics" in SUMOylation and Ubiquitination: Current and Emerging Concepts, V. G. Wilson, Ed. (Caister Academic Press, U.K., 2019), https://doi.org/10.21775/9781912530120.10 chap. 10, pp. 147-160.
15. Y. Perez-Riverol *et al.*, The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic acids research* **47**, D442-D450 (2019).



4

SUMO Activated Target Traps (SATTs) enable the identification of a comprehensive E3-specific SUMO proteome

Daniel Salas-Lloret[1] , Nicolette S. Jansen[1,8], Easa Nagamalleswari[2,8], Ekaterina Gracheva[1], Coen van der Meulen[1], Arnoud H. de Ru[3] , H. Anne Marie Otte[3] , Peter A. van Veelen[3] , Andrea Pichler[2,4], Joachim Goedhart[5] , Alfred C.O. Vertegaal[1] , Román González-Prieto[1,6,7,9]

[1] Cell and Chemical Biology, Leiden University Medical Center, Leiden, The Netherlands.

[2] Max Plank Institute for Immunobiology and Epigenetics, Freiburg, Germany.

[3] Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, The Netherlands.

[4] Institute of Biochemistry, ETH Zürich, Zürich, Switzerland.

[5] University of Amsterdam, Amsterdam, The Netherlands

[6] Genome Proteomics laboratory, Department of Genome biology, Andalusian Center for Molecular Biology and regenerative Medicine (CABIMER), University of Seville, Seville, Spain.

[7] Department of Cell Biology, University of Seville, Seville, Spain.

[8] Equal contribution

[9] Corresponding author

This chapter has been published in Science Advances. 9, eadh2073 (2023).

Abstract

Ubiquitin and ubiquitin-like conjugation cascades consist of dedicated E1, E2 and E3 enzymes with E3s providing substrate specificity. Mass spectrometry-based approaches have enabled the identification of more than 6,500 SUMO2/3 target proteins. The limited number of SUMO E3s provides the unique opportunity to systematically study E3-substrate wiring. We developed SUMO Activated Target Traps (SATTs) and systematically identified substrates for eight different SUMO E3s, PIAS1, PIAS2, PIAS3, PIAS4, NSMCE2, ZNF451, LAZSUL(ZNF451-3) and ZMIZ2. SATTs enabled us to identify 427 SUMO1 and 961 SUMO2/3 targets in an E3-specific manner. We found pronounced E3 substrate preference, even at the substrate isoform level. Quantitative proteomics enabled us to measure substrate specificity of E3s, quantified using the SATT index. Furthermore, we developed the Polar SATTs web-based tool (https://amsterdamstudygroup.shinyapps.io/polarVolcaNoseR_revised/) to browse the dataset in an interactive manner, increasing the accessibility of this resource for the community. Overall, we uncover E3-to-target wiring of 1681 SUMO substrates, highlighting unique and overlapping sets of substrates for eight different SUMO E3 ligases.

Keywords: Mass-Spectrometry; SUMO; E3s; SATTs

INTRODUCTION

Protein fate and function is controlled by numerous Post-Translational Modifications (PTMs). Among them, ubiquitination is the second most important PTM after phosphorylation (1) and controls virtually every process in eukaryotic cells in a dynamic manner. Ubiquitination consists of the covalent attachment of the small 76 amino acids ubiquitin protein to acceptor proteins. It is performed by an enzymatic cascade in which ubiquitin-activating enzymes (E1) activate ubiquitin and transfer it to ubiquitin-conjugating enzyme (E2) which conjugates ubiquitin to the substrate assisted by a ubiquitin-ligase enzyme (E3). E3s are responsible for determining substrate specificity. The human genome encodes for two ubiquitin E1s, 30-40 E2s and more than 600 E3s (2).

Similar to ubiquitin, other ubiquitin-like (Ubl) modifiers exist, which have dedicated E1-E2-E3 enzymatic cascades. Among these Ubls, Small Ubiquitin-like Modifiers (SUMOs) are the most abundant after ubiquitin. In vertebrates, there are three different types of active SUMOs: SUMO1, SUMO2 and SUMO3. Mature SUMO2 and SUMO3 differ only in a couple of amino acids and are commonly referred to as SUMO2/3. In contrast to ubiquitin, vertebrates express a single E1, a single E2 and less than a dozen *bona fide* E3s for SUMOs (2).

Recent advances in mass spectrometry technologies and the optimization of sample preparation methodologies (3) have enabled the identification of several tens of thousands of acceptor sites on thousands of proteins in human cells both for ubiquitin and SUMOs (4-10). However, our knowledge on E3-substrate wiring is still very limited. Determining which E3 modifies which substrate is a major challenge.

For ubiquitin, given the high number of E3s, solving the E3-to-target wiring in a proteome-wide manner is virtually impossible. However, for SUMOs, the E3 complexity is limited, simplifying this task. A proposed approach has been the quantification of changes on the SUMO proteome after SUMO E3 overexpression(11), which in principle, is an indirect measure. Another applied approach has been the performance of SUMOylation assays on protein array-based screens (12), which is an *ex vivo* system that misses out on the restricted subcellular localization of proteins and lacks protein-protein complexes that are abundant in cells.

Here, we took advantage of our previous experience in the systematic identification of ubiquitination substrates using Ubiquitin Activated Interaction Traps (UbAITs) (13) in the Targets of Ubiquitin Ligase Identified by Proteomics (TULIP) methodologies (14-16) and applied it for the identification of SUMO E3-specific substrates in a systematic manner for SUMO E3s in a proteome-wide approach.

RESULTS

SUMO E3 overexpression causes SUMO2/3 depletion in an RNF4-dependent manner.

Aiming to identify putative E3-specific SUMOylation substrates, we employed a similar approach as previously done with PIAS1(11). We made GFP-tagged constructs for different SUMO E3s, including NSMCE2, PIAS1, PIAS2, PIAS3, PIAS4, ZNF451, the LAP2α isoform of the ZNF451 SUMO Ligase (LAZSUL) and, additionally, another PIAS-like enzyme, ZMIZ2(17-20) (**Figure 1A**), which , we previously tested for *in vitro* SUMO E3 activity both for SUMO1 and SUMO2 ligase activity (**Figure 1B**). As a result, we observed that ZMIZ2 has E3 enzymatic activity for SUMO2 but not for SUMO1. Next, we transfected the GFP-tagged constructs of the E3s indicated in Figure 1A in U2OS cells. To evaluate the transfection efficiency of our constructs, we analyzed our cells by fluorescence microscopy after immunostaining for SUMO2/3 (**Figure 1C, Supplementary Figure 1A**). GFP-positive cells could be observed for every construct at different efficiencies, except for GFP-PIAS2, which transfection did not lead to the appearance of GFP-positive cells. Unexpectedly, the immunofluorescence SUMO2/3 signal was highly reduced in GFP-positive cells for NSMCE2, PIAS1, ZNF451 and LAZSUL (**Figure 1C, Supplementary Figure 1A**). Therefore, we quantified the SUMO2/3 signal by immunofluorescence for GFP-positive and -negative cells from three independent experiments (**Figure 1D**). While GFP-NSMCE2, -PIAS1, -PIAS3, -ZNF451 and -LAZSUL reduced the average SUMO2/3 nuclear signal by 41%, 63%, 19%, 77% and 82% respectively, GFP-PIAS4 positive cells presented a slight increase of 4% in SUMO2/3 signal. Interestingly, GFP-ZMIZ2 positive cells had a remarkable 46% increase in SUMO2/3 signal.

In a previous screen for targets of the SUMO Targeted Ubiquitin Ligase (STUbL) RNF4, we had observed that SUMO E3s were targets of RNF4 for ubiquitination and subsequent degradation by the proteasome, with ZNF451 and PIAS1 being the strongest RNF4 ubiquitination targets, and PIAS4 the weakest. ZMIZ2 was not a substrate for RNF4 (15, 16). We hypothesized that overexpression of these E3s was promoting their hyperactivation leading to their auto-SUMOylation and increased SUMOylation of their substrates and subsequent degradation in an RNF4-dependent manner. Thus, proteasome inhibition should rescue the effect on SUMO2/3 levels in response to the different E3s overexpression. Therefore, we compared the effect on SUMO2/3 levels of the different SUMO E3s overexpression in the presence or absence of the proteasome inhibitor MG132 for 5 hours (**Supplementary Figure 1B**). The 5 hours MG132 treatment could rescue the effect on SUMO2/3 levels of the E3s which overexpression had milder phenotypes, namely NSMCE2, PIAS3 and PIAS4, but was not sufficient to rescue strong effects of PIAS1, ZNF451, LAZSUL overexpression. Furthermore, proteasome inhibition by MG132 has many pleiotropic effects.

To further test our hypothesis, we made stable inducible U2OS cells for GFP-LAZSUL, which was the E3 with the strongest phenotype (**Figure 1C-D, Supplementary Figure 1 A-B**). Cells were treated with a control or an RNF4-targetting siRNA, GFP-LAZSUL was dox induced and analyzed by immunostaining (**Figure 1E-F**) and immunoblotting (**Figure 1G**). RNF4 knock down caused an increase in the fraction of GFP-LAZSUL positive cells and rescued the SUMO2/3 depletion phenotype. Consistently, RNF4 knockdown increased the levels of both modified and non-modified GFP-LAZSUL.

RNF4 knockdown increases cellular SUMO2/3 levels but does not affect SUMO1 levels (21). Thus, we also decided to investigate the effect of NSMCE2, PIAS1, ZNF451, and LAZSUL transient overexpression on SUMO1 levels by immunofluorescence (Supplementary Figure 1C-D). Accordingly, the overexpression of these E3s did not cause SUMO1 depletion as previously observed for SUMO2/3 (Figure 1C-F).

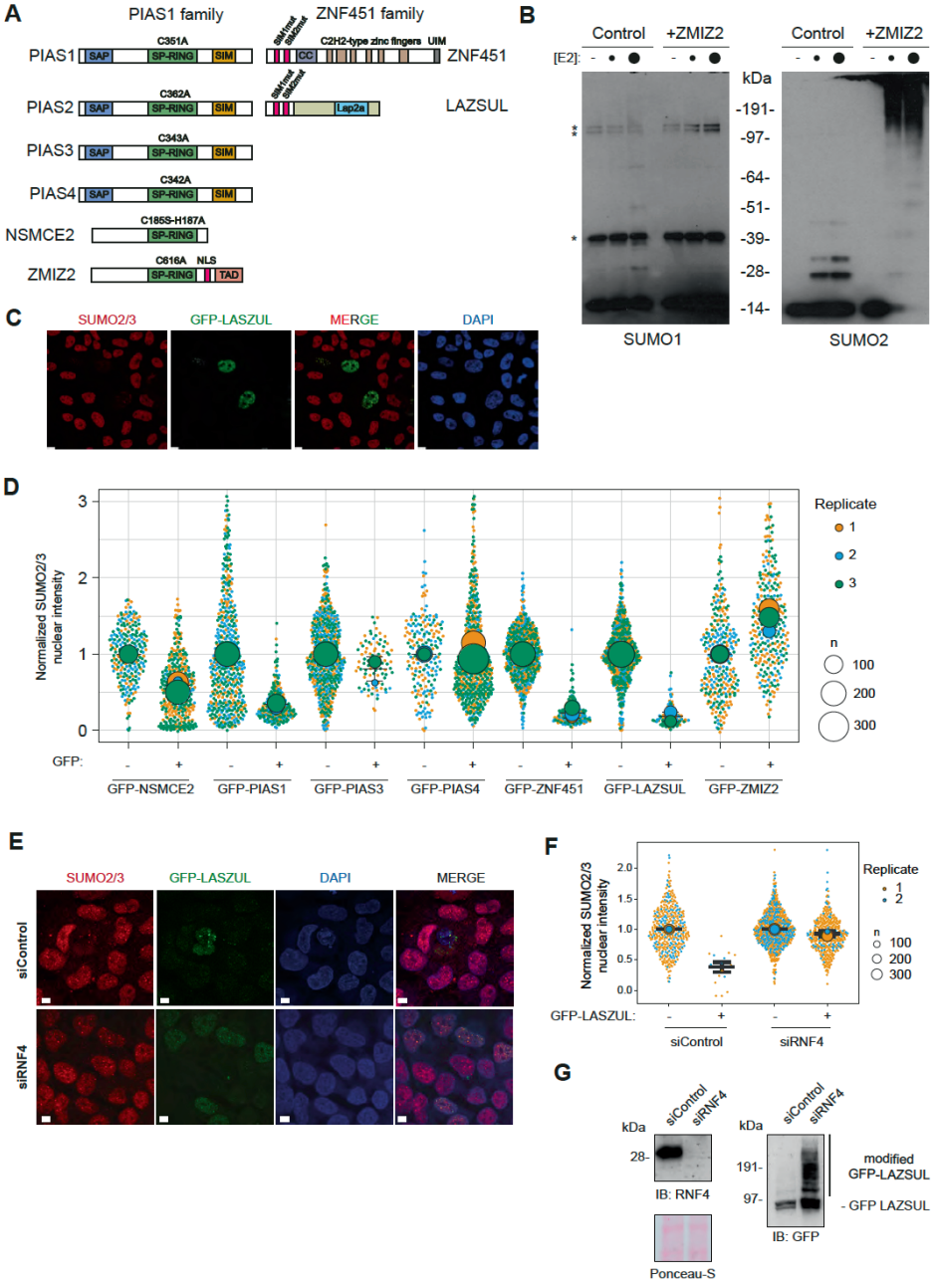


Figure 1. (A) E3s studied in this article. The mutations performed on each E3 to construct the catalytic-dead mutant controls are indicated **(B))** *In vitro* sumoylation assays including or ZMIZ2 SUMO E3 enzyme and different concentrations of the SUMO E2 (0, 100 and 200 ng). Assays were carried out using either SUMO1 or SUMO2. **(C)** Representative immunofluorescence image of U2OS cells transiently transfected with GFP-LAZSUL immunostained for SUMO2/3. **(D)** Superplot depicting relative SUMO2/3 nuclear intensities after immunostaining of individual U2OS cells transiently transfected with GFP-tagged constructs of different E3s. Values were normalized to the average SUMO2/3 nuclear intensity of GFP-negatives from each individual experiment. Values from 3 independent experiments are depicted. **(E)** Stable inducible GFP-LAZSUL expressing U2OS cells were treated with control or RNF4-targeting siRNAs. 36h after siRNA transfection, GFP-LAZSUL expression was induced with 20 µg/mL doxycycline. Cells were fixe 48h after siRNA transfection and analyzed by immunostaining. **(F)** Quantification of the normalized nuclear SUMO2/3 intensities from the cells in (E). Independent values from two independent experiments are depicted. **(G)** Analysis by immunoblotting of the cells in (C). Size bars in fluorescence microscopy images represent 10 µm.

4

We conclude that SUMO E3 overexpression-based screens to identify SUMOylation substrates could potentially be misleading due to a negative control loop mediated by RNF4. This loop is activated upon SUMO E3 overexpression and leads to SUMO2/3 depletion in cells. Therefore, SUMO E3 overexpression screens must be carefully evaluated.

SUMO Activated Target Traps (SATTs) to identify the E3-specific SUMO proteome.

Previously, in an effort to identify E3-specific ubiquitin substrates, Ubiquitin Activated Interaction Traps (UbAITs) were engineered (13), which we later adopted and optimized for systematic screening in the Targets for Ubiquitin Ligases Identified by Proteomics (TULIP) methodology (15, 16). However, due to the high number of Ubiquitin E3 enzymes in the human proteome, performing the TULIP methodology on each E3 is an incredibly challenging task.

In contrast to Ubiquitin, the number of *bona fide* SUMO E3 enzymes is more limited, comprising the Siz/Pias Really Interesting New Gene (S-P RING) family, the ZNF451 family and RANBP2 (2). Therefore, addressing the E3-substrate wiring for SUMO E3s is a more manageable challenge.

Thus, similar to the TULIP2 methodology (15), we designed the SUMO Activated Target Traps (SATTs) approach, in which lentiviral doxycycline-inducible plasmids consisting of 10xHIS tag and a Gateway cloning sequence, followed by 10xHIS and either mature SUMO1 or mature $SUMO2_{Q87R}$ were constructed. **(Figure 2A**) The Gateway sequence enables the straightforward shuttling of any SUMO E3 of interest. The $SUMO2_{Q87R}$ mutation facilitates the identification of SUMO acceptor sites by mass spectrometry-based proteomics (7, 8). Consistently, the rationale behind this approach is that if we generate a linear fusion between an E3 and activated SUMO, the E3 will be prone to use the attached SUMO moiety to modify its substrate, enabling the co-purification of the E3 together with its substrate and subsequent identification by mass spectrometry-based proteomics. In line with TULIP2 methodology (15), we included two different negative controls in our screens. The first control is a ΔGG construct where the SUMO moiety lacks the C-terminal di-Gly motif and thus cannot be conjugated to a substrate. The second control is a catalytic dead mutant where the interaction with the SUMO E2 enzyme is abolished, thus the transfer of the SUMO moiety from the E2 to the substrate cannot be catalyzed **(Figure 2B**).

Accordingly, we built SATTs for the E3s indicated in Figure 1A. SUMO1 SATTs for the S-P RING SUMO E3 enzymes, PIAS1, PIAS2, PIAS3, PIAS4 and NSMCE2. Additionally, for SUMO2 SATTs, in addition to the S-P RING SUMO E3 enzymes, we also included ZNF451, LAZSUL and ZMIZ2 as they are exclusive for SUMO2/3 (22, 23) (**Supplementary Figure 2A**). Other characterized SUMO E3s were not included for different reasons. RANBP2 was left out from our screen due to the size of the protein (3,224 amino acids). Also, the ZNF451 family E3 KIAA1586 was left out because it is exclusively found in primates, and not in other vertebrates (23). To generate the catalytic-dead mutant controls, we introduced specific mutations in each E3 (**Figure 1A**). For the S-P RING family E3s, we mutated cysteines in the S-P RING domain, for the ZNF451 family E3s, we mutated the SUMO Interaction Motifs (SIMs) by substituting the long hydrophobic amino acids into alanines (23).

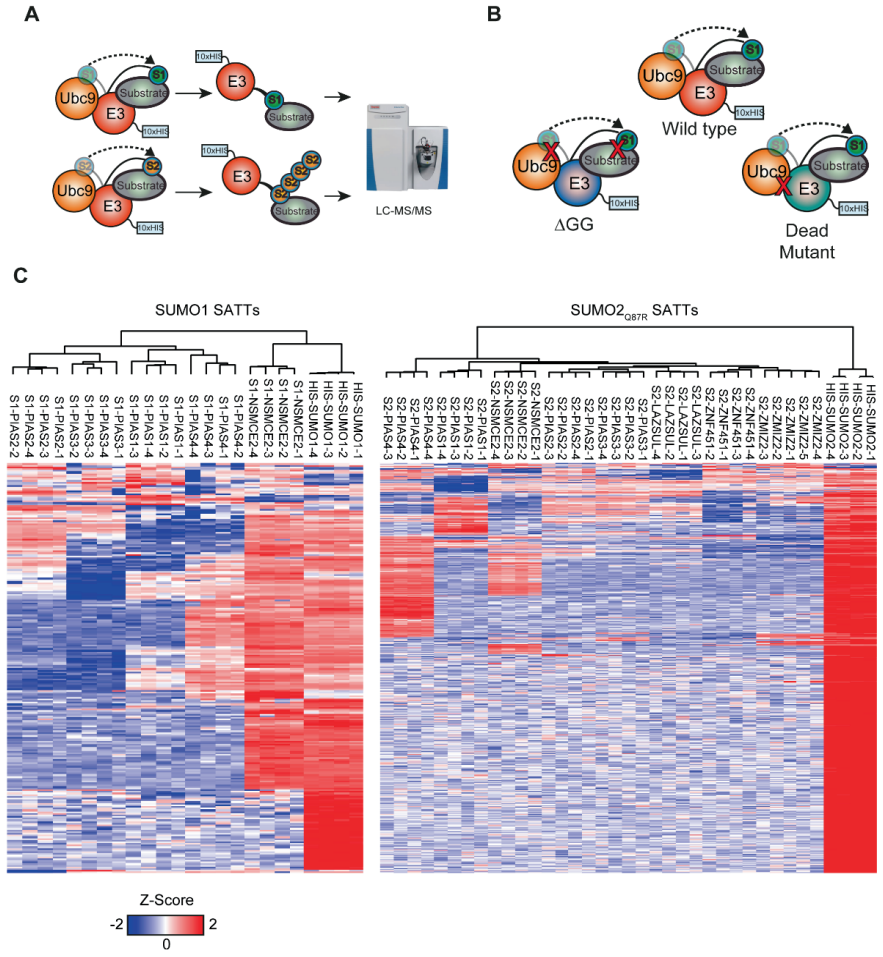


Figure 2. (A) SATTs screen rationale. SUMO moieties fused to the C term of an E3 of interest will be attached to E3 substrates, enabling the co-purification of the E3 together with the SUMOylation target, which will be later identified by mass spectrometry. **(B)** SATTs negative controls rationale. While ΔGG SATTs lack the C-terminal SUMO diGly motif, unable to conjugation to the substrate, catalytic-dead mutants prevent

interaction with the SUMO E2. **(C**) Heatmap of Z-scores for different SATT targets. Only HIS-SUMO1 and HIS-SUMO2$_{Q87R}$ targets are included.

Next, we constructed U2OS cells expressing HIS-SUMO1 or HIS-SUMO2$_{Q87R}$ in a constitutive manner or stably expressing the inducible E3 SATTs constructs indicated in Supplementary Figure 2A, including the ΔGG and catalytic-dead mutant negative controls. We induced the expression of the constructs for 24h, lysed the cells in denaturing conditions and purified the SATT conjugates from four independent biological repeats (five in the case of ZMIZ2) yielding a total of 171 samples. In order to avoid RNF4-mediated degradation of the SATTs due to auto-SUMOylation (**Figure 1**), and to increase the number of SUMOylation conjugates (8), the proteasome was inhibited for 5h with MG132.

Sample analysis by immunoblotting (**Supplementary Figure 2B-C**) showed that the expression levels of the SATTs were below or close to endogenous counterparts for every construct. Moreover, signal could be observed in a higher molecular weight smear for the wild type and catalytic-dead mutant construct for every SATT corresponding to E3-SUMO-target conjugates. This smear was absent in the ΔGG constructs. Consistently, the catalytic-dead SATT smears had different profiles than their wild type counterparts, indicating that the SUMO moieties in the mutant SATTs could still be used for conjugation by other endogenous E3s.

For the cell lines expressing either HIS-SUMO1 or HIS-SUMO2$_{Q87R}$ in a constitutive manner, mass spectrometry analysis of the samples enabled the identification of 244 SUMO1 targets and 1509 SUMO2 targets (**Supplementary Datasets 1-2**) after 5h of proteasome inhibition with MG132. Among the 244 SUMO1 targets, 171 could be considered a SUMO1 SATT conjugate for at least one SUMO E3. In the case of SUMO2 SATT conjugates, the numbers were 570 out of 1509 which were preferential or specific for the different E3s (**Figure 2C, Supplementary Datasets 3-4**). When compared with the biggest SUMO2 target study to date in U2OS cells (7) the number of SUMO2 SATT targets that had been identified as SUMO2 targets increased to 656 (**Supplementary Dataset 5**).

However, the SUMO proteome is highly diverse depending on the cell type and experimental conditions of study (7). Therefore, we considered the possibility that the SATTs could enable the modification and enrichment of SUMOylation substrates which are not constitutively modified in the cell line and condition of study or not detectable when a total SUMO proteome purification is performed. Accordingly, we also analyzed the mass spectrometry data from the SATTs in an unbiased manner. Namely, proteins not considered as His-SUMO targets were included when identified as SATT substrates (**Supplementary Figure 3, Supplementary Datasets 6-7**). This way, we could identify 302 extra putative substrates for SUMO1 SATTs, mainly for NSMCE2 and PIAS2 and 459 new putative substrates for SUMO2 SATTs, mainly for PIAS2 and LAZSUL.

The SATT Index.

Although the substrates we identified for each tested SUMO E3 were relatively specific for every E3 when comparing a wild type SATT with its ΔGG counterpart, all the substrates did not remain equally significant compared to their catalytic dead mutant counterpart (**Supplementary datasets 3,4,6,7**), indicating that, as previously shown for UbAITs (13, 15, 16), the SUMO moiety attached to the mutant SATT can also be conjugated to a substrate by another endogenous SUMO E3.

Therefore, we used the relation between the differences of the enrichment of a substrate for a specific E3 comparing both the ΔGG and the mutant counterpart to wildtype, which we termed SATT index:

$$SATT_i = \frac{[\log_2 SATT^{WT}] - [\log_2 SATT^{Mut}]}{[\log_2 SATT^{WT}] - [\log_2 SATT^{\Delta GG}]}$$

Values close to 1 and higher are considered very specific and values close to 0 and lower are considered not specific.

Different E3s have different preferences towards SUMO1 or SUMO2/3.

It could be argued that making a SUMO1 SATT with an E3 which normally catalyzes SUMO2/3 conjugation might force SUMO1 conjugation on SUMO2/3 substrate. Thus, we decided to investigate if SUMO E3s could discriminate substrate specificity depending on the SUMO type they were conjugating. First, we looked at the overlap between SUMO1 and SUMO2 substrates (**Figure 3A**). Overall, and similar to previous studies, where SUMO2 is the most abundant and important SUMO(10, 24) , while most of the identified SUMO1 targets (87%) can be also modified by SUMO2/3, only 14% of the SUMO2 substrates can be also modified by SUMO1. Next, we looked at the substrate preference and overlap for the different S-P RING E3s that had been investigated both for SUMO1 and SUMO2 SATTs (**Figure 3B**). On one side, in contrast to SUMO proteome data (**Figure 3A**), NSMCE2, PIAS1and PIAS2 data indicated a preference for SUMO1 modification. For PIAS1, only 9% of the SUMO1 substrates were also SUMO2/3 substrates (87% in SUMO proteome), and 45% of the SUMO2/3 were also substrates for SUMO1 (14% in SUMO proteome). This preference was milder for NSMCE2, where the numbers were 63% and 57%, respectively, and PIAS2, 69% for SUMO1 and 23% for SUMO2. For PIAS3 and PIAS4 the preference for SUMO2 modification was more acute than in total SUMO proteome analysis. For PIAS3 there was only one protein identified for SUMO1 conjugation. This protein was also found for SUMO2 conjugation (100%). For PIAS4, 93% of SUMO1 conjugates were also modified by SUMO2. These values indicate that PIAS3 and PIAS4 are mainly a SUMO2/3 E3 enzyme which is consistent with previous studies on SUMO specificity, and, consistently, SUMO1 ligase activity has also been reported to be higher for PIAS1 and PIAS2 (12).

Gene Ontology analysis.

Gene Ontology analysis for biological processes of the SUMOylation substrates for the different E3s indicated that different E3s are involved in different processes (**Figure 3C, Supplementary Dataset 8**). As expected, PIAS1, PIAS4, NSCME2 and ZNF451 substrates are enriched in Gene Ontology terms relative to genome biology (25-31), and PIAS3 substrates are enriched for maintenance of proteins at the nucleus (32-34). LAZSUL highest enrichment term was protein SUMOylation. ZMIZ2 substrates were not enriched for specific cellular processes. Interestingly, PIAS2 substrates are enriched for membrane translocation and ADP/ATP mitochondrial transport.

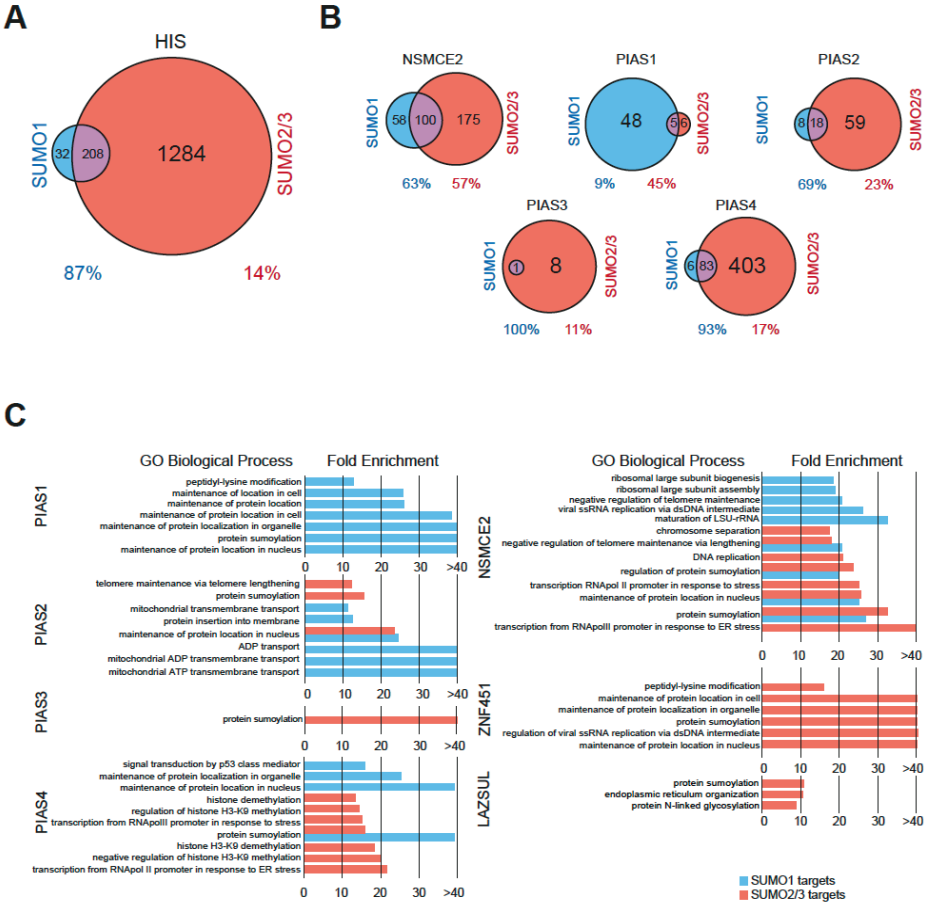


Figure 3. (A) Overlap between HIS-SUMO1 and HIS- SUMO2$_{Q87R}$ targets. **(B)** Overlap between SUMO1- and SUMO2-SATT substrates for the indicated E3s. **(C)** Gene ontology analysis for the SUMOylation substrates of the different E3 SATTs analyzed in this study.

PIAS4 and NSMCE2 make hybrid SUMO1-SUMO2/3 chains.

$SUMO2_{Q87R}$ SATTs leave a QQTGG remnant after tryptic digestion on acceptor lysines which can be identified by mass spectrometry-based proteomics (8). Although K11 is known to be the canonical site to make SUMO2/3 chains (35), several other SUMO2/3 sites at the endogenous level have been identified (6). Therefore, in addition to K11- SUMO2/3 chains, other chain types exist.

Mass spectrometry analysis of our samples enabled us to obtain MS/MS spectra in which the QQTGG remnant could be localized on SUMOs in an unambiguous manner (**Supplementary Figure 4**) and the intensity of these SUMOylation sites could be quantified (**Figure 4**). SUMO2/3 K11 chains, were found with every E3. As expected, no QQTGG-modified peptides were found in ΔGG SATTs samples. In contrast, signal for the modification with SUMO2/3 on K11 either on SUMO2 or SUMO3 could be detected in every SATT and, at less intense level, on K5. This included both wild type and catalytic-dead mutant SATTs. Importantly, only NSCME2 and PIAS4 wild type SATTs were able to modify SUMO1 with SUMO2/3 at K7, being completely dependent on the catalytic activity of the SATT. Similarly, SUMO3 K7 chains were also formed with NSCME2 and PIAS4, but only depending on the catalytic activity of NSMCE2. Finally, SUMO2/3/4 K32/33/33 chains were only detected for NSCME2 and were completely depending on NSMCE2 catalytic activity.

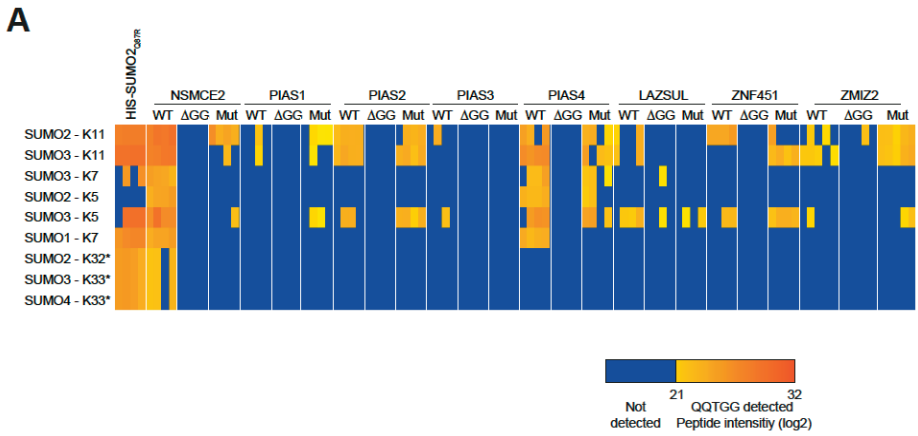


Figure 4 (A). Total peptide intensities of the different peptides corresponding to different acceptor lysines in SUMO1, SUMO2, SUMO3 and SUMO4 depicted in Supplementary Figure 4 in the different SUMO2Q87R SATT samples. * Indicates that corresponds to a peptide present in the three indicated SUMO types, thus, not enabling distinction.

SATTs complement analysis

Finally, we made statistical comparisons of proteins that were enriched or depleted in the wild type SATTs samples using all the other wild type SATTs as control. This was done both for considering and not considering exclusively the SUMO1 or SUMO2-3 substrates identified in Supplementary Datasets 1 and 2 (**Supplementary Datasets 9-12, Figure 5 and Supplementary Figure 5**).

Altogether, these analyses enable the identification of very high confidence E3-specific SUMOylation substrates in which wild type SATTs target proteins were statistically enriched when

compared to their ΔGG and catalytic dead mutant counterparts and compared to all the other wild type SATTs (**Supplementary Dataset 13**).

Polar SATTs – A user-friendly site to browse the dataset.

Most proteomic screens, including this one, usually consist of large spreadsheet datasets full of gene/protein names and values and comparisons. The interpretation of these datasets can be daunting for researchers from other disciplines. To overcome these potential outreach hurdles, we developed an online web app tool to browse the dataset, which is freely accessible (https://amsterdamstudygroup.shinyapps.io/polarVolcaNoseR_revised/). This tool enables users to select a protein of interest, and, if present in this study, will pop up in a polar plot in the sectors corresponding to the relevant E3s, indicating enrichment in terms of p-value and difference between wild type and either ΔGG SATTs, or all the other wild type SATTs as complement control. Additionally, for the ΔGG SATTs, how relatively specific the substrate is for the E3 in terms of SATT index is depicted with a color scale.

Moreover, the app can be used to customize the data visualization by enabling adjustment of the p-value and differences, choosing to hide the values that exceed the limits. The size of the datapoints can also be adjusted to facilitate visualization, and the resulting figure highlighting the desired substrates can be exported in .pdf or .png. A “Dark Theme” is also available.

Finally, if users prefer browsing in independent Volcano Plots instead of the default Polar Plot, that is also enabled. An example of visualization of different SUMOylation substrates for different E3s using the app is shown in Figure 6.

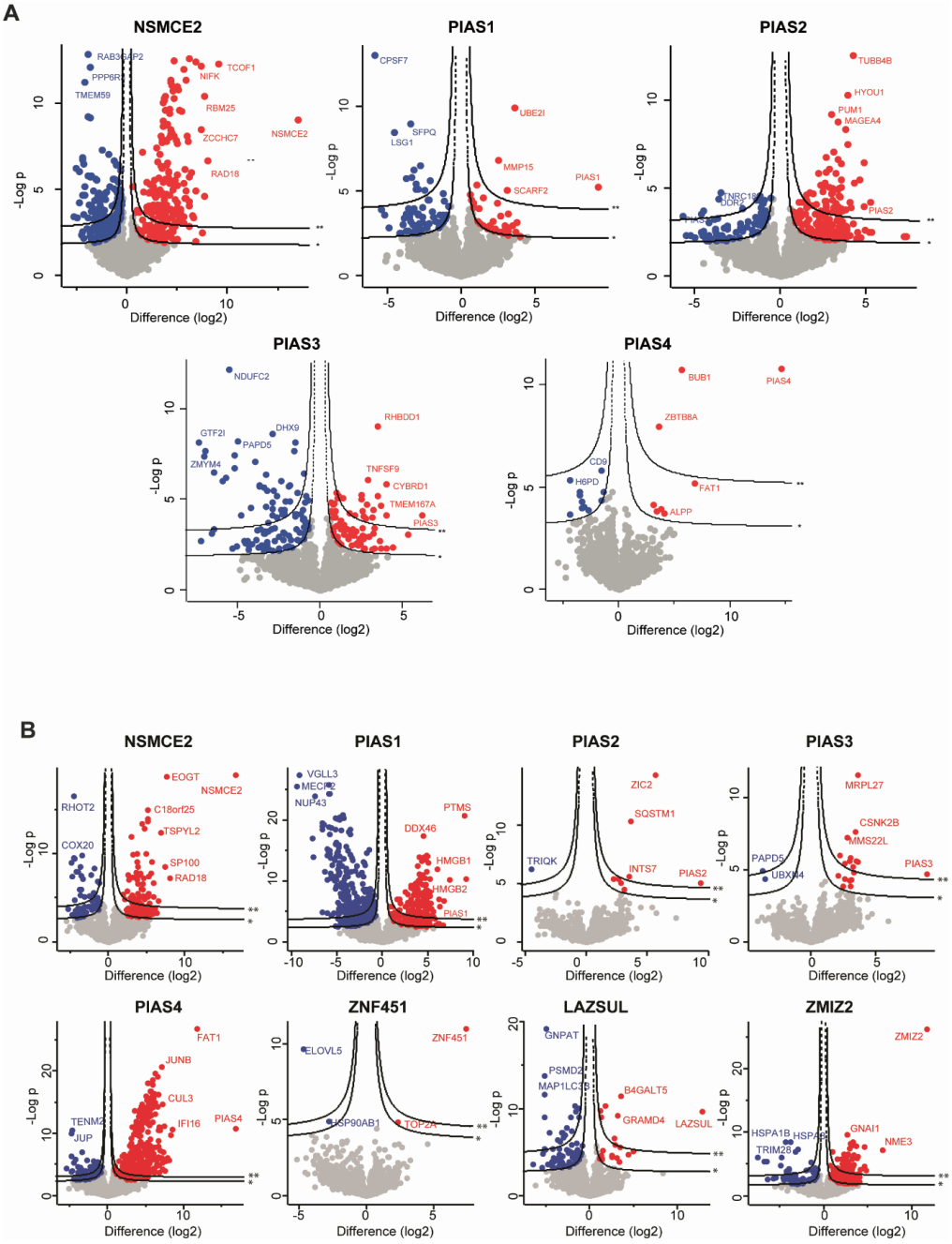


Figure 5. Multiple volcano plots (Hawaii plot) depicting statistical differences of different unbiased SUMO1- **(A)** or SUMO2- **(B)** SATTs targets using the values from the rest of wild type SATTs as complement negative control. Cutoffs represent a Pearson of 100 and an FDR=0.05 (*) or 0.01 (**) and an S0=0.1.

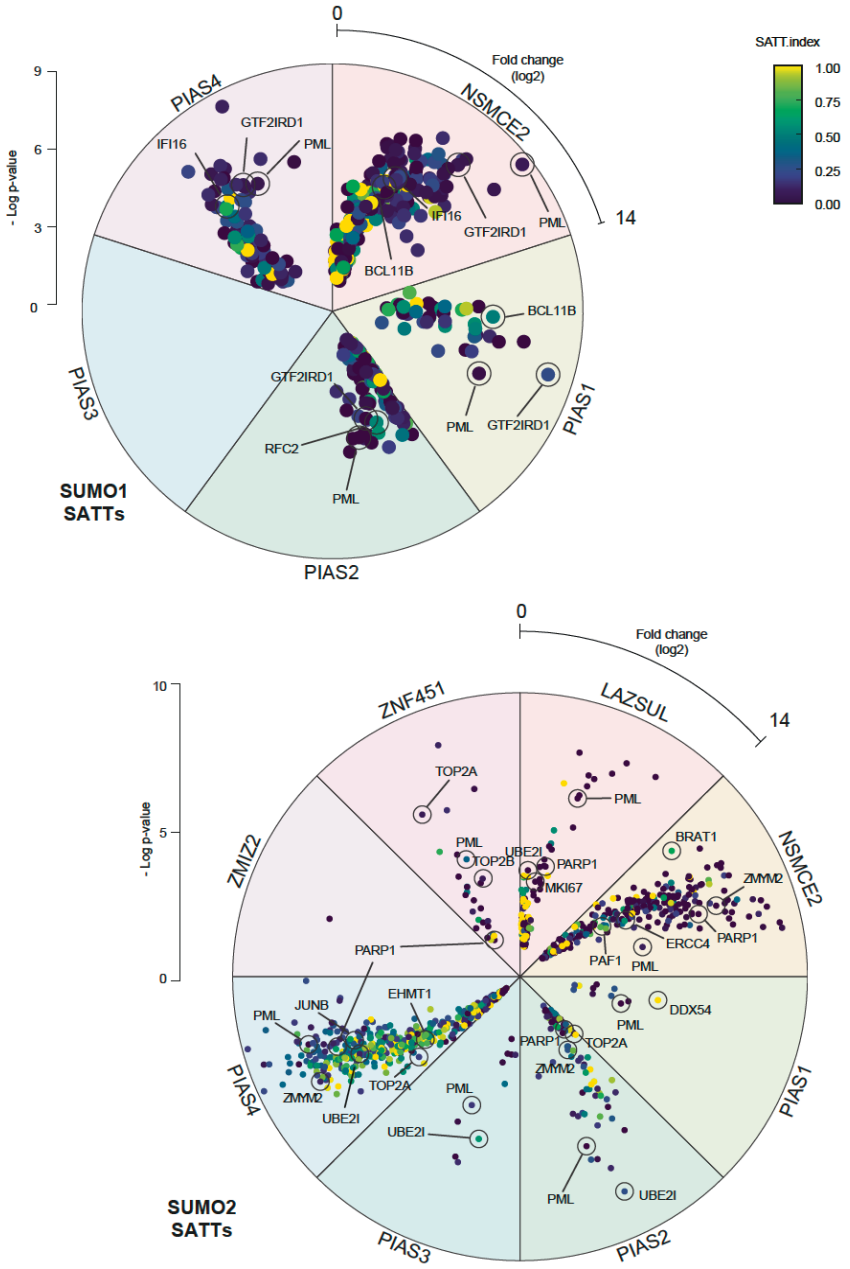


Figure 6. SATT Polar plots extracted from the polarVolcaNoseR web app. Some of the most prominent or specific substrates from different E3s are indicated. Depicted plots use ΔGG as control, includes the SATT index and only consider HIS-SUMO targets.

Orthogonal Validation

Finally, we performed orthogonal validation of the newly identified substrates for some E3 enzymes, including NSMCE2, LAZSUL and PIAS4.

For orthogonal validation of NSMCE2 substrates, we stably expressed or not HIS-SUMO2Q87R in parental and NSMCE2 knockout (NSMCE2-KO) cells (36) and rescued NSMCE2-KO with either wild type NSMCE2 or with a catalytic dead mutant $NSMCE2_{C185S/H187A}$ (37). Next, cells were cultured, treated with proteasome inhibitor MG132 as previously done for the SATTs and lysed, for subsequent Ni-NTA purification of the HIS-SUMO2Q87R proteome. Mass spectrometry based proteomics analysis was performed and the obtained data processed together with the NSCME2-SUMO2-SATT data (**Figure 7A-B, Supplementary dataset 14**). Among the statistically significant NSMCE2-SUMO2-SATT substrates compared to both their ΔGG and catalytic mutant counterparts and both in this analysis and analysis from Supplementary Dataset 4, 31 proteins were also considered as putative SUMOylation substrates for both U2OS and NSCME2-KO parental cells. Of those 31, 10 proteins were statistically reduced in their SUMOylation levels depending on NSMCE2 catalytic activity. Among the 21 proteins which SUMOylation status was not affected by the lack of NSMCE2 catalytic activity, 17 of them were also strong substrates for other E3 SATTs. Two proteins, namely, ERCC4 and PAF1, were strong substrates for NSMCE2-SUMO2-SATT and had their SUMOylation levels affected by lack of NSMCE2 catalytic activity and were not strong E3 substrates for other SATTs. We decided to analyze these two candidates by immunoblotting (**Figure 7C**). Consistently, a big reduction in the SUMOylation signal could be observed when the catalytic activity of NSMCE2 was not present.

Additionally, we generated LAZSUL knockout U2OS cells (**Supplementary Figure 6**), introduced HIS-SUMO2 and validated the proliferation marker protein Ki-67 as a target which SUMOylation was considerably reduced when LAZSUL was not present (**Figure 7D**).

In an alternative approach, we performed an shRNA-mediated knockdown of PIAS4 in parental, or HIS-SUMO2Q87R U2OS cells. Analysis by immunoblotting of the HIS-SUMOQ87R proteome validated strong PIAS4-S2-SATT targets with high SATT index value, which had not been previously described in the literature as the Histone-lysine N-methyltransferase EHMT1, or the transcription factor jun-B (**Figure 7E**).

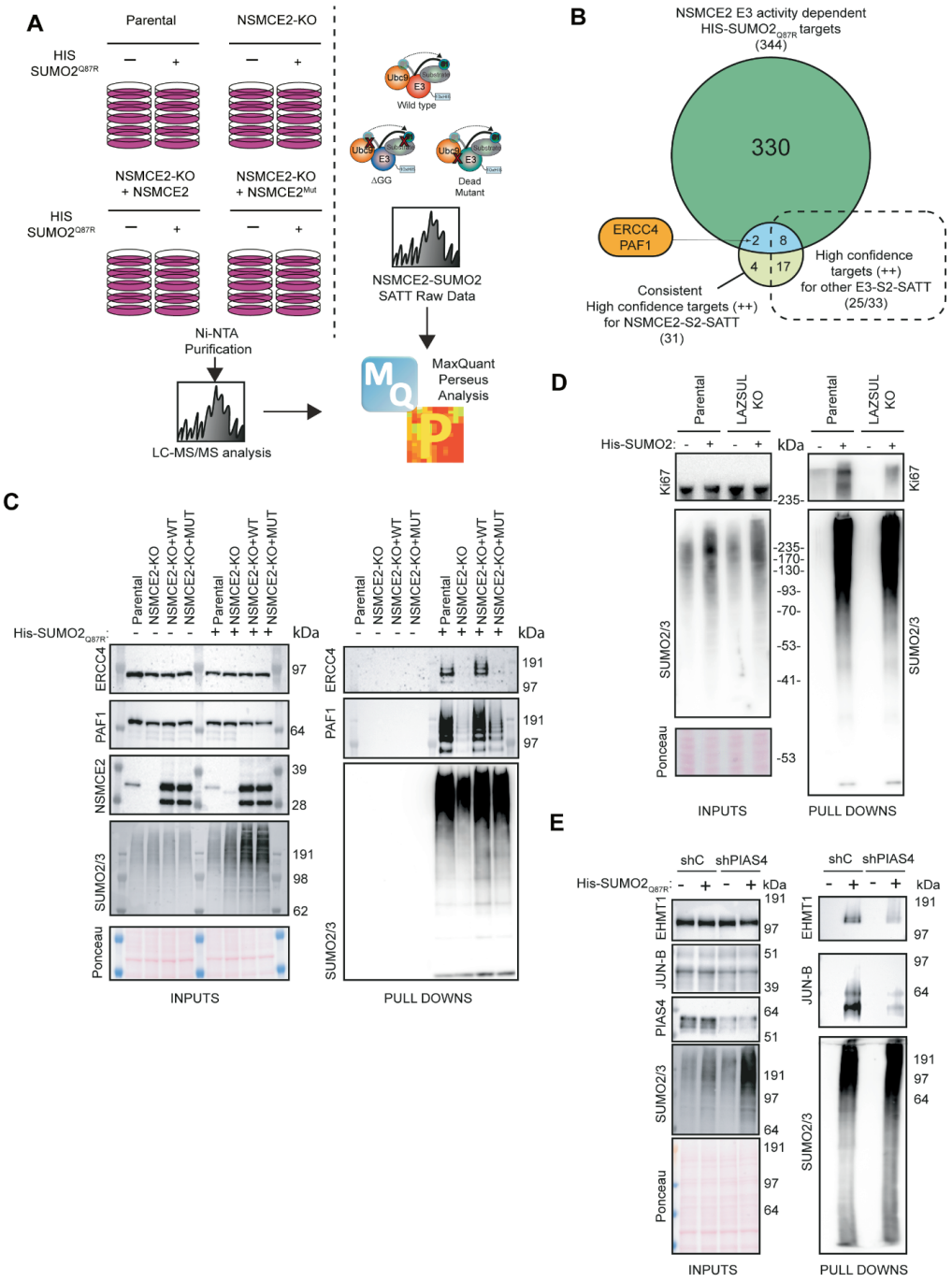


Figure 7. Orthogonal Validation. **(A)** Experimental setup. Parental and NSMCE2-KO U2OS cells rescued or not with either wild type NSMCE2 or a catalytic dead mutant and expressing or not HIS-SUMO2Q87R were cultured, incubated for 5h with MG132, lysed and analyzed by LC-MS/MS. Next, data was processed together with the NSMCE2-SUMO2-SATT data by MaxQuant and Perseus. **(B)** Venn Diagram depicting

overlap between NSMCE2-SUMO2-SATT substrates, HIS-SUMO2Q87R substrates affected by lack of NSMCE2 catalytic activity, and NSMCE2-SUMO2-SATT substrates shared with other E3-SATTs. (C-E) Immunoblot analysis of HIS-SUMO2 substrates decreasing upon NSMCE2 catalytic activity **(C)**, LAZSUL knockout **(D)** or PIAS4 knockdown **(E)**.

DISCUSSION

This resource identifies 427 SUMO1 and 961 SUMO E3 targets in an E3-specific manner. Nevertheless, while the number of SUMOylation substrates that have been identified for SUMO1 and SUMO2 in U2OS cells is higher (7, 9, 10) **(Supplementary Datasets 1-2**), proteins that have not been previously described as SUMOylation substrates upon proteasome inhibition with MG132, were identified as E3 substrates for SUMO conjugation in this work. It is noteworthy, that the screens performed in this project only comprised a single condition of 5h of proteasome inhibition with MG132 in a single cell line, U2OS, for 8 different SUMO ligases.

The results obtained from this screen also corroborate previous observations on protein arrays regarding E3 preferences for SUMO1 or SUMO2/3(12). PIAS3 and PIAS4 had been proposed to have a big preference for SUMO2. Consistently, for PIAS3, enrichments are much higher in SUMO2 SATTs than in SUMO1 SATTs (**Figure 3B**), and for PIAS4, 93% of SUMO1 substrates are shared with SUMO2. PIAS1 was observed to have a more balanced preference, in line with our SATT screen. For PIAS2 and NSMCE2 there is no data available from protein arrays to make this comparison, but they show a more balanced preference towards SUMO1.

Additionally, substrates that are highly SUMOylated and shared by every E3, such as PML, RANGAP, RNF216 or ZBTB33 have a very low SATT index for every E3, which indicates that either are redundant substrates, can be modified by different E3s or require no E3 at all for SUMOylation (**Supplementary Datasets 1-2**). Moreover, some substrates with high SATTi values for specific E3s have been previously independently identified substrates in other laboratories for the relevant E3s. Namely, PARP1 (38), ZMYM2 (39), and TOP2a (27, 30) for PIAS4 and SUMO2/3. (**Figure 6, Supplementary Datasets 4,7**). SMC5 for NSMCE2 and SUMO1 (26) (Supplementary Datasets 3,6)

SUMO2/3 is not recycled at the proteasome

Whereas the classical model suggests that SUMO2/3 is deconjugated and recycled at the proteasome (40), we have observed that when overexpressing PIAS1, ZNF451 or LAZSUL, SUMO2/3 is depleted from the nucleus in an RNF4-dependent manner (**Figure 1**). This indicates that SUMO2/3 moieties attached to STUbL targets are also degraded. Whether the proteasome discriminates between mixed SUMO2/3-Ubiquitin chains on substrates and substrates that are co-modified independently by SUMO2/3 and Ubiquitin chains requires further investigation. We previously showed that the oncogene c-Myc, which SUMOylated form highly accumulates upon proteasome inhibition and is an RNF4 target (8, 15, 41), is SUMOylated and ubiquitinated on different residues in an RNF4-dependent manner. Based on these results, we favor the hypothesis that independent ubiquitination of the SUMOylated substrate is sufficient for the degradation of SUMO moieties attached to the substrate without the need for mixed chains. Nevertheless, we also found that mixed SUMO/ubiquitin polymers are efficiently stabilized by proteasome inhibitors, indicating that these mixed polymers also constitute an efficient proteasomal degradation signal (5).

In contrast to GFP-PIAS1, -ZNF451, -LAZSUL and, partially, -PIAS3, which overexpression caused nuclear SUMO2/3 depletion, GFP-PIAS4 overexpression was linked to a slight increase in nuclear SUMO2/3 levels (**Figure 1**). Interestingly, PIAS4 and NSMCE2 were the only E3s capable of assembling SUMO2/3-SUMO1 mixed chains (**Figure 4**). We therefore hypothesize that SUMO2/3-SUMO1 mixed chains are poor substrates for RNF4. Similarly, SUMO1-capped SUMO2/3 chains are poor substrates for RNF4 but efficient substrates for RNF111/Arkadia, another STUbL (42).

Perspective and implications

SUMOylation of proteins occur in response to many different types of cellular stresses, such as DNA damage or replication stress and heat shock among others (31, 43). The E3 that modifies a specific target may vary depending on the scenario, and SATTs should be screened for specific targets at specific conditions. One E3 can be specific for the SUMOylation of a protein in a certain context, and another in a different context. In this regard, PIAS4 and NSCME2 shared many SUMOylation substrates from the DNA damage response. We speculate that, while NSMCE2 SUMOylates these substrates in response to DNA damage in the context of DNA replication as part of the SMC5/6 complex (44), PIAS4 SUMOylates these substrates in a replication-independent manner.

Moreover, altering the expression levels of different components of the SUMO pathway affects the SUMOylation landscape (7, 11, 45-47). Although here we expressed the SATTs at close-to-endogenous or below levels, this might affect both the SUMOylation status of target proteins, E3 specificity, or SUMO type modification, as proposed in Figure 3A.

Previously, it was shown that both SUMO1 and SUMO2 were recruited to sites of DNA damage, with SUMO1 recruitment depending on PIAS4 (25). Accordingly, DNA damage repair pathways are significantly enriched among PIAS4 substrates for SUMO1 **(Supplementary Dataset 8**). However, we did not detect affinity of DNA Damage Response-related proteins for SUMO1 moieties (10). Which might indicate that SUMO1 moieties incorporated at DNA damage sites in a PIAS4-dependent manner correspond to hybrid SUMO1-SUMO2/3 chains. Our results indicate that different types of mixed SUMO polymers and mixed SUMO/ubiquitin polymers constitute differential signals.

Nonetheless, results from this screen might open new lines of investigation. Gene ontology analysis revealed that PIAS2 substrates are significantly enriched for serine import into mitochondrion (**Figure 3B**) as biological function, which deregulation has been very recently described as causative for Parkinson disease (48). Interestingly, PIAS2 has also been recently linked to Parkinsonism (49). This suggest a potential role for the identified PIAS2 substrates in the development of this neurological disorder. Future mining of the resource presented here could improve our understanding of the biological functions of the different SUMO E3 ligases.

MATERIALS AND METHODS

Antibodies

Antibodies are listed in Supplementary Table 1 with working dilutions. For the generation of the Rabbit anti-LAZSUL antibody, recombinant ZNF451-3 (LAZSUL) was produced as in (23) and supplied to Cambridge Research Biochemicals (United Kingdom) for antibody production. Two rabbits were injected and the antibody affinity-purified for LAZSUL.

Generation of SATT toolbox

For the generation of the SATTs plasmids AgeI-10HIS-SUMO1-XmaI, AgeI-10HIS-SUMO1ΔGG-XmaI, AgeI-10HIS-SUMO2Q87R-XmaI, AgeI-10HIS-SUMO1ΔGGQ87R-XmaI restriction fragments from PCR products amplified with primers FW-AgeI-10HIS-SUMO1, FW-AgeI-10HIS-SUMO2, RV-XmaI-SUMO1, RV-XmaI-SUMO1noGlyGly, RV-XmaI-SUMO2-Q87R and RV-XmaI-SUMO2-Q87R-noGlyGly, were cloned into AgeI-SpeI sites of pCW57.1-nonStop plasmid (16). The N-terminal 10xHIS tag was cloned as previously done for the H-TULIP2 plasmids (15). Primer sequences are listed in Table T2.

Generation of lentiviral plasmids

SUMO1 and SUMO2$_{Q87R}$ SATT plasmids containing a SUMO E3 ligase of interest were generated using Gateway® cloning LR reaction (Thermo Fisher Scientific). LR reactions were performed using a donor plasmid containing an E3 enzyme cDNA without stop codon and a SATT plasmid as destination vector. We used different donor plasmids: pENTR223-PIAS1, pDNOR221-PIAS2, pENTR223-PIAS3, pDNOR221-PIAS4, pENTR223- NSMCE2, and pENTR233-ZMIZ2 were obtained from DNASU repository with the following IDs: HsCD00505402, HsCD00042133, HsCD00514170, HsCD00041383, HsCD00287670 and HsCD00505806 respectively. pDNOR207-ZNF451-1 (isoform 1), pDNOR207-ZNF451-3 (LAZSUL) and pDNOR207-GPF-LAZSUL were generated using the Gateway® cloning BP reaction (Thermo Fisher Scientific) upon cDNA amplification using BP-tailed primers and pDNOR207 as donor vector. Catalytic dead mutants of each SUMO E3 ligase were generated by site-directed mutagenesis on donor plasmids and subsequent LR reaction into SATT plasmids. Primer sequences are listed in Table T2.

Stable-inducible GFP-LAZSUL construct was generated by LR Gateway cloning between pDNOR207-GFP-ZNF451-3 and pCW57.1 (Addgene#41393). NSMCE2 rescued constructs were generated by LR between either pENTR223- NSMCE2-WT or pENTR223- NSMCE2C185S-H187A and pLX303 (Addgene#25897).

Transfection of GFP-E3 constructs

For the transient transfection experiments of GFP-tagged SUMO E3 ligases in Figure 1 and Supplementary Figure 1, 1x10^4 U2OS cells were transferred to 6-well plates containing 18 mm coverslips and left to attach overnight. The next day, 300 µL of transfection mixture consisting of 150 mM NaCl containing 1 µg of plasmid DNA and 6 µg of Polyethyleneimine were added to the cells. 24h after transfection, culture medium was replaced for fresh medium. Cells were fixed with 1% paraformaldehyde 48h after transfection for immunofluorescence analysis.

Generation of LAZSUL-KO cell lines

Three different gRNAs targeting LAZSUL exon 4 start and LAP2α domain (**table T2**) were cloned, independently, into a Cas9-GFP-containing pX458 backbone plasmid (AddGene #48138). U2OS cells were seeded in 15 cm diameter plates at 10% confluency and left to attach overnight. The next day, 2 mL of transfection mixture consisting of 150 mM NaCl containing 5 µg DNA of each plasmid and 100 µg of Polyethylenimine were added to the cells. Transfection medium was replaced by fresh culture medium after 24 h. 48 h after transfection, cells were GFP FACS-sorted by FACSAria III (BD Biosciences) and seeded for monoclonal expansion.

Selection of positive clones was performed by genomic PCR and immunoblotting (Supplementary Figure 6). Two primers targeting the LAP2α coding sequence were employed for clones validation by genomic PCR (**table T2**). 6 out of 30 clones were found positive with a 972 bp deletion in the LAP2α coding sequence (**Supplementary Figure 6A**). Parental and clone 27 LAZSUL-KO cell lines were subjected to immunoblotting against an in-house produced LAZSUL antibody (**table T1**) (**Supplementary Figure 6B**).

siRNA transfection for RNF4 knockdown

siRNA-mediated knockdowns were performed as previously described (16). DharmaFect 1 Transfection Reagent (GE Lifesciences) was used, according to the manufacturer's instructions, using on-target plus RNF4 siRNAs (J-006557-08) and the non-targeted control siGENOME non-Targeting siRNA #1 (GE Lifesciences).

Lentivirus production

293T HEK cells were seeded at 30% confluency in a T175 flask containing 16 mL of DMEM + 10% FBS. The following day, a 2 mL transfection mixture containing lentiviral packaging plasmids 7.5 µg pMD2.G (#12259, Addgene), 11.4 µg pMDLg-RRE (#12251, Addgene), 5.4 µg pRSV-REV (#12253, Addgene) and 13.7 µg SATT plasmid with 114 µL of 1 mg/mL Polyethyleneimine (PEI) was prepared in 150 mM NaCl. After vortexing, solutions were incubated 10 min at room temperature before adding to the HEK cells. The day after transfection, culture medium was changed by fresh DMEM/FBS/Pen/Strep. three days after transfection, lentiviral suspension was harvested by filtering through a 0.45 µm syringe filter (PN4184, Pall Corporation). Lentiviral particle concentration was determined using the HIV Type 1 p24 antigen ELISA Kit (ZeptoMetrix Corporation).

Generation of SATT, GFP-LAZSUL and NSMCE2 rescued cell lines

U2OS cells were seeded in a 15 cm diameter plates at 10% confluency with DMEM + 10% FBS. The next day, cell culture medium was replaced with either lentiviral SATT, GFP-LAZSUL or NSMCE2 rescued constructs containing medium and polybrene 8 µg/mL. Lentiviral containing medium was replaced by fresh DMEM/FBS/Pen/Strep culture medium after 24h. After two days in fresh medium, 3 µg/mL puromycin was added to the medium for selection of SATT positive clones.

shPIAS4 transduction for PIAS4 knockdown

Transductions were performed in DMEM containing 8 µg/mL polybrene either with PIAS4 or a control non-targeting shRNA. Cells were infected with a multiplicity of infection (MOI) of 3 with third generation lentiviruses encoding shRNA. DMEM containing virus was replaced after 24 hours of infection. Cells were harvested and lysed 3 days post infection. Plasmids used for shRNA mediated knockdown were derived from the MISSION ® shRNA library (Sigma) with number TRCN0000004115 (shPIAS4), and SHC-002 (shControl).

Cell culture

293T HEK, U2OS and RPE1 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10 % Fetal Bovine Serum (FBS) and 100 U/mL penicillin 100 µg/mL

streptomycin at 37°C and 5% CO2 unless specified. Cells were regularly tested for mycoplasma contamination.

Anti-SUMO1 and SUMO2/3 immunostaining

Cells were grown on 9 mm coverslips and fixed with 1% paraformaldehyde (PFA), 0.3% Triton X-100 for 20 min at RT. A second round was performed with 1% PFA, 0.3% Triton X-100 and 0.5% methanol for 20 minutes at Room Temperature (RT). Next, cells were washed three times with PBS and then blocked for 30 minutes with 0.5% blocking reagent (Roche) in 0.1 M Tris, pH 7.5 and 0.15 M NaCl (TNB). Cell were then incubated with either anti-SUMO1 or anti-SUMO2/3 antibody in TNB for one hour. Coverslips were washed five times with PBS and incubated with the secondary antibody (Goat anti-mouse coupled to Alexa-594) in TNB for one hour. Next, coverslips were washed five times with PBS- and mounted onto a microscopy slide using citifluor/DAPI solution (500 ng/mL). Immunofluorescence image analysis was performed using the FiJi – ImageJ distribution(50).

Generation of His-SUMO1 and His-SUMO2 Q87R U2OS cell lines

U2OS cells were infected using a bicistronic lentivirus encoding either a 10xHis-SUMO1-IRES-GFP or a 10xHis-SUMO2Q87R-IRES-GFP separated by an IRES, which was modified from previously described 10xHis-SUMO2-WT (51). Following infection, U2OS cells were sorted in an FACSAria III (BD Biosciences) for low GFP levels.

Purification of 10xHis-SUMO1, 10xHis-SUMO2 Q87R and SATT conjugates

Following the TULIP2 methodology (15), five 15 cm diameter plates of U2OS control cells or expressing either 10xHis-SUMO1, 10xHis-SUMO2Q87R or a particular SATT, were grown up to 60% to 80% confluence. Expression of SATTs constructs was induced with 1 µg/mL doxycycline once 60-80% confluency was reached. 24 h after doxycycline induction, cells were treated with proteasome inhibitor MG132 (Sigma Aldrich) at 10 µM for 5 h. After proteasome inhibition, cells were washed twice with ice-cold PBS and scraped. Cells were spun down and collected in 2 mL ice-cold PBS, 100 µL of sample was taken as input and lysed in 200 µL SNTBS buffer (2% SDS, 1% NP-40, 50mM TRIS pH 7.5, 150 mM NaCl). After additional centrifugation, cells were lysed in 10 mL Guanidinium buffer (6M guanidine-HCl, 0.1M Sodium Phosphate, 10mM TRIS, pH 7.8) and snap frozen in liquid nitrogen.

After thawing, lysates were homogenized at room temperature by sonication at 80% amplitude during 5 s using a tip sonicator (Q125 Sonicator, QSonica, Newtown, USA). Sonication was performed twice. Subsequently, protein concentration was determined by BiCinchoninic Acid (BCA) Protein Assay Reagent (Thermo Scientific). Total protein in cell lysates was equalized accordingly. After equalization, cell lysates were supplemented with 5 mM β-mercaptoethanol and 50 mM Imidazole pH 8.0. 100 µL of dry nickel-nitrilotriacetic acid-agarose (Ni-NTA) beads (QIAGEN), were equilibrated with Guanidinium buffer supplemented with 5 mM β-mercaptoethanol and 50 mM Imidazole pH 8.0. Equilibrated Ni-NTA beads were added to the cell lysates and incubated overnight at 4°C under rotation.

After lysate-bead incubation, Ni-NTA beads were spun down and transferred with Wash Buffer 1 (6 M Guanidine-HCl, 0.1M Sodium Phosphate, 10 mM Tris, 10 mM Imidazole, 5 mM β-mercaptoethanol, 0.2 % Triton X-100, pH 7.8) to an Eppendorf LoBind tube (Eppendorf). After

mixing and spinning down again, the beads were moved to a new LoBind tube with Wash buffer 2 (8 M Urea, 0.1M Sodium Phosphate, 10 mM Tris, 10 mM imidazole, 5mM β-mercaptoethanol, pH 8). This procedure was repeated with Wash buffer 3 (8 M urea, 0.1M Sodium Phosphate, 10 mM TRIS, 10 mM imidazole, 5 mM β-mercaptoethanol, pH 6.3). Ultimately, beads were washed twice with Wash buffer 4 (8 M urea, 0.1M Sodium Phosphate, 10 mM TRIS, 5 mM β-mercaptoethanol, pH 6.3). When washing with wash buffer 3 and 4, beads were allowed to equilibrate with the buffer for 15 min under rotation.

For the analysis by immunoblotting the beads were boiled at 99 oC for 10 min with 2x LDS loading buffer.

Trypsin digestion of SATT-purified conjugates

After the final wash with Wash buffer 4, Ni-NTA beads were resuspended in 7 M urea, 0.1 M NaH2PO4/Na2HPO4, 0.01 M Tris/HCl, pH 7 and digested with 500 ng recombinant Lys-C (Promega) at RT while shaking at 1,400 rpm. After 5h with Lys-C, urea buffer was diluted to <2M by adding 50 mM ABC. A second digestion was performed o/n at 37°C while shaking at 1,400 rpm using 500 ng of sequencing grade modified trypsin (Promega). Trypsin digested peptides were separated from Ni-NTA beads by filtering through a 0.45 µm filter Ultrafree-MC-HV spin column (Merck-Millipore).

Mass Spectrometry sample preparation

Digested peptides were acidified by adding 2% TriFlourAcetic (TFA) acid. Subsequently, peptides were desalted and concentrated on triple-disc C18 Stage-tips as previously described (52). Stage-tips were in-house assembled using 200 µL micro pipet tips and C18 matrix. Stage-tips were activated by passing through 100 µL of methanol. Subsequently 100 µL of Buffer B (80% acetonitrile, 0.1% formic acid), 100 µL of Buffer A (0.1% formic acid), the acidified peptide sample, and two times 100 µL Buffer A were passed through the Stage-tip. Elution was performed in 50 µL of 32.5% acetonitrile, 0.1% formic acid.

Samples were vacuum dried using a SpeedVac RC10.10 (Jouan, France) and stored at -20°C. Prior to mass spectrometry analysis, samples were reconstituted in 10 µL 0.1% formic acid and transferred to autoload vials.

LC-MS/MS analysis

All the experiments were analyzed by on-line C18 nanoHPLC MS/MS with a system consisting of an Ultimate3000nano gradient HPLC system (Thermo, Bremen, Germany), and an Exploris480 mass spectrometer (Thermo, Bremen, Germany). Samples were injected onto a cartridge precolumn (300 µm × 5 mm, C18 PepMap, 5 µm, 100 A) with a flow of 10 µl/min for 3 minutes (Thermo, Bremen, Germany) and eluted via a homemade analytical nano-HPLC column (50 cm × 75 µm; Reprosil-Pur C18-AQ 1.9 µm, 120 A) (Dr. Maisch, Ammerbuch, Germany). The gradient was run from 2% to 38% solvent B (80% acetonitrile, 0.1% formic acid) in 120 min. The nano-HPLC column was drawn to a tip of ~10 µm and acted as the electrospray needle of the MS source. The temperature of the nano-HPLC column was set to 50°C (Sonation GmbH, Biberach, Germany). The mass spectrometer was operated in data-dependent MS/MS mode for a cycle time of 3 seconds, with a HCD collision energy at 28 V and recording of the MS2 spectrum in the orbitrap, with a quadrupole isolation width of 1.2 Da. In the master scan (MS1) the resolution was 120,000, the

scan range 350-1600, at a standard AGC target with maximum fill time of 50 ms. A lock mass correction on the background ion m/z=445.12 was used. Precursors were dynamically excluded after n=1 with an exclusion duration of 45 s, and with a precursor range of 10 ppm. Charge states 2-5 were included. For MS2 the scan range mode was set to automated, and the MS2 scan resolution was 30,000 at a normalized AGC target of 100% with a maximum fill time of 60 ms.

Mass Spectrometry data analysis

All raw data was analyzed using MaxQuant (version 1.6.14) as previously described (53). We performed two searches, first one analyzing the U2OS samples, HIS-SUMO1 and SUMO1 SATTs, and the other one analyzing U2OS, HIS-SUMO2Q87R and SUMO2Q87R SATTs. We performed the search against an in silico digested UniProt reference proteome for Homo sapiens including canonical and isoform sequences (5th July 2021). Database searches were performed according to standard settings with the following modifications: Digestion with Trypsin/P was used, allowing 4 missed cleavages. Oxidation (M), Acetyl (Protein N-term), Phospho (S, T) and, in the SUMO2Q87R SATTs analysis, also QQTGG (K) (for SUMOylation sites) were allowed as variable modifications with a maximum number of 3. Carbamidomethyl (C) was disabled for SATTs analysis as a fixed modification. Label-Free Quantification was enabled, not allowing Fast LFQ. At least 2 peptides needed to be identified to calculate LFQ for a protein. All peptides were used for protein quantification.

Output from the analysis in MaxQuant was further processed in the Perseus computational platform version 1.6.14 (54) for statistical analysis. LFQ values were log2 transformed and contaminants, proteins identified by site and reverse peptides were excluded from the analysis. Next, samples were separated in two packages. The first package consisted of U2OS control samples and HIS-SUMO samples whereas the second package consisted of the SATT samples. The first package was analyzed to determine which proteins could be statistically considered a HIS-SUMO target. Proteins that were not identified in at least four replicates of at least one condition were removed. Then, missing values were imputed from a normal distribution with a width of 0.3 and a downshift of 2.5. Which resulted in a percentage of imputed valid values of 29.7% (Supplementary Dataset 3), 39,7 (Supplementary Dataset 4), 7.6% (Supplementary Dataset 6) and 11.8% (Supplementary Dataset 7). We performed t-test corrected with an FDR=0.05 and S0=0.1. Statistically enriched proteins were considered HIS-SUMO substrates (Supplementary Datasets 1-2). Next, we proceeded only considering proteins that were HIS-SUMO substrates and merge the table with package 2 samples (left sided), so only proteins which were HIS-SUMO substrates remained in the datasets. For the unbiased analysis, this step was omitted. Each SATT set was independently analyzed as performed with U2OS and HIS-SUMO samples. Finally, all the analyses were merged and exported and further processed in Microsoft Excel 365 for comprehensive data visualization and calculation of SATT indexes. For the heatmaps, the ΔGG and Mut samples were removed and Z-score was calculated for heatmap visualization.

Similarly, for the NSMCE2-KO orthogonal validation, LFQ values were log2 transformed and contaminants, proteins identified by site and reverse peptides were excluded from the analysis. Proteins that were not identified in at least four replicates of at least one condition were removed. Then, missing values were imputed from a normal distribution with a width of 0.3 and a downshift

of 2.5. Groups comparisons were performed by t-test with an FDR=0.05 and S0=0.1 and data exported into MS Excel 365 for comprehensive data browsing.

Gene Ontology Analysis

Gene Ontology analyses from the SATT substrates were performed using the PANTHER overrepresentation test (released 20221013) from the Gene Ontology Consortium (55). The Gene Ontology Database used was released on 2022-07-11 and the test was a Fisher test with Bonferroni correction. The whole human proteome was used as reference for comparison.

Electrophoresis and immunoblotting

Samples were separated on Novex 4-12% gradient gels (Thermo Fisher Scientific) using NuPAGE® MOPS SDS running buffer (50mM MOPS, 50mM Tris-base, 0.1% SDS, 1mM EDTA pH 7.7) and transferred onto Amersham Protran Premium 0.45 NC Nitrocellulose blotting membranes (GE Healthcare) using a Bolt Mini-Gel system (Thermo Fisher Scientific), which was used for both the gel electrophoresis and the protein transfer to the membrane according to vendor instructions. Membranes were stained with Ponceau-S (Sigma Aldrich) to determine total amount of protein loaded. Next, membranes were blocked with blocking solution (8% Elk milk, 0.1% Tween-20 in PBS) for 1h prior to primary antibody incubation. Chemiluminescence reaction was initiated with Western Bright Quantum Western blotting detection kit (Advansta – Isogen) and measured in a ChemiDocTM imaging system (BIO-RAD, Hercules, CA, USA).

In vitro SUMOylation assay

Recombinant SUMOylation machinery (56), including 300 ng E1, 0/100/200 ng E2, 4 µg SUMO1/SUMO2 and 0/300 ng E3 was incubated for 3h at 37ºC in in vitro SUMOylation buffer (50mM Tris pH 7.5, 2 mM ZnCl2 with Protease Inhibitors cocktail without EDTA (11836170001, Roche) and Energy Regeneration Solution 1:25 (B-10, Boston Biochem) in a total volume of 25 µL. SUMO chain formation was analyzed by immunoblotting.

Data Representation

Super plots were constructed using SuperPlotsOfData (57).

For the SATT polar plots, a web app for the display of multiple volcano plots side-by-side, named polarVolcaNoseR, was made with R/Shiny.

The code was written using R (https://www.r-project.org) and Rstudio (https://www.rstudio.com). To run the app, several freely available packages are required: shiny, ggplot2, magrittr, dplyr, ggrepel, htmlwidgets, ggiraph, glue, scales. The web app is freely accessible at: https://amsterdamstudygroup.shinyapps.io/PolaRVolcaNoseR/ and the code is available at github: https://github.com/ScienceParkStudyGroup/polarVolcaNoseR. In the default ‘polar’ representation, the volcano plots are plotted in a circle, where the radius depicts the -Log10(p-value) and the circumference reflects the positive Log2(Fold-change). Labels of proteins can be added by point-and-click and the data of individual dots are displayed when the cursor hovers over a datapoint. A customized, interactive plot can be exported as a HTML file.

ACKNOWLEDGEMENTS

Reagents and Funding statement

Authors would like to thank Dr. Geert Hamer for sharing Parental and NSMCE2-KO U2OS cell lines This work was supported by a Young Investigator Grant from the Dutch Cancer Society (KWF-KIG 11367/2017-2) the EMERGIA program from the Andalusian Regional Government, Spain (Junta de Andalucia – EMERGIA20_00276), the Spanish Ministry of Science and Innovation, the Spanish Research Agency, and the European Regional Development Fund (Proyecto PID2021-122361NA-I00 financiado por MCIN /AEI /10.13039/501100011033 / FEDER, UE) to RG-P. Work in the laboratory of A.C.O.V. has been supported by the European Research Council (ERC; grant 310913) and the Dutch Research Council (NWO; grant 724.016.003).

Competing interest

The authors declare that they have no competing interests.

Author contributions

RG-P and ACOV conceived the project. DSL, CvdM and RG-P performed the GFP-tagged E3 overexpression experiments. DS-L prepared all the resource reagents and samples. AdR, AO and PvV injected mass spectrometry samples. DS-L and RG-P analyzed mass spectrometry data. EG made recombinant ZMIZ2. RG-P and EN performed in vitro SUMOylation assays. DSL, NSJ and RG-P preformed orthogonal validation. EN raised the anti-LAZSUL antibody, JG programmed the polarVolcanoseR web App. AP supervised EN. RG-P supervised DS-L and CvdM. ACOV supervised EG. PvV supervised AdR and AO. RG-P lead the project and wrote the manuscript together with DS-L and input from all the authors.

Data availability statement

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (58) with the dataset identifiers PXD038326.

SUPLEMENTARY FIGURES AND DATASET LEGENDS

Supplementary Dataset 1. His-SUMO1 targets dataset. Spreadsheet including total identified SUMO1 targets, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 2. His-SUMO2 targets dataset. Spreadsheet including total identified SUMO2 targets, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 3. SUMO1 SATTs among HIS-SUMO1 targets. Spreadsheet including identified SUMO1 targets for the different SATTs among SUMO1 targets identified in Supplementary Dataset 1, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 4. SUMO2 SATTs among HIS-SUMO2 targets. Spreadsheet including identified SUMO2 targets for the different SATTs among SUMO2 targets identified in Supplementary Dataset 2, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 5. SUMO2 SATTs targets found in Hendriks et al, 2017. Spreadsheet including identified SUMO2 targets for the different SATTs among SUMO2 targets identified in Hendriks et al, 2017, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 6. Total SUMO1 SATTs. Spreadsheet including identified SUMO1 targets for the different SATTs, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 7. Total SUMO2 SATTs. Spreadsheet including identified SUMO2 targets for the different SATTs, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 8. Gene Ontology spreadsheet.

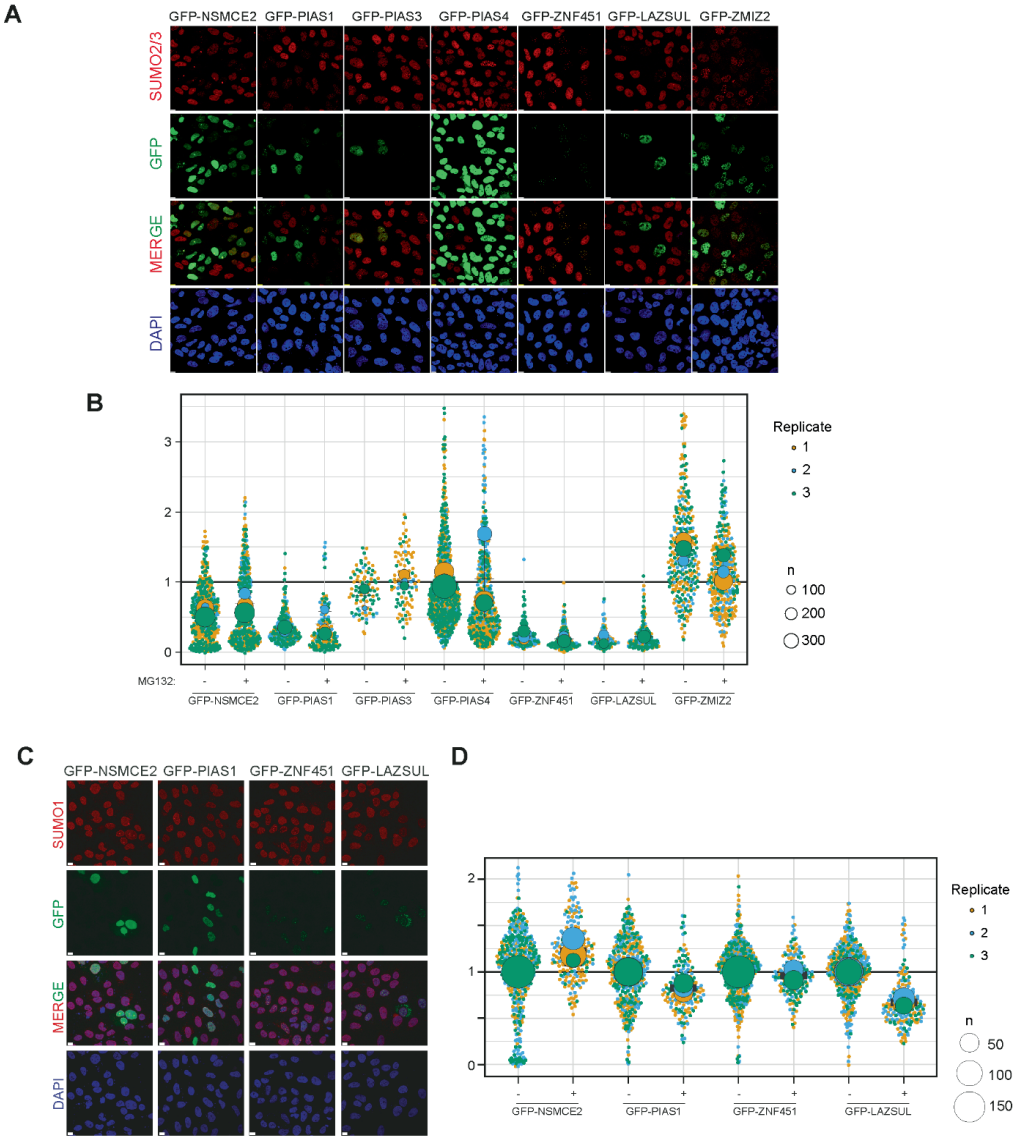
Supplementary Dataset 9. SATTs complement analysis for SUMO1 targets among HIS-SUMO1 targets. Spreadsheet including identified SUMO1 targets identified in Supplementary Dataset 1 after complementation with all the E3 ligases in this study for the different SATTs, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 10. SATTs complement analysis for SUMO2 targets among HIS-SUMO2 targets. Spreadsheet including identified SUMO2 targets identified in Supplementary Dataset 2 after complementation with all the E3 ligases in this study for the different SATTs, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

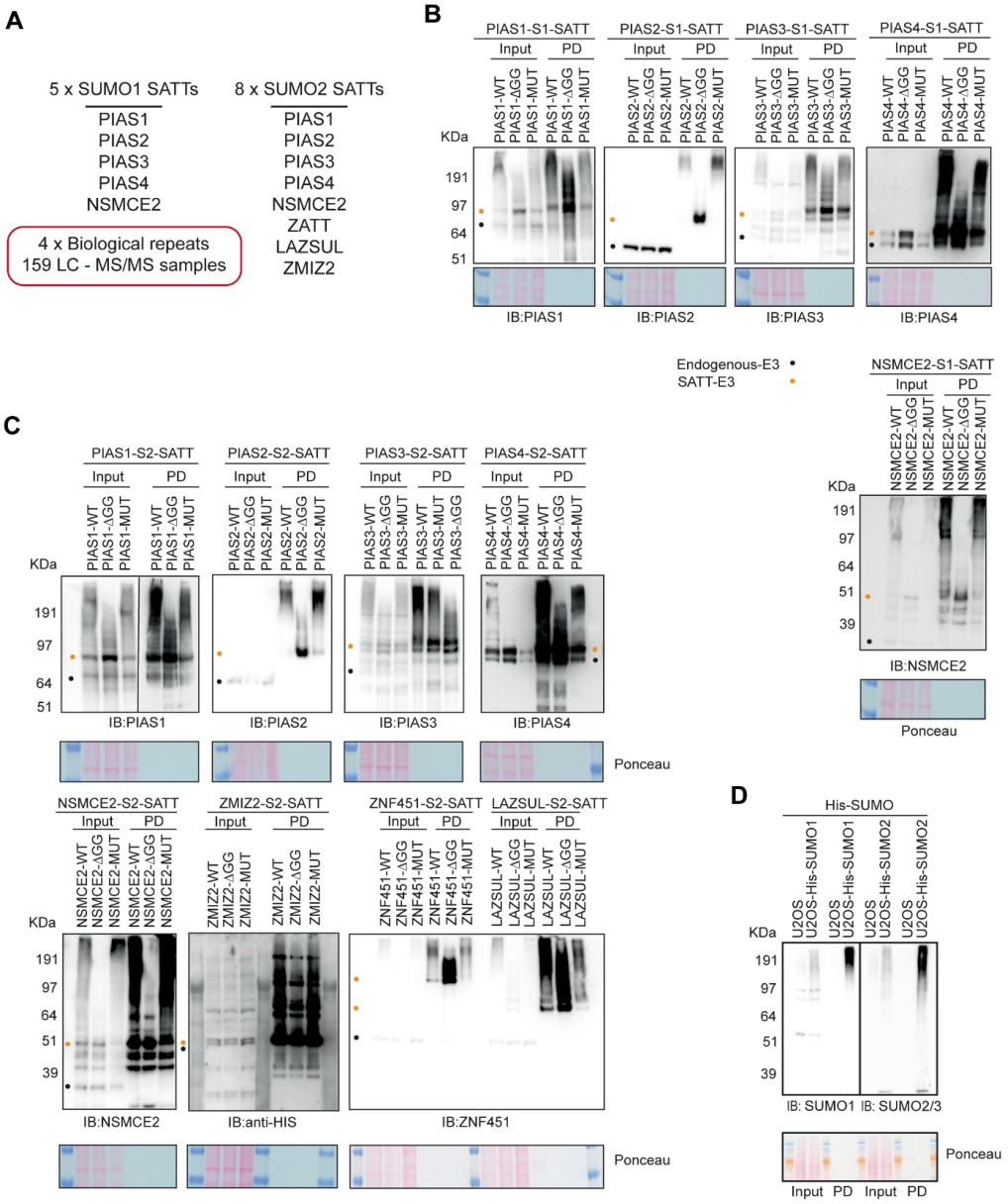
Supplementary Dataset 11. SATTs complement analysis for SUMO1 targets. Spreadsheet including identified SUMO1 targets after complementation with all the E3 ligases in this study for the different SATTs, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

Supplementary Dataset 12. SATTs complement analysis for SUMO2 targets. Spreadsheet including identified SUMO2 targets after complementation with all the E3 ligases in this study for the different SATTs, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.

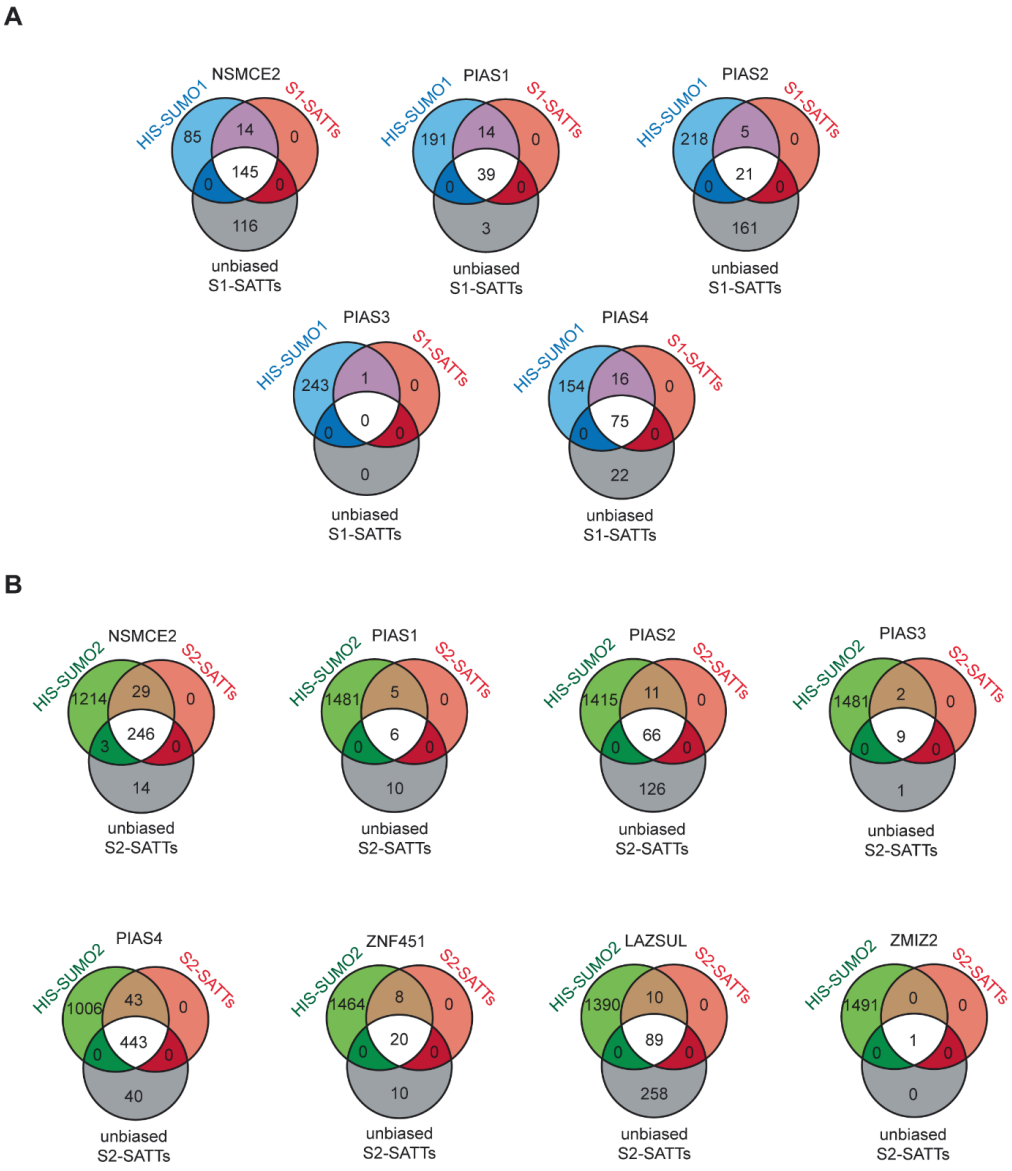
Supplementary Dataset 13. Very high confidence preferential targets. Spreadsheet including high confidence identified SUMO1 and SUMO2 targets for the different SATTs, SATTi, Statistic comparisons and miscellaneous MS/MS proteomics values.



Supplementary Figure 1. (A) Immunofluorescence images of U2OS cells transiently transfected with GFP-tagged constructs of E3 enzymes used in this study. Cells were stained for SUMO2/3 signal. **(B)** Superplot depicting the relative intensity compared to not-treated GFP negative cells of GFP positive cells for the different E3s used in this study after being treated or not with the proteasome inhibitor MG132. **(C)** Immunofluorescence images of U2OS cells transiently transfected with GFP-tagged constructs of indicated E3 enzymes. Cells were stained for SUMO1 signal. **(D)** Superplot depicting the relative SUMO1 intensity of GFP-positive or -negative cells. Intensity was normalized to the average SUMO1 nuclear intensity of GFP-negative cells for each replicate.



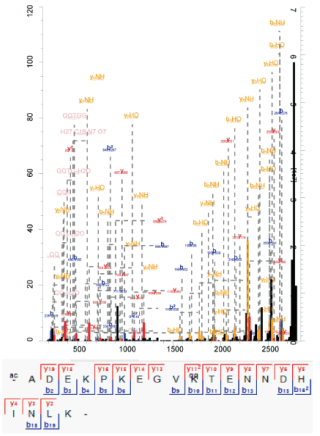
Supplementary Figure 2. (A) SUMO E3 SATTs investigated in this resource. **(B-D)** Immunoblotting analysis of the SUMO1 **(B)** or SUMO2Q87R **(C)** SATTs samples and HIS-SUMO samples **(D)**. Input and pulldown (PD) samples are included. Antibodies used are indicated, Ponceau staining is included as loading control.



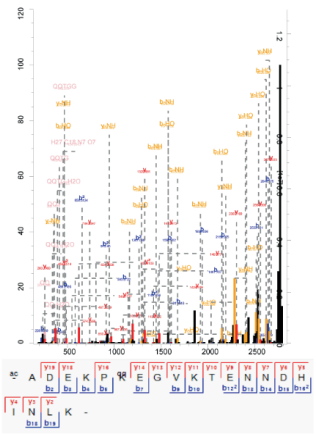
Supplementary Figure 3. Venn diagrams depicting the overlap between the identified **(A)** SUMO1 (blue) or **(B)** SUMO2Q87R (green) identified substrates, the SATTs targets only considering HIS-SUMO targets (red) and the SATTs targets identified in an unbiased manner (grey).

4

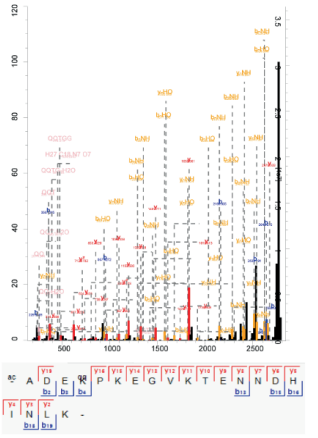
SUMO3 - K11 - SUMO2/3ylation
Score 267.21



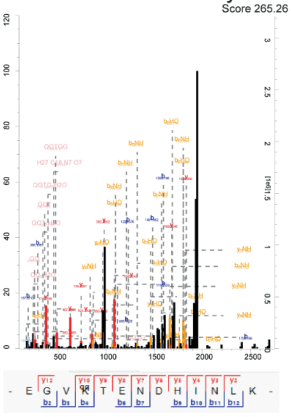
SUMO3 - K7 - SUMO2/3ylation
Score 279.1



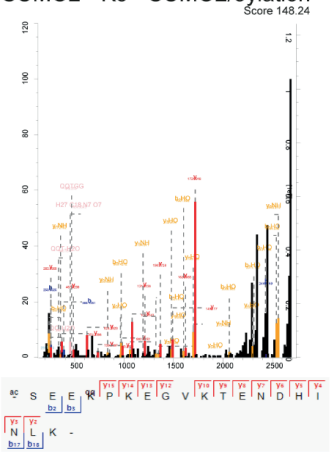
SUMO3 - K5 - SUMO2/3ylation
Score 216.32



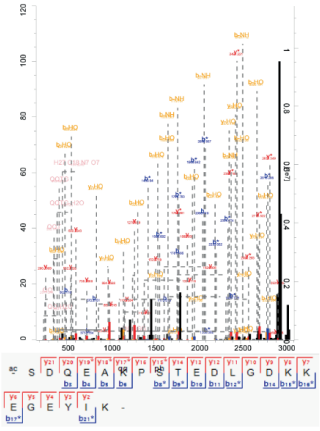
SUMO2 - K11 - SUMO2/3ylation
Score 265.26



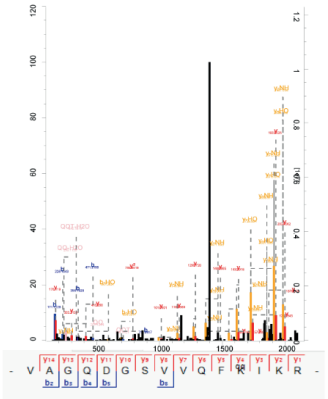
SUMO2 - K5 - SUMO2/3ylation
Score 148.24



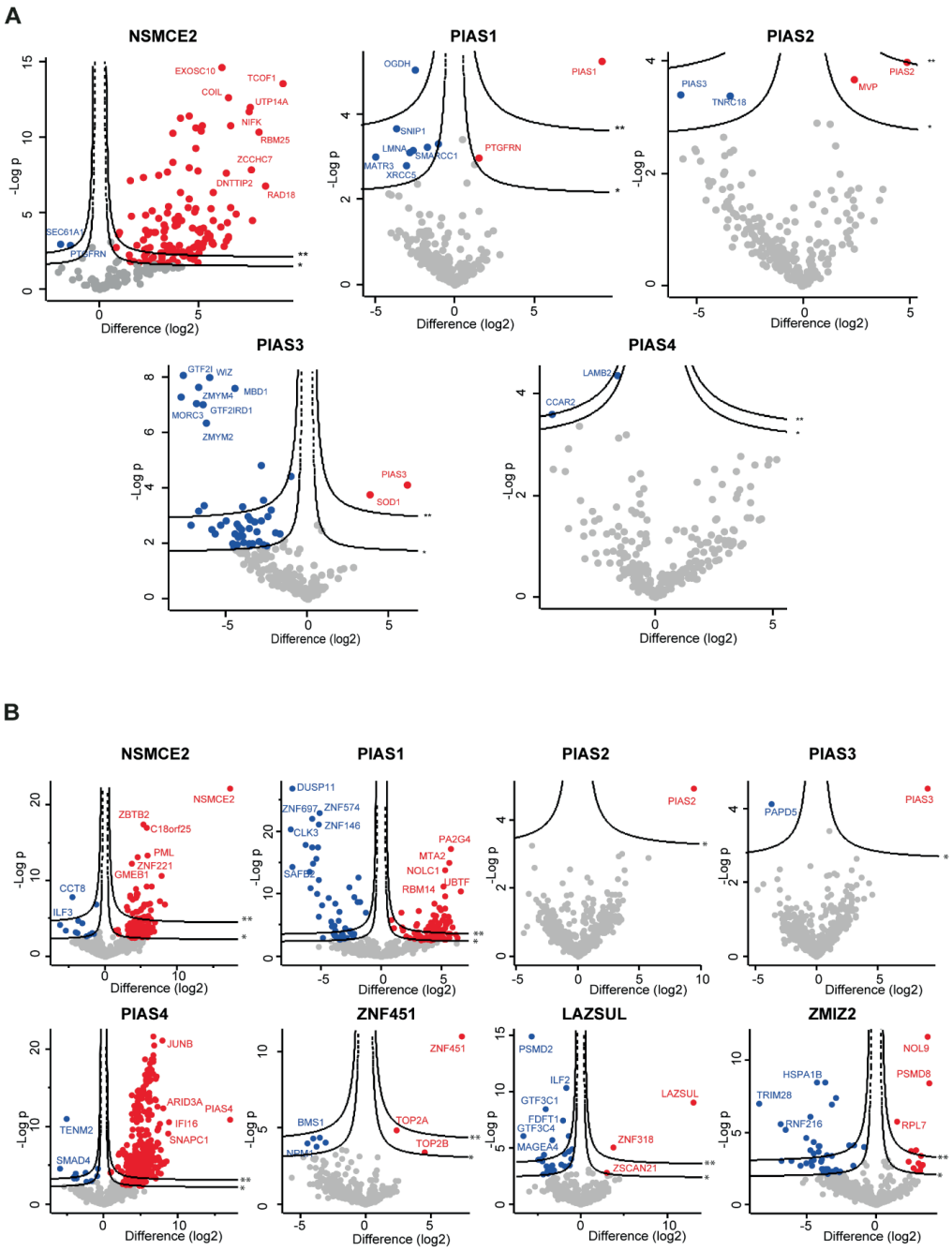
SUMO1 - K7 - SUMO2/3ylation
Score 325.24



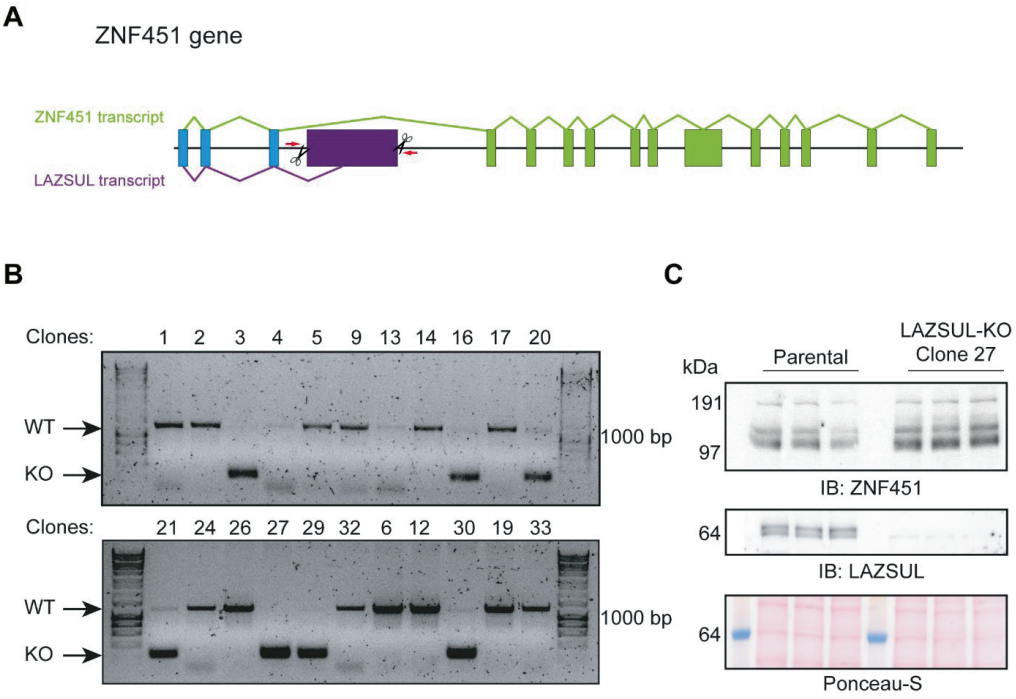
SUMO2/3 - K32/33 - SUMO2/3ylation
Score 192.54



Supplementary Figure 4. Best MS/MS spectra corresponding to the QQTGG sites on SUMOs identified in an unambiguous manner.



Supplementary Figure 5. Multiple volcano plots depicting statistical differences of different SUMO1 **(A)** or SUMO2 **(B)** SATTs targets, only considering HIS-SUMO substrates using the values from the rest of wild type SATTs as complement negative control. Cutoffs represent a Pearson of 100 and an FDR=0.05 (*) or 0.01 (**) and an S0=0.1.



Supplementary Figure 6. U2OS LAZSUL knockout cell line construction. **(A)** Diagram depicting the ZNF451 gene, in blue, common exons for ZNF451 and LAZSUL, in green, ZNF451 exons and in purple the LAP2α exon corresponding to LAZSUL. Scissors represent CRISPR cutting sites employed in the knockout strategy, arrows represent primer sites for clone testing. **(B)** Genomic PCR amplification of LAP2α loci different clones. Size fragments of 1227 bp correspond to LAP2α (WT), while 255 bp fragments correspond to mutated LAP2α (KO). **(C)** Immunoblotting of ZNF451 isoform 1 (ZNF451) and 3 (LAZSUL) of parental and clone 27 LAZSUL-KO cell lines. LAZSUL is detected in parental U2OS cell lines at 63 kDa and is absent in LAZSUL-KO clone 27.

Suplementary Table 1. Antibodies used in this study.

Antibody	Target	Dilution	Company
Primary antibodies			
Rabbit anti-PIAS1 (D33A7)	PIAS1	1:500	Cell Signaling Technology
Rabbit anti-PIAS2 (NBP2-19819)	PIAS2	1:1000	Novus Biologicals
Rabbit anti-PIAS3 (D5F9) 9042S	PIAS3	1:500	Cell Signaling Technology
Rabbit anti-PIAS4 (D2F12) 4392S	PIAS4	1:500	Cell Signaling Technology
Mouse anti-NSMCE2 (13627-1-AP)	NSMCE2	1:5000	SanBio
Rabbit anti-ZNF451	ZNF451	1:1000	Gift from A. Pichler
Mouse anti-His (H1029)	His	1:1000	Sigma-Aldrich
Rabbit anti-histone H1 (PA530055)	histone H1	1:1000	Invitrogen
Mouse anti-SUMO 8A2	SUMO2/3	WB: 1:1000 IF: 1:200	8A2, obtained from Developmental Studies Hybridoma Bank (DSHB), University of Iowa, *in-house* produced
Mouse anti-SUMO1 (4930S)	SUMO1	WB: 1:1000 IF: 1:200	Cell signaling Technology
Rabbit anti-GFP (NB600-308)	GFP	IF	Novus Biologicals
Rabbit anti-JUN-B (210)	JUN-B	1:1000	Santa Cruz
Mouse anti-c-JUN (G-4)	c-JUN	1:1000	Santa Cruz
Rabbit anti-EHMT1/GLP	EHMT1	1:1000	Millipore
Mouse anti-XPF (3F2/3)	ERCC4	1:1000	Santa Cruz
Rabbit anti-PAF1 (A300-172A)	PAF1	1:1000	Bethyl Laboratories
Rabbit anti-LAZSUL	LAZSUL	1:1000	Taylor made A. Pichler (this study)
Secundary antibodies			
HRP-conjugated Donkey anti-Rabbit (31458)	Anti-Rabbit	1:5000	Thermo Fisher Scientific
HPR-conjugated Goat anti-Mouse IgG (H+L) (31432)	Anti-Mouse	1:5000	Thermo Fisher Scientific
HRP-conjugated anti-Rabbit	Anti-Rabbit	1:1000	Jackson ImmunoResearch Laboratories

Supplementary table 2. Primers used in this study.

Primer Name	Sequence	Use
FW-AgeI-10HIS-SUMO1	ACCGGTATGGCTCACCATCACCACCATCATCATCATCATCATGGTGGATCCATGTCTGACCAGGAGGCAAAAC	Generation of SATTs plasmids toolbox SUMO1
FW-AgeI-10HIS-SUMO3	ACCGGTATGGCTCACCATCACCACCATCATCATCATCATCATGGTGGATCC	Generation of SATTs plasmids toolbox SUMO3
RV-XmaI-SUMO1	CCCGGGTCAACCCCCCGTTTGTTCCT	Generation of SATTs plasmids toolbox SUMO1
RV-XmaI-SUMO1noGlyGly	CCCGGGTCACGTTTGTTCCTGATAAACTTCAATCACATC	Generation of SATTs plasmids toolbox SUMO1DGG
RV-XmaI-SUMO3-Q87R	CCCGGGCTAACCTCCCGTCTGCTGC	Generation of SATTs plasmids toolbox SUMO3
RV-XmaI-SUMO3-Q87R-noGlyGly	CCCGGGCTACGTCTGCTGCCGGAACAC	Generation of SATTs plasmids toolbox SUMO3DGG
BP-FW-ZNF451	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGGAGACCCGGGGTCG	Generation of pDNOR ZNF451- and pDNOR-ZNF451-3ZNF451
BP-RV-ZNF451-1	GGGGACCACTTTGTACAAGAAAGCTGGGTACATTTCCTCAAGACTTCTTCTAATAGCTTCTTCTA	Generation of pDNOR ZNF451ZNF451
BP-RV-ZNF451-3	GGGGACCACTTTGTACAAGAAAGCTGGGTAATTCTTGTACCCATGGCGTTCTGG	Generation of pDNOR ZNF451-3 and pDNOR-GFP-ZNF451-3
BP-FW-GFP	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGTGAGCAAGGGCGAGGA	Generation of pDNOR-GFP-ZNF451-3
FW-PIAS1-C351A	TTCCGTGTCGGGCCCTTACAgcTTCTCATCTACAATGTTTTGACGCAACTC	Generation of pDNOR PIAS1 Catalytic dead mutant
RV-PIAS1-C351A	GAGTTGCGTCAAAACATTGTAGATGAGAAgcTGTAAGGGCCCGACACGGAA	Generation of pDNOR PIAS1 Catalytic dead mutant
FW-PIAS2-C362A	CCATGCCGTGCAGTGACTgcTACACATCTGCAGTGTTTTGATGCT	Generation of pDNOR PIAS2 Catalytic dead mutant
RV-PIAS2-C362A	AGCATCAAAACACTGCAGATGTGTAgcAGTCACTGCACGGCATGG	Generation of pDNOR PIAS2 Catalytic dead mutant
FW-PIAS3-C343A	TGTCGTGCCCTCACCgcCGCCCACCTGCAGAGCTTC	Generation of pDNOR PIAS3 Catalytic dead mutant
RV-PIAS3-C343A	GAAGCTCTGCAGGTGGGCGgcGGTGAGGGCACGACA	Generation of pDNOR PIAS3 Catalytic dead mutant
FW-PIAS4-C342A	CCTGCCGGGCAGAGACCgcCGCCCACCTGCAGTGCTTC	Generation of pDNOR PIAS4 Catalytic dead mutant
RV-PIAS4-C342A	GAAGCACTGCAGGTGGGCGgcGGTCTCTGCCCGGCAGG	Generation of pDNOR PIAS4 Catalytic dead mutant
FW- NSMCE2-C185S-H187A	GGAAATGAAGAAGCCAGTGAAAAATAAAGTGtctGGCgccACCTATGAAGAGGACGCCATTG	Generation of pDNOR NSMCE2 Catalytic dead mutant
RV- NSMCE2-C185S-H187A	CAATGGCGTCCTCTTCATAGGTggcGCCagaCACTTTATTTTTCACTGGCTTCTTCATTTCC	Generation of pDNOR NSMCE2 Catalytic dead mutant
FW-ZNF451- SIM1	CGGATGAAAATGAAGACGACGCTCAGGCTGCCAGTGAAGGACCATTACGACCTG	Generation of pDNOR ZNF451 and LAZSUL Catalytic dead mutants

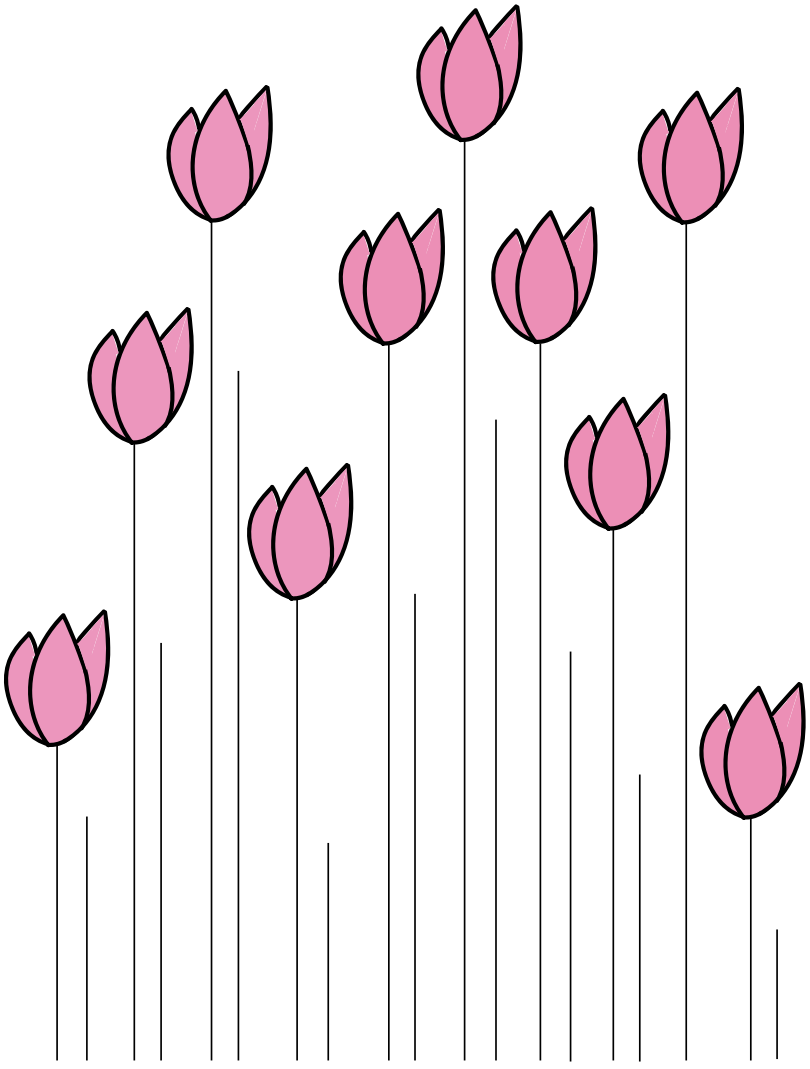
RV-ZNF451- SIM1	CAGGTCGTAATGGTCCTTCACTGGCAG CCTGAGCGTCGTCTTCATTTTCATCCG	Generation of pDNOR ZNF451 and LAZSUL Catalytic dead mutants
FW-ZNF451- SIM2	CCATTACGACCTGTTCTTGAATACGCTG ATGCGGCCAGCAGTGATGATGAAGAG CCTAGCACC	Generation of pDNOR ZNF451 and LAZSUL Catalytic dead mutants
RV-ZNF451- SIM2	GGTGCTAGGCTCTTCATCATCACTGCT GGCCGCATCAGCGTATTCAAGAACAG GTCGTAATGG	Generation of pDNOR ZNF451 and LAZSUL Catalytic dead mutants
LAZSUL gRNA #1-FW	caccgGAGGATGGCACCTTAGATTC	Generation of LAZSUL-KO cell line
LAZSUL gRNA #1-RV	aaacGAATCTAAGGTGCCATCCTCc	Generation of LAZSUL-KO cell line
LAZSUL gRNA #2-FW	caccgAGCCCGAGCTGCTAGCTACA	Generation of LAZSUL-KO cell line
LAZSUL gRNA #2-RV	aaacTGTAGCTAGCAGCTCGGGCTc	Generation of LAZSUL-KO cell line
LAZSUL gRNA #3-FW	caccgACGCTGATATTGAGTCATAT	Generation of LAZSUL-KO cell line
LAZSUL gRNA #3-RV	aaacATATGACTCAATATCAGCGTc	Generation of LAZSUL-KO cell line
ZNF451-3 KO seq - FW	ACTTCTCTGATTCCTTTTGCTGACC	Validation of LAZSUL-KO
ZNF451-3 KO seq - RV	CTGGGTCATGAACCAGATCTTCC	Validation of LAZSUL-KO

REFERENCES

1. G. Duan, D. Walther, The roles of post-translational modifications in the context of protein interaction networks. *PLoS Comput Biol* **11**, e1004049 (2015).
2. D. Salas-Lloret, R. Gonzalez-Prieto, Insights in Post-Translational Modifications: Ubiquitin and SUMO. *Int J Mol Sci* **23**, (2022).
3. F. Trulsson, A. C. O. Vertegaal, Site-specific proteomic strategies to identify ubiquitin and SUMO modifications: Challenges and opportunities. *Semin Cell Dev Biol*, (2021).
4. V. Akimov *et al.*, UbiSite approach for comprehensive mapping of lysine and N-terminal ubiquitination sites. *Nature structural & molecular biology* **25**, 631-640 (2018).
5. F. Trulsson *et al.*, Deubiquitinating enzymes and the proteasome regulate preferential sets of ubiquitin substrates. *Nature communications* **13**, 2736 (2022).
6. I. A. Hendriks *et al.*, Site-specific characterization of endogenous SUMOylation across species and organs. *Nature communications* **9**, 2456 (2018).
7. I. A. Hendriks *et al.*, Site-specific mapping of the human SUMO proteome reveals co-modification with phosphorylation. *Nat Struct Mol Biol* **24**, 325-336 (2017).
8. I. A. Hendriks *et al.*, Uncovering global SUMOylation signaling networks in a site-specific manner. *Nat Struct Mol Biol* **21**, 927-936 (2014).
9. F. Impens, L. Radoshevich, P. Cossart, D. Ribet, Mapping of SUMO sites and analysis of SUMOylation changes induced by external stimuli. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 12432-12437 (2014).
10. R. Gonzalez-Prieto *et al.*, Global non-covalent SUMO interaction networks reveal SUMO-dependent stabilization of the non-homologous end joining complex. *Cell reports* **34**, 108691 (2021).
11. C. Li *et al.*, Quantitative SUMO proteomics identifies PIAS1 substrates involved in cell migration and motility. *Nature communications* **11**, 834 (2020).
12. I. Uzoma *et al.*, Global Identification of Small Ubiquitin-related Modifier (SUMO) Substrates Reveals Crosstalk between SUMOylation and Phosphorylation Promotes Cell Migration. *Molecular & cellular proteomics : MCP* **17**, 871-888 (2018).
13. H. F. O'Connor *et al.*, Ubiquitin-Activated Interaction Traps (UBAITs) identify E3 ligase binding partners. *EMBO reports* **16**, 1699-1712 (2015).
14. R. González-Prieto, A. C. O. Vertegaal, in *SUMOylation and Ubiquitination: Current and Emerging Concepts,* V. G. Wilson, Ed. (Caister Academic Press, U.K., 2019), chap. 10, pp. 147-160.
15. D. Salas-Lloret, G. Agabitini, R. Gonzalez-Prieto, TULIP2: An Improved Method for the Identification of Ubiquitin E3-Specific Targets. *Front Chem* **7**, 802 (2019).
16. R. Kumar, R. Gonzalez-Prieto, Z. Xiao, M. Verlaan-de Vries, A. C. O. Vertegaal, The STUbL RNF4 regulates protein group SUMOylation by targeting the SUMO conjugation machinery. *Nat Commun* **8**, 1809 (2017).
17. R. Moreno-Ayala, D. Schnabel, E. Salas-Vidal, H. Lomeli, PIAS-like protein Zimp7 is required for the restriction of the zebrafish organizer and mesoderm development. *Dev Biol* **403**, 89-100 (2015).
18. Y. Peng, J. Lee, C. Zhu, Z. Sun, A novel role for protein inhibitor of activated STAT (PIAS) proteins in modulating the activity of Zimp7, a novel PIAS-like protein, in androgen receptor-mediated transcription. *The Journal of biological chemistry* **285**, 11465-11475 (2010).
19. J. Beliakoff, Z. Sun, Zimp7 and Zimp10, two novel PIAS-like proteins, function as androgen receptor coregulators. *Nucl Recept Signal* **4**, e017 (2006).
20. H. Lomeli, ZMIZ proteins: partners in transcriptional regulation and risk factors for human disease. *J Mol Med (Berl)*, (2022).
21. M. H. Tatham *et al.*, RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nat Cell Biol* **10**, 538-546 (2008).
22. L. Cappadocia, A. Pichler, C. D. Lima, Structural basis for catalytic activation by the human ZNF451 SUMO E3 ligase. *Nature structural & molecular biology*, (2015).
23. N. Eisenhardt *et al.*, A new vertebrate SUMO enzyme family reveals insights into SUMO-chain assembly. *Nature structural & molecular biology*, (2015).
24. L. Wang *et al.*, SUMO2 is essential while SUMO3 is dispensable for mouse embryonic development. *EMBO Rep* **15**, 878-885 (2014).
25. Y. Galanty *et al.*, Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks. *Nature* **462**, 935-939 (2009).

26. N. Varejao *et al.*, DNA activates the Nse2/Mms21 SUMO E3 ligase in the Smc5/6 complex. *The EMBO journal* **37**, (2018).
27. M. Agostinho *et al.*, Conjugation of human topoisomerase 2 alpha with small ubiquitin-like modifiers 2/3 in response to topoisomerase inhibitors: cell cycle stage and chromosome domain specificity. *Cancer research* **68**, 2409-2418 (2008).
28. M. J. Schellenberg *et al.*, ZATT (ZNF451)-mediated resolution of topoisomerase 2 DNA-protein cross-links. *Science* **357**, 1412-1416 (2017).
29. J. Liang *et al.*, SUMO E3 ligase Mms21 prevents spontaneous DNA damage induced genome rearrangements. *PLoS Genet* **14**, e1007250 (2018).
30. Y. Sun *et al.*, A conserved SUMO pathway repairs topoisomerase DNA-protein cross-links by engaging ubiquitin-mediated proteasomal degradation. *Sci Adv* **6**, (2020).
31. P. Sarangi, X. Zhao, SUMO-mediated regulation of DNA damage repair and responses. *Trends in biochemical sciences* **40**, 233-242 (2015).
32. S. Dabir, A. Kluge, A. Dowlati, The association and nuclear translocation of the PIAS3-STAT3 complex is ligand and time dependent. *Mol Cancer Res* **7**, 1854-1860 (2009).
33. M. Sundvall *et al.*, Protein inhibitor of activated STAT3 (PIAS3) protein promotes SUMOylation and nuclear sequestration of the intracellular domain of ErbB4 protein. *The Journal of biological chemistry* **287**, 23216-23226 (2012).
34. J. H. Man *et al.*, PIAS3 induction of PRB sumoylation represses PRB transactivation by destabilizing its retention in the nucleus. *Nucleic acids research* **34**, 5552-5566 (2006).
35. M. H. Tatham *et al.*, Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. *The Journal of biological chemistry* **276**, 35368-35374 (2001).
36. D. E. Verver *et al.*, Non-SMC Element 2 (NSMCE2) of the SMC5/6 Complex Helps to Resolve Topological Stress. *Int J Mol Sci* **17**, (2016).
37. A. Jacome *et al.*, NSMCE2 suppresses cancer and aging in mice independently of its SUMO ligase activity. *The EMBO journal* **34**, 2604-2619 (2015).
38. D. B. Krastev *et al.*, The ubiquitin-dependent ATPase p97 removes cytotoxic trapped PARP1 from chromatin. *Nature cell biology* **24**, 62-73 (2022).
39. D. Lee *et al.*, ZMYM2 restricts 53BP1 at DNA double-strand breaks to favor BRCA1 loading and homologous recombination. *Nucleic acids research* **50**, 3922-3943 (2022).
40. J. Schimmel *et al.*, The ubiquitin-proteasome system is a key component of the SUMO-2/3 cycle. *Molecular & cellular proteomics : MCP* **7**, 2107-2122 (2008).
41. R. Gonzalez-Prieto, S. A. Cuijpers, R. Kumar, I. A. Hendriks, A. C. Vertegaal, c-Myc is targeted to the proteasome for degradation in a SUMOylation-dependent manner, regulated by PIAS1, SENP7 and RNF4. *Cell Cycle* **14**, 1859-1872 (2015).
42. A. M. Sriramachandran *et al.*, Arkadia/RNF111 is a SUMO-targeted ubiquitin ligase with preference for substrates marked with SUMO1-capped SUMO2/3 chain. *Nature communications* **10**, 3678 (2019).
43. A. Seifert, P. Schofield, G. J. Barton, R. T. Hay, Proteotoxic stress reprograms the chromatin landscape of SUMO modification. *Sci Signal* **8**, rs7 (2015).
44. J. J. Palecek, SMC5/6: Multifunctional Player in Replication. *Genes (Basel)* **10**, (2018).
45. C. Li *et al.*, SUMO Proteomics Analyses Identify Protein Inhibitor of Activated STAT-Mediated Regulatory Networks Involved in Cell Cycle and Cell Proliferation. *Journal of proteome research* **22**, 812-825 (2023).
46. I. A. Hendriks, R. C. D'Souza, J. G. Chang, M. Mann, A. C. Vertegaal, System-wide identification of wild-type SUMO-2 conjugation sites. *Nature communications* **6**, 7289 (2015).
47. J. Schimmel *et al.*, Uncovering SUMOylation dynamics during cell-cycle progression reveals FoxM1 as a key mitotic SUMO target protein. *Molecular cell* **53**, 1053-1066 (2014).
48. P. Gonzalez-Rodriguez *et al.*, Disruption of mitochondrial complex I induces progressive parkinsonism. *Nature* **599**, 650-656 (2021).
49. J. Magalhaes *et al.*, PIAS2-mediated blockade of IFN-beta signaling: a basis for sporadic Parkinson disease dementia. *Mol Psychiatry* **26**, 6083-6099 (2021).
50. J. Schindelin *et al.*, Fiji: an open-source platform for biological-image analysis. *Nature methods* **9**, 676-682 (2012).

51. Z. Xiao *et al.*, System-wide Analysis of SUMOylation Dynamics in Response to Replication Stress Reveals Novel Small Ubiquitin-like Modified Target Proteins and Acceptor Lysines Relevant for Genome Stability. *Molecular & cellular proteomics : MCP* **14**, 1419-1434 (2015).
52. J. Rappsilber, M. Mann, Y. Ishihama, Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* **2**, 1896-1906 (2007).
53. S. Tyanova, T. Temu, J. Cox, The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **11**, 2301-2319 (2016).
54. S. Tyanova *et al.*, The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731-740 (2016).
55. H. Mi *et al.*, PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic acids research* **49**, D394-D403 (2021).
56. K. Eifler *et al.*, SUMO targets the APC/C to regulate transition from metaphase to anaphase. *Nature communications* **9**, 1119 (2018).
57. J. Goedhart, SuperPlotsOfData-a web app for the transparent display and quantitative comparison of continuous data from different conditions. *Molecular biology of the cell* **32**, 470-474 (2021).
58. Y. Perez-Riverol *et al.*, The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic acids research* **47**, D442-D450 (2019).



5

BRCA1/BARD1 ubiquitinates PCNA in unperturbed conditions to promote replication fork stability and continuous DNA synthesis

Daniel Salas-Lloret[1], Néstor García-Rodríguez[2,3], Lisanne Giebel[1], Arnoud de Ru[4], Peter A. van Veelen[4], Pablo Huertas[2,3], Alfred C.O. Vertegaal[1] and Román González-Prieto[1,3,5,6]

[1]Department of Cell and Chemical Biology, Leiden University Medical Centre, Leiden, The Netherlands

[2]Departamento de Genética, Facultad de Biología, Universidad de Sevilla, Sevilla, Spain

[3]Andalusian Centre for Regenerative Medicine and Molecular Biology (CABIMER), Universidad de Sevilla-CSIC-Universidad Pablo de Olavide-Junta de Andalucía, Sevilla, Spain.

[4]Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, The Netherlands.

[5]Departamento de Biología Celular, Facultad de Biología, Universidad de Sevilla, Sevilla, Spain

[6]correspondence: roman.gonzalez@cabimer.es

This chapter is under revision in Molecular Cell (2023)

Abstract

Deficiencies in the BRCA1 tumor suppressor gene are the main cause of hereditary breast and ovarian cancer. BRCA1 is involved in the Homologous Recombination DNA repair pathway, and, together with BARD1, forms a heterodimer with ubiquitin E3 activity. The relevance of the BRCA1/BARD1 ubiquitin E3 activity for tumor suppression and DNA repair remains controversial and most efforts aimed to identify BRCA1/BARD1 ubiquitination substrates rely on indirect evidence. Here, we observed that the BRCA1/BARD1 ubiquitin E3 activity was not required for Homologous Recombination or resistance to Olaparib. Using TULIP2 methodology, which enables the direct identification of E3-specific ubiquitination substrates, we identified substrates for BRCA1/BARD1. We found that PCNA is ubiquitinated by BRCA1/BARD1 in unperturbed conditions independently of RAD18. PCNA ubiquitination by BRCA1/BARD1 avoids the formation of ssDNA gaps during DNA replication and promotes replication fork stability epistatically to BRCA1 S114 phosphorylation, addressing the controversy about the function of BRCA1/BARD1 E3 activity in Homologous Recombination.

Keywords: TULIP2; Mass-Spectrometry; BRCA1-BARD1; PCNA

INTRODUCTION

Breast cancer susceptibility type 1 (BRCA1) binds its partner BRCA1-Associated RING Domain 1 (BARD1) to form an obligated and multifunctional heterodimer with ubiquitin E3 ligase activity (1-3). The BRCA1/BARD1 heterodimer acts as tumor suppressor and maintains genome stability, generally, by repairing deleterious double-strand DNA breaks (DSBs) via error-free homologous recombination (HR) (4, 5). DSBs can originate either from endogenous agents, such as reactive oxygens, replication fork progression and single-stranded DNA (ssDNA), or from environmental exposure to chemicals, ionizing radiation and ultraviolet light (UV) (6).

Germline mutations in BRCA1 and BARD1 are the main cause of hereditary breast and ovarian cancers, conferring a life-time probability of up to 90% for developing breast cancer and 50% for ovarian cancer (7-9). Mice studies suggested that the E3 ligase activity is not essential to prevent tumor development (10). However, mutations in BARD1 that impair Histone 2A (H2A) ubiquitination have been identified in familial breast cancer (11).

BRCA1/BARD1 has a very well established histone H2A ubiquitin ligase activity on lysines K127/K129 (12-15). This ubiquitination has been related to the maintenance of heterochromatin integrity, genetic stability and senescence (2, 12, 15).

Up to date, the relevance of the BRCA1/BARD1 E3 activity for DNA Double Strands Break repair, tumor suppression and resistance to PARP inhibitors and platinum-based compounds is still controversial. Histone H2A K127/129 ubiquitination is required for DNA DSB repair by homologous recombination and RAD51 foci formation (2). However, deficiencies in histone H2A ubiquitination by BRCA1/BARD1 can be rescued by other ubiquitin E3s such as RNF168 (16, 17). While a study from the Morris group (2) using siRNA-based knockdowns in HeLa cells showed that the BRCA1/BARD1 E3 activity was required for resistance to agents such as the PARP inhibitor Olaparib and cisplatin, a degrON-based strategy on HCT116 cells showed that the BRCA1/BARD1 E3 activity was dispensable for resistance to Olaparib (18), consistent with most of the published literature (10, 16, 18-20).

Moreover, many research groups have addressed the challenge of identifying BRCA1/BARD1 ubiquitination substrates. However, all these studies have relied in indirect evidence to identify putative BRCA1/BARD1 ubiquitination substrates, by either overexpressing or depleting BRCA1 and identification of changes in the ubiquitin proteome by mass spectrometry-based proteomics.(15, 21-25). However, since BRCA1/BARD1 plays an important role in several signaling pathways including cell cycle regulation, replication fork protection, gene transcription regulation and DNA damage repair (5, 26), these approaches could indirectly affect the ubiquitination state of proteins. Recently, we developed the TULIP and TULIP2 methodologies (27, 28) which enabled the unambiguous identification of direct ubiquitination substrates for an E3 of interest.

Here, we investigated the role of the BRCA1/BARD1 heterodimer E3 activity in the non-tumoral cell line RPE1 and applied the TULIP2 methodology to identify novel substrates for ubiquitination by BRCA1/BARD1. Combined, our data provided novel insights in BRCA1/BARD1 function and enabled us to solve the controversy in the field about the relevance of the ubiquitin E3 function of BRCA1/BARD1 for homologous recombination.

RESULTS

BRCA1/BARD1 ubiquitin E3 activity is dispensable for DSB repair and resistance to treatment with Olaparib.

The relevance of the BRCA1/BARD1 ubiquitin E3 activity for homologous recombination in terms of DNA DSB repair and treatment resistance remains contradictory in current literature applying different approaches in different cancer cell lines and models (2, 10, 18-20). To address this controversy, we employed RPE1 TP53-/- (Parental) and RPE1 TP53-/- BRCA1-/- (BRCA1-KO) cells (3) rescued or not with GFP-tagged BRCA1 wild type or I26A mutant. The BRCA1 I26A mutant abrogates the BRCA1/BARD1 E3 activity without disrupting the formation of the BRCA1/BARD1 heterodimer (29, 30). We decided to test the proficiency in homologous recombination of the ubiquitin E3 activity BRCA1 mutant I26A in non-tumoral cells both in terms of resistance to Olaparib (**Figure 1A**) and RAD51 foci formation in response to Ionizing Irradiation (IR) (**Figure 1B-C**). As expected, BRCA1-KO cells were deficient in resistance to Olaparib and RAD51 foci formation in response to ionizing radiation, and BRCA1-GFP rescued cells could restore both phenotypes to the Parental levels. Importantly, BRCA1I26A-GFP rescued cells could also restore Olaparib resistance and RAD51 foci to the Parental levels, which is consistent with most of the published literature about the relevance of BRCA1 ubiquitin E3 activity in homologous recombination in DSB repair and resistance to Olaparib and Platinum-based compounds (10, 16, 18-20).

BRCA1/BARD1 ubiquitinates PCNA in a constitutive manner.

Despite proficiency in DNA DSB repair, cancer predisposition has been observed in BRCA1 E3 mutant mouse models (19), and previous studies aiming to identify BRCA1/BARD1 ubiquitination substrates are based on indirect evidence (13, 15, 31). Moreover, the ubiquitination of histone H2A-K127/K129, which is the best studied BRCA1/BARD1 ubiquitination substrate (13, 15), can be rescued by RNF168, another E3 enzyme, upon BRCA1 deficiencies (17). Thus, we hypothesized that BRCA1/BARD1 ubiquitination substrates other than histone H2A isoforms exist.

We recently developed the TULIP2 methodology for the systematic application of Ubiquitin Activated Interaction Traps (27, 28, 32), which enables the identification of E3-specific targets in a direct manner **(Supplementary Figure 1A; Figure 1D**). The rationale behind this approach is that if we have a linear fusion between an E3 and ubiquitin, this E3 could use its attached ubiquitin to modify its substrate, enabling the co-purification and subsequent identification by mass spectrometry-based proteomics of the E3 together with ubiquitin and the covalently modified substrate. Therefore, we decided to apply TULIP2 methodology to identify BRCA1/BARD1-specific ubiquitination substrates. Three different TULIP2 constructs were generated to identify BRCA1 direct ubiquitination targets: BRCA1-WT-TULIP2, BRCA1-I26A-TULIP2 and BRCA1-WT-TULIP2ΔGG. Both BRCA1-I26A and BRCA1-WT-TULIP2ΔGG served as negative controls. While BRCA1-I26A is a catalytic dead mutant without E3 activity, BRCA1-WT-TULIP2ΔGG fused ubiquitin cannot be conjugated to target proteins due to lack of the C-terminal diGly motif (**Supplementary Figure 1A,B**).

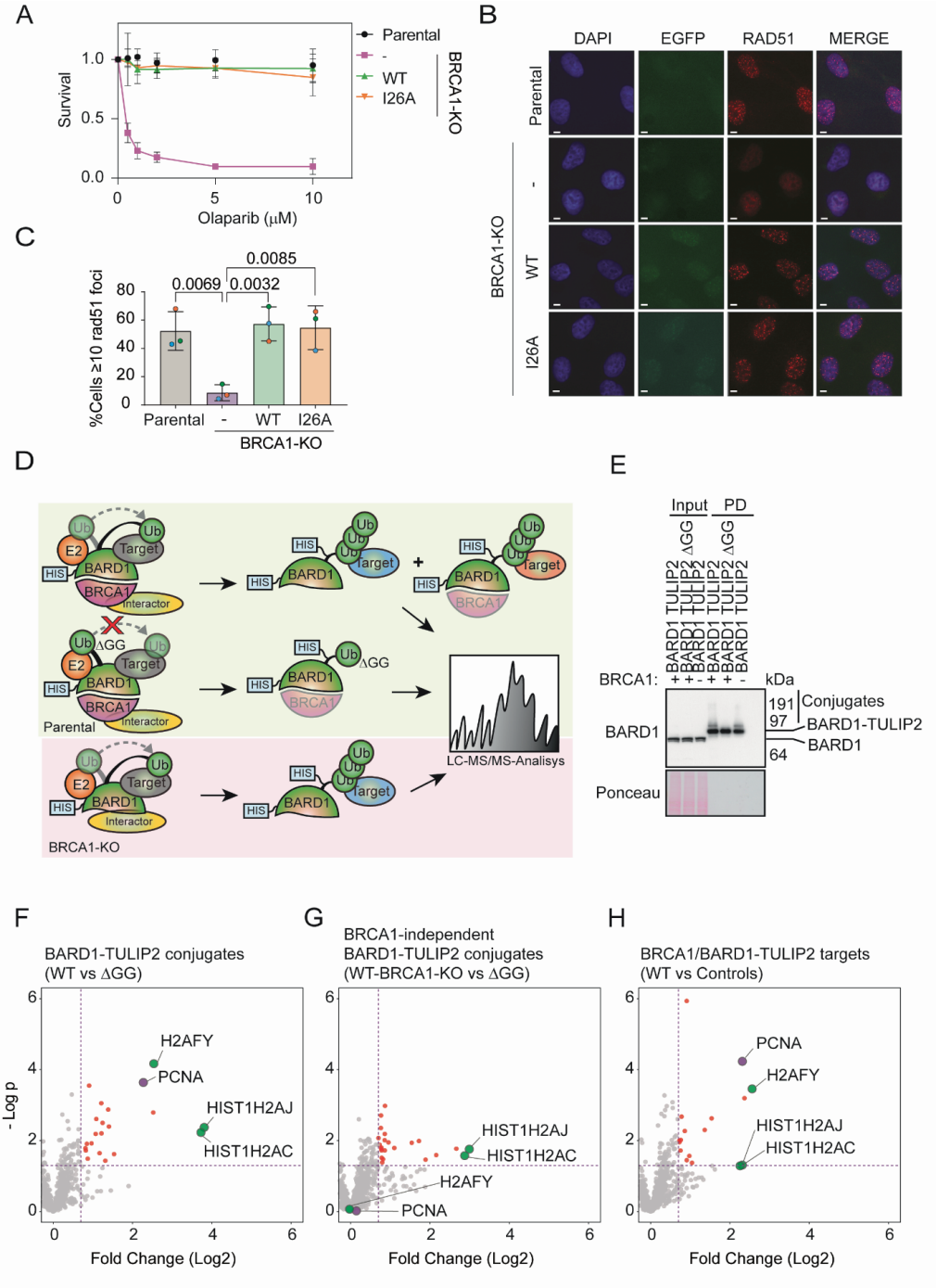


Figure 1. BRCA1 E3 activity is not required for homologous recombination. **A**. Clonogenic survival assay of Parental and BRCA1-KO cells complemented or not with either BRCA1-WT-GFP or BRCA1-I26A-GFP in response to different concentrations of Olaparib. Average and standard deviation of three independent

experiments with three technical repeats are depicted (N=3). **B-C**. Immunofluorescence analysis Parental and BRCA1-KO cells complemented or nor with either BRCA1-WT-GFP or BRCA1-I26A-GFP after exposure to 10 Gy ionizing irradiation. Representative images (B) and quantification (C) of cells with RAD51 foci is shown. Size bars in fluorescence microscopy images represent 10 µm. Data corresponds to three independent experiments. The average of each individual experiment is shown by an orange, green or blue circle respectively. Unpaired t-tests were performed with p-values shown in the figure **D**. Cartoon depicting BARD1-TULIP2 rationale. **E**. Analysis by immunoblotting of BARD1-TULIP2 samples. F-H. Volcano plots depicting differences between each of the specified BARD1-TULIP2 constructs. Histones H2A and macro-H2A and PCNA are labelled. Each dot represents a protein.

First, we confirmed the functionality of the BRCA1-TULIP2 constructs for homologous recombination by introducing them in BRCA1-KO cells in a stable inducible manner and testing their resistance to Olaparib treatment (**Supplementary Figure 1C**), and RAD51 foci formation in response to IR (**Supplementary Figure 1D,E**). As previously seen for the GFP-tagged constructs, BRCA1-KO cells were hypersensitive to Olaparib and failed to form RAD51 foci in response to IR. However, BRCA1-KO cells stably expressing either BRCA1-WT-TULIP2 or BRCA1-I26A-TULIP2 completely rescued both Olaparib sensitivity and deficiency in RAD51 foci formation (**Supplementary Figure 1C-E**). Besides, BRCA1-TULIP2 constructs co-localized with RAD51 similarly to endogenous BRCA1 (**Supplementary Figure 1D**). These results corroborate the functionality of BRCA1-TULIP2 constructs and indicate a minor role of BRCA1/BARD1 E3 activity on HR pathway for DNA DSB repair and Olaparib resistance.

Next, we performed the TULIP2 assay to identify BRCA1 ubiquitin E3 activity-specific substrates using our BRCA1-TULIP2 cell lines (**Supplementary Figure 1A**) and performed mass spectrometry-based proteomics analysis (**Supplementary Dataset 1; Supplementary Figure 1F-G**). Analysis by immunoblotting failed to show a smear up from the wild type and I26A corresponding to BRCA1-TULIP2 conjugates as expected from functional TULIPs (**Supplementary Figure 1B**) (27, 28, 33). Moreover, proteomics analysis was unable to identify neither histone H2A nor macro-H2A as BRCA1-TULIP2 substrates, which would had served as positive controls, and the top hit was RAB43, which is not a nuclear protein. We concluded that the BRCA1-TULIP2 constructs were functional for BRCA1 activity but not regarding the TULIP2 assay, likely due to steric hindrance.

Nevertheless, a recent cryo-EM study has shown that it is BARD1, and not BRCA1, the heterodimer partner which positions the E2 enzyme and directs the ubiquitination of histone H2A (14). Therefore, we made stable inducible BARD1-TULIP2 constructs and introduced them in Parental cells, including the wild type and ΔGG TULIP2 constructs. As an additional negative control, in this case, we introduced wild type BARD1-TULIP2 constructs in BRCA1-KO cells (**Figure 1D**). Analysis by immunoblotting of the BARD1-TULIP2-expressing cells (**Figure 1E**) revealed that: 1) The endogenous BARD1 levels were similar in parental and BRCA1-KO cells; 2) the BARD1-TULIP2 construct levels were not detectable compared to endogenous BARD1 levels in the input samples, which is important to avoid overexpression artefacts; and 3) The BARD1-TULIP2 constructs were very efficiently enriched in the His-pulldown following TULIP2 methodology and a smear up from the wild type BARD1-TULIP2 construct corresponding to BARD1-TULIP2 conjugates, which was absent in the ΔGG control, could be detected. Mass spectrometry analysis of BARD1-TULIP2 samples (**Supplementary Dataset 2; Figure 1F-H**) identified both histones H2A

and macro-H2A as BARD1 ubiquitination substrates, serving as internal positive controls. Interestingly, we could identify another top hit, PCNA, as a BRCA1/BARD1 specific ubiquitination substrate using TULIP2 methodology (**Figure 1H**).

BRCA1/BARD1 and RAD18 ubiquitinates PCNA-K164 in distinct signaling pathways

The proteomics spectra obtained in our BARD1-TULIP2 experiments did not enable the identification of the ubiquitin acceptor lysine in PCNA for BRCA1/BARD1. K164 is the main ubiquitin acceptor site in PCNA, and this ubiquitination promotes the recruitment of Trans-Lesion Synthesis (TLS) DNA polymerases (34). However, other acceptor lysines for ubiquitin have been identified in PCNA, which physiological relevance are unknown (35). Thus, there was the possibility that the BRCA1/BARD1 heterodimer was ubiquitinating PCNA on a lysine other than K164. Therefore, we performed the BARD1-TULIP2 assay on RPE1 TP53-/- PCNAK164R/K164R (PCNA-K164R) cells (36) (**Supplementary Dataset 3; Figure 2A**). In this assay, we could identify histones H2A and macroH2A as BARD1 ubiquitination substrates. However, PCNA was lost as a BARD1-specific substrate, which indicates that BRCA1/BARD1 ubiquitinates PCNA on lysine K164.

BRCA1 has been shown to promote the recruitment of RAD18 to chromatin in response to replication-blocking lesions (37). Since RAD18 is considered the canonical ubiquitin ligase for PCNA in response to replication barriers (38-40) and protein-protein interactions between BRCA1 and RAD18 have been described (37), there was the possibility that it was RAD18 the E3 enzyme using the ubiquitin moiety in the BARD1-TULIP2 construct to modify PCNA. To exclude this possibility, we made RAD18-TULIP2 constructs, introduced them in Parental cells and performed the TULIP2 assay with and without UV light irradiation (**Supplementary dataset 4; Figure 2B**). As expected, immunoblotting analysis (**Supplementary Figure 2A**) showed that the wild type RAD18-TULIP2 samples formed a smear up from the RAD18-TULIP2 construct, indicative of its functionality, that was suppressed in the ΔGG negative controls. Mass spectrometry-based proteomics analysis of the non-irradiated samples revealed that among the identified RAD18 substrates through TULIP2 methodology was not PCNA, which indicated that BRCA1/BARD1, and not RAD18, was the preferred ubiquitin E3 enzyme for PCNA under unperturbed conditions (**Figure 2C**). However, there was still the possibility that the absence of PCNA among the RAD18-specific substrates, regardless of the formation of the smear, was due to steric hindrance, similarly as previously observed for BRCA1 (**Supplementary Figure 1**).

To serve as positive control, we irradiated with UV light both BARD1-TULIP2 and RAD18-TULIP2 expressing Parental cells, performed the TULIP2 methodology and identified the substrates for both constructs after UV irradiation (**Supplementary Dataset 4; Figure 2D**). While PCNA enrichment as a BARD1-TULIP2 substrate was not affected by UV irradiation, in the case of RAD18-TULIP2, PCNA became now a substrate while it was not in unperturbed conditions, confirming the functionality of the RAD18-TULIP2 construct. Accordingly, PCNA ubiquitination levels were lower in BRCA1 knockout cells compared to parental cells in unperturbed conditions (Supplementary Figure 2B). But, after UV irradiation, ubiquitination kinetics of PCNA were not affected by BRCA1 knockout (**Supplementary Figure 2C**). Additionally, we confirmed by RAD18 knockdown, that PCNA ubiquitination after UV exposure was RAD18-dependent and BRCA1-independent (**Supplementary Figure 2D**). Finally, we performed the RAD18-TULIP2 assay in BRCA1-KO cells

after UV irradiation (**Supplementary Figure 2E**). Mass spectrometry analysis revealed an additional enrichment of PCNA as a RAD18 substrate compared to parental cells (**Supplementary Figure 2F**).

Combined, these data allowed us to hypothesize that BRCA1/BARD1 and RAD18 share PCNA as substrate but act in different scenarios. While RAD18 would ubiquitinate PCNA when the replication fork encounters a replication fork barrier, BRCA1/BARD1 would perform PCNA ubiquitination at low levels to facilitate continuous DNA synthesis in unperturbed conditions (Figure 2E).

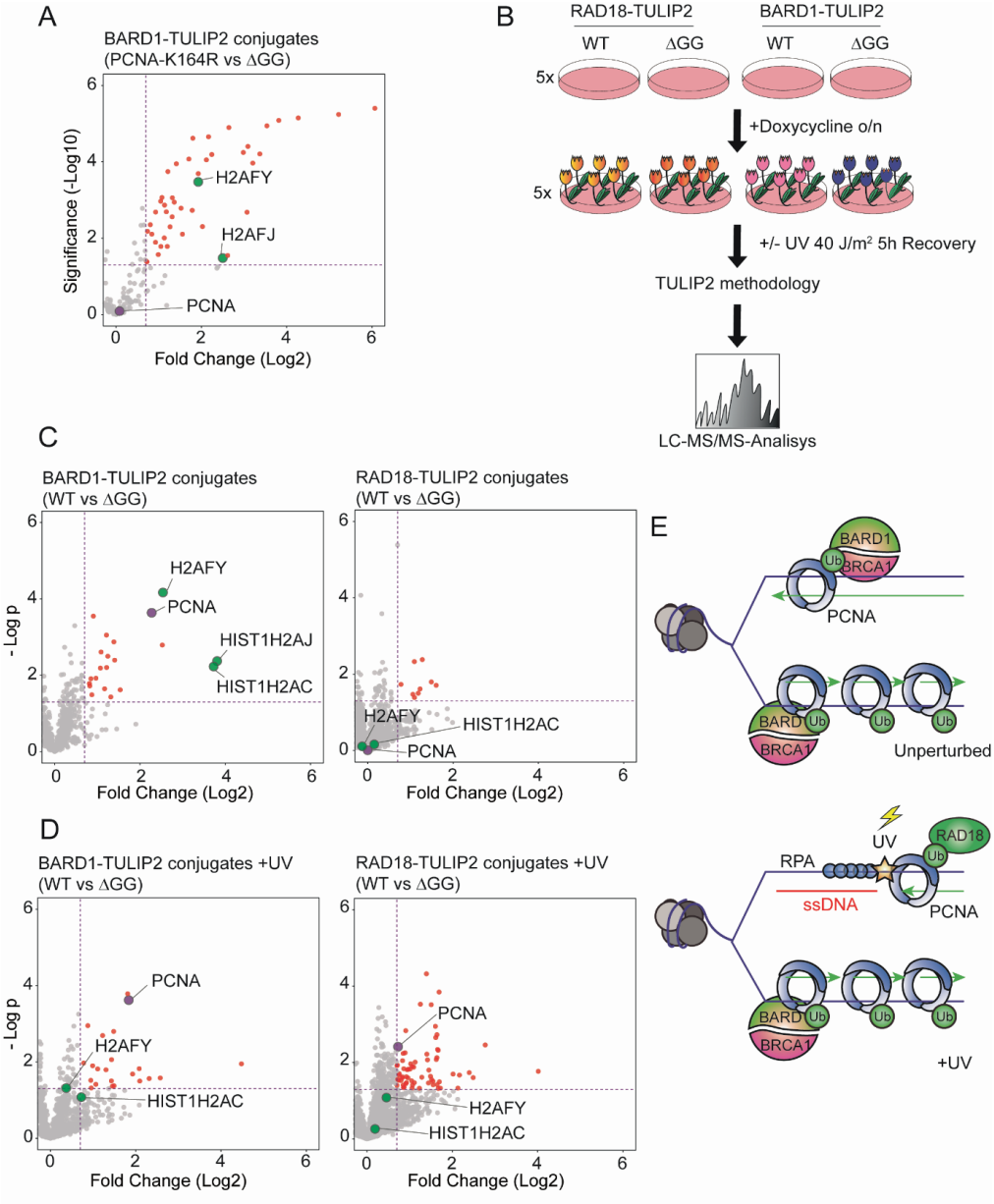


Figure 2. RAD18 and BARD1 ubiquitinates PCNA in different pathways. **A**. Volcano plot depicting differences between BARD1-TULIP2 conjugates in PCNA-K164R cell line versus □GG control. Histones H2A and macro-H2A and PCNA are labelled. **B**. Experimental setup. The expression of specified BARD1-TULIP2 and RAD18-TULIP2 constructs were induced overnight. Cells were irradiated with 40 J/m2 of UV light and allowed cells to recover for 5h prior to cell lysis, protein purification and identification by LC-MS/MS of the TULIP2 conjugates **C, D**. Volcano plots depicting differences between wild type BARD1-TULIP2 or RAD18-TULIP2 with their respective ΔGG counterparts in unperturbed conditions (C) or 5h after UV irradiation (D). Each dot represents a protein, PCNA, histone H2A and macro-H2A are labeled. E. Proposed model. During unperturbed conditions BRCA1/BARD1 ubiquitinates PCNA at low levels and upon UV irradiation RAD18 ubiquitinates PCNA in response to DNA lesions that represent a barrier for the replication fork while BRCA1/BARD1 still ubiquitinates PCNA.

BRCA1/BARD1 ubiquitinates PCNA to prevent ssDNA accumulation under unperturbed conditions.

PCNA K164 ubiquitination has recently been shown to participate in replication fork protection (36) and in preventing the accumulation of ssDNA gaps during DNA replication in unperturbed conditions (41), two functions that have been previously also attributed to BRCA1 (42-45). Different groups have shown that BRCA1-deficient cancer cells accumulate ssDNA gaps and thus can be therapeutically exploited with PARP or REV1 inhibitors (43, 46). We hypothesized that BRCA1/BARD1 mediates PCNA ubiquitination to protect replication forks and avoid the accumulation of ssDNA gaps in unperturbed conditions. Therefore, mutations on either PCNA or BARD1-BRCA1 that compromise its E3 activity could lead to ssDNA accumulation and genetic instability (**Figure 3A**).

To test BRCA1/BARD1 E3 activity on ssDNA gaps formation, we cultured different RPE1 cell lines with 10 μM CldU for 48 h and analyzed by immunofluorescence in non-denaturing conditions using an anti-BrU antibody that will only recognize CldU present in ssDNA gaps (42) (**Figure 3B**). In line with previous research, we observed an increased number of ssDNA gaps in BRCA1-KO and PCNA K164R mutant RPE1 cells under normal growth conditions (36, 43). BRCA1 cells reconstituted with BRCA1-GFP suppressed the formation of ssDNA gaps to the parental levels. However, complementation with the BRCA1-I26A-GFP mutant construct could not suppress ssDNA gaps formation and showed similar levels of ssDNA foci to PCNA-K164R mutant cells. Alternatively, to corroborate that the E3 activity of the BRCA1/BARD1 heterodimer prevents the formation of ssDNA gaps, we performed the S1-nuclease assay on DNA fibers (**Figure 3C**) in the presence of mild replication stress (42, 46). In this assay, DNA fibers containing ssDNA gaps would be shortened after treatment with S1 nuclease. Consistently, after mild replication stress (0.5 mM hydroxyurea for 2 hours), BRCA1-KO cells accumulated ssDNA gaps behind the replication fork, and this accumulation was suppressed after complementation with wild-type BRCA1-GFP but not by BRCA1-I26A-GFP.

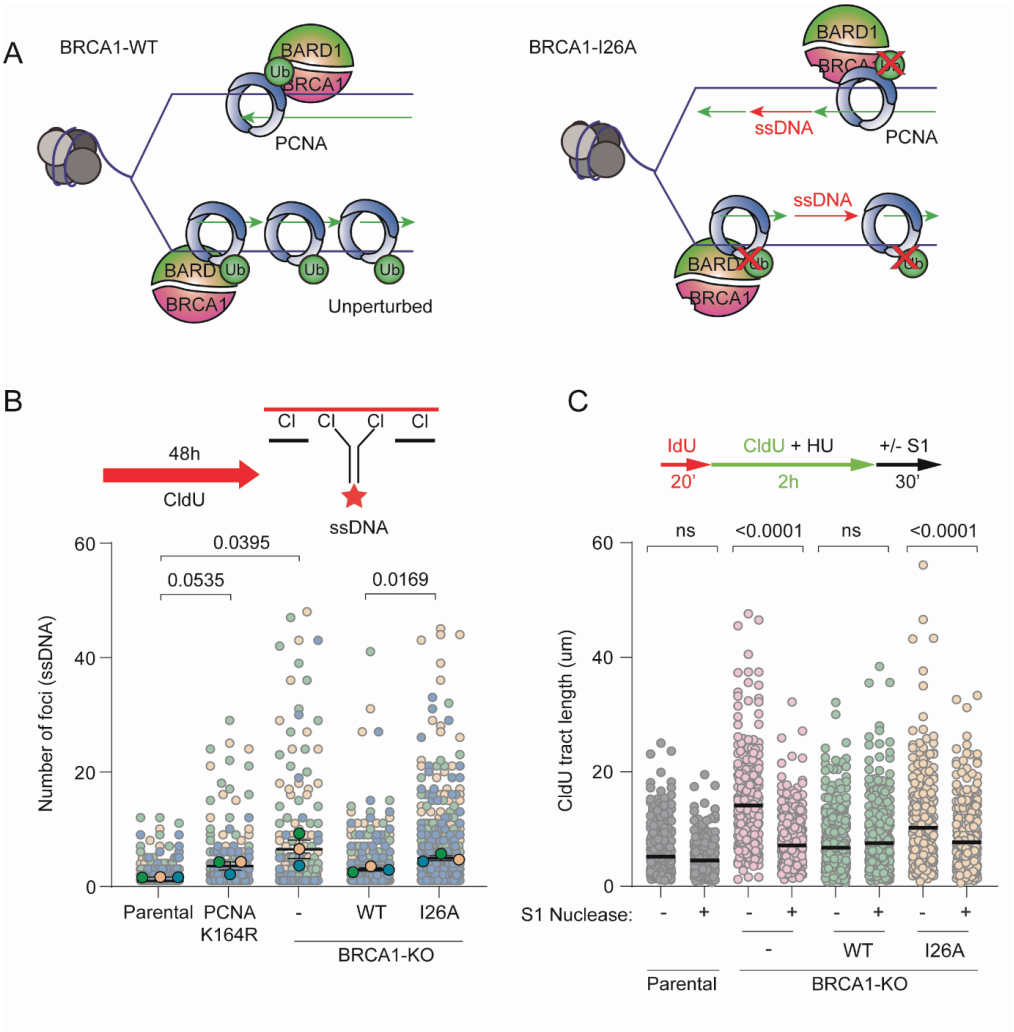


Figure 3. BRCA1-BARD1 mediated PCNA ubiquitination prevents ssDNA gaps. **A**. Proposed working model. While in wild type BRCA1 cells, BRCA1 would ubiquitinate PCNA in unperturbed conditions to facilitate continuous DNA synthesis, in an E3-deficient BRCA1-I26A mutant, deficiencies in PCNA ubiquitination in unperturbed conditions would promote the accumulation of ssDNA gaps. **B**. Cells were cultured in presence of CldU for 48h, fixed and analyzed by immunostaining against ssDNA. Number of ssDNA foci per cell is quantified. Each circle represents a cell. Average and standard deviation are depicted. Data corresponds to three independent experiments (N=3). p-values corresponds to two-tailed unpaired t-tests. **C**. Top: Scheme of the IdU/CldU pulse-labelling protocol, followed by S1 nuclease treatment. Bottom: CldU tract lengths in the indicated cell lines with and without S1 nuclease treatment. Each dot represents one fiber and the black bar represents the mean. At least 300 fibers were measured from two biological independent experiments (N=2). p-values were calculated using the Mann-Whitney test (ns: p>0.05).

BRCA1/BARD1 E3 activity is required for replication fork protection and BRCA1/BARD1 isomerization

PCNA K164 ubiquitination protect stalled replication forks from nuclease activity and subsequent accumulation of ssDNA (36). We hypothesized that upon replication fork blockade, deficiencies in PCNA ubiquitination by BRCA1/BARD1 will also result in fork de-protection and, due to nucleases activity, an increase in the exposed ssDNA, which can be measured by immunofluorescence. Therefore, we cultured Parental or BRCA1-KO cells complemented or not with either wild type BRCA1-GFP or BRCA1-I26A-GFP in the presence of CldU and treated them for 5 hours with 5 mM hydroxyurea to produce a complete replication fork blockade. Next, we analyzed the cells by immunofluorescence against RPA and exposed ssDNA (BrdU) and quantified the ssDNA signal (**Figure 4A**). Consistent with our hypothesis, BRCA1-deficient cells showed an increase in the intensity of ssDNA foci, and this effect was rescued by complementation with the wild type BRCA1-GFP construct but not with the BRCA1-I26A-GFP mutant construct.

Replication fork protection upon replication stress by BRCA1/BARD1 requires BRCA1 phosphorylation at S114 which promotes isomerization of BRCA1/BARD1. BRCA1 S114 phosphorylation enhances the interaction of the BRCA1/BARD1 heterodimer with RAD51 and increases its loading at stalled forks, and consequently, BRCA1-S114 phosphorylation can be used as a reporter for BRCA1/BARD1 isomerization (44). BRCA1-S114A mutants are deficient in fork protection (44) similar to PCNA-K164R mutants (36), and the BRCA1 ubiquitin E3-mutant I26A (**Figure 4A**). Therefore, we revisited the acquired mass spectrometry spectra of phosphorylation sites from our BRCA1-TULIP2 experiments (**Supplementary Figure 1**). In our experiment, we obtained good quality spectra for the tryptic peptide containing BRCA1-S114 both in its phosphorylated and non-phosphorylated form (**Supplementary Figure 3A**). However, while the signal for the unmodified peptide was similar for all three BRCA1-TULIP2 constructs, the I26A mutation prevented S114 phosphorylation from occurring (**Figure 4B-C**). This suggests that BRCA1 ubiquitin E3 activity is required for BRCA1 S114 phosphorylation and, as a consequence, BRCA1/BARD1 isomerization. Thus, we decided to test the sensitivity to hydroxyurea of the Parental and BRCA1-KO cells complemented or not with GFP-tagged constructs of either wild type or I26A mutant BRCA1. PCNA-K164R mutant was also included (**Figure 4D**). In contrast to what we observed for Olaparib (**Figure 1A**), where sensitivity of BRCA1-deficient cells could be rescued by complementation with the BRCA1-I26A mutant, in the case of hydroxyurea, while the complementation with the wild type GFP-BRCA1 construct could rescue the sensitivity of the BRCA1-deficient cells, the BRCA-I26A complemented cells were more sensitive than their wild-type counterpart. PCNA-K164R mutant cells were more sensitive than the I26A mutant, similar to BRCA1-deficient cells.

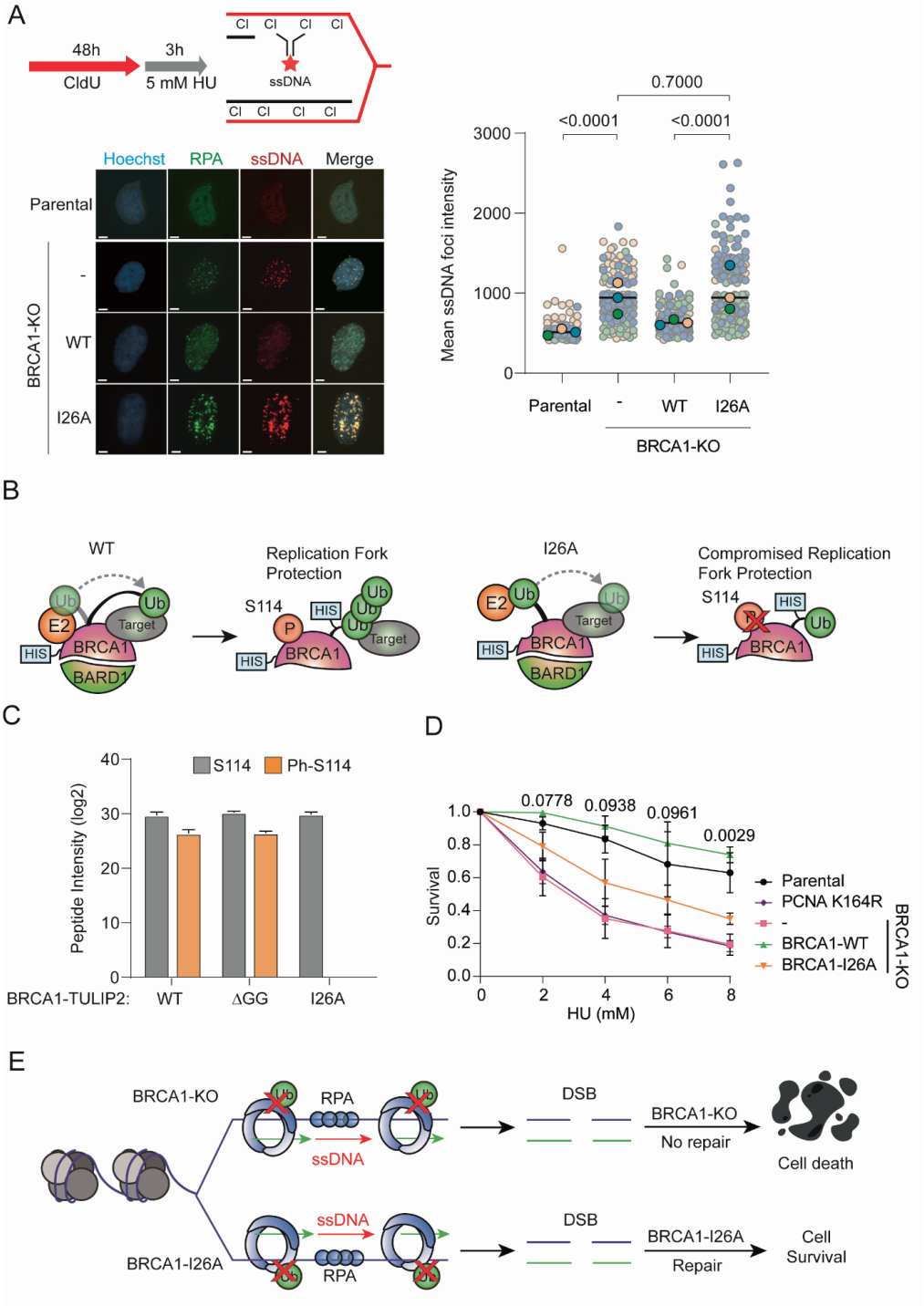

A
48h
CldU
3h
5 mM HU
Cl
ssDNA
Hoechst
RPA
ssDNA
Merge
Parental
BRCA1-KO
-
WT
I26A
0.7000
<0.0001
<0.0001
Mean ssDNA foci intensity
3000
2000
1000
0
Parental
-
WT
I26A
BRCA1-KO
B
WT
Replication Fork Protection
I26A
Compromised Replication Fork Protection
Ub
E2
Target
HIS
BRCA1
BARD1
S114
P
C
S114
Ph-S114
Peptide Intensity (log2)
40
30
20
10
0
BRCA1-TULIP2:
WT
ΔGG
I26A
D
0.0778
0.0938
0.0961
0.0029
Survival
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
HU (mM)
Parental
PCNA K164R
-
BRCA1-WT
BRCA1-I26A
BRCA1-KO
E
BRCA1-KO
RPA
ssDNA
DSB
No repair
Cell death
BRCA1-I26A
Repair
Cell Survival

Figure 4. BRCA1/BARD1 E3 activity is required for replication fork protection and isomerization. **A**. ssDNA intensity in Parental and BRCA1-KO cell lines rescued with either BRCA1-WT or BRCA1-I26A. Cells were cultured in presence of CldU for 48 h and treated with 5 mM hydroxyurea for 3 h before fixation and immunofluorescence analysis against ssDNA. Size bars in fluorescence microscopy images represent 10 µm. Each point represents a cell. Three independent experiments were performed (N=3), average and standard deviation are shown. Average of each independent experiment is displayed with an orange, green or blue circle respectively for experiment 1, 2 and 3. p-values were calculated using the Mann-Whitney test. **B**. Schematic representation of S114 phosphorylation identified by MS in BRCA1-WT-TULIP2 and BRCA1-I26A-TULIP2. **C**. Intensities of tryptic peptide containing BRCA1 S114 in its unmodified or phosphorylated forms. Average of four independent experiments is displayed. Error bars correspond to Standard deviation. **D**. Survival assay of Parental, PCNA-K164R mutant and BRCA1KO cell lines rescued with either BRCA1-WT-GFP or BRCA1-I26A-GFP against different hydroxyurea concentrations. Three independent experiments with two technical repeats were performed (N=3). Standard deviation and p-values from unpaired t-tests comparing BRCA1-WT-GFP and BRCA1-I26A-GFP complemented cells are displayed. **E**. Model representing the relevance of BRCA1/BARD1 E3 activity in preventing ssDNA gap accumulation and promoting replication fork stability by PCNA ubiquitination. In the absence of BRCA1/BARD1 E3 activity, ssDNA gaps accumulate, which can eventually lead to the formation of DSBs. While in the absence of BRCA1, these DSBs would lead to cell death, in E3-deficient mutants, the DSBs can still be repaired by Homologous Recombination.

DISCUSSION

Here, we provided novel insight in the role of the BRCA1/BARD1 ubiquitin E3 activity in the homologous recombination DNA repair pathway. Our results (**Figure 1A-C**) are in accordance with most studies and show that the E3 activity of BRCA1/BARD1 is not relevant for Homologous Recombination and Olaparib resistance (10, 18, 20). However, we noticed that cells which were deficient in the BRCA1/BARD1 ubiquitin E3 activity were sensitive to hydroxyurea (**Figure 4D**) and accumulated more ssDNA in response to high concentrations of hydroxyurea that produce a total replication fork blockade (**Figure 4A**), which indicated that deficiencies in the E3 activity of BRCA1/BARD1 did not affect the DSB repair function of HR but do affect the replication fork protection function of HR (47).

Using TULIP2 methodology in unperturbed conditions, we identified PCNA-K164 as a ubiquitination substrate for BRCA1/BARD1 but not for RAD18 (**Figure 1D-H; Figure 2C**). Accordingly, previous studies show synthetic lethality in BRCA1 deficient cells with either RAD18 loss or the use of REV1 inhibitors (46, 48). PCNA is known to be monoubiquitinated at low levels in mammalian cells in unperturbed conditions (36, 49, 50), and PCNA ubiquitination levels in unperturbed conditions were reduced in BRCA1-deficient cells (**Supplementary Figure 2B**), although still some PCNA ubiquitination could be detected. E3s other than RAD18 and BRCA1 have been described, like for example Cullin 4 RING ligase, CRL4 (51). However, while CRL4 works synergistically with RAD18 and CRL4 knockdowns affect the accumulation of mono-ubiquitinated PCNA in response to UV irradiation (51), no differences were observed in PCNA ubiquitination in response to UV light in BRCA1-deficient cells (**Supplementary Figure 2C-D**), which supports the model that BRCA1/BARD1 and RAD18 modify PCNA in alternative pathways (**Figure 2E**).

Recently, it has been proposed that the synthetic lethality between BRCA1-KO and Olaparib treatments come from and increased replication fork speed in BRCA1 deficient cells and

subsequent hyper accumulation on ssDNA gaps behind replication forks (42, 43). The accumulation of ssDNA gaps behind the fork have been also detected in PCNA-K164R mutants (36), and also when treating BRCA1-deficient cells with mild replicative stress like hydroxyurea concentrations not sufficient to produce a complete nucleotide depletion (42, 46). In this study, we show that the E3 activity BRCA1-I26A mutant show similar levels of ssDNA gaps formation as PCNA-K164R mutant and BRCA1-KO cells (**Figure 3B-C**), which seems, in principle, contradictory with the E3-deficient mutant being resistant to Olaparib treatment (Figure 1A; (10, 18, 20)). We propose a model (**Figure 4E**) in which upon replication stress the BRCA1/BARD1 E3 activity would promote replication fork stability and continuous, ssDNA-free, DNA synthesis. Deficiencies in the E3 activity would eventually promote the appearance of DNA DSBs. While in BRCA1-deficient cells these DSBs would lead to cell death, in the E3-deficient mutants could still be repaired through homologous recombination-mediated DSB repair.

Noteworthy, the excessive replication fork speed observed in BRCA1-deficient mutants ((42, 43, 46), **Figure 3C**), is, at least partially, rescued in the BRCA1-I26A mutant but not the ssDNA gaps formation. Thus, BRCA1/BARD1 inhibits the advance of replications fork in an E3-independent manner. We can speculate that this is mediated by the interaction of BRCA1 with the checkpoint protein ATR (52). Consistently, PARP inhibitors-resistant tumors, including BRCA1/2-revertant tumors which do not reconstitute RAD51 foci formation and Cyclin E-amplified tumors, are sensitive to ATR inhibition (53). Cyclin E overexpression also promotes ssDNA gap formation (48). Accordingly, fork slowing is an adaptive mechanism to persistent ATR inhibition (54).

Finally, some PARP inhibitor-sensitive BRCA1 tumors which mutations lead to disruption of the RING domain and a premature stop codon can revert by restoring the expression of a BRCA1 protein lacking the RING domain functionality but expressing the rest of the domains (19). Thus, treating these patients with other chemo therapeutic agents such as hydroxyurea, for which BRCA1/BARD1 E3-deficient cells are still sensitive (**Figure 4D**) might open a therapeutic opportunity for cancer treatment.

ACKNOWLEDGMENTS

Authors thank Daniel Durocher and Sylvie Noordermeer for sharing the TP53 -/- and TP53 -/- BRCA1 -/- hTERT RPE1 cell lines (3) used in this work. This work was supported by a Young Investigator Grant from the Dutch Cancer Society (KWF-KIG 11367/2017-2) and the EMERGIA 2020 program (EMERGIA20_00276) from the Andalusian Regional Government- Junta de Andalucía, Spain to RG-P. Work in the laboratory of A.C.O.V. has been supported by the European Research Council (ERC; grant 310913) and the Dutch Research Council (NWO; grant 724.016.003).

DECLARATION OF INTEREST

Authors declare no conflict of interests.

AUTHOR CONTRIBUTION

RG-P initiated the project. DS-L conducted most of the experimental work. LG assisted DS-L in RAD18-TULIP2 sample preparation. DS-L, LG and RG-P performed experiments in ACOV laboratory, NG-R performed the S1-nuclease fiber assay in PHS laboratory. AdR and PvV provided mass

spectrometry data acquisition support. RG-P and AdR acquired mass spectrometry data. DS-L and RG-P analyzed mass spectrometry data. LG was supervised by DS-L, DS-L was supervised by RG-P. DS-L and RG-P proposed and designed experiments and wrote the manuscript with input from other authors.

METHODS

Generation of TULIP2 constructs

BRCA1 and BARD1 genes were amplified and STOP codon removed by PCR using BP-tailed primers and cloned into pDNOR207 by Gateway® cloning BP reaction (Thermo Fisher Scientific). RAD18-pDNOR223 plasmid was obtained from Open Biosystems (Clone 1782) and the STOP codon was removed by site-directed mutagenesis. TULIP2 constructs were generated by Gateway® cloning LR reaction between donor plasmids containing BARD1, BRCA1 or RAD18 and acceptor TULIP2 plasmids (27). Primers are listed in supplementary table 1.

Cell culture

293T HEK and RPE1 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10 % Fetal Bovine Serum (FBS) and 100 U/mL penicillin/100 µg/mL streptomycin at 37°C and 5% CO2 unless specifically specified. Cells were regularly tested for mycoplasma contamination.

TULIP2 Lentivirus production

293T HEK cells were seeded at 30% confluency in a T175 flask containing 17 mL of DMEM + 10% FBS. After 24 hours, the transfection mixture was prepared by combining lentiviral packaging plasmids 7.5 µg pMD2.G (#12259, Addgene), 11.4 µg pMDLg-RRE (#12251, Addgene), 5.4 µg pRSV-REV (#12253, Addgene) and 13.7 µg TULIP2 plasmid with 114 µL of 1 mg/mL Polyethylenimine (PEI) in 2mL 150 mM NaCl. The mixture was vortexed and incubated 15 min at room temperature before adding to the HEK cells. Next day, culture medium was refreshed with DMEM 10% FBS. 72 hours post-transfection, lentiviral suspension was harvested by filtering through a 0.45 µm syringe filter (PN4184, Pall Corporation) and kept at -20 ºC for further use. Lentiviral particle concentration was determined using the HIV Type 1 p24 antigen ELISA Kit (ZeptoMetrix Corporation).

Generation of TULIP2 and stable 10xHis-Ub RPE1 cell lines

RPE1-hTERT TP53-/- and RPE1-hTERT TP53-/- BRCA1-/- cells were kindly provided by Dr. Sylvie Noordermeer (55) and were seeded in 15 cm diameter plates at 10% confluency with DMEM 10% FBS. Next day, cell culture medium was replaced with lentiviral TULIP2 constructs containing medium with 8 µg/mL polybrene. After 24 hours, medium was refreshed with DMEM 10% FBS, 1% Pen/Strep. 72 hours post-infection, TULIP2 positive clones were selected on puromycin.

RPE1-hTERT TP53-/- and RPE1-hTERT TP53-/- BRCA1-/- cells were infected using a bi-cistronic lentivirus encoding 10xHis-Ub and Puromycin resistant gene separated by an IRES (56) and selected on puromycin after 72 h.

Purification of TULIP2 conjugates

Following the TULIP2 methodology (27), five 15 cm diameter plates of RPE1 cells containing a TULIP2 construct, were grown up to 60% to 80% confluence. Expression of TULIP2 constructs was induced with 1µg/mL doxycycline once 60-80% confluence was reached. 24 h after doxycycline induction, cells were non-treated (NTC) or treated with different treatments. Proteasome inhibitor MG132 (Sigma Aldrich) was used at 10 µM for 5h, olaparib (Bio-connect) and bleomycin (Millipore) for 2 h at 10 µM and 5 µg/mL respectively. Treatment was performed at 40 J/m2 UV with 5 hours recovery.

5

After a particular treatment, cells were washed twice with ice-cold PBS and scraped. Next, cells were spun down and collected in 5 mL ice-cold PBS, 100 µL of sample was taken as input and lysed in 200 µL SNTBS buffer (2% SDS, 1% NP-40, 50mM TRIS pH 7.5, 150 mM NaCl). After additional centrifugation, cells were lysed in 10 mL Guanidinium buffer (6M guanidine-HCl, 0.1M Sodium Phosphate, 10mM TRIS, pH 7.8) and snap frozen in liquid nitrogen. After thawing, lysates were homogenized at room temperature by sonication at 80% amplitude during 5 s using a tip sonicator (Q125 Sonicator, QSonica, Newtown, USA). Sonication was performed twice. Subsequently, protein concentration was determined by BiCinchoninic Acid (BCA) Protein Assay Reagent (Thermo Scientific). After equalization, lysates were supplemented with 5 mM β-mercaptoethanol and 50 mM Imidazole pH 8.0. 100 µL of dry nickel-nitrilotriacetic acid-agarose (Ni-NTA) beads (QIAGEN), were equilibrated with Guanidinium buffer supplemented with 5 mM β-mercaptoethanol and 50 mM Imidazole pH 8.0. Equilibrated Ni-NTA beads were added to the cell lysates and incubated overnight at 4°C under rotation.

After lysate-beads incubation, Ni-NTA beads were transferred with Wash Buffer 1 (6 M Guanidine-HCl, 0.1M Sodium Phosphate, 10 mM Tris, 10 mM Imidazole, 5 mM β-mercaptoethanol, 0.2 % Triton X-100, pH 7.8) to an Eppendorf LoBind tube (Eppendorf). Subsequently, beads were washed with Wash buffer 2 (8 M Urea, 0.1M Sodium Phosphate, 10 mM Tris, 10 mM imidazole, 5mM β-mercaptoethanol, pH 8) and transferred to a new LoBind tube with Wash buffer 3 (8 M urea, 0.1M Sodium Phosphate, 10 mM TRIS, 10 mM imidazole, 5 mM β-mercaptoethanol, pH 6.3). Ultimately, beads were washed twice with Wash buffer 4 (8 M urea, 0.1M Sodium Phosphate, 10 mM TRIS, 5 mM β-mercaptoethanol, pH 6.3). After last wash, Ni-NTA beads were resuspended in 100 µL of 7 M urea, 0.1 M NaH_2PO_4/Na_2HPO_4, 0.01 M Tris/HCl, pH 7 and 10% of the sample was taken as pull down for immunoblotting.

Lys-C and trypsin digestion of TULIP2-purified conjugates

Ni-NTA beads were firstly digested with 500 ng recombinant Lys-C (Promega) at RT while shaking at 1,400 rpm. After 5h with Lys-C, urea buffer was diluted to <2M by adding 50 mM ABC. A second digestion was performed o/n at 37°C while shaking at 1,400 rpm using 500 ng of sequencing grade modified trypsin (Promega). Trypsin digested peptides were separated from Ni-NTA beads by filtering through a 0.45 µm filter Ultrafree-MC-HV spin column (Merck-Millipore).

Mass Spectrometry sample preparation

Digested peptides were acidified by adding 2% TriFlourAcetic (TFA) acid. Subsequently, peptides were desalted and concentrated on triple-disc C18 Stage-tips as previously described

(57). Stage-tips were in-house assembled using 200 µL micro pipet tips and C18 matrix (Sigma Aldrich). Stage-tips were activated by passing through 100 µL of methanol. Subsequently 100 µL of Buffer B (80% acetonitrile, 0.1% formic acid), 100 µL of Buffer A (0.1% formic acid), the acidified peptide sample, and two times 100 µL Buffer A were passed through the Stage-tip. Elution was performed twice with 25 µL of Elution buffer (32,5% acetonitrile, 0.1% formic acid solution).

Samples were vacuum dried using a SpeedVac RC10.10 (Jouan, France) and stored at -20°C. Prior to mass spectrometry analysis, samples were reconstituted in 10 µL 0.1% formic acid and transferred to autoload vials.

LC-MS/MS data acquisition

Mass spectrometry data was acquired either by and nanoLC Easy 1000 (Proxeon, Odense, Germany) coupled to a Q-Exactive mass spectrometer (Thermo, Bremen, Germany) (BARD1-TULIP2 samples) or Ultimate 3000 nano-gradient HPLC system (Thermo, Bremen, Germany), coupled to an Exploris480 mass spectrometer (Thermo, Bremen, Germany) (BRCA1-TULIP2; RAD18-TULIP2 and BARD1-TULIP2 with UV irradiation).

For the Q-Exactive, chromatography was performed as in (27), peptides were separated in an in-house packed with Reprosil-Pur C18-AQ 1.9 µm (Dr. Maisch, Ammerbuch, Germany) 20 cm analytical column in a 45 minutes gradient from 0% to 30% acetonitrile gradient in 0.1% Formic Acid followed of 20 minutes of column re-equilibration. The mass spectrometer was operated in a Data-Dependent Acquisition (DDA) mode with a top-7 method and a scan range of 300–1,600 m/z. Full-scan MS spectra were acquired at a target value of 3 × 106 and a resolution of 70,000, and the Higher-Collisional Dissociation (HCD) tandem mass spectra (MS/MS) were recorded at a target value of 1 × 105 and with a resolution of 35,000, an isolation window of 2.2 m/z, and a normalized collision energy (NCE) of 25%. The minimum AGC target was 1 ×104. The maximum MS1 and MS2 injection times were 250 and 120 ms, respectively. The precursor ion masses of scanned ions were dynamically excluded (DE) from MS/MS analysis for 20 s. Ions with charge 1, and >6, were excluded from triggering MS2 analysis.

For the Exploris480, samples were injected as in (58) onto a cartridge precolumn (300 µm × 5 mm, C18 PepMap, 5 µm, 100 A) with a flow of 10 µl/min for 3 minutes (Thermo, Bremen, Germany) and eluted via a homemade analytical nano-HPLC column (50 cm × 75 µm; Reprosil-Pur C18-AQ 1.9 µm, 120 A) (Dr. Maisch, Ammerbuch, Germany). The gradient was run from 2% to 38% solvent B (80% acetonitrile, 0.1% formic acid) in 120 min. The nano-HPLC column was drawn to a tip of ∼10 µm and acted as the electrospray needle of the MS source. The temperature of the nano-HPLC column was set to 50°C (Sonation GmbH, Biberach, Germany). The mass spectrometer was operated in data-dependent MS/MS mode for a cycle time of 3 seconds, with a HCD collision energy at 28 V and recording of the MS2 spectrum in the orbitrap, with a quadrupole isolation width of 1.2 Da. In the master scan (MS1) the resolution was 120,000, the scan range 350-1600, at an standard AGC target with maximum fill time of 50 ms. A lock mass correction on the background ion m/z=445.12 was used. Precursors were dynamically excluded after n=1 with an exclusion duration of 45 s, and with a precursor range of 10 ppm. Charge states 2-5 were included. For MS2 the scan range mode was set to automated, and the MS2 scan resolution was 30,000 at a normalized AGC target of 100% with a maximum fill time of 60 ms.

Mass Spectrometry data analysis

All raw data was analyzed using MaxQuant (version 1.6.7.0) as previously described (59). Search was performed against an in-silico digested UniProt reference proteome for Homo sapiens including canonical and isoform sequences (24th January 2022). Database searches were performed according to standard settings with the following modifications: Digestion with Trypsin/P was used, allowing 4 missed cleavages. Oxidation (M), acetyl (protein N-term), phospho (S,T) and GlyGly (K) for ubiquitination sites were allowed as variable modifications with a maximum number of 3. Label-Free Quantification (LFQ) was enabled, not allowing Fast LFQ while permitting iBAQ and matching between runs.

Output from MaxQuant Data were exported and processed for statistical analysis in the Perseus computational platform version 1.6.7.0 (60). LFQ intensity values were log2 transformed and potential contaminant and proteins either identify by site or only reverse peptides were removed. Samples were grouped in experimental categories and proteins not identified in 4 out of 4 replicates in at least one group were removed. Missing values were imputed using normally distributed values with 0.3 width and 1.8 down shift separately for each column. After imputation, statistical analysis was performed using two-sided Student's t tests. Results were exported into in MS Excel for a comprehensive browsing and visualization of the datasets. Volcano plots were constructed for data visualization using the VolcaNoseR web app (61) (https://huygens.science.uva.nl/VolcaNoseR2/).

LC-MS/MS data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (62)partner repository with the dataset identifier PXD039167

Purification of 10xHis-Ubiquitin conjugates

10xHis-Ub conjugates were purified using Ni-NTA beads as previously described (63). In brief, RPE1 cells were grown in three 15 cm dishes and were scraped and lysed in 6M guanidinium buffer. A small fraction of cells was separately lysed in SNTBS buffer as input control. After homogenization by sonication, lysates were incubated with Ni-NTA beads o/n at 4 °C. Next day, Ni-NTA beads were washed with buffers 1-4 and eluted in 7 M urea, 0.1 M NaH2PO4/Na2HPO4, 0.01 M Tris/HCl, pH 7.0, 500 mM imidazole pH7.

Electrophoresis and immunoblotting

Samples were separated on Novex 4-12% gradient gels (Thermo Fisher Scientific) using NuPAGE® MOPS SDS running buffer (50mM MOPS, 50mM Tris-base, 0.1% SDS, 1mM EDTA pH 7.7) and transferred onto Amersham Protran Premium 0.45 NC Nitrocellulose blotting membranes (GE Healthcare) using a Bolt Mini-Gel system (Thermo Fisher Scientific), which was used for both the gel electrophoresis and the protein transfer to the membrane according to vendor instructions. Membranes were stained with Ponceau-S (Sigma Aldrich) to determine total amount of protein loaded. Next, membranes were blocked with blocking solution (8% Elk milk, 0.1% Tween-20 in PBS) for 1 h prior to primary antibody incubation. Chemiluminescence reaction was initiated with Western Bright Quantum Western blotting detection kit (Advansta-Isogen) and measured in a

ChemiDocTM imaging system (BIO-RAD, Hercules, CA, USA). Antibodies are listed in supplementary table 2.

Clonogenic survival assay

RPE1 cells lines were seeded at 3000 cells/well in 6-well plates and allowed to attach overnight. TULIP2 constructs were induced with 1 µg/mL doxycycline prior treatment. Olaparib (Bio-connect) was added at different concentrations for 24 hours. After treatment, medium was refreshed. Hydroxyurea (Sigma) was added at different concentrations and medium was refreshed after 16 hours treatment. Subsequently, cells were allowed to grow for 10 days and fixed for 20 minutes in 4% paraformaldehyde (PFA) in PBS. Cells were stained with Crystal Violet 0.05% for 30 minutes and washed with water. Afterward, Crystal Violet was re-solubilized in methanol and O.D.595 was measured in the VICTOR X3 Multilabel Plate Reader 2030-0030 (Perkin Elmer). GraphPad was used for statistical analysis and the value of untreated cells was set at 100% survival.

Immunofluorescence

RPE1 cell lines were seeded at 20% confluency on 8 mm coverslips in 12-well plates and allowed to attach overnight. 1 µg/mL doxycycline was added to required cell lines to induce TULIP2 constructs 24 hours prior treatment. For RAD51 foci experiments, cells were treated with 10 Gy with 4 h recovery before fixation. For single-stranded DNA (ssDNA) gaps, medium was supplemented with 10 □M CldU (Sigma). After 48 hours, medium was replaced with either 10 □M olaparib or 5 mM hydroxyurea and cells were fixed with 1% PFA, 0.3% Triton X-100, 0.5% methanol after 3 hours treatment. RAD51 experiments were performed using a Leica SP8 confocal microscope taking 8 frames per image. ssDNA gaps experiments were performed in a ZEISS fluorescent microscope. ImageJ and GraphPad were used for quantification and statistical analysis.

Quantification and statistical analysis

Quantification of microscopy data was performed using Fiji-ImageJ and the statistical analysis was performed in GraphPad Prism 8. Statistical details of individual experiments can be found in figure legends, including the statistical test performed and definition of center and dispersion representation. For every analysis, N represents the number of values considered in the statistical analysis.

S1 nuclease assay

Cells were pulse-labeled with 20 µM IdU (20 min), washed twice with PBS and pulse-labeled with 200 µM CldU in the presence of 0.5 mM HU for 2h. Cells were then washed twice with PBS and permeabilized with CSK buffer (100 mM NaCl, 10 mM MOPS pH 7, 3 mM MgCl2, 300 mM Sucrose and 0.5% Triton X-100 in water) for 8 min at RT. Permeabilized cells were treated with S1 nuclease buffer (30 mM Sodium acetate pH 4.6, 10 mM Zinc acetate, 5% glycerol, 50 mM NaCl in water) with or without 20 U/ml S1 nuclease (Invitrogen, 18001-016) for 30 min at 37ºC. Cells were then scrapped in PBS + 0.1% BSA, pelleted and resuspended in PBS + 0.1% BSA at a final concentration of 1-2x103 cells/µl. 2.5 µl of cell suspension were spotted on a positively charged slide and lysed with 7.5 µl of spreading buffer (200 mM Tris-HCl pH 7.5, 50 mM EDTA, 0.5% SDS). After 8 min, slides were tilted at 45 degrees to allow the DNA to spread. Slides were then air-dried,

fixed with ice-cold methanol/acetic acid (3:1) for 5 mins, air-dried and stored at 4ºC. Slides were rehydrated with PBS, denatured with 2.5 M HCl for 1h, washed with PBS twice, and blocked with blocking buffer (3% BSA, 0.1% Triton X-100 in PBS) for 40 min. Next, slides were incubated with primary antibody mix of mouse anti-BrdU which recognizes IdU (Becton Dickinson #347580, 1:250) and rat anti-BrdU which recognizes CldU (Abcam #6326, 1:250) diluted in blocking buffer for 2.5h at RT in a dark humid chamber. Slides were washed 3 times with PBS for 5 min each and incubated with secondary antibodies anti-mouse Alexa fluor 594 and anti-rat Alexa fluor 488 (1:250, Invitrogen #A11005 and #A11006, respectively) in blocking buffer for 1h at RT in a dark humid chamber. After washing 3 times with PBS and air-drying, slides were mounted with Prolong gold antifade reagent (Invitrogen, P36930) and stored at 4ºC until imaging. Images were acquired using a AF6000 Leica Fluorescence microscope equipped with a HCX PL APO 63x (NA =1.4) oil objective. At least 300 fibers per condition were measured using the segmented line tool on ImageJ FIJI software (https://fiji.sc).

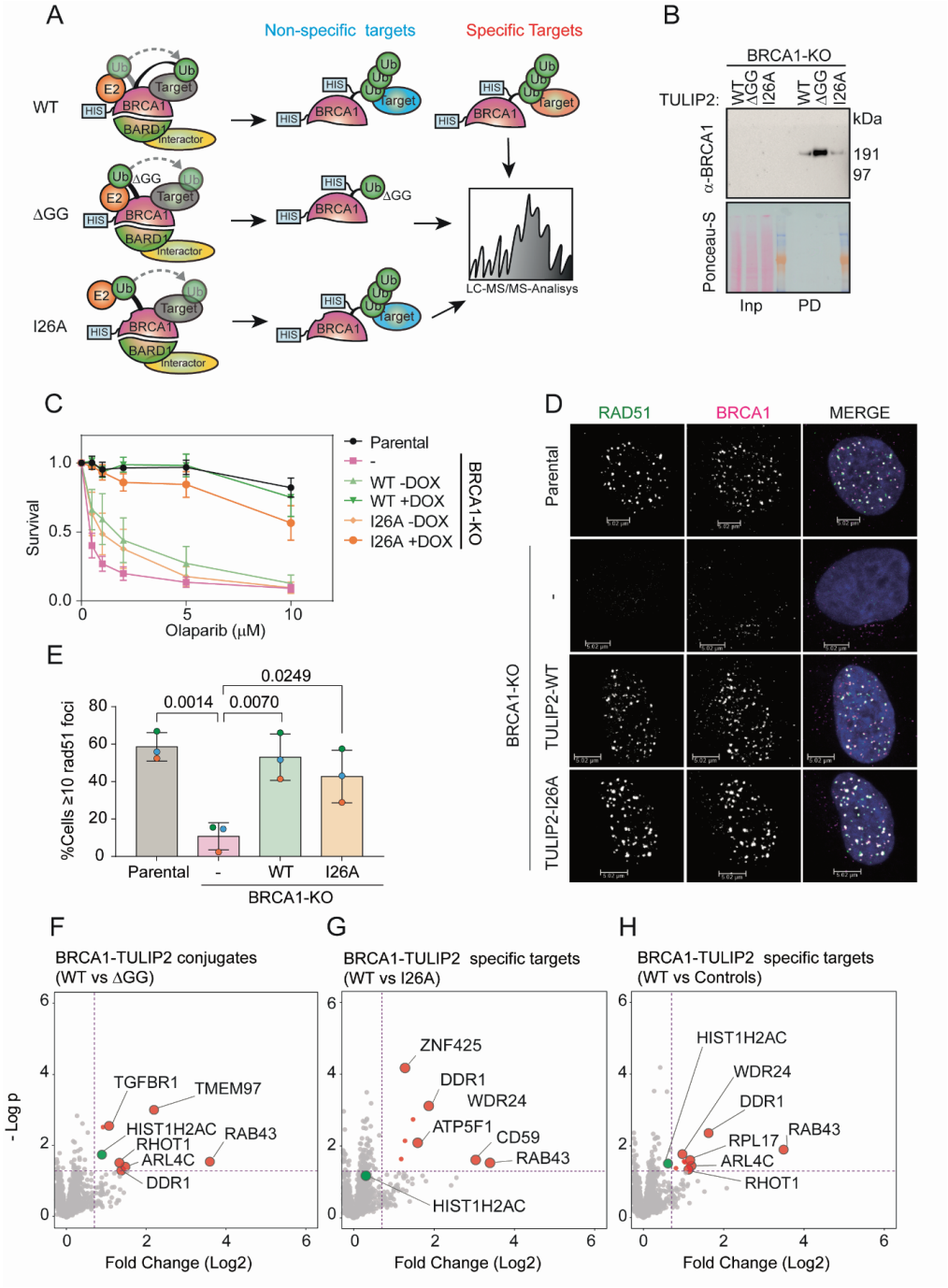
SUPLEMENTARY FIGURES AND DATASET LEGENDS

Dataset 1. BRCA1-TULIP2 RPE1 dataset. Spreadsheet including identified BRCA1-TULIP2 ubiquitination targets and proteins in the vicinity. Statistic comparisons and miscellaneous MS/MS proteomics values.

Dataset 2. RPE1 BARD1-TULIP2 dataset. Spreadsheet including identified BARD1-specific targets and ubiquitinated proteins in the vicinity. Statistic comparisons and miscellaneous MS/MS proteomics values.

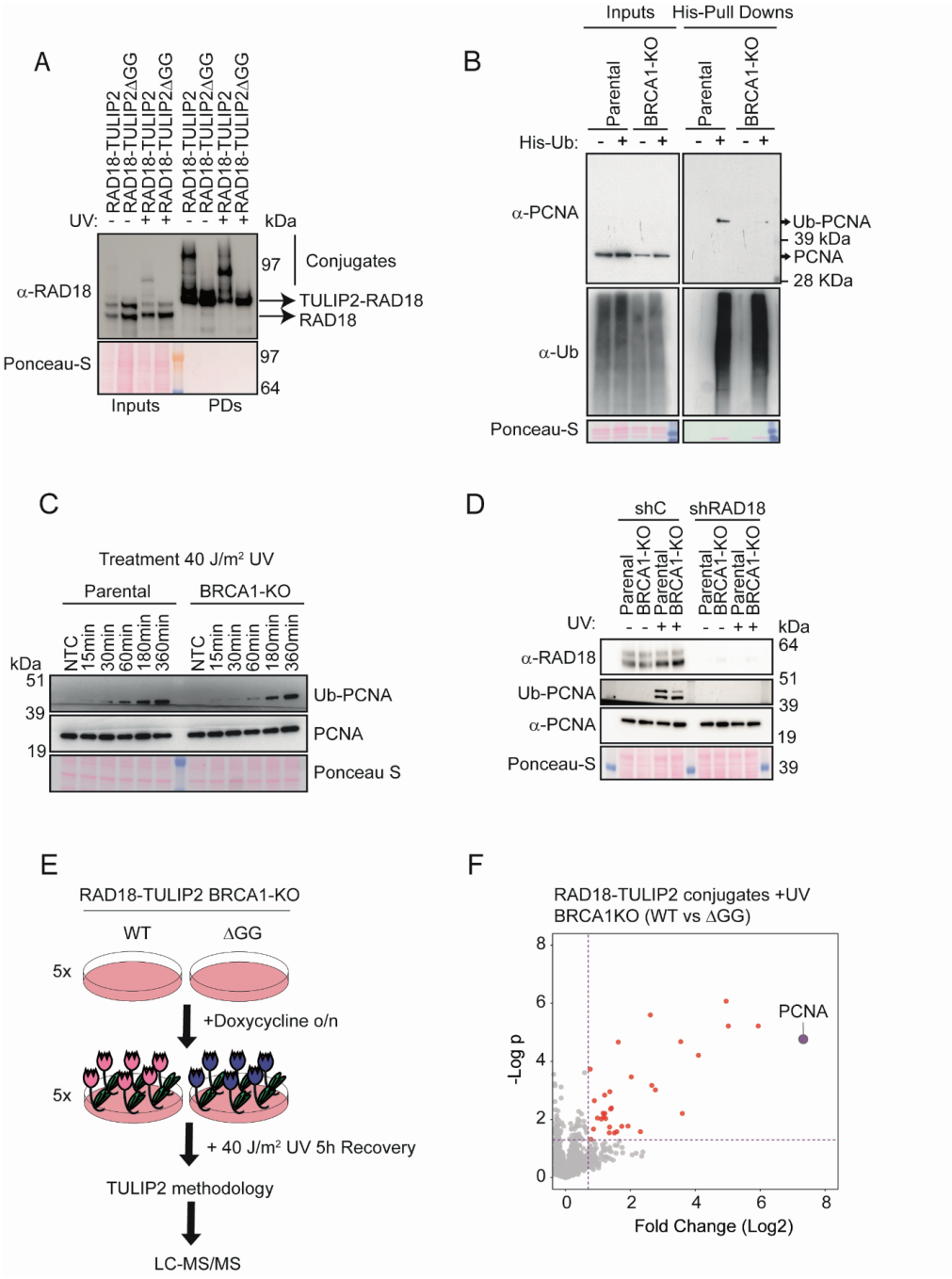
Dataset 3. PCNA-K164R dataset. Spreadsheet including identified BARD1-TULIP2 targets employing a PCNA-K164R mutant cell line. Statistic comparisons and miscellaneous MS/MS proteomics values.

Dataset 4. RAD18-TULIP2 dataset. Spreadsheet including identified RAD18-TULIP2 ubiquitinated proteins with and without UV damage. Statistic comparisons and miscellaneous MS/MS proteomics values.



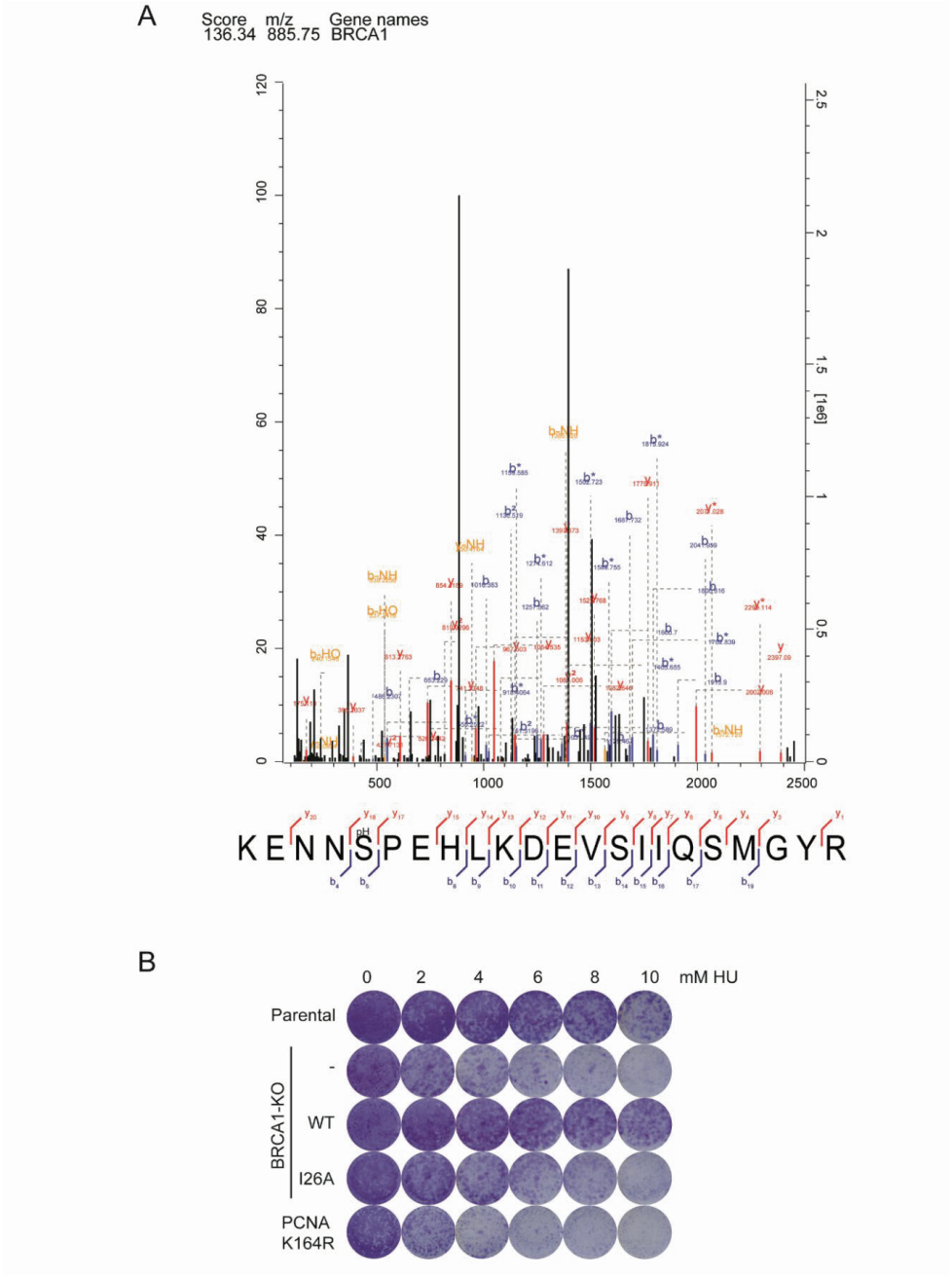
Supplementary Figure 1. **BRCA1-TULIP2 Characterization**. **A**. BRCA1-TULIP2 Rationale to find BRCA1 specific targets using Mass Spectrometry. **B**. Analysis by immunoblotting of TULIP2 pull downs in BRCA1-KO cells rescue with BRCA1-TULIP2 constructs (WT, ΔGG and I26A). **C**. Survival assay after treatment with Olaparib

of Parental and BRCA1-KO cells resued or not with either WT or I26A mutant BRCA1-TULIP2 constructs. Four independent experiments with 5 technical repeats were performed (N=4). Average and standard deviations are displayed. **D, E**. Analysis by immunofluorescence against RAD51 and BRCA1 of Parental and BRCA1-KO cells rescued or not with either WT or I26A BRCA1-TULIP2 constructs. Quantification of percentage (%) of cells with equal or more than 10 RAD51 foci (D) and representative images (E) are provided. Size bars in fluorescence microscopy images represent 10 µm. Three independent experiments were performed per condition displaying the average and standard deviation (N=3). The average of each independent experiment is represented by an orange, green or purple circle. P-values correspond to two-tailed unpaired t-tests. **F**. Volcano plot depicting statistical differences between BRCA1-WT and DGG TULIP2 constructs. Each dot represents a protein **G**. Volcano plot depicting statistical difference between BRCA1-WT and BRCA1-I26A TULIP2 constructs. Each point represents a protein. **H**. Volcano plot depicting statistical difference between BRCA1-WT-TULIP2 samples compared to ΔGG and BRCA1-I26A TULIP2 samples as controls. Each dot represents a protein.



Supplementary Figure 2. (A) Analysis by immunoblotting of RAD18-TULIP2 samples with and without 40 J/m^2 UV treatment. (B) Analysis by immunoblotting of the HIS-Ubiquitin proteome in Parental and BRCA1-KO cells. (C) Analysis by immunoblotting of PCNA ubiquitination after 40 J/m^2 UV irradiation in a time course manner in Parental and BRCA1-KO cells. (D) Analysis by immunoblotting against PCNA 5h after 40 J/m^2 UV

irradiation in Parental and BRCA1-KO cell after treating with a control or RAD18-targetting shRNA. (E) Experimental setup for RAD18-TULIP2 methodology in BRCA1KO cells 5h after 40 J/m^2 UV irradiation. (F). Volcano plot depicting statistical differences between RAD18-TULIP2 samples and RAD18-ΔGG-TULIP2 samples in BRCA1KO cells 5h after 40 J/m^2 UV irradiation. Each dot represents a protein, PCNA is highlighted.



Supplementary Figure 3. A. Accquired spectrum corresponding to the best localized oeotude fir BRCA1-S114 phosphorylation. B. Survival assay of Parental, PCNA-K164R mutant and BRCA1-KO cell lines rescued with either BRCA1-WT or BRCA1-I26A against different hydroxyurea concentrations. Three independent experiments with two technical repeats were performed (N=3).

Table 1. Primers used in this chapter.

Primer Name	Sequence	Use
BP-FW-BRCA1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCatggatttatctgctcttcgcgtt	BRCA1 TULIP2 Cloning
BP-RV-BRCA1-no-STOP	GGGGACCACTTTGTACAAGAAAGCTGGGTgtagtggcTgtggggatct	BRCA1 TULIP2 Cloning
BARD1		BARD1 TULIP2 Cloning
BARD1		BARD1 TULIP2 Cloning
BP-FW-RAD18	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCatggactccctggccgag	RAD18 TULIP2 Cloning
BP-RV-RAD18	GGGGACCACTTTGTACAAGAAAGCTGGGTaattcctattacgcttgtttcttggttcaatc	RAD18 TULIP2 Cloning
RemoveStop-RAD18-FW	GCTGAGATTGAACCAAGAAACAAGCGTAATAGGAATCCAACTTTCTTGTACAAAGTTGGC	RAD18 TULIP2 Cloning
RemoveStop-RAD18-RV	GCCAACTTTGTACAAGAAAGTTGGATTCCTATTACGCTTGTTTCTTGGTTCAATCTCAGC	RAD18 TULIP2 Cloning

Table 2. Antibodies used in this chapter.

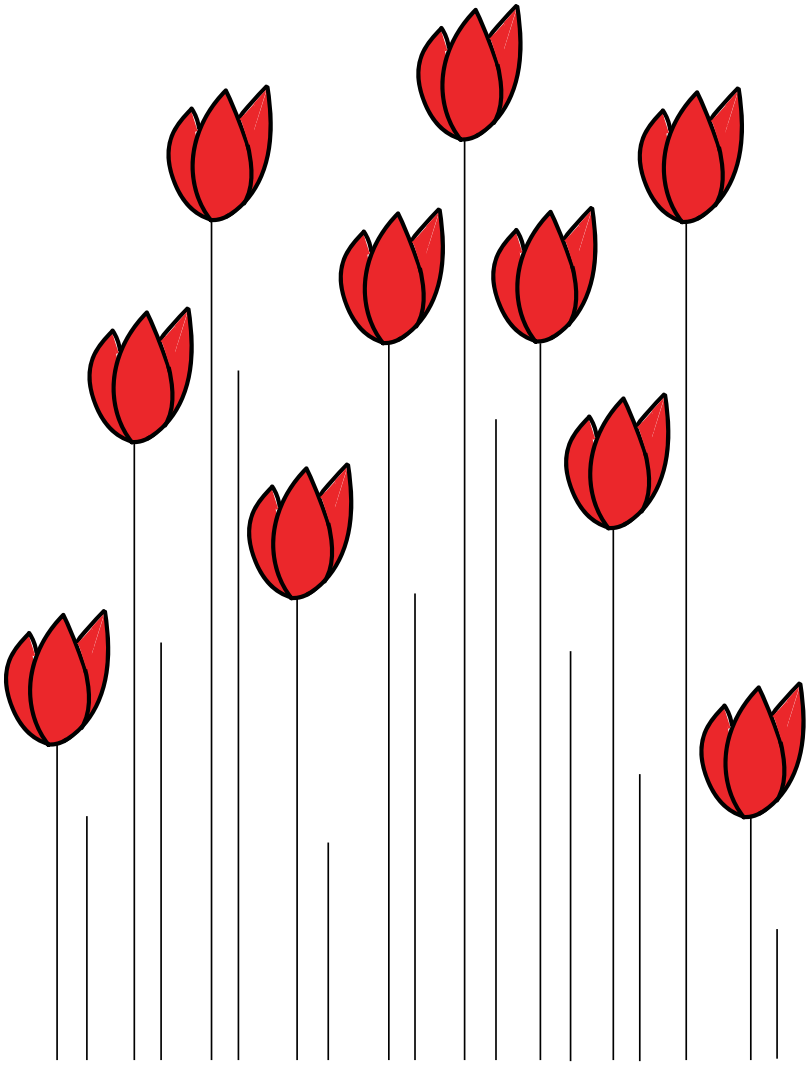
Antibody	Target	Dilution	Company
Primary Antibodies			
Rabbit anti-BARD1 (A300-263A)	BARD1	1:1000	Bethyl
Rabbit anti-BRCA1 (9010S)	BRCA1	1:1000	Cell Signaling Technology
Mouse anti-BRCA1 (OP92)	BRCA1	1:1000	Millipore
Mouse anti-PCNA (Sc-56)	PCNA	1:5000	Santa Cruz
Rabbit anti-RAD18 (9040S)	RAD18	1:1000	Cell Signaling Technology
Rabbit anti-RAD51 (70-001)	RAD51	1:15000	BioAcamedia
Rabbit anti-RPA1 (NA18-100UG)	RPA1	1:1000	Millipore
Mouse anti-Ubiquitin (Sc-8017)	Ubiquitin	1:1000	Santa Cruz
Rat anti-BrdU (ab6326)	BrdU	1:200	Abcam
Secondary Antibodies			
HRP-conjugated Donkey anti-Rabbit (31458)	Anti-Rabbit	1:5000	Thermo Fisher Scientific
HPR-conjugated Goat antiMouse IgG (H+L) (31432)	Anti-Mouse	1:5000	Thermo Fisher Scientific
Secondary Alexa Flour 594	Anti-Rat	1:1000	Thermo Fisher Scientific
Secondary Alexa Flour 488	Anti-Rabbit	1:1000	Thermo Fisher Scientific
Secondary Alexa Flour 594	Anti-Mouse	1:1000	Thermo Fisher Scientific

REFERENCES

1. P. S. Brzovic, P. Rajagopal, D. W. Hoyt, M. C. King, R. E. Klevit, Structure of a BRCA1-BARD1 heterodimeric RING-RING complex. Nat Struct Biol 8, 833-837 (2001).
2. R. M. Densham et al., Human BRCA1-BARD1 ubiquitin ligase activity counteracts chromatin barriers to DNA resection. Nature structural & molecular biology 23, 647-655 (2016).
3. S. M. Noordermeer et al., The shieldin complex mediates 53BP1-dependent DNA repair. Nature 560, 117-121 (2018).
4. W. Zhao et al., BRCA1-BARD1 promotes RAD51-mediated homologous DNA pairing. Nature 550, 360-365 (2017).
5. M. Tarsounas, P. Sung, The antitumorigenic roles of BRCA1-BARD1 in DNA repair and replication. Nat Rev Mol Cell Biol 10.1038/s41580-020-0218-z (2020).
6. A. Tubbs, A. Nussenzweig, Endogenous DNA Damage as a Source of Genomic Instability in Cancer. Cell 168, 644-656 (2017).
7. M. C. King, J. H. Marks, J. B. Mandell, G. New York Breast Cancer Study, Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2. Science 302, 643-646 (2003).
8. R. Scully, N. Puget, BRCA1 and BRCA2 in hereditary breast cancer. Biochimie 84, 95-102 (2002).
9. N. Weber-Lassalle et al., Germline loss-of-function variants in the BARD1 gene are associated with early-onset familial breast cancer but not ovarian cancer. Breast Cancer Res 21, 55 (2019).
10. R. Shakya et al., BRCA1 tumor suppression depends on BRCT phosphoprotein binding, but not its E3 ligase activity. Science 334, 525-528 (2011).
11. M. D. Stewart et al., BARD1 is necessary for ubiquitylation of nucleosomal histone H2A and for transcriptional regulation of estrogen metabolism genes. Proc Natl Acad Sci U S A 115, 1316-1321 (2018).
12. Q. Zhu et al., BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. Nature 477, 179-184 (2011).
13. R. Kalb, D. L. Mallery, C. Larkin, J. T. Huang, K. Hiom, BRCA1 is a histone-H2A-specific ubiquitin ligase. Cell Rep 8, 999-1005 (2014).
14. S. R. Witus et al., BRCA1/BARD1 site-specific ubiquitylation of nucleosomal H2A is directed by BARD1. Nature structural & molecular biology 28, 268-277 (2021).
15. B. J. Kim et al., The Histone Variant MacroH2A1 Is a BRCA1 Ubiquitin Ligase Substrate. Cell Rep 19, 1758-1766 (2017).
16. A. Sherker et al., Two redundant ubiquitin-dependent pathways of BRCA1 localization to DNA damage sites. EMBO reports 22, e53679 (2021).
17. D. Zong et al., BRCA1 Haploinsufficiency Is Masked by RNF168-Mediated Chromatin Ubiquitylation. Mol Cell 73, 1267-1281 e1267 (2019).
18. K. Nakamura et al., H4K20me0 recognition by BRCA1-BARD1 directs homologous recombination to sister chromatids. Nat Cell Biol 21, 311-318 (2019).
19. R. Drost et al., BRCA1 RING function is essential for tumor suppression but dispensable for therapy resistance. Cancer cell 20, 797-809 (2011).
20. L. J. Reid et al., E3 ligase activity of BRCA1 is not essential for mammalian cell viability or homology-directed repair of double-strand DNA breaks. Proc Natl Acad Sci U S A 105, 20876-20881 (2008).
21. M. Song, K. Hakala, S. T. Weintraub, Y. Shiio, Quantitative proteomic identification of the BRCA1 ubiquitination substrates. J Proteome Res 10, 5191-5198 (2011).
22. A. Matsuzawa et al., The BRCA1/BARD1-interacting protein OLA1 functions in centrosome regulation. Mol Cell 53, 101-114 (2014).
23. J. B. Heidelberger, S. A. Wagner, P. Beli, Mass Spectrometry-Based Proteomics for Investigating DNA Damage-Associated Protein Ubiquitylation. Frontiers in genetics 7, 109 (2016).

24. C. M. Eakin, M. J. Maccoss, G. L. Finney, R. E. Klevit, Estrogen receptor alpha is a putative substrate for the BRCA1 ubiquitin ligase. Proceedings of the National Academy of Sciences of the United States of America 104, 5794-5799 (2007).
25. J. K. Barrows, G. Fullbright, D. T. Long, BRCA1-BARD1 regulates transcription through BRD4 in Xenopus nucleoplasmic extract. Nucleic Acids Res 49, 3263-3273 (2021).
26. C. X. Deng, BRCA1: cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution. Nucleic Acids Res 34, 1416-1426 (2006).
27. D. Salas-Lloret, G. Agabitini, R. Gonzalez-Prieto, TULIP2: An Improved Method for the Identification of Ubiquitin E3-Specific Targets. Front Chem 7, 802 (2019).
28. R. Kumar, R. Gonzalez-Prieto, Z. Xiao, M. Verlaan-de Vries, A. C. O. Vertegaal, The STUbL RNF4 regulates protein group SUMOylation by targeting the SUMO conjugation machinery. Nature communications 8, 1809 (2017).
29. P. S. Brzovic et al., Binding and recognition in the assembly of an active BRCA1/BARD1 ubiquitin-ligase complex. Proceedings of the National Academy of Sciences of the United States of America 100, 5646-5651 (2003).
30. D. E. Christensen, P. S. Brzovic, R. E. Klevit, E2-BRCA1 RING interactions dictate synthesis of mono- or specific polyubiquitin chain linkages. Nature structural & molecular biology 14, 941-948 (2007).
31. X. Yu, S. Fu, M. Lai, R. Baer, J. Chen, BRCA1 ubiquitinates its phosphorylation-dependent binding partner CtIP. Genes & development 20, 1721-1726 (2006).
32. H. F. O'Connor et al., Ubiquitin-Activated Interaction Traps (UBAITs) identify E3 ligase binding partners. EMBO Rep 16, 1699-1712 (2015).
33. R. González-Prieto, A. C. O. Vertegaal, "TULIP: Targets of Ubiquitin Ligases Identified by Proteomics" in SUMOylation and Ubiquitination: Current and Emerging Concepts, V. G. Wilson, Ed. (Caister Academic Press, U.K., 2019), 10.21775/9781912530120.10 chap. 10, pp. 147-160.
34. N. Garcia-Rodriguez, R. P. Wong, H. D. Ulrich, Functions of Ubiquitin and SUMO in DNA Replication and Replication Stress. Front Genet 7, 87 (2016).
35. V. Akimov et al., UbiSite approach for comprehensive mapping of lysine and N-terminal ubiquitination sites. Nature structural & molecular biology 25, 631-640 (2018).
36. T. Thakar et al., Ubiquitinated-PCNA protects replication forks from DNA2-mediated degradation by regulating Okazaki fragment maturation and chromatin assembly. Nat Commun 11, 2147 (2020).
37. F. Tian et al., BRCA1 promotes the ubiquitination of PCNA and recruitment of translesion polymerases in response to replication blockade. Proceedings of the National Academy of Sciences of the United States of America 110, 13558-13563 (2013).
38. K. Y. Lee, K. Myung, PCNA modifications for regulation of post-replication repair pathways. Mol Cells 26, 5-11 (2008).
39. K. Watanabe et al., Rad18 guides poleta to replication stalling sites through physical interaction and PCNA monoubiquitination. The EMBO journal 23, 3886-3896 (2004).
40. K. N. Choe, G. L. Moldovan, Forging Ahead through Darkness: PCNA, Still the Principal Conductor at the Replication Fork. Mol Cell 65, 380-392 (2017).
41. W. Leung et al., PCNA-K164 ubiquitination facilitates origin licensing and mitotic DNA synthesis. bioRxiv 10.1101/2020.06.25.172361, 2020.2006.2025.172361 (2020).
42. N. J. Panzarino et al., Replication Gaps Underlie BRCA Deficiency and Therapy Response. Cancer research 81, 1388-1397 (2021).
43. K. Cong et al., Replication gaps are a key determinant of PARP inhibitor synthetic lethality with BRCA deficiency. Molecular cell 81, 3128-3144 e3127 (2021).
44. M. Daza-Martin et al., Isomerization of BRCA1-BARD1 promotes replication fork protection. Nature 571, 521-527 (2019).

45. D. Billing et al., The BRCT Domains of the BRCA1 and BARD1 Tumor Suppressors Differentially Regulate Homology-Directed Repair and Stalled Fork Protection. Molecular cell 72, 127-139 e128 (2018).
46. A. Taglialatela et al., REV1-Polzeta maintains the viability of homologous recombination-deficient cancer cells through mutagenic repair of PRIMPOL-dependent ssDNA gaps. Molecular cell 81, 4008-4025 e4007 (2021).
47. S. Tye, G. E. Ronson, J. R. Morris, A fork in the road: Where homologous recombination and stalled replication fork protection part ways. Semin Cell Dev Biol 113, 14-26 (2021).
48. S. Nayak et al., Inhibition of the translesion synthesis polymerase REV1 exploits replication gaps as a cancer vulnerability. Sci Adv 6, eaaz7808 (2020).
49. H. Arakawa et al., A role for PCNA ubiquitination in immunoglobulin hypermutation. PLoS biology 4, e366 (2006).
50. I. Unk et al., Human HLTF functions as a ubiquitin ligase for proliferating cell nuclear antigen polyubiquitination. Proceedings of the National Academy of Sciences of the United States of America 105, 3768-3773 (2008).
51. K. Terai, T. Abbas, A. A. Jazaeri, A. Dutta, CRL4(Cdt2) E3 ubiquitin ligase monoubiquitinates PCNA to promote translesion DNA synthesis. Molecular cell 37, 143-149 (2010).
52. M. Chiappa et al., Combinations of ATR, Chk1 and Wee1 Inhibitors with Olaparib Are Active in Olaparib Resistant Brca1 Proficient and Deficient Murine Ovarian Cells. Cancers (Basel) 14 (2022).
53. H. Kim et al., Combining PARP with ATR inhibition overcomes PARP inhibitor and platinum resistance in ovarian cancer models. Nature communications 11, 3726 (2020).
54. D. Dibitetto, A. Sanchi, E. J. Sanford, M. Lopes, M. B. Smolka, Fork Slowing and Reversal as an Adaptive Response to Chronic ATR Inhibition. bioRxiv 10.1101/2021.05.18.444697, 2021.2005.2018.444697 (2021).
55. D. E. Verver et al., Non-SMC Element 2 (NSMCE2) of the SMC5/6 Complex Helps to Resolve Topological Stress. Int J Mol Sci 17 (2016).
56. S. A. G. Cuijpers, E. Willemstein, A. C. O. Vertegaal, Converging Small Ubiquitin-like Modifier (SUMO) and Ubiquitin Signaling: Improved Methodology Identifies Co-modified Target Proteins. Molecular & cellular proteomics : MCP 16, 2281-2295 (2017).
57. J. Rappsilber, M. Mann, Y. Ishihama, Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat Protoc 2, 1896-1906 (2007).
58. D. Salas-Lloret et al., SUMO Activated Target Traps (SATTs) enable the identification of a comprehensive E3-specific SUMO proteome. bioRxiv 10.1101/2022.06.22.497173, 2022.2006.2022.497173 (2022).
59. S. Tyanova, T. Temu, J. Cox, The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat Protoc 11, 2301-2319 (2016).
60. S. Tyanova et al., The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods 13, 731-740 (2016).
61. J. Goedhart, M. S. Luijsterburg, VolcaNoseR is a web app for creating, exploring, labeling and sharing volcano plots. Sci Rep 10, 20560 (2020).
62. Y. Perez-Riverol et al., The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic acids research 47, D442-D450 (2019).
63. I. A. Hendriks, A. C. Vertegaal, Label-Free Identification and Quantification of SUMO Target Proteins. Methods Mol Biol 1475, 171-193 (2016).



6

Ubiquitinome profiling reveals in vivo UBE2D3 targets and implicates UBE2D3 in protein quality control

Zeliha Yalcin[1], **Daniel Salas-Lloret**[3,#], Karel Bezstarosti[2,#], Onno B. Bleijerveld[4,#], Vera Boersma[1], Román González-Prieto[3,5,6], Maarten Altelaar[4,7], Jeroen A.A. Demmers[2] and Jacqueline J.L Jacobs[1,*]

[#]These authors contributed equally

[1]Division of Oncogenomics, The Netherlands Cancer Institute, Amsterdam, The Netherlands

[2]Proteomics Center, Erasmus Medical Center, Rotterdam, The Netherlands

[3]Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, the Netherlands.

[4]Proteomics Facility, The Netherlands Cancer Institute, Amsterdam, The Netherlands

[5]Genome Proteomics Laboratory, Andalusian Center for Molecular Biology and Regenerative Medicine (CABIMER), University of Seville, Seville, Spain.

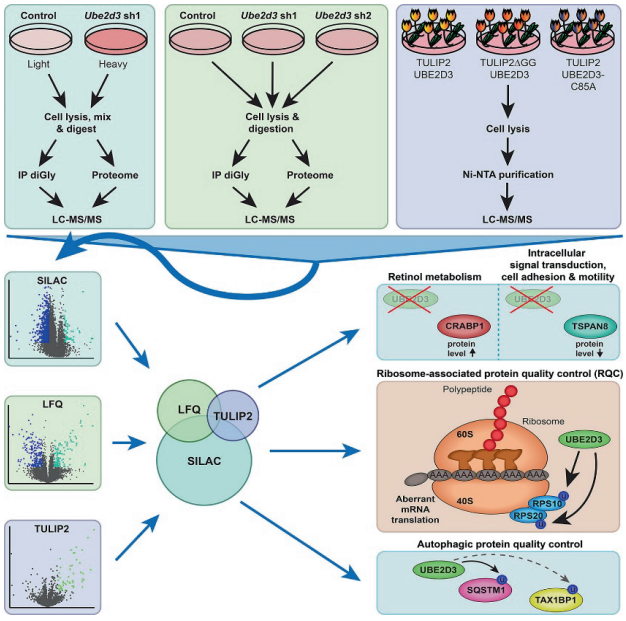
[6]Department of Cell Biology, University of Seville, Seville, Spain.

[7]Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, University of Utrecht, and Netherlands Proteomics Center, Utrecht, The Netherlands

This chapter has been published in Molecular & Cellular Proteomics. 22(6), 100548 (2023)

Abstract

Ubiquitination has crucial roles in many cellular processes and dysregulation of ubiquitin machinery enzymes can result in various forms of pathogenesis. Cells only have a limited set of ubiquitin-conjugating (E2) enzymes to support the ubiquitination of many cellular targets. As individual E2 enzymes have many different substrates and interactions between E2 enzymes and their substrates can be transient, it is challenging to define all *in vivo* substrates of an individual E2 and the cellular processes it affects. Particularly challenging in this respect is UBE2D3, an E2 enzyme with promiscuous activity *in vitro* but less defined roles *in vivo*. Here, we set out to identify *in vivo* targets of UBE2D3 by using SILAC-based and label-free quantitative ubiquitin diGly proteomics to study global proteome and ubiquitinome changes associated with UBE2D3 depletion. UBE2D3 depletion changed the global proteome, with the levels of proteins from metabolic pathways, in particular retinol metabolism, being the most affected. However, the impact of UBE2D3 depletion on the ubiquitinome was much more prominent. Interestingly, molecular pathways related to mRNA translation were the most affected. Indeed, we find that ubiquitination of the ribosomal proteins RPS10 and RPS20, critical for ribosome-associated protein quality control (RQC), is dependent on UBE2D3. We show by TULIP2 methodology that RPS10 and RPS20 are direct targets of UBE2D3 and demonstrate that UBE2D3's catalytic activity is required to ubiquitinate RPS10 *in vivo*. In addition, our data suggest that UBE2D3 acts at multiple levels in autophagic protein quality control (PQC). Collectively, our findings show that depletion of an E2 enzyme in combination with quantitative diGly-based ubiquitinome profiling is a powerful tool to identify new *in vivo* E2 substrates, as we have done here for UBE2D3. Our work provides an important resource for further studies on the *in vivo* functions of UBE2D3.



INTRODUCTION

Ubiquitination is a post-translational modification in which the conserved 76 amino-acid protein ubiquitin is attached to a lysine residue of a target protein by the combined actions of ubiquitin-activating (E1) enzymes, ubiquitin-conjugating (E2) enzymes and ubiquitin (E3) ligases. This mostly occurs through the formation of an isopeptide bond between the C-terminal glycine carboxyl moiety of ubiquitin and the ε-amino group of a lysine residue on the substrate protein, but ubiquitin can also be attached to the amino terminus of substrates. Deubiquitinating enzymes (DUBs) remove ubiquitin from targets, to maintain cellular ubiquitin homeostasis[1,2]. Ubiquitination of a target protein can have various consequences, such as proteasomal degradation or changes in protein activity, but it can also act in signal transduction or affect the cellular localization of the target protein. Proteins can be monoubiquitinated, in which one ubiquitin molecule is attached to one lysine residue of the target, or multi-monoubiquitinated, in which the target is modified with one ubiquitin molecule at multiple lysines. In addition, proteins can be polyubiquitinated, in which a ubiquitin-chain is formed on the initially attached ubiquitin. The linkage type can determine the consequence of the ubiquitination. For example, K48-linked ubiquitin chains are known to promote proteasomal degradation of the ubiquitinated protein, while K63-linked ubiquitin chains mainly play a role in signaling events[1,3]. Mammalian cells contain two E1 enzymes, ± 40 E2 enzymes, > 600 E3 enzymes and ± 100 DUBs[2,3]. Together, these enzymes need to coordinate all ubiquitination reactions in the cell and therefore have many substrates each. A major challenge is to identify all substrates of individual enzymes. Not only because of the number of substrates, but also because interactions between these enzymes and their substrates are often very transient and therefore difficult to detect. Identifying as many targets of ubiquitin-system enzymes as possible increases knowledge on the cellular processes affected by the ubiquitin-system, helps predicting possible secondary or adverse effects of targeting these enzymes and potentially helps the development of more specific drugs for treatment of pathologies.

In this study we focus on the discovery of new *in vivo* targets of UBE2D3 (also known as UBCH5C), a member of the UBE2D (or UBCH5) family of E2 enzymes[4]. *In vitro* UBE2D3 is a very promiscuous E2 enzyme. It is frequently used in *in vitro* ubiquitination assays, because it is often the most active E2 enzyme in such assays, functioning with almost every E3 ligase in ubiquitination of target proteins[5-8]. Therefore, many *in vitro* substrates and interacting E3 ligases have been identified for this E2. However, *in vivo* ubiquitination is highly controlled, and much less is known about the *in vivo* targets and E3 enzymes that UBE2D3 partners with to ubiquitinate these target proteins. Nevertheless, UBE2D3 has been linked to multiple important cellular processes *in vivo*. For instance, UBE2D3 contributes to keeping p53 tumor suppressor levels low in unstressed cells by acting with the E3 ligase MDM2 to ubiquitinate p53 and thereby target it for proteasomal degradation[9]. Also, UBE2D3 contributes to early stages of DNA repair by the homologous recombination repair machinery by acting with RNF138 to promote ubiquitination and accrual of the DNA end-resection promoting factor CtIP[10]. Moreover, aberrant UBE2D3 activity is associated with pathology. *UBE2D3* is mutated and its expression level is altered in a wide variety of cancers, including breast, ovarian, cervical, head and neck and esophageal cancer, melanoma, leukemia and multiple myeloma (Oncomine, EMBL-EBI Expression Atlas, Cosmic and ICGC databases). Furthermore, UBE2D3 levels affect responses to cancer cell treatment with radiation or all-trans

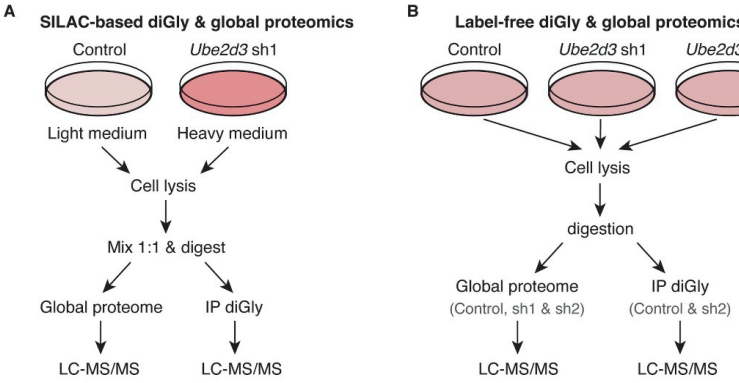
retinoic acid (RA)[11,12]. To better understand the function of UBE2D3, the biological processes it affects, and the potential consequences of targeting this enzyme in a disease setting, it is relevant to expand knowledge on the direct and indirect *in vivo* targets of this E2.

In the past years, different proteomic approaches have been developed to characterize the ubiquitinome and identify substrates of E3 ligases[13-15]. These are based on specific antibody-mediated enrichment of diGly peptides, which result from the trypsinization of ubiquitinated proteins[16,17]. Here, we combined ubiquitin diGly proteomics with stable isotope labeling of amino acids in cell culture (SILAC), allowing for quantitative identification of ubiquitinated proteins and mapping of modified lysines on the substrates[16-20]. We integrated this with RNA interference-mediated depletion of the E2 UBE2D3 to identify new *in vivo* substrates for UBE2D3, and used a modified TULIP2 (Targets of Ubiquitin Ligases Identified by Proteomics 2) methodology to confirm substrates as direct targets of UBE2D3[21]. We show how UBE2D3 affects the proteome and, much more prominently, the ubiquitinome. Our analysis reveals important roles for UBE2D3 in metabolic pathways, cell adhesion, cell signaling, mRNA translation and protein quality control. Most notably, we show that UBE2D3 regulates CRABP1 and TSPAN8 protein levels and that ubiquitination of the ribosomal proteins RPS10 and RPS20 by ZNF598, important for functional ribosome-associated protein quality control (RQC)[22-24], is dependent on UBE2D3 *in vivo*. We identify RPS10 and RPS20 as direct targets of UBE2D3 and show that the catalytic activity of UBE2D3 is indeed required for the ubiquitination of RPS10 *in vivo*.

Collectively, the quantitative ubiquitinome profiling in combination with UBE2D3 depletion that we performed here is a powerful source for insights into the spectrum of UBE2D3 substrates and the biological processes affected by UBE2D3, including its role in mRNA translation.

RESULTS

To identify *in vivo* UBE2D3 substrates, we investigated the effect of RNA interference-mediated UBE2D3 depletion on the global cellular proteome and ubiquitinome by combining affinity enrichment for ubiquitin diGly-modified peptides with SILAC-based proteomics. Control and *Ube2d3* shRNA1-transduced p53-deficient mouse embryonic fibroblasts (MEFs) were differentially labelled by culturing them in SILAC media. Cells were cultured for two weeks in either the light medium with normal lysine and arginine (control cells) or the medium with heavy lysine and arginine isotopes (*Ube2d3* sh1 cells), to completely label all proteins with heavy amino acids. Subsequently, cells were harvested, lysed and mixed at a 1:1 protein content ratio. The samples were digested and used for both global proteome analysis and ubiquitinome analysis after ubiquitin diGly-peptide enrichment (**Fig. 1A**). Experiments were performed as six biological replicates, all in a forward fashion and after confirming the reduction in UBE2D3 protein levels in heavy-labeled *Ube2d3* sh1 cells compared to light-labeled control cells (**Fig. S1A**). To exclude potential off-target effects of the shRNA used against *Ube2d3* (shRNA1) in the SILAC-based diGly proteomics analysis, we performed an independent validation including a second independent shRNA (shRNA2) targeting *Ube2d3*, in a label-free setting (LFQ). P53-deficient MEFs were transduced with Control shRNA, *Ube2d3* shRNA1 or *Ube2d3* shRNA2 and used for global proteome analysis and for ubiquitinome analysis after diGly-peptide enrichment (**Fig. 1B and Fig. S1B)**. Experiments were performed as three biological replicates.



C

	Numbers of identified proteins				
	Total # proteins	Total Up	Significantly Up	Total Down	Significantly Down
SILAC sh1	7389	98	35	171	91
LFQ sh1	3694	44	8	78	19
LFQ sh2	3703	72	24	79	21

D

	% of proteins affected in abundance by Ube2d3 depletion		
	Total Sign. (Up & Down)	Significantly Up	Significantly Down
SILAC sh1	1.7	0.5	1.2
LFQ sh1	0.7	0.2	0.5
LFQ sh2	1.2	0.6	0.6

E

diGly modified peptides SILAC Ube2d3 sh1		
Total in at least 2/6 exp.	Significantly Up	Significantly Down
13404	77	1186
% of diGly mod. peptides affected by Ube2d3 sh1		
Total Sign. (Up & Down)	Significantly Up	Significantly Down
9.4	0.6	8.8
diGly modified peptides LFQ Ube2d3 sh2		
Total in at least 2/3 exp.	Significantly Up	Significantly Down
5831	101	183
% of diGly mod. peptides affected by Ube2d3 sh2		
Total Sign. (Up & Down)	Significantly Up	Significantly Down
4.9	1.7	3.1

F

diGly modified proteins SILAC Ube2d3 sh1		
Total in at least 2/6 exp.	Significantly Up	Significantly Down
4912	67	665
% of diGly mod. proteins affected by Ube2d3 sh1		
Total Sign. (Up & Down)	Significantly Up	Significantly Down
14.9	1.4	13.5
diGly modified proteins LFQ Ube2d3 sh2		
Total in at least 2/3 exp.	Significantly Up	Significantly Down
2378	90	164
% of diGly mod. proteins affected by Ube2d3 sh2		
Total Sign. (Up & Down)	Significantly Up	Significantly Down
10.7	3.8	6.9

Figure 1. Ubiquitinome and global proteome profiling to identify novel *in vivo* targets of UBE2D3. **A**, Overview of the experimental approach used for SILAC-based quantitative ubiquitin diGly proteomics. Control and UBE2D3-depleted MEFs were cultured in SILAC medium with heavy or light isotopes. Cells were lysed, lysates were mixed in a 1:1 ratio (on total protein content) and digested. Both global proteome analysis and analysis for enriched ubiquitinated (diGly) peptides were performed for 6 biological replicates. ***B***, Setup LFQ global and diGly proteomics. Control and UBE2D3-depleted MEFs were lysed and digested. Global proteome analysis and analysis for enriched diGly peptides was performed for 3 biological

replicates. ***C***, Table with numbers of proteins identified in the global proteomics experiments for SILAC *Ube2d3* sh1 (in at least 2 out of 6 replicates) and LFQ *Ube2d3* sh1 and sh2 (in at least 2 out of 3 replicates), including numbers of proteins that are up- or downregulated in abundance in UBE2D3-depleted cells. Protein abudance is significantly up- or downregulated when $p \leq 0.05$. *D*, Table with percentage of proteins affected in abundance by UBE2D3 depletion for SILAC *Ube2d3* sh1 and LFQ *Ube2d3* sh1 and sh2. The total percentage of proteins significantly upregulated, downregulated and up- and downregulated in UBE2D3-depleted cells is shown. Significant when $p \leq 0.05$. Upregulated for SILAC sh1: $\log2 \geq 0.585$ and downregulated for SILAC sh1: $\log2 \leq -0.585$. Upregulated for LFQ sh1 and sh2: $\log2 \geq 1.0$; downregulated for LFQ sh1 and sh2: $\log2 \leq -1.0$). ***E,*** Table with diGly modified peptides and percentages of diGly modified peptides affected by UBE2D3 depletion. The top part shows the total number of identified diGly modified peptides in at least 2 out of 6 SILAC *Ube2d3* sh1 experiments and which of these are significantly up- or downregulated in their ubiquitination upon UBE2D3 depletion (significant: $p \leq 0.05$; up: $\log2 \geq 0.585$; down: $\log2 \leq -0.585$), and the total percentage of diGly modified peptides significantly upregulated, downregulated or up- and downregulated in UBE2D3-depleted cells for SILAC *Ube2d3* sh1. The bottom part shows the total number of identified diGly modified peptides in at least 2 out of 3 LFQ *Ube2d3* sh2 experiments and which of these are significantly up- or downregulated in their ubiquitination upon UBE2D3 depletion (significant: $p \leq 0.05$; up: $\log2 \geq 1.0$; down: $\log2 \leq -1.0$), and the total percentage of diGly modified peptides significantly upregulated, downregulated or up- and downregulated in UBE2D3-depleted cells for LFQ *Ube2d3* sh2. ***F***, Table with diGly modified proteins and percentages of diGly modified proteins affected by UBE2D3 depletion. The top part shows the total number of identified diGly modified proteins in at least 2 out of 6 SILAC *Ube2d3* sh1 experiments and which of these are significantly up- or downregulated in their ubiquitination upon UBE2D3 depletion (significant: $p \leq 0.05$; up: $\log2 \geq 0.585$; down: $\log2 \leq -0.585$), and the total percentage of diGly modified proteins significantly upregulated, downregulated or up- and downregulated in UBE2D3-depleted cells for SILAC *Ube2d3* sh1. The bottom part shows the total number of identified diGly modified proteins in at least 2 out of 3 LFQ *Ube2d3* sh2 experiments and which of these are significantly up- or downregulated in their ubiquitination upon UBE2D3 depletion (significant: $p \leq 0.05$; up: $\log2 \geq 1.0$; down: $\log2 \leq -1.0$), and the total percentage of diGly modified proteins significantly upregulated, downregulated or up- and downregulated in UBE2D3-depleted cells for LFQ *Ube2d3* sh2.

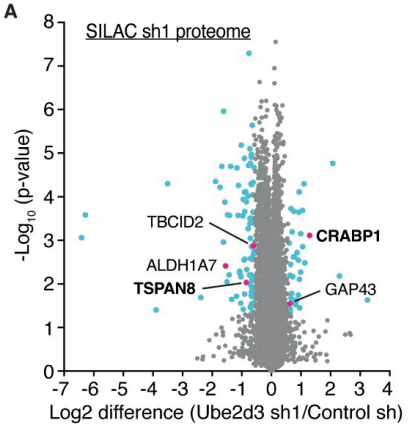
The experiments were performed without addition of proteasome inhibitors, causing some targets of UBE2D3 to be potentially missed due to proteasomal degradation, but allowing assessment of the dynamics of protein abundance upon UBE2D3 depletion. For the SILAC-based proteomics, typically over 6300 proteins were identified per experiment (**Fig. S1C and Table S1**). Overall, on average over the 6 replicates, 130 proteins were $\geq$ 1.5 fold increased in abundance, while 195 proteins were $\geq$ 1.5 fold decreased in abundance upon UBE2D3 depletion, per experiment (**Fig. S1C**). For 35 of these proteins the increase was statistically significant ($p \leq 0.05$), and for 91 proteins their decrease was statistically significant (**Fig. 1C, Fig. 2A and Table S2**). This translates into UBE2D3 significantly affecting the levels of ± 1.7% of the total number of detected proteins, of which ± 0.5% increased in abundance and ± 1.2% decreased in abundance (**Fig. 1D**). For the LFQ proteomics, the results show that with both shRNAs around 3700 proteins were identified in at least 2 out of 3 experiments (**Fig. 1C and Table S4, S5**). A total of 8 proteins were significantly increased in abundance and 19 proteins were significantly decreased in abundance in *Ube2d3* sh1-transduced cells. In *Ube2d3* sh2-transduced cells, 24 proteins were significantly increased in abundance and 21 proteins were significantly decreased in abundance (**Fig. 1C, Fig. 2B, C and Table S6**). Consistent with the SILAC-based proteome data obtained with *Ube2d3* sh1, also these LFQ-proteome data show that UBE2D3 only regulates the abundance of a small percentage of the total number of identified proteins (**Fig. 1D**). Combined, the results of the SILAC-

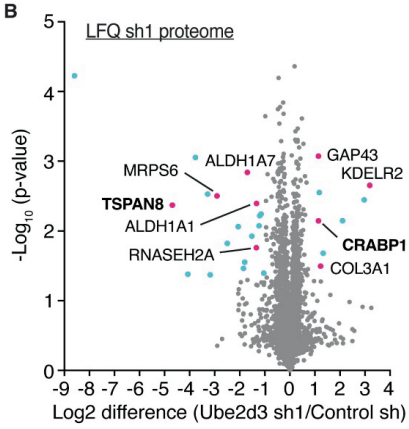
based and label-free approaches indicate that UBE2D3 not only decreases the levels of certain proteins by for instance promoting their degradation, but also (indirectly) increases the levels of certain other proteins. However, the low numbers of proteins affected indicate that UBE2D3 only has a modest role in regulating protein abundance.

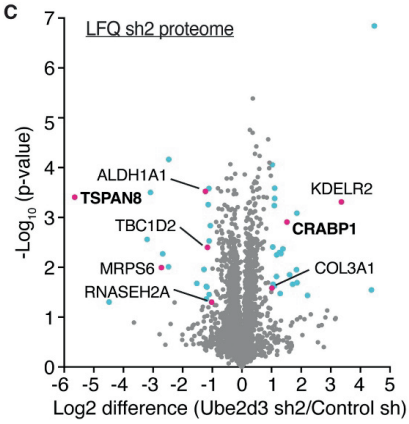
SILAC-based ubiquitinome analysis revealed over 19,000 diGly peptides over all experiments. In total, almost 500 peptides were ≥ 1.5 fold increased in ubiquitination, while over 3,500 peptides were ≥ 1.5 fold decreased in ubiquitination upon UBE2D3 depletion (**Fig. S1D**). As can be expected with depleting an E2 enzyme that would most likely result in loss of ubiquitination of its targets, the number of peptides that decreased in ubiquitination is substantially higher than the number of peptides that increased in ubiquitination. Over 13,000 peptides were diGly modified in at least 2 out of 6 experiments, of which 77 peptides were significantly ($p \leq 0.05$) increased in their ubiquitination and 1186 peptides were significantly decreased in their ubiquitination in UBE2D3-depleted cells (**Fig. 1E and Table S7-9**). This corresponds to ± 4,900 diGly modified proteins, of which 67 were significantly increased (≥ 1.5 fold; $p \leq 0.05$) in their ubiquitination and 665 were significantly decreased in their ubiquitination in UBE2D3-depleted cells (**Fig. 1F**). Thus, we found UBE2D3 to significantly affect the ubiquitination status of ± 9.5% of all detected diGly modified peptides, corresponding to ± 15% of the total amount of diGly modified proteins. Of these, 0.6% of the diGly modified peptides and 1.4% of the diGly modified proteins were increased in their ubiquitination, and 8.8% of the diGly modified peptides and 13.5% of the diGly modified proteins were decreased in their ubiquitination upon depletion of UBE2D3 (**Fig. 1E, F**). Label-free ubiquitinome analysis revealed 5831 diGly modified peptides over all experiments (**Fig. 1D**). A total of 101 peptides, representing 90 unique proteins, were significantly increased in their ubiquitination upon UBE2D3 depletion, while 183 peptides, representing 164 unique proteins, were decreased in their ubiquitination (≥ 2.0 fold; $p \leq 0.05$) (**Fig. 1E, F and Table S10, S11**). These results show that UBE2D3 significantly affects the ubiquitination of ± 5% of all detected diGly modified peptides, corresponding to ± 11% of all diGly modified proteins, in at least 2 out of 3 experiments (**Fig. 1E, F**). Thus, in contrast to the promiscuous behavior of UBE2D3 seen under *in vitro* conditions, *in vivo* UBE2D3 is a quite selective E2 enzyme, only affecting a small portion of the total number of identified diGly-modified proteins and peptides.

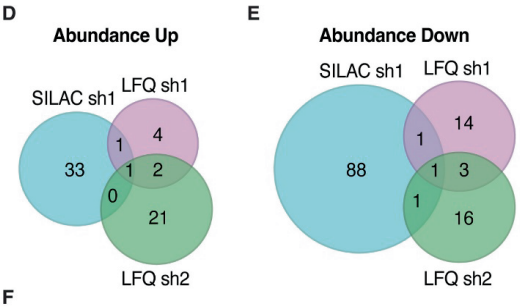
Global proteome changes upon UBE2D3 depletion

Ubiquitination of proteins by UBE2D3 might change the stability of its direct or indirect target proteins by affecting their proteasomal degradation. Direct targets of UBE2D3 are proteins that are directly ubiquitinated through association of UBE2D3 with an E3 ligase and subsequent transfer of the ubiquitin to the target protein, while indirect targets of UBE2D3 are not directly subject of UBE2D3-mediated ubiquitination, but are indirectly affected in their ubiquitination and/or stability as a result of UBE2D3-mediated ubiquitination of other proteins. Therefore, we first focused on the 126 proteins that significantly changed ≥ 1.5-fold in abundance upon UBE2D3 depletion in the SILAC-based global proteomics (**Fig. 2A and Table S2**). As the quantitative accuracy of SILAC mass spectrometry is higher than that of label-free mass spectrometry, we increased the cut-off for the LFQ-proteome data and selected proteins with a ≥ 2-fold and significant (p-value ≤ 0.05) change in abundance for further analysis (**Fig. 2B, C and Table S6**).









F

Overlapping hits					
	Name	SILAC sh1	LFQ sh1	LFQ sh2	Biological function
Up	COL3A1		X	X	Alpha1 chain of collagen type III
Up	**CRABP1**	X	X	X	Retinoic acid-binding protein
Up	GAP43	X	X		Neuronal growth cone protein
Up	KDELR2		X	X	ER resident protein-binding receptor
Down	ALDH1A1		X	X	Retinol dehydrogenase
Down	ALDH1A7	X	X		Aldehyde dehydrogenase
Down	MRPS6		X	X	28S Ribosomal protein
Down	RNASEH2A		X	X	Catalytic subunit of RNase H II
Down	TBC1D2	X		X	RAC1 effector
Down	**TSPAN8**	X	X	X	Cell surface glycoprotein

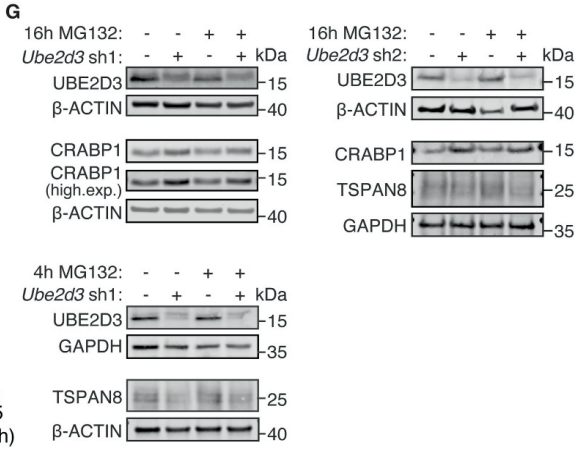


Figure 2. UBE2D3 regulates the abundance of CRABP1 and TSPAN8. ***A***, Volcano plot showing changes in total protein levels upon UBE2D3 depletion with shRNA1 in the SILAC-based proteomics (*n*=6). Blue dots

represent proteins that are at least 1.5-fold up- or downregulated in abundance (log2 ≥ 0.585 and log2 ≤ -0.585 respectively) and have a p-value of ≤ 0.05. Red dots indicate hits overlapping with LFQ proteomics with sh1 and/or sh2. Bold proteins overlap between SILAC *Ube2d3* sh1, LFQ *Ube2d3* sh1 and LFQ *Ube2d3* sh2. ***B***, Volcano plot showing changes in total protein levels upon UBE2D3 depletion with shRNA1 in label-free quantification (LFQ; *n*=3). Blue dots represent proteins that are at least 2-fold up- or downregulated in abundance (log2 ≥ 1.0 and log2 ≤ -1.0 respectively) and have a p-value of ≤ 0.05. Red dots indicate hits overlapping with SILAC proteomics with *Ube2d3* sh1 and/or LFQ proteomics with *Ube2d3* sh2. Bold proteins overlap between SILAC *Ube2d3* sh1, LFQ *Ube2d3* sh1 and LFQ *Ube2d3* sh2. ***C***, Volcano plot showing changes in total protein levels upon UBE2D3 depletion with shRNA2 in LFQ (*n*=3). Blue dots represent proteins that are at least 2-fold up- or downregulated in abundance (log2 ≥ 1.0 and log2 ≤ -1.0 respectively) and have a p-value of ≤ 0.05. Red dots indicate hits overlapping with SILAC with proteomics *Ube2d3* sh1 and/or LFQ proteomics with *Ube2d3* sh1. Bold proteins overlap between SILAC *Ube2d3* sh1, LFQ *Ube2d3* sh1 and LFQ *Ube2d3* sh2. ***D***, Venn diagrams illustrating the overlap between proteins that go up in abundance in SILAC *Ube2d3* sh1, LFQ *Ube2d3* sh1 and LFQ *Ube2d3* sh2. ***E***, Venn diagrams illustrating the overlap between proteins that go down in abundance in SILAC *Ube2d3* sh1, LFQ *Ube2d3* sh1 and LFQ *Ube2d3* sh2. ***F***, Table of overlapping proteins in B and C, including their biological functions. Bold proteins overlap between SILAC *Ube2d3* sh1, LFQ *Ube2d3* sh1 and LFQ *Ube2d3* sh2. Orange colored proteins are part of retinol metabolism and signaling pathways. ***G***, Immunoblot analysis of CRABP1, TSPAN8 and UBE2D3 protein levels in MEFs with or without depletion of *Ube2d3* and untreated (DMSO) or treated with 10 µM of proteasome inhibitor MG132. For *Ube2d3* shRNA1 (sh1) cells were treated with 10 µM MG132 for 16h for assessment of CRABP1 protein levels and for 4h for assessment of TSPAN8 protein levels. For *Ube2d3* shRNA2 (sh2) cells were treated with 10 µM MG132 for 16 hours for the assessment of CRABP1 and TSPAN8 protein levels. B-ACTIN and GAPDH serve as loading controls. Representative blots of 3 independent experiments are shown (2 additional biological replicates for CRABP1 and TSPAN8 levels in *Ube2d3* sh1 cells and 3 additional biological replicates in *Ube2d3* sh2 cells can be found in Supplementary Figure 2A).

When we overlaid the hits from the SILAC proteomics (SILAC sh1), the label-free quantification (LFQ) with *Ube2d3* sh1 (LFQ sh1) and the LFQ with *Ube2d3* sh2 (LFQ sh2), we found only CRABP1 to be significantly increased in abundance in all 3 experiments, and only TSPAN8 to be significantly decreased in abundance in all 3 experiments (**Fig. 2D-F**). When assessing the overlap between 2 out of 3 SILAC or LFQ experiments, we found 3 additional proteins that increased in abundance and 5 additional proteins that decreased in abundance (**Fig. 2D-F and Table S3**). In general, hits from the SILAC-based proteomics with *Ube2d3* sh1 that were not overlapping with the LFQ sh1 and LFQ sh2 experiments, were either not identified in the LFQ experiments or were altered just below the cut-offs that we set.

CRABP1 is a protein that responds to retinol (vitamin A) signaling by binding to RA (a metabolite of retinol). In the nucleus, RA binds and activates RA receptors, eventually resulting in transcription of target genes and thereby regulating differentiation and proliferation[36]. CRABP1 sequesters RA in the cytoplasm and thereby prevents activation of gene expression[36-38]. As CRABP1 levels increase upon UBE2D3 depletion in the SILAC and LFQ diGly proteomics data (**Fig. 2A-C, F**), UBE2D3 might promote degradation of CRABP1 and thereby impact RA signaling. To address this hypothesis, we assessed CRABP1 protein levels in immunoblots of protein lysates from control cells and UBE2D3-depleted cells, treated with and without the proteasome inhibitor MG132 (**Fig. 2G, Fig. S2A)**. Indeed, we found CRABP1 protein levels to be increased upon UBE2D3 depletion, thereby validating the proteomics results. However, MG132 treatment of up to 16 hours, which

was the maximum duration tolerated by the cells, did not lead to clearly visible CRABP1 protein stabilization, which precluded us from confirming that UBE2D3 promotes proteasomal degradation of CRABP1 (**Fig. 2G, Fig. S2A,B**). Nevertheless, q-RT PCR analysis indicated that the increased CRABP1 protein levels upon UBE2D3 depletion are not accompanied by increased *Crabp1* mRNA levels, indicating that UBE2D3 regulates CRABP1 on protein level (**Fig. S2C**).

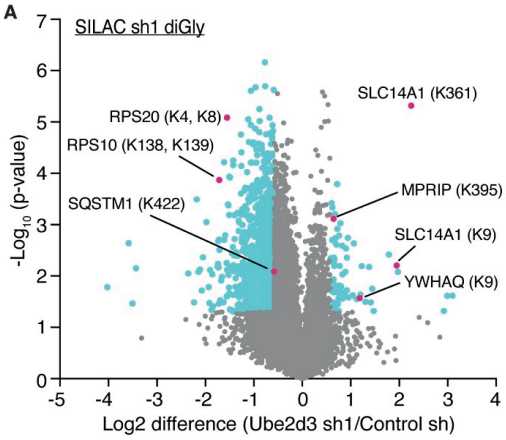
TSPAN8 is a cell surface glycoprotein belonging to the transmembrane protein family of tetraspanins that function in intracellular signal transduction, cell adhesion and motility. It forms a complex with integrins, which promote cell-cell and cell-matrix adhesion[39-43]. A decrease in TSPAN8 levels in UBE2D3-depleted cells suggests that UBE2D3 indirectly regulates TSPAN8 levels (**Fig. 2A-C, F**). To independently validate the impact of UBE2D3 on TSPAN8 protein levels and to investigate if this is dependent on the proteasome, we assessed TSPAN8 protein levels in immunoblots of protein lysates from control cells and UBE2D3-depleted cells, treated with and without the proteasome inhibitor MG132 (**Fig. 2G, Fig. S2A**). Indeed, UBE2D3-depleted cells express lower levels of TSPAN8 protein, confirming the proteomics results. However, MG132 barely affected TSPAN8 protein levels over the duration of treatment, prohibiting us from determining the involvement of proteasomal degradation in the regulation of TSPAN8 levels by UBE2D3 (**Fig. 2G, Fig. S2A**). However, q-RT PCR analysis indicated that the decreased TSPAN8 protein levels upon UBE2D3 depletion are not due to decreased *Tspan8* mRNA levels. Upon the expression of *Ube2d3* shRNA1 TSPAN8 transcript levels were even significantly increased over 3 independent biological replicates. Thus, the reduced TSPAN8 protein levels in UBE2D3-depleted cells seem to result from regulation at the protein level (**Fig. S2D**).

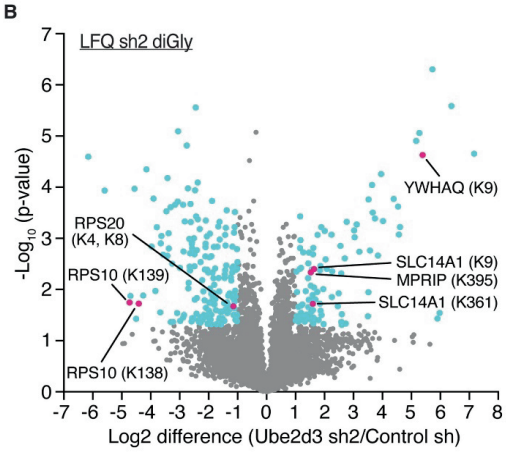
Ubiquitinome analysis identifies increased protein ubiquitination in UBE2D3-depleted cells

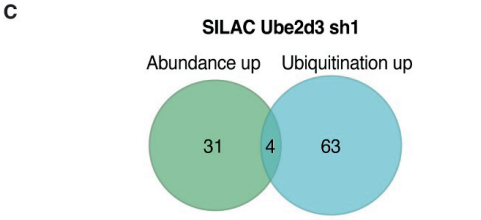
To identify novel direct and indirect targets of UBE2D3 *in vivo*, we subsequently focused on the ubiquitinome changes occurring upon UBE2D3 depletion in our SILAC-based diGly proteomics and LFQ diGly proteomics data, represented by diGly peptides with a normalized ratio of $\geq$ 1.5-fold (SILAC) or $\geq$ 2-fold (LFQ) that were significantly changed (p-value $\leq$ 0.05) (**Fig. 3A, B and Table S7-11**).

First, we examined the 67 (SILAC *Ube2d3* sh1) and 90 (LFQ *Ube2d3* sh2) proteins that were significantly increased $\geq$ 1.5-fold (SILAC) or $\geq$ 2-fold (LFQ) in their ubiquitination upon UBE2D3 depletion and most likely represent indirect targets of UBE2D3 (**Fig. 1F and Table S9, S11**). Of these, 4 proteins (PTGS2, SLC14A1, SORBS2 and TUBB3) increased in abundance in SILAC *Ube2d3* sh1 and 4 proteins (COL6A2, CRABP1, LAMC1 and PEG10) increased in abundance in LFQ *Ube2d3* sh2 (**Fig. 3C, D and Table S3**). Hence, the observed increase in their ubiquitination could be a consequence of their increased abundance. However, it could also reflect a functional role in which for instance the ubiquitination of these proteins interferes with their degradation, causing an increase in their abundance. CRABP1 was also found increased in ubiquitination and abundance in the SILAC diGly proteomics with shRNA1, but the change in ubiquitination was not statistically significant (**Table S2, S7**). Vice versa, TUBB3 was increased both in ubiquitination and abundance in the SILAC sh1 proteomics, while in the LFQ sh2 proteomics TUBB3 was also increased in

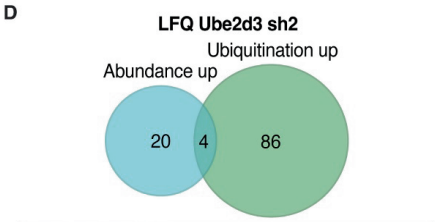
abundance, but its ubiquitination status remains unassessed as no diGly peptides were detected for TUBB3 (**Table S2, S6, S9 and S10**).



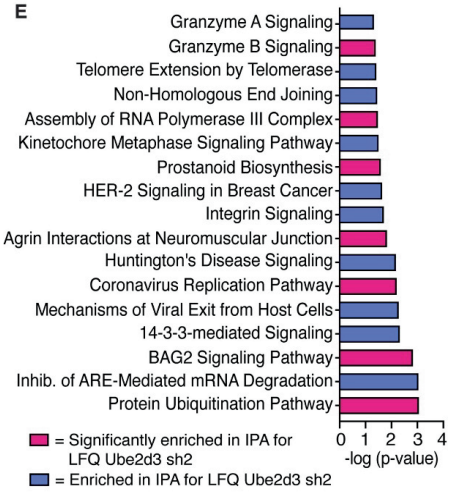


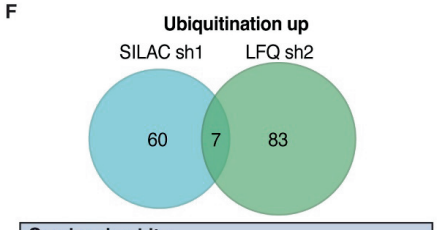


Overlapping hits	
Name	**Biological function**
PTGS2	Prostaglandin-endoperoxide synthase
SLC14A1	Urea transporter
SORBS2	Adaptor protein
TUBB3	Beta tubulin class III protein



Overlapping hits	
Name	**Biological function**
COL6A2	Alpha 2 chain of collagen type VI
CRABP1	Retinoic acid-binding protein
LAMC1	Gamma chain of laminin
PEG10	Gag-like & gag/pol-like protein





Overlapping hits		
Name	**SILAC sh1 lysine**	**LFQ sh2 lysine**
CTTN	K79	K359
LAMC1	K800	K1402
LMNB1	K533	K390
MPRIP	**K395**	**K395**
PEG10	K191	K303, K813
SLC14A1	**K9, K361**	**K9, K361**
YWHAQ	**K9**	**K9**

Figure 3. DiGly sites upregulated upon UBE2D3 depletion. ***A***, Volcano plot showing the log2 fold changes of diGly (ubiquitinated) peptides upon UBE2D3 depletion with sh1 in SILAC-based diGly proteomics (*n*=6). Blue dots represent peptides that are at least 1.5x up- or downregulated in their ubiquitination (log2 ≥ 0.585 and log2 ≤ -0.585) and have a p-value of ≤ 0.05. Red dots represent top hits, including the modified lysine(s), that are overlapping with LFQ *Ube2d3* sh2 hits. ***B***, Volcano plot showing the log2 fold changes of diGly peptides upon UBE2D3 depletion with sh2 in LFQ proteomics (*n*=3). Blue dots represent peptides that are at least 2x up- or downregulated in their ubiquitination (log2 ≥ 1.0 and log2 ≤ -1.0) and have a p-value of ≤ 0.05. Red dots represent top hits, including the modified lysine, that are overlapping with SILAC *Ube2d3* s1 hits. ***C*,** Venn diagram illustrating the overlap between proteins that go up in abundance and ubiquitination in SILAC *Ube2d3* sh1. The table below shows the overlapping hits and their biological function. ***D***, Venn diagram illustrating the overlap between proteins that go up in abundance and ubiquitination in LFQ with *Ube2d3* sh2. The table below shows the overlapping hits and their biological function. ***E***, Ingenuity analysis of diGly peptides upregulated upon UBE2D3 depletion. Bar plots show the top significantly enriched pathways from SILAC *Ube2d3* sh1 that overlap with LFQ *Ube2d3* sh2. ***F,*** Venn diagram illustrating the overlap between proteins that go up in ubiquitination in SILAC *Ube2d3* sh1 and LFQ *Ube2d3* sh2. The table below shows the overlapping hits, including on which lysine they are ubiquitinated. Bold hits are ubiquitinated on the same lysine in the SILAC *Ube2d3* sh1 experiments as in the LFQ *Ube2d3* sh2 experiments.

Furthermore, while we don't necessarily expect proteins that become more ubiquitinated upon UBE2D3 depletion to be direct targets of UBE2D3, we did find 5 of 67 proteins identified in the SILAC-based diGly proteomics to be known UBE2D3 interactors (from BioGRID, IntAct, MINT, STRING and UniProt databases) (**Fig. S3A and Table S3)**. These are the E2 enzyme UBE2S, the E3 ligases ARIH1, DZIP3 and UHRF1, and the F-box protein FBXW11 (or □-TrCP2), a substrate binding subunit of an SCF (Skp1-cullin-F-box) E3 ligase. UBE2S and UBE2D3 were shown to cooperate in the ubiquitination of APC/C substrates, such as cyclin B, *in vitro*[44]. Furthermore, UBE2D3 was shown to function together with the E3 ligases ARIH1 and SCF (SCF$^{□\text{-TrCP}}$ and SCF^{FBW7}) in ubiquitination of targets such as DCNL1, □-catenin, Myc and I□B□[44-49]□ The abundance of these 5 proteins (UBE2S, ARIH1, DZIP3, UHRF1, FBXW11) was not changed upon UBE2D3 depletion (**Table S1**), suggesting that their increased ubiquitination in UBE2D3-depleted conditions has no effect on their proteasomal degradation, but might have a regulatory consequence or be without effect. We also found 3 proteins with increased ubiquitination upon UBE2D3 knockdown in the LFQ diGly proteomics to be among known UBE2D3 interactors: the two E3 ligases HERC2 and UBE4B and the ubiquitin domain-containing protein UBTD1 (**Fig. S3B and Table S3**). HERC2 and UBE2D3 ubiquitinate BRCA1 *in vitro*[50], while UBE4B and UBE2D3 were shown to ubiquitinate the HTLV-1 Tax oncoprotein and FEZ1, which functions in axonal outgrowth and fasciculation[51,52]. In addition, UBTD1 interacts with UBE2D3 and □-TrCP in the □-TrCP/UBE2D3/□-catenin complex and promotes □-TrCP-YAP interaction and YAP ubiquitination[53]. How UBE2D3 inhibits ubiquitination of these E3 ligases, UBE2S and UBTD1, and with what consequences, is unclear at this point and would be interesting to further address.

There is no overlap between known UBE2D3 interactors with increased ubiquitination in the SILAC *Ube2d3* sh1 experiments and those in the LFQ *Ube2d3* sh2 diGly experiments, because for the UBE2D3 interactors with increased ubiquitination in the SILAC sh1 experiments no diGly-modified peptides were detected in the LFQ sh2 experiments and vice versa. Therefore, we could not evaluate potential changes in ubiquitination. Except for UHRF1, which was increased in

ubiquitination at K196 in both the SILAC *Ube2d3* sh1 and LFQ *Ube2d3* sh2 experiments, but which did not meet the cut-off for statistical significance in LFQ sh2, and was therefore not included in further analysis.

Ingenuity pathway analysis (IPA) of the 67 and 90 proteins that were more ubiquitinated upon depletion of UBE2D3, in the SILAC-based and LFQ diGly proteomics respectively, revealed the ubiquitin signaling, cell adhesion and signaling, and apoptosis pathways as the top enriched overlapping pathways between SILAC sh1 and LFQ sh2 hits (**Fig. 3E and Table S12-14**). The cytoskeletal proteins alpha-actinin 4 (ACTN4), tubulin beta-3 chain (TUBB3) and tubulin beta-4B chain (TUBB4B), the regulators of the actin cytoskeleton Rho-associated protein kinase 1 (ROCK1) and myosin phosphatase Rho-interacting protein (MPRIP), and Src substrate cortactin (CTTN) are the main factors from our data that act in cell adhesion pathways. Cell adhesion is critical in cell migration and tissue development[54,55]. Disruption could lead to pathologies and plays an important role in cancer metastasis[56]. By regulating these factors, next to TSPAN8 which is also involved in cell adhesion, UBE2D3 could potentially play a role in cell adhesion-regulated pathologies.

Of the 90 proteins that were significantly increased in their ubiquitination in the LFQ sh2 experiments, 7 proteins were also identified as significantly increased in ubiquitination in the SILAC sh1 experiments (**Fig. 3F**). Of these 7 proteins, 3 were found to be modified on the same lysine in the SILAC and LFQ diGly experiments, further supporting that these proteins are potential new indirect targets of UBE2D3. Among these, SLC14A1 was also increased in abundance upon UBE2D3 depletion in the SILAC sh1 condition (**Fig. 3C**; it was not picked up in the LFQ sh2 global proteomics), therefore the observed increase in ubiquitination could potentially be a consequence of increased abundance. However, for MPRIP and 14-3-3 (YWHAQ) no change in abundance was observed, further supporting that these proteins are new indirect ubiquitination targets of UBE2D3. MPRIP is a filamentous actin-binding protein that can affect the integrity of the actin cytoskeleton[57]. MPRIP ubiquitination at K395 (K396 human) is increased upon UBE2D3 depletion in both the SILAC sh1 and the LFQ sh2 condition. This lysine is located in the second Pleckstrin Homology (PH) domain of MPRIP. PH-domains are known to promote recruitment of proteins to membranes. Potentially, by regulating ubiquitination of MPRIP at K395, UBE2D3 could affect interaction of MPRIP with other proteins and their recruitment to certain cellular compartments. 14-3-3 proteins are phosphoserine and phosphothreonine binding proteins that regulate almost every process in the cell. 14-3-3 proteins generally function as adaptors or chaperones and interact with many different proteins, such as kinases, phosphatases and receptors[58,59]. UBE2D3 depletion results in increased ubiquitination of 14-3-3q at K9, which is located in the first alpha-helix of the protein. 14-3-3 proteins can form dimers through the first two helices of one monomer and the fourth helix of the second monomer[58]. Possibly, by affecting ubiquitination of 14-3-3 q at K9, UBE2D3 could interfere with homo-or heterodimerization of 14-3-3q. However, more research will be needed to determine how these indirect targets of UBE2D3 are ubiquitinated and what the potential consequences are.

Ubiquitinome analysis identifies decreased protein ubiquitination in UBE2D3-depleted cells

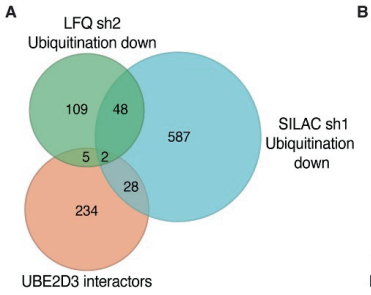
To identify novel direct targets of UBE2D3 we focused on the peptides that were less ubiquitinated in *Ube2d3* knockdown cells. From the 665 proteins that were on average reduced in their ubiquitination over all SILAC-based diGly proteomics experiments, 50 proteins were also found less ubiquitinated in the LFQ sh2 proteomics and 30 proteins overlap with previously identified UBE2D3 interactors (**Fig. 4A-C and Table S3, S9**). From the 164 proteins that were on average reduced in their ubiquitination over all three LFQ *Ube2d3* sh2 diGly proteomics experiments, 7 proteins are known UBE2D3 interactors (**Fig. 4A, B and Table S3, S11**). Most of these proteins are E3 ligases. One of these being CBL, for which UBE2D3 was previously shown to both promote autoubiquitination and substrate ubiquitination *in vitro*[60]. Our data suggest that this autoubiquitination occurs on lysine 519, as ubiquitination on this lysine is significantly decreased upon UBE2D3 depletion (**Table S9**). Furthermore, CBL was reported to ubiquitinate the epidermal growth factor receptor (EGFR) together with UBE2D2/UBE2D3[60-62] and depletion of CBL or UBE2D3 resulted in decreased EGFR ubiquitination[60]. Indeed, the SILAC sh1 data also show significantly decreased ubiquitination of EGFR upon UBE2D3 depletion (**Table S9**), thus supporting the previous findings. By promoting (auto)ubiquitination of E3 ligases such as CBL, UBE2D3 could potentially be involved in regulating their activity.

We also identified sequestome-1 (SQSTM1/p62), which is known to interact with UBE2D3 and to be ubiquitinated at K420 by UBE2D3 to activate its function as an autophagy receptor[63]. Human and mouse SQSTM1 share ± 90% amino acid identity, with lysine 420 in human SQSTM1 corresponding to lysine 422 in mouse SQSTM1. In our SILAC-based diGly proteomics dataset SQSTM1(K422) ubiquitination is ± 2-fold decreased in UBE2D3-depleted cells (**Fig. 3A and Table S9**), indicating that our data is in line with, and supports, previous findings on SQSTM1 ubiquitination by UBE2D3. Unfortunately, in the LFQ diGly proteomics no diGly peptide with lysine 422 was detected. In addition, we also identified the chaperone HSPA5 as one of the significant hits that is decreased in its ubiquitination (**Table S9**). SQSTM1 and HSPA5 are both part of the autophagic protein quality control (PQC) pathway, a pathway that removes misfolded proteins and their aggregates via lysosomal degradation. HSPA5 binds to these protein aggregates and is recruited by SQSTM1. SQSTM1 delivers HSPA5 and its cargo to autophagosomes, which subsequently fuse with lysosomes to induce lysosomal degradation[64]. By regulating the ubiquitination of these factors and the activity of SQSTM1, UBE2D3 might affect degradation of misfolded proteins via autophagic PQC.

Of the known UBE2D3 interactors, 2 proteins, TAX1BP1 and BIRC2, are reduced in their ubiquitination in both the SILAC *Ube2d3* sh1 and LFQ *Ube2d3* sh2 diGly proteomics (**Fig. 4B**). TAX1BP1 and BIRC2 are both involved in regulation of NF-κB signaling[65-68]. Interestingly, TAX1BP1 is an autophagy receptor that has also been proposed to function downstream of SQSTM1, promoting autophagosome formation and degradation of protein aggregates[69,70]. This adds another potential layer of regulation by UBE2D3, in which it might also ubiquitinate TAX1BP1, next to SQSTM1 and HSPA5, and thereby play a role in degradation of ubiquitinated proteins/protein aggregates by autophagy. Thus, our approaches not only successfully identified and confirmed previously known targets of UBE2D3, but are also a useful resource for establishing to what extent

in vitro demonstrated ubiquitination activities of UBE2D3 also represent ubiquitination events with dependency on UBE2D3 *in vivo*. Moreover, they identified > 700 novel direct or indirect *in vivo* targets of ubiquitination by UBE2D3.

Intriguingly, we also found 5 and 8 proteins, in respectively the SILAC *Ube2d3* sh1 and LFQ *Ube2d3* sh2 diGly proteomics, that both increased and decreased in their ubiquitination (on different lysine residues) following UBE2D3 depletion. These proteins have different biological functions, for example in protein degradation, post-translational protein modification and cytoskeleton support (**Fig. 4D, E and Table S3**). Potentially, ubiquitination of these proteins on one lysine residue by UBE2D3 could inhibit or block ubiquitination by other enzymes on another lysine residue. However, further investigation is needed to understand the mechanisms and biological relevance associated with this differential ubiquitination.



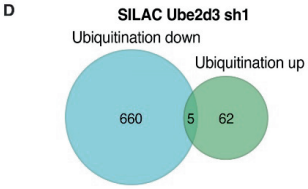
B

Known UBE2D3 interactors with reduced ubiquitination in SILAC sh1 and LFQ sh2 diGly				
SILAC Ube2d3 sh1				LFQ sh2
AMFR	MIB1	RNF146	UBE3C	**BIRC2**
BIRC2	NEDD4	RNF31	UBR5	RNF168
BIRC3	NEDD4L	SQSTM1	VPRBP	**TAX1BP1**
CBL	PJA1	STAM2	WWP1	TRAF2
EPS15	PJA2	**TAX1BP1**	WWP2	UBE2C
HECTD1	RABGEF1	TRIM25	XIAP	UBE2V2
ITCH	RBCK1	UBC		UHRF1
MGRN1	RNF13	UBE2O		

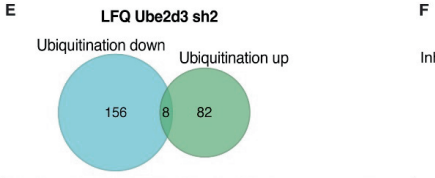
C

Overlapping hits SILAC sh1 & LFQ sh2 ubiquitination down			
ACO1	DYNC1H1	PCM1	SLC6A6
ACTN4	EIF3A	PLEC	SNRNP200
ADRM1	EIF3C	PLIN2	SOAT1
ARHGAP1	EPHA2	POR	SPTAN1
ASCC2	HIP1	PSMD13	TAX1BP1
BAG2	HNRNPU	PSMD6	TEX10
BIRC2	HSPD1	RAB1B	TIMM8A1
BSG	IDE	RDH11	TXN
CISD1	KLHL9	RPL14	UBE2D3
CISD2	MYH9	RPS10	UBXN4
CUL7	MYO1E	RPS20	UBXN7
DIABLO	NAA50	RPS21	
DNMT1	NIPAL1	SLC12A7	

D



Overlapping hits			
Name	Lysine Up	Lysine Down	Biological function
ACTN4	K422	K55	Actin-binding protein
AHNAK	K536	Multiple	Neuroblast differentiation
EFEMP1	K396	K207, 453	Extracellular matrix glycoprotein
PSMD12	K197	K212	26S proteasome subunit
STT3A	K640	K289	Oligosaccharyltransferase subunit



Overlapping hits LFQ sh2 ubiquitination up and down			
Name	Lysine Up	Lysine Down	Biological function
ALDOA	K99	K14	Fructose-bisphosphate aldolase
DDX21	K61	K847	RNA helicase
DNAJA1	K37, K296	K32, K37	Co-chaperone of HSPA8
EPRS1	K794	K599	Aminoacyl-tRNA synthetase
HSPA4	K687	K477	Heat shock protein
PLEC	K3031	K2671	Interlinks cytoskeleton elements
RANBP2	K1677, K2120	K1964	SUMO E3 ligase
UBFD1	K196	K304	Ubiquitin-binding protein

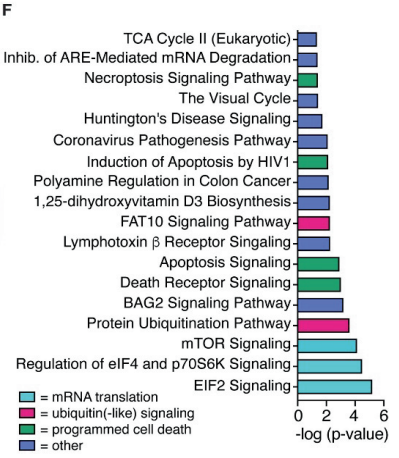


Figure 4. DiGly sites downregulated upon UBE2D3 depletion. ***A***, Venn diagrams illustrating the overlap between known UBE2D3 interactors (*BioGRID, IntAct, MINT, STRING and UniProt* databases*)* and proteins significantly downregulated in their ubiquitination in the SILAC *Ube2d3* sh1 and LFQ *Ube2d3* sh2 experiments. ***B***, Table showing the overlap between known UBE2D3 interactors and proteins significantly reduced in their ubiquitination upon UBE2D3 depletion in the SILAC sh1 and LFQ sh2 data sets. In bold are the overlapping proteins found in both the SILAC *Ube2d3* sh1 and LFQ *Ube2d3* sh2 data. ***C***, Table showing the overlapping hits between proteins that go down in ubiquitination in SILAC with *Ube2d3* sh1 and LFQ with *Ube2d3* sh2. ***D***, Venn diagram illustrating the overlap between proteins that go up and down in ubiquitination in SILAC *Ube2d3* sh1. The table below shows the overlapping hits, the lysines on which they are modified and the biological function of these proteins. ***E***, Venn diagram illustrating the overlap between

proteins that go up and down in ubiquitination in LFQ *Ube2d3* sh2. The table below shows the overlapping hits, the lysines on which they are modified and the biological function of these proteins. ***F,*** Ingenuity analysis of the 50x overlapping proteins between SILAC *Ube2d3* sh1 and LFQ *Ube2d3* sh2 that were significantly downregulated in their ubiquitination upon UBE2D3 depletion. Bar plots show the top 18 significantly enriched pathways.

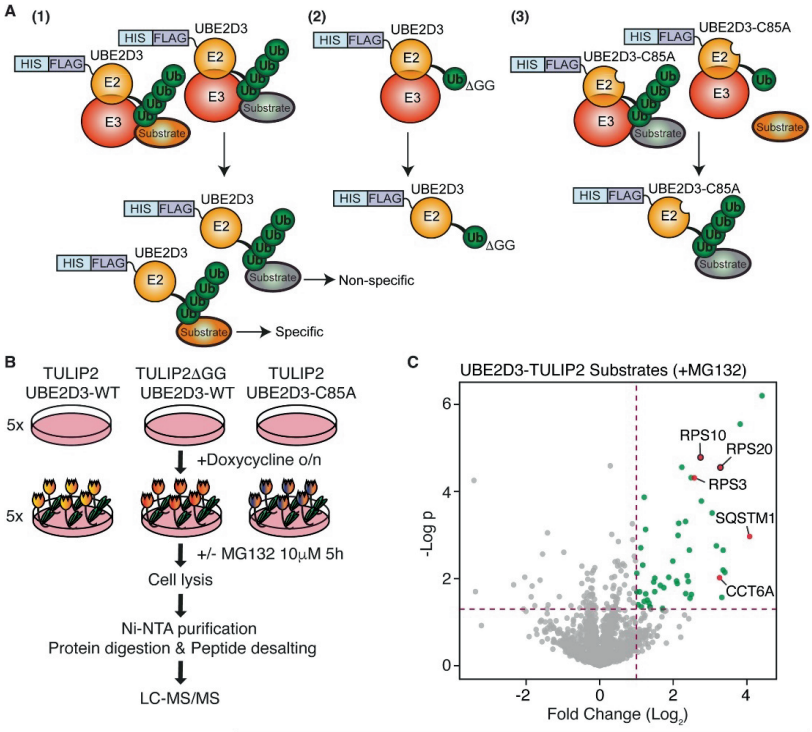
UBE2D3 depletion impairs ubiquitination of factors in mRNA translation pathways

As we identified a considerable number of proteins that were significantly reduced in their ubiquitination upon UBE2D3 depletion, we used IPA to understand in which biological processes these proteins work and whether specific processes were impacted. IPA analysis on the 50 proteins that were significantly less ubiquitinated in both SILAC *Ube2d3* sh1 and LFQ *Ube2d3* sh2 diGly proteomics revealed 18 significantly enriched pathways (**Fig. 4A, C, F and Table S15**). These mostly belong to mRNA translation and degradation, protein degradation/ubiquitin(-like) signaling and programmed cell death processes. Furthermore, STRING analysis on the 50 proteins reduced in ubiquitination and overlapping between SILAC sh1 and LFQ sh2, also identified a network of proteins involved in mRNA translation and degradation (**Fig. S4A**). We also evaluated if the decreased ubiquitination of these 50 proteins upon UBE2D3 depletion could be a consequence of decreased abundance. However, we found no indication for this as there was no overlap between these 50 proteins and the proteins decreasing in abundance in the corresponding global proteome data from SILAC sh1 and LFQ sh2 (**Fig. S4B, C**). The top hits that were affected by *Ube2d3* depletion with both sh1 and sh2, and are related to mRNA translation, are the 40S ribosomal proteins RPS10 and RPS20. Both were > 2.5-fold reduced in their ubiquitination upon *Ube2d3* knockdown (**Fig. 3A, B** and **Table S9, S11**). RPS10 and RPS20 ubiquitination plays an important role in the ribosome-associated protein quality control (RQC) pathway[22-24]. Also, other proteins acting in the RQC pathway, such as the key components Listerin (LTN1) and VCP/p97 were significantly decreased in their ubiquitination in the SILAC sh1 data (**Table S9**). Despite these changes not reaching statistical significance in the LFQ sh2 data (**Table S10**), this may point towards a broader impact of UBE2D3 on the RQC pathway.

TULIP2 confirms RPS10 and RPS20 as direct substrates of UBE2D3

The diGly proteomics approaches that we used here successfully identified a considerable number of potential novel targets for UBE2D3, but are not able to distinguish between direct and indirect substrates of UBE2D3. Therefore, we used the TULIP2 methodology (Targets of Ubiquitin Ligases Identified by Proteomics 2)[21], originally set up to identify substrates of E3 ligases through E3 ligase-ubiquitin fusions, to confirm substrates as direct targets of UBE2D3. TULIP2 methodology makes use of lentiviral doxycycline-inducible expression constructs with an activated ubiquitin C-terminally fused to the E3 ligase of interest, a linker containing a 10X HIS-tag in between the E3 ligase and ubiquitin, and a 10x HIS-FLAG tag N-terminal of the E3 ligase. This allows for nickel-NTA purification and the use of harsh buffers to distinguish between substrates and interactors[21]. Here, we used TULIP2 for the first time to identify targets of an E2 enzyme (**Fig. 5A**). After generation of HeLa cells with stably integrated TULIP2 constructs with UBE2D3 WT, UBE2D3 C85A (catalytically inactive mutant) and UBE2D3 WT with a ubiquitin lacking the diGly motif (ΔGG) as a negative control, expression was induced with doxycycline and the cells were untreated or treated with proteasome inhibitor (MG132) to prevent degradation of ubiquitinated substrates of UBE2D3.

Subsequently, substrates covalently attached to TULIP2-UBE2D3 constructs, through interaction of UBE2D3 with different E3 ligases, were purified using Ni-NTA beads and peptides were analyzed by mass spectrometry (**Fig. 5A, B, Fig. S5A**).



Overlap SILAC sh1 & LFQ sh2 ubiquitination down hits with TULIP2 hits			
Name	Lysine SILAC sh1	Lysine LFQ sh2	Biological function
CCT6A	K5	K251*	Component of TRiC complex; functions in protein folding
CTNNA1	K216, K474, **K478**, K747,	**K478***, K683* K695*	Catenin alpha 1 protein
PAM	K922	-	Catalyzes conversion of neuroendocrine peptides to alpha-amidated products
PCM1#	**K259**, **K727**, **K760**	**K259**, **K727**, **K760***	Component of centriolar satellites
PLEC	K129, K488, K3391, K4006, K4212, K4266	K2460, K2644,* K2671*	Interlinks cytoskeleton elements
RPS10	**K138**, **K139**	**K138**, **K139**	Component of 40S ribosomal subunit
RPS20	**K4**, **K8**	**K4**, **K8**	Component of 40S ribosomal subunit
RPS3	K214	-	Component of 40S ribosomal subunit
SQSTM1	K157, K314, K422, K437	-	Adaptor protein regulating activation of the NF-kB signaling pathway

Figure 5. TULIP2 for identification of direct substrates of UBE2D3. ***A***, UBE2D3-TULIP2 rationale. Ubiquitin (Ub) fused to UBE2D3 can be conjugated to a substrate protein following the canonical ubiquitination cascade and pulled down in harsh denaturing conditions together with UBE2D3 and the conjugated substrate (specific and non-specific substrates in the vicinity (1)). Ub fused to TULIP2-ΔGG cannot be conjugated, thus there will be no conjugated substrates (2). UBE2D3-C85A lacks catalytic activity and is not

able to conjugate Ub to its specific targets. However, other E2/E3 enzymes in the surrounding are still able to use the fused Ub for their substrates, named non-specific substrates (3). ***B***, Overview experimental approach of TULIP2 methodology. HeLa cells containing UBE2D3-TULIP2, UBE2D3-TULIP2ΔGG and UBE2D3-C85A-TULIP2 expression cassettes were induced o/n with doxycycline, not treated or treated with proteasome inhibitor (MG132) and lysed. TULIP2 conjugates were Ni-NTA purified, proteins were digested, and peptides were desalted and analyzed by LC-MS/MS. ***C***, Volcano plot depicting the UBE2D3 conjugates (*n*=4). Green dots represent statistically enriched proteins in the UBE2D3-TULIP2 samples compared to UBE2D3-TULIP2ΔGG samples after proteasome inhibition for p-value = 0.05 and S0 = 0.1. Red dots correspond to top hits significantly downregulated in their ubiquitination in SILAC *Ube2d3* sh1 experiments. Red dots with a black stroke around them also overlap with top hits significantly downregulated in their ubiquitination in LFQ *Ube2d3* sh2 experiments. ***D***, Venn diagrams illustrating the overlap between TULIP2 hits (+MG132 and no MG132) and proteins significantly downregulated in their ubiquitination in the SILAC *Ube2d3* sh1 and LFQ *Ube2d3* sh2 diGly proteomics. Table on the right shows overlapping hits between TULIP2 and SILAC sh1 and LFQ sh2, including modified lysines and the biological functions of these proteins. Bold lysines overlap between SILAC sh1 and LFQ sh2. Asterisk indicates that these sites were decreased in their ubiquitination, but not significantly (p-value > 0.05). "#" indicates that only lysines that overlap between SILAC sh1 and LFQ sh2 were selected, as a large number of peptides significantly decreased in their ubiquitination were found for PCM1 in SILAC sh1 experiments.

The TULIP2 experiments identified 62 targets in total (with and without proteasome inhibitor combined) (**Table S16**). TULIP2 confirmed 9 potential substrates of UBE2D3 with decreased ubiquitination in the SILAC diGly proteomics as direct targets of UBE2D3 (**Fig. 5C, D, Fig. S5B**). These include the top hits RPS10 and RPS20, but also SQSTM1, which is a known target of UBE2D3[63]. In addition, TULIP2 identified PCNA, another previously identified target of UBE2D3[71], demonstrating that this methodology succeeds in identifying targets of E2 enzymes. Moreover, at least 4 of these 9 targets, PCM1, PLEC, RPS10 and RPS20, were also independently confirmed by the LFQ *Ube2d3* sh2 diGly proteomics (**Fig. 5C, D, Fig. S5B and Table S16**). The TULIP2 data confirm the top hits from our SILAC sh1 and LFQ sh2 diGly proteomics experiments, RPS10 and RPS20, as direct targets of UBE2D3 and emphasize the importance of UBE2D3 in regulation of mRNA translation, in line with the IPA and STRING results. Therefore, we continued with further validating RPS10 and RPS20.

RPS10 and RPS20 are ubiquitinated in vivo by UBE2D3

Stalling of ribosomes during protein synthesis activates the RQC pathway to eliminate aberrant polypeptides that are potentially deleterious. Stalling of translation can occur for multiple reasons, such as truncated or damaged mRNA, excessive mRNA secondary structure, prematurely poly-adenylated mRNA and insufficient amounts of an amino acid or tRNA[72]. Upon encountering a prematurely poly-adenylated mRNA, the ribosome stalls and the E3 ligase ZNF598 ubiquitinates the 40S ribosomal proteins RPS3, RPS3A, RPS10 and RPS20 at lysine 214 (K214), lysine 249 (K249), lysines 138/139 (K138/139) and lysines 4/8 (K4/8), respectively[22-24]. This results in ribosome disassembly and activation of the RQC pathway, and subsequent ubiquitination and degradation of the nascent polypeptide[72,73]. ZNF598 was previously shown to interact with UBE2D3 and depletion of UBE2D3 partially impaired RQC, suggesting that UBE2D3 may be involved in ubiquitination of ZNF598 substrates[22]. We found significantly reduced RPS3-K214 ubiquitination in the SILAC *Ube2d3* sh1 diGly proteomics, but no change in RPS3-K214 ubiquitination in the LFQ

Ube2d3 sh2 diGly proteomics. We found RPS3A-K249 ubiquitination reduced in the LFQ sh2 experiments, but this was not statistically significant and therefore not considered as a hit, while RPS3A peptides with Ub modification on K249 were not identified in SILAC sh1 experiments (**Table S7, S9-11**). Given the lack of reproducibility in RPS3 and RPS3A modification between *Ube2d3* sh1 and sh2 experiments, we decided to not continue with these proteins. However, from the different diGly-peptides recovered for RPS10, all peptides ubiquitinated on K138 and/or K139 were decreased in *Ube2d3* knockdown cells, with both shRNAs (**Fig. 6A, B**). Of these, peptides ubiquitinated on both K138 and K139 were more frequently observed than peptides with only one of these sites modified. From the diGly-peptides recovered for RPS20, all peptides ubiquitinated on K4 and/or K8 were decreased in *Ube2d3* knockdown cells, with both shRNAs (**Fig. 6A, B**). Also here, mainly peptides modified on both K4 and K8 were found. Importantly, the abundance of RPS10 or RPS20 was not affected by UBE2D3 depletion, thus the observed decrease in their ubiquitination is not caused by lower protein abundance (**Table S1, S4**). This indicates that RPS10(K138/139) and RPS20(K4/8) ubiquitination by ZNF598 *in vivo* is dependent on UBE2D3.

A

RPS10 & RPS20 ubiquitination in SILAC Ube2d3 sh1					
Protein	**Peptide sequence**	**#GlyGly (K)**	**Lysine(s)**	**H/L ratio (Log2)**	**-Log10 (p-value)**
RPS10	PKGPEGERPAR	1	K107	-0.768	1.897
RPS10	SAVPPGADKK	1	**K138**	**-1.306**	**3.047**
RPS10	KAEAGAGSATEFQFR	1	**K139**	**-1.084**	**2.242**
RPS10	SAVPPGADKKAEAGAGSATEFQFR	2	**K138, 139**	**-1.715**	**3.869**
RPS20	AFKDTGKTPVEPEVAIHR	1	**K4**	**-1.697**	**1.527**
RPS20	AFKDTGK	1	**K4**	**-1.236**	**2.991**
RPS20	DTGKTPVEPEVAIHR	1	**K8**	**-1.258**	**4.408**
RPS20	AFKDTGKTPVEPEVAIHR	2	**K4, 8**	**-1.555**	**5.085**

B

RPS10 & RPS20 ubiquitination in LFQ Ube2d3 sh2				
Protein	**#GlyGly (K)**	**Lysine**	**Log2 (Ube2d3 sh2/Control)**	**-Log10 (p-value)**
RPS10	2	**K138**	**-4.411**	**1.723**
RPS10	2	**K139**	**-4.729**	**1.746**
RPS10	1	**K139**	**-2.284**	**1.458**
RPS20	2	**K4**	**-1.149**	**1.672**
RPS20	2	**K8**	**-1.149**	**1.672**

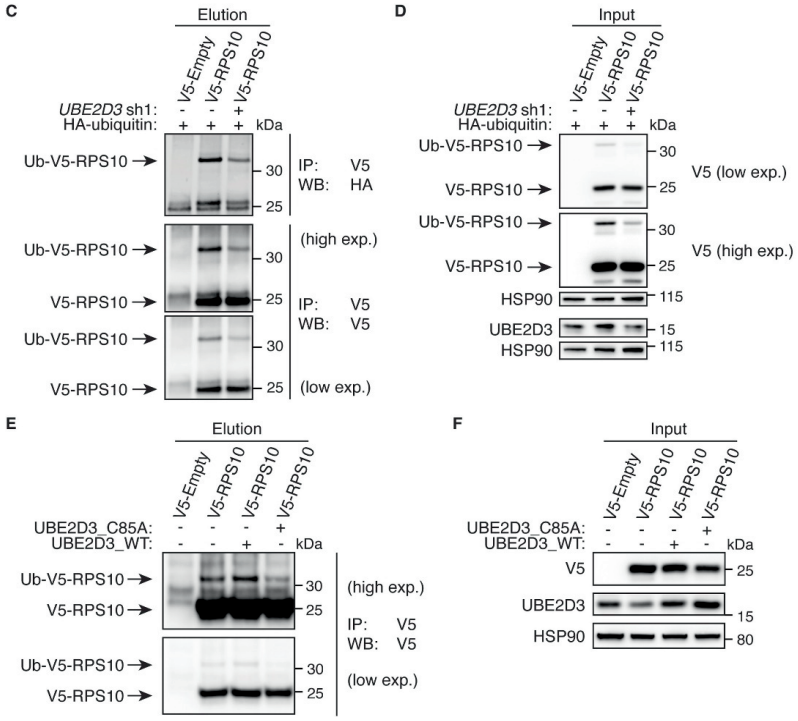


Figure 6. UBE2D3 regulates ubiquitination of RPS10. ***A***, Table with significantly changed diGly-modified peptides found for RPS10 and RPS20 in the SILAC-based diGly proteomics, including the number (#) and sites of modified lysines, average H/L normalized log2 ratios and statistical significance (hits are considered significant when -log10 p-value ≥ 1.3 or p-value ≤ 0.05). Bold lysines overlap between SILAC sh1 and LFQ sh2. ***B***, Table with significantly changed diGly-modified peptides found for RPS10 and RPS20 from the LFQ diGly proteomics, including the number (#) and sites of modified lysines, average log2 ratios and statistical significance (p-value). Bold lysines overlap between SILAC sh1 and LFQ sh2. ***C-D***, Immunoprecipitation assay (IP) in 293T cells +/- *UBE2D3* sh1 transfected with V5-tagged RPS10 and HA-tagged ubiquitin. Immunoblots representing eluates of the V5-IP, showing decreased ubiquitination of RPS10 in UBE2D3-depleted cells (*C*) and immunoblots of input samples (*D*). Representative blots of *n* = 3 are shown. ***E-F***, IP in control 293T cells,

or overexpressing either UBE2D3 wild-type (WT) or a catalytically inactive mutant of UBE2D3 (C85A), transfected with V5-tagged RPS10. Immunoblots representing eluates of the V5-IP, showing increased ubiquitination of RPS10 in UBE2D3 WT cells and decreased ubiquitination of RPS10 in UBE2D3 C85A cells (*E*) and immunoblots of input samples (*F*). Representative blots of *n* = 2 are shown.

Ubiquitination of RPS10 on K138/139 is critical for proper RQC[22,23]. To further address the ubiquitination of RPS10 *in vivo* and validate the results from the ubiquitin diGly proteomics and TULIP2 analyses with independent *UBE2D3* shRNAs and cell line, we transfected wild-type (WT) or *UBE2D3*-depleted HEK 293T cells with V5-tagged RPS10 and HA-tagged ubiquitin and immunoprecipitated V5-RPS10. Indeed, when *UBE2D3* was depleted with two independent shRNAs, V5-RPS10 ubiquitination was decreased, indicating that ubiquitination of RPS10 *in vivo* is to a considerable extent dependent on UBE2D3 (**Fig. 6C, D, Fig. S6A-C**). In addition, we also confirmed decreased RPS20 ubiquitination upon *UBE2D3* depletion in HEK 293T cells (**Fig. S6D, E**). Moreover, we show that ubiquitination of RPS10 by UBE2D3 *in vivo* is dependent on the catalytic activity of UBE2D3, as overexpression of UBE2D3 WT, but not UBE2D3 C85A, resulted in increased ubiquitination of RPS10 (**Fig. 6E, F**). Hence, despite redundancy between UBE2D family members and E2 enzymes in general, UBE2D3 seems to be the main E2 that ubiquitinates RPS10 in the mouse embryo fibroblasts used in our proteomics experiments as well as in the human HeLa cells used in TULIP2 experiments, and in HEK 293T cells used for *in vivo* ubiquitination experiments. Thus, here we demonstrate that UBE2D3 ubiquitinates RPS10 and RPS20 *in vivo*, supporting a role for UBE2D3 in the degradation of aberrant nascent polypeptides through the RQC pathway.

DISCUSSION

Identification of a complete set of targets of ubiquitin machinery enzymes is challenging. Different methods have been developed in the past to try tackle this problem, such as functional genomics using siRNA libraries, global protein stability profiling and affinity-based proteomics[14,74-76].This has resulted in the identification of targets of different E3 ligases, including HUWE1, PARKIN, SCF^{Saf1} and HRD1, and of DUBs such as USP32 and CYLD[13-15,75,77,78]. For the limited set of E2 enzymes that exist in cells for the ubiquitination of all proteins targeted by E3s, it is even more challenging to identify the probably larger number of targets of a particular E2. Here, we combined quantitative ubiquitinome profiling by SILAC-based or label-free ubiquitin diGly proteomics with depletion of UBE2D3 to identify novel *in vivo* targets of UBE2D3, one of the most promiscuous E2 enzymes *in vitro*. In addition, we used a modified TULIP2 approach to confirm potential substrates of UBE2D3 as direct targets. We show that UBE2D3-depleted cells undergo global proteome and ubiquitinome changes, of which the ubiquitinome changes are most prominent. Our analysis revealed that UBE2D3 affects the abundance and/or ubiquitination of proteins involved in a wide variety of processes, in particular proteins belonging to molecular pathways related to metabolism and mRNA translation. Interestingly, our results indicate that in contrast to the very promiscuous activity of UBE2D3 *in vitro*, UBE2D3 is a rather selective E2 enzyme *in vivo*, affecting approximately 2% of the global proteome and ± 10% of the recovered diGly-modified peptides. Although it is possible that some targets or UBE2D3 are not detected in our approaches due to incomplete ablation of UBE2D3 activity or redundant activities of other E2 enzymes.

In our global proteome data we identified new potential targets of UBE2D3 that function in retinol metabolism and respond to retinol signaling. In particular, both the SILAC-based and label-free proteomics data identified increased abundance of CRABP1 upon *Ube2d3* depletion with 2 independent shRNAs, which we confirmed in Western blots of cell lysates (**Fig. 2G**). Interestingly, ALDH1A1, identified as a hit with decreased abundance in both LFQ experiments with 2 independent *Ube2d3* shRNAs, is also involved in retinol signaling (**Fig. 2B, C, F**). In the SILAC-based proteomics ALDH1A1 also decreased in abundance, but was not selected as a hit from that data because its change did not meet the threshold for significance (-log10 p-value = 1.27). ALDH1A1 is a retinol dehydrogenase that converts retinaldehyde to RA and thereby plays an important role in retinol metabolism[79,80]. Moreover, RBP1 (Retinol-binding protein 1) was also significantly increased in abundance upon UBE2D3 depletion in the SILAC-based proteomics (**Table S1, S2**). Although RBP1 peptides were unfortunately not identified in the label-free proteomics, by regulating the abundance of CRABP1, ALDH1A1 and RBP1, UBE2D3 might be an important multi-level regulator of retinol metabolism and downstream responses to retinoids. This is interesting in the light of cancer treatment with all-trans retinoic acid (ATRA), a derivative of retinol. ATRA regulates gene expression and cell differentiation and is used to treat acute promyelocytic leukemia (APL), in which it induces cell cycle arrest and results in clinical remission. Interestingly, UBE2D3 was found upregulated in ATRA-treated NB4 APL cells, leading to UBE2D3-mediated degradation of cyclin D1 and cell cycle arrest[11]. This hints at a potential feedback mechanism in which UBE2D3, as regulator of retinol metabolism and responses, is itself responding to ATRA-mediated retinoid signaling, perhaps to contribute to retinoid signaling homeostasis. Interestingly, when UBE2D3 is depleted, NB4 cells were found to acquire resistance to ATRA-treatment. It is tempting to speculate that this could relate to increased retinol metabolism and retinoid response proteins caused by reduced UBE2D3 activity.

Furthermore, we found significantly decreased protein levels of TSPAN8 in UBE2D3-depleted cells, both in our proteomics data and during validation by Western blotting of cell lysates. Analysis of TSPAN8 transcript levels indicate that this decrease is due to regulation of TSPAN8 at the protein level. The mechanism underlying this regulation by UBE2D3, whether it involves regulation of degradation and/or of translation, and what functional consequences it has, are all unclear, but could be interesting directions for further research. TSPAN8 is overexpressed (on protein and/or mRNA level) in various cancers, including pancreatic, lung, breast, ovarian, colon, liver and gastric cancers[40]. High TSPAN8 levels in tumors promote proliferation, migration, angiogenesis and metastasis, and correlate with poor survival. On the other hand, TSPAN8 depletion was shown to reduce cell viability, proliferation and invasion of different tumors and increase apoptosis[40]. It would be interesting to address whether this is regulated by UBE2D3 and whether UBE2D3 levels are potentially also increased in these cancers. Interestingly, a potential mechanism through which UBE2D3 might affect TSPAN8 levels is suggested by our observation that UBE2D3 depletion results in increased ubiquitination of 14-3-3θ on lysine 9 (K9), both in SILAC sh1 as well as LFQ sh2 experiments. 14-3-3 proteins are adaptor and chaperone proteins that function in many different processes such as signal transduction, protein trafficking and apoptosis[58,81]. Interestingly, 14-3-3θ was recently reported to promote TSPAN8 translocation into the nucleus in MDA-MB-231 cells (human breast adenocarcinoma)[82]. Although the consequences of TSPAN8 translocation to the nucleus are as yet unexplored, it could potentially rescue TSPAN8 from being degraded in the

cytoplasm. An interesting direction for further research would be to address whether increased ubiquitination of 14-3-3θ in UBE2D3-depleted cells changes the activity or localization of 14-3-3θ and thereby interferes with 14-3-3θ-mediated TSPAN8 translocation to the nucleus, and thereby its potential protection from degradation in the cytoplasm.

We particularly focused on the role of UBE2D3 in the RQC pathway, a pathway that acts upon ribosome stalling, including at prematurely poly-adenylated mRNAs, and senses 60S ribosomal subunits obstructed with peptidyl-tRNA[72,73]. RQC is associated with multiple layers of ubiquitin-mediated control. These include ubiquitination of the 40S ribosomal protein RPS10 at K107 by the E3 ligase MKRN1 at the start of poly(A) sequences to stall the ribosome[83], followed by ubiquitination of RPS10 at K138/139 and RPS20 at K4/8 by the E3 ligase ZNF598, which recognizes collision of the next ribosome with the stalled ribosome[22-24,84]. Subsequently, RACK1 facilitates RPS2 and RPS3 ubiquitination[24,85], the ribosome disassembles and the RQC pathway is activated, and the nascent polypeptide is polyubiquitinated by the E3 ligase LTN1. This is followed by recruitment of the AAA-ATPase VCP that extracts the ubiquitinated peptide, that is then degraded by the proteasome[72,73]. Here we found that ubiquitination of the 40S ribosomal proteins RPS10 and RPS20, on respectively lysines K138/139 and K4/8, is governed by UBE2D3 *in vivo*, and through direct immunoprecipitation of RPS10 and RPS20 from cells we confirmed that their ubiquitination is indeed dependent on UBE2D3. In addition, through TULIP2 methodology we identified RPS10 and RPS20 as direct substrates of UBE2D3. Moreover, immunoprecipitation experiments in 293T cells confirmed that UBE2D3's catalytic activity as an E2 enzyme is required for ubiquitination of RPS10 *in vivo*, further supporting that RPS10 is a direct target of UBE2D3. In absence of ZNF598-mediated ubiquitination of RPS10-K138/139 and RPS20-K4/8 execution of RQC is compromised[22-24]. Garzia *et al.,* demonstrated before that depletion of UBE2D3 partially impairs RQC and that ZNF598 and UBE2D3 interact with eachother, which points at UBE2D3 as the potential candidate E2 enzyme that functions together with ZNF598 in the ubiquitination of RPS10 and RPS20[22]. Our findings confirm that UBE2D3 is indeed the E2 that is required for the ubiquitination of these residues by ZNF598 *in vivo*. UBE2D1, a closely related and partially redundant family member of UBE2D3, was shown to be able to act *in vitro* with ZNF598 in the ubiquitination of RPS3A, RPS10 and RPS20[23]. However, our results showing that RPS10/20 ubiquitination significantly drops in cells with impaired UBE2D3 activity and that RPS10 and RPS20 are direct targets of UBE2D3, suggest that, although UBE2D1 may contribute, UBE2D3 is the main E2 enzyme ubiquitinating these proteins *in vivo,* in the mouse and human cell lines we used.

Our ubiquitin diGly proteomics data also showed decreased ubiquitination of RPS2, RPS3 and RPS3A at the previously reported ubiquitinated lysines (K275 for RPS2, K214 for RPS3 and K249 for RPS3A) in UBE2D3-depleted cells (**Table S9, S10**). While this suggests RPS2 and RPS3 as potential targets of UBE2D3, it could also reflect an indirect consequence of the reduced RPS10 ubiquitination in UBE2D3-depleted cells, as defective RPS10 ubiquitination was found associated with partially decreased RPS2 and RPS3 ubiquitination[85]. UBE2D3 depletion only slightly (± 1.7-fold change) affects K107 ubiquitination on RPS10 (**Fig. 6A** and **Table S9**), suggesting that UBE2D3 does not operate at the level of MKRN1, preceding RQC and prior to ZNF598 activity. However, besides RPS10 and RPS20 and the effects on RPS2, RPS3 and RPS3A, we also observed impaired ubiquitination of other, later, components of the RQC pathway when UBE2D3 was depleted. Both LTN1 and VCP showed reduced ubiquitination at multiple lysines throughout these proteins in

both SILAC sh1 and LFQ sh2 data sets. However these modifications only reached statistical significance in the SILAC sh1 experiments and their functional importance is thus far unclear (**Table S9, S10**). Nevertheless, while these results clearly implicate UBE2D3 in RQC at the level of RPS10 and RPS20 ubiquitination by ZNF598, UBE2D3 might have a more wide-spread role in RQC by also affecting the ubiquitination of other proteins acting in this pathway.

As RQC is one of the main mechanisms preventing the production and accumulation of defective proteins[86,87], by functioning in this pathway UBE2D3 could have an important role in preserving cellular functions and in preventing diseases linked to protein aggregate formation. Moreover, UBE2D3 also functions in the protein quality control (PQC) pathway that targets misfolded proteins for proteasomal degradation[88] and as we observed in our own data, UBE2D3 might also regulate autophagic PQC via SQSTM1, HSPA5 and TAX1BP1. This indicates that next to preventing the production of defective proteins via RQC, UBE2D3 can also contribute to the removal of dysfunctional proteins and their aggregates via PQC. SQSTM1 and HSPA5 dysregulation has been implicated in neurodegenerative disorders, such as Alzheimer's disease and Parkinson's disease[89,90]. By regulating SQSTM1 and HSPA5 ubiquitination and regulating SQSTM1 activity, UBE2D3 could potentially contribute to clearance of aggregates and thereby help in reducing the chance of development of neurodegenerative diseases. Interesting to note, as it supports a potential role for UBE2D3 in pathology related to its functions in RQC and (autophagic) PQC, the expression of UBE2D3 was found decreased at early stages of Parkinson's disease in mouse brains[91].

Altogether, by combining UBE2D3 depletion with SILAC-based and label-free quantitative ubiquitin diGly proteomics we successfully identified proteins whose ubiquitination and/or abundance *in vivo* is affected by UBE2D3. This gives novel insights into the spectrum of potential *in vivo* UBE2D3 targets and exposes important *in vivo* roles for UBE2D3, in particular in retinol metabolism and signaling, ribosome-associated protein quality control (RQC) and autophagic protein quality control (PQC). We established multiple new links, and for some roles, where indications already existed from previous work that UBE2D3 might be affecting key factors in these pathways, we were able to confirm such links *in vivo* and demonstrate how UBE2D3 changes the levels or ubiquitination of these factors. Most notably, we found that UBE2D3 is the primary E2 enzyme that ubiquitinates RPS10 and RPS20 *in vivo*. It will be interesting to further explore how UBE2D3 affects different cellular processes and pathologies in which retinol metabolism, RQC and PQC play an important role and which E3 enzymes act together with UBE2D3 to ubiquitinate proteins involved in these processes.

EXPERIMENTAL PROCEDURES

Cell culture

All cells were grown in DMEM with 100 U penicillin, 0.1 mg ml^{-1} streptomycin, 2 mM L-glutamine and 10% FBS. *Trf2*$^{-/-}$;*p53*$^{-/-}$;Trf2^{Ile468Ala} mouse embryonic fibroblasts (TRF2ts MEFs) were described before and maintained at 32°C[25]. This cellular model was chosen as in the context of other work we have been using this same p53-deficient cell line to study the potential impact of UBE2D3 in cells exposed to conditional TRF2 inactivation. This cell line allows for comparison between conditions of wild-type or perturbed TRF2 function. However in the work presented in

this manuscript we do not make use of this property. HeLa, Phoenix and HEK 293T cells (ATCC) were grown at 37°C.

Retroviral and lentiviral transductions

For the production of retrovirus, ecotropic phoenix producer cells were seeded on a 10 cm dish and transfected with 20 μg of retroviral vector DNA using $CaPO_4$ precipitation. Medium was refreshed at 16 and 24 h after transfection and viral supernatants were collected at approximately 48, 62 and 72 h post-transfection. Viral supernatants were either frozen in liquid nitrogen and stored at -80°C until use or used immediately. For retroviral infection, cells were overlaid with viral supernatant supplemented with 4 μg ml^{-1} polybrene. Cells were infected a minimum of three times to achieve 100 percent transduction efficiency, which was confirmed by acquired resistance to selection drugs (4 μg ml^{-1} puromycin or 5 μg ml^{-1} blasticidin (Invitrogen)). Cells were transduced with pRetrosuper-puro or pRetrosuper-blast retroviruses encoding shRNA targeting mouse and human *UBE2D3* (*Ube2d3* sh1): 5'-TGGGTTTGGATCACATATC-3', for which we confirmed by q-RT-PCR that it specifically depletes *Ube2d3* and no other member of the UBE2D family. Control cells were transduced with empty pRetrosuper-puro or pRetrosuper-blast retrovirus. For the production of lentivirus, HEK 293T cells were transfected by $CaPO_4$ precipitation with 10 μg of pLKO-puro shRNA lentiviral vectors obtained from Mission library clones (Sigma). Medium was refreshed at 16 h after adding the DNA/$CaPO_4$ precipitate and viral supernatants were collected at approximately 48 h post-transfection. Viral supernatants were either frozen in liquid nitrogen and stored at -80°C until use or used immediately. For lentiviral infection, target cells were incubated for 16 h with viral supernatant that was supplemented with 4 □g ml^{-1} polybrene. Viral transduction was confirmed by acquired resistance to 2 μg ml^{-1} puromycin (Invitrogen). The following lentiviral shRNAs were used: human *UBE2D3* shRNA #1 TRCN0000038792: 5'-CCAGAGATTGCACGGATCTAT-3', human *UBE2D3* shRNA #2 TRCN0000038790: 5'-GCCTGCTTTAACAATTTCTAA-3', mouse *Ube2d3* shRNA #2 TRCN0000039469: 5'-ACAACAGAATATCTCGGGAAT-3' and scrambled non-targeting shRNA: 5'-CAACAAGATGAAGAGCACCAA-3'. Human UBE2D3 wild-type (WT) and C85A mutant cDNA were cloned into a pCDH-puro backbone using traditional cloning and were used for lentiviral infections.

SILAC labeling

Control (pRetrosuper) or *Ube2d3* shRNA-transduced MEFs were grown in either light or heavy SILAC medium (SILAC DMEM Lysine(6) Arginine(10) Kit, 282986434, Silantes), supplemented with 10% FBS, 2 ml Glutamine, 100 U penicillin and 0.1 mg ml^{-1} streptomycin. Control MEFs were grown in light medium (Lys0, Arg0; $^{12}C_6$ lysine and $^{12}C_6$, $^{14}N_4$ arginine). *Ube2d3* shRNA1-transduced MEFs were grown in heavy medium (Lys6, Arg10; $^{13}C_6$ lysine and $^{13}C_6$, $^{15}N_4$ arginine). Cells were cultured for at least 2 weeks in SILAC light or heavy medium before the experiments were done to accomplish complete labelling. The SILAC experiments were performed and analysed as 6 replicates.

SILAC sample preparation

Cells grown in SILAC light or heavy medium were harvested for global proteome and ubiquitinome analyses. Cells were washed three times with ice-cold PBS and lysed in 200 μl of an

8 M urea / 50 mM Tris-HCl pH 8.0 / 50 mM NaCl lysis buffer. Lysates were incubated on ice for 10 min and sonicated, debris was removed by centrifugation and protein concentrations were determined using a BCA assay (Pierce). Control and UBE2D3-depleted samples were mixed in a 1:1 ratio based on total protein content. 20 mg protein was used for diGly peptide enrichment for ubiquitinome analyses; for global proteome analyses 0.5 mg lysate was used.

Protein digestion and fractionation (SILAC)

Protein lysates were reduced with 10 mM dithiothreitol (DTT) for 1 h at room temperature followed by alkylation with 55 mM chloroacetamide (CAM) for 1 h in the dark. The mixture was diluted four times with 50 mM ammonium bicarbonate buffer before the addition of CaCl2 (1 mM final concentration). Proteins were digested with sequencing grade trypsin (1:100 (w:w), Roche) overnight at room temperature. Alternatively, proteins were digested with LysC (1:100 (w:w), Wako Chemicals) for 1 h at room temperature before trypsinization. Protein digests were then desalted using a Sep-Pak tC18 Vac cartridge (Waters) and eluted with 80% acetonitrile (AcN). Tryptic peptides were fractionated by HILIC on an Agilent 1100 HPLC system using a 5 µm particle size 4.6 x 250 mm TSKgel amide-80 column (Tosoh Biosciences). 200 µg of tryptic digest in 80% AcN was loaded onto the column. Peptides were eluted using a non-linear gradient from 80% B (100% AcN) to 100% A (20 mM ammonium formate in water) with a flow of 1 ml/min. Sixteen 6 ml fractions were collected, lyophilized and pooled into 8 final fractions. Each fraction was then analyzed by nanoflow LC-MS/MS.

DiGly peptide enrichment (SILAC)

DiGly-modified peptides were enriched by immunoprecipitation using PTMScan® ubiquitin remnant motif (K-ε-GG) antibody bead conjugate (Cell Signaling Technology) starting from 20 mg total protein, essentially according to the manufacturer's protocol. Unbound peptides were removed by washing and the captured peptides were eluted with a low pH buffer. Eluted peptides were analyzed by nanoflow LC-MS/MS.

Nanoflow LC-MS/MS (SILAC)

Nanoflow LC-MS/MS was performed on an EASY-nLC system (Thermo) coupled to a Fusion Lumos Tribrid Orbitrap mass spectrometer (Thermo), operating in positive mode and equipped with a nanospray source. Peptide mixtures were trapped on a ReproSil C18 reversed phase column (Dr Maisch GmbH; column dimensions 1.5 cm × 100 µm, packed in-house) at a flow rate of 8 µl/min. Peptide separation was performed on ReproSil C18 reversed phase column (Dr Maisch GmbH; column dimensions 15 cm × 50 µm, packed in-house) using a linear gradient from 0 to 80% B (A = 0.1% FA; B = 80% (v/v) AcN, 0.1 % FA) in 70 or 120 min and at a constant flow rate of 250 nl/min. The column eluent was directly sprayed into the ESI source of the mass spectrometer. All mass spectra were acquired in profile mode. The resolution in MS1 mode was set to 70,000 (AGC: 3E6), the m/z range 350-1700. Fragmentation of precursors was performed in data-dependent mode by HCD (Top15) with a precursor window of 3.0 m/z and a normalized collision energy between 26.0-28.0; MS2 spectra were recorded in the ion trap. Singly charged precursors were excluded from fragmentation. Dynamic exclusion was set to 20 seconds and the intensity

threshold was set to 8.0E3. For the ubiquitinome analysis, a single LC-MS/MS run was performed for all immunoprecipitated peptide material from one sample.

Data analysis (SILAC)

Mass spectrometric raw data were analyzed using the MaxQuant software suite (version 1.6.15.0)[26] for identification and relative quantification of proteins. Precursor mass tolerances were set to 10 ppm and fragment mass tolerance was set to 20 ppm for Orbitrap, 0.6 Da for ion-trap fragmentation data. A false discovery rate (FDR) of 0.01 for proteins and peptides, and a minimum peptide length of 6 amino acids were required. The Andromeda search engine was used to search the MS/MS spectra against the Mus musculus Uniprot database (63,756 entries; version: up_mouse_2017_09.fasta) concatenated with the reversed versions of all sequences and a contaminant database listing typical background proteins. A maximum of two missed cleavages were allowed. MS/MS spectra were analyzed using MaxQuant's default settings for Orbitrap and ion trap spectra. The maximum precursor ion charge state used for searching was 7 and the enzyme specificity was set to trypsin. Further modifications were Cysteine carbamidomethylation (fixed) as well as Methionine oxidation and Lysine ubiquitination (variable). The minimum number of peptides for positive protein identification was set to 2. Heavy-to-light (H:L) ratios were calculated using MaxQuant's default settings, including a minimum ratio count of 2 for label based protein quantification and of 1 for diGly peptide quantification. The minimum number of razor and unique peptides was set to 1. Only unique and razor non-modified, methionine oxidized and protein N-terminal acetylated peptides were used for protein quantitation. The minimal score for modified peptides was set to 40 (default value). The 'requantify' option was selected in all cases. Only proteins that were identified and quantified in both duplicates and with consistent ratios, were considered for further analysis. The proteingroups.txt, modificationspecificpeptides.txt and GlyGly_(K)sites.txt output tables were all processed in R (version 4.1.0). Reverse hits (decoys) and potential contaminants were filtered excluded from analysis. Normalized ratios of all replicate experiments were Log2 transformed and these values were used in a one sample two-sided t-test (mu = 0) to calculate p-values. The modificationspecificpeptides.txt table was further adjusted by adding the diGly sites probability column from the GlyGly_(K)sites.txt table. Protein sets were further analyzed with Perseus (version 1.6.0.7)[27], through the use of Ingenuity Pathway Analysis (IPA; QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis)[28], STRING analysis (https://string-db.org/)[29] and in-house developed software[30]. For statistical testing of quantitative data, we used a one-sample t-test in Perseus using 250 randomizations and further settings FDR=0.05 and S0=0.5. Zipped MaxQuant data have been deposited to the PRIDE repository with the data identifier PXD035045.

Sample preparation and diGly peptide enrichment (LFQ)

Six 15 cm diameter plates of MEFs were grown and harvested for each condition (control shRNA, *Ube2d3* shRNA1 and *Ube2d3* shRNA2). Cells were washed twice with ice-cold PBS. Lysis, digestion and enrichment of ubiquitinated peptides was performed according to the PTMScan HS K-ε-GG Remnant Magnetic Immunoaffinity Beads Kit protocol (Cell Signaling Technology, catalog #34608). Briefly, cell pellets were lysed in 1x S-Trap lysis buffer, sonicated with a probe and lysates were cleared by centrifugation. Equal protein amounts (2mg per sample) were reduced with DTT,

alkylated with chloroacetamide and digested overnight at 37°C with trypsin (Sigma-Aldrich, 1:10) using S-Trap Midi cartridges (ProtiFi, NY, United States). Peptides were eluted with elution buffers 1 to 3, after which 20µg aliquots were taken for proteome analysis, the remainder being reserved for ubiquitinated peptide enrichment. All samples were dried in a Speedvac and stored at -80°C until LC-MS/MS or further processing. Immmunoaffinity purification of ubiquitinated peptides was performed with K-ε-GG Remnant Magnetic Immunoaffinity Beads using the manufacturer's instructions. Ubiquitinated peptides were eluted twice with 50 µL 0.15% TFA and concentrated with C18 stage tips (Thermo Scientific, SP301).

Nanoflow LC-MS/MS (LFQ)

Prior to mass spectrometry analysis, peptides were reconstituted in 2% formic acid. Peptide mixtures were analyzed by nanoLC-MS/MS on an Q Exactive HF-X Hybrid Quadrupole-Orbitrap Mass Spectrometer equipped with an EASY-NLC 1200 system (Thermo Scientific). Samples were directly loaded onto the analytical column (ReproSil-Pur 120 C18-AQ, 1.9µm, 75 µm × 500 mm, packed in-house). Solvent A was 0.1% formic acid/water and solvent B was 0.1% formic acid/80% acetonitrile. Samples were eluted from the analytical column at a constant flow of 250 nl/min. For single-shot proteome analysis, a 210-min gradient containing a linear 194-min increase from 6 to 26% solvent B was used, whereas a 144-min gradient containing a linear 124-min increase from 4 to 24% solvent B was used for ubiproteome analysis. MS settings were as follows: full MS scans (375-1500 m/z) were acquired at 60,000 resolution with an AGC target of 3 × 106 charges and max injection time of 45 ms. Loop count was set to 20 and only precursors with charge state 2-7 were sampled for MS2 using 15,000 resolution (or 30,000 for ubiproteome), MS2 isolation window of 1.4 m/z, 1 × 105 AGC target, a max injection time of 22 ms (or 54 ms for ubiproteome) and a normalized collision energy of 26.

Data analysis of proteome and ubiquitinated peptides (LFQ)

Proteome and ubiproteome data (RAW files) were analyzed by label-free quantitation (LFQ) using MaxQuant (version 1.6.17.0)[31] with standard settings. Precursor mass tolerances were set to 20 ppm in the first search and 4.5 ppm in the main search, and fragment mass tolerances was set to 20 ppm. MS/MS data were searched against the mouse Swissprot reviewed database (17,042 entries, release 2020_07). Carbamidomethylation on cysteine was specified as fixed modification, whereas methionine oxidation and protein N-terminal acetylation were set as variable modifications. Trypsin/p was specified as cleavage specificity, false discovery rates were set to 1% for both protein and peptide level and in the case of diGly peptide enrichment experiments, and GG(K) was set as additional variable modification. LFQ intensities were Log2-transformed in Perseus (ver. 1.6.14.0)[27], after which proteins (in the case of proteome analysis) or ubiquitination sites (in the case of ubiproteome analysis) were filtered for at least three valid values (out of 3 total) in at least one condition. For ubiquitination sites, localization probability > 0.75 was used as additional filter and intensities were normalized by median subtraction. Missing values were replaced by an imputation-based normal distribution using a width of 0.3 and a downshift of 1.8. Differential proteins or ubiquitination sites were determined using a t-test (cutoffs: $p<0.05$ and LFQ differences $\log2 \geq 1.0$ and $\log2 \leq -1.0$). Annotated spectra have been deposited to MS-Viewer: For proteome, the link is: https://msviewer.ucsf.edu/prospector/cgi-

bin/mssearch.cgi?report_title=MS-Viewer&search_key=3opswum57r&search_name=msviewer (key: 3opswum57r) and for the diGly data set the link is: https://msviewer.ucsf.edu/cgi-bin/mssearch.cgi?report_title=MS-Viewer&search_key=kapdokzmqp&search_name=msviewer (key: kapdokzmqp). Zipped MaxQuant data have been deposited to the PRIDE repository with the data identifier PXD035045.

UBE2D3 TULIP2 sample preparation

UBE2D3-TULIP2 samples were prepared as previously described [21], UBE2D3 wild type and C85A mutant were cloned in the TULIP2 and TULIP2ΔGG plasmids by Gateway® cloning (Thermo Fisher Scientific). TULIP2 constructs were introduced in HeLa cells by lentiviral transduction and positive clones were selected on 3 mg/ml puromycin.

6

Five 15 cm diameter plates of HeLa cells were grown up to 60-80% confluency. Expression of UBE2D3-TULIP2 constructs was induced with 1 µg/ml doxycycline for 24 h and cells were untreated or treated for 5 h with 10 µM proteasome inhibitor MG132 (Sigma Aldrich). Next, cells were washed twice with ice-cold PBS, scraped and lysed in 10 ml Guanidinium buffer (6M guanidine-HCl, 0.1 M Sodium Phosphate, 10 mM TRIS, pH 7.8). Samples were homogenized at room temperature by sonication using a tip sonicator (Q125 Sonicator, QSonica, Newtown, USA). Protein concentration was determined by using BiCinchoninic Acid (BCA) Protein Assay Reagent (Thermo Scientific) and equalized accordingly.

Once equalized, lysates were supplemented with 5 mM β-mercaptoethanol and 50 mM Imidazole pH 7.8. 100 µl of nickel-nitrilotriacetic acid-agarose (Ni-NTA) beads (QIAGEN), were equilibrated with Guanidinium buffer supplemented with 5 mM β-mercaptoethanol and 50 mM Imidazole pH 7.8, added to the cell lysates and incubated overnight at 4°C under rotation. Next, Ni-NTA beads were transferred with Wash buffer 1 (6 M Guanidine-HCl, 0.1 M Sodium Phosphate, 10 mM TRIS, 10 mM Imidazole, 5 mM β-mercaptoethanol, 0.2% Triton X-100, pH 7.8) to an Eppendorf LoBind tube (Eppendorf) and sequentially washed with Wash buffer 2 (8 M Urea, 0.1 M Sodium Phosphate, 10 mM TRIS, 10 mM imidazole, 5 mM β-mercaptoethanol, pH 8), Wash buffer 3 (8 M urea, 0.1 M Sodium Phosphate, 10 mM TRIS, 10 mM imidazole, 5 mM β-mercaptoethanol, pH 6.3) and twice with Wash buffer 4 (8 M urea, 0.1 M Sodium Phosphate, 10 mM TRIS, 5 mM β-mercaptoethanol, pH 6.3). In every wash step, beads were allowed to equilibrate with the buffer for 15 min under rotation.

Ni-NTA beads were resuspended in one bead-volume of elution buffer without imidazole (7 M urea, 0.1 M Sodium Phosphate and 10 mM TRIS pH 7.0) and incubated with 375 ng of rLys-C (Promega) for 5h at 37 °C on a shaker at 1400 rpm. Subsequently, 4 volumes of 50 mM ammonium bicarbonate were added to the samples and a second digestion using 500 ng Trypsin (Promega) was done overnight at 37°C and 1400 rpm for a complete digestion. Resulting peptides were purified and desalted using C-18 StageTips[32].

Mass Spectrometry Data acquisition (UBE2D3-TULIP2)

UBE2D3 samples were analyzed in an Orbitrap Exploris 480 (Thermo Fisher Scientific) mass spectrometer coupled to an Ultimate 3000 UHPLC (Dionex). Digested peptides were separated using a 50 cm long fused silica emitter (FS360-75-15-N-5-C50, New Objective, Massachusetts, US)

in-house packed with 1.9 µm C18-AQ beads (Reprospher-DE, Pur, Dr. Maisch, Ammerburch-Entringen, Germany) and heated to 50ºC in a Column Oven for ESI/Nano Spray (Sonation, Germany). Peptides were separated by liquid chromatography using a gradient from 2% to 32% acetonitrile with 0.1% formic acid for 100 min followed by column re-conditioning for 20 minutes. A Lock Mass of 445.12003 (polysiloxane) was used for internal calibration. Data was acquired in a Data Dependent Acquisition (DDA) mode with a TopSpeed method with cycle time of 3 s with a scan range of 300-1600 m/z and resolutions of 60,000 and 30,000 for MS1 and MS2, respectively. For MS2, an isolation window of 1.6 m/z and an HCD collision energy of 28% was applied. Precursors with a charge of 1 and higher than 6 were excluded from triggering MS2 as well as previously analyzed precursors with a dynamic exclusion window of 30s.

Mass Spectrometry Data analysis (UBE2D3-TULIP2)

Mass spectrometry data was analyzed using MaxQuant v1.6.14.0 according to Tyanova et al[26] with the following modifications: Maximum missed cleavages by trypsin was set to 4. The initial precursor maximum mass tolerances were set to 20 ppm in the first search and 4.5 ppm in the main search, and the fragment mass tolerances was set to 20 ppm.Searches were performed against an *in silico* digested database from the human proteome including isoforms and canonical proteins (96,850 entries; Uniprot, 8th June 2020). Oxidation (M), Acetyl (Protein N-term), GlyGly (K) and Phospho (STY) were set as variable modifications with a maximum of 3. Carbamidomethyl (C) was disabled as fixed modification. Label-free quantification was activated not enabling Fast LFQ. False Discovery Rates at the peptide and protein level was 0.01. The match between runs feature was activated with default parameters.

MaxQuant output data were further processed in the Perseus Computational Platform v1.6.14 according to Tyanova et al[27]. LFQ intensity values were log2 transformed and potential contaminants and proteins identified by site only or reverse peptide were removed. Samples were grouped in experimental categories and proteins not identified in 4 out of 4 replicates in at least one group were also removed. Missing values were imputed using normally distributed values with a 1.8 downshift (log2) and a randomized 0.3 width (log2) considering whole matrix values. Two-sided t-tests were performed to compare groups. Analysed data were exported from Perseus and further processed in Microsoft Excel 365 for comprehensive visualization. Volcano plots were generated using VolcanoseR[33]. Annotated spectra have been deposited to MS-Viewer:

https://msviewer.ucsf.edu/cgi-bin/mssearch.cgi?report_title=MS-Viewer&search_key=lqf88o5m68&search_name=msviewer (key: lqf88o5m68).

Quantitative real-time PCR

To determine gene expression levels, total RNA was isolated using TRIzol reagent (Ambion), and reverse transcribed into cDNA using the Tetro cDNA Synthesis kit (BIO-65043, Bioline) with Oligo(dT) primers. Quantiative real-time PCR was performed using the SensiFAST SYBR No-ROX kit (BIO-98020, Bioline) on a QuantStudio 5 Flex real-time PCR system (Thermo Fisher Scientific). *Hprt* was used as a reference for transcript expression. Transcripts were amplified using the following primers for mouse *Tspan8* Fw: 5'-CTGACTGTGCAACTTATCA GG-3' and mouse *Tspan8* Rev: 5'-GCCAGTCCAAAAGCAATTCC-3'; for mouse *Crabp1* Fw: 5'-CGGAGATCAACTTCAAGGTCGG-3' and mouse *Crabp1* Rev: 5'-CCCTCAAGAAGTGTCT GTGTGC-3'; for mouse *Hprt* Fw: 5'-

CTGGTGAAAGGACCTCTCG-3' and mouse *Hprt* Rev: 5'-TGAAGTACTCATTATAGTCAAGGGCA-3'. Data was analyzed according to the 2-DCt methodology.

Immunoprecipitation

For immunoprecipitation with ectopically expressed V5-tagged constructs, control (non-targeting shRNA) and UBE2D3-depleted HEK 293T cells were co-transfected with pcDNA3.1(+)-HA-ubiquitin[34] and either pLX304-blast-V5 or pLX304-blast-RPS10-V5 (CCSB-Broad Lentiviral Expression Library hORFeome v8.1)[35]. Or HEK 293T cells expressing pCDH-puro, pCDH-puro-UBE2D3 WT or pCDH-puro-UBE2D3 C85A were transfected with pcDNA3.1(+)-HA-ubiquitin and either pLX304-blast-V5 or pLX304-blast-RPS10-V5. 48 h post-transfection, cells were washed in ice-cold PBS and scraped in 450 μl lysis buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl and 0.5% Triton X), freshly supplemented with protease inhibitors (cOmplete, 4693124001, Roche), phosphatase inhibitors (PhosSTOP™, 4906837001, Roche), 0.1 mM PMSF, 1 mM DTT and 10 mM Iodoacetamide. 150 μl B3 buffer (100 mM Tris-HCl pH 8.0, 1 mM EDTA, 2% SDS) was added and samples were sonicated. 900 μl lysis buffer was added and samples were centrifuged for 20 minutes at maximum speed at 4°C. Lysates were added to V5-conjugated Protein G Dynabeads (Invitrogen) and incubated overnight while rotating at 4°C. Beads were washed 4 times with lysis buffer and bound proteins were eluted with 2x sample buffer (100 mM Tris-HCl pH 6.8, 4% SDS and 20% glycerol) containing 10 mM DTT and boiling for 15 minutes at 100°C. For endogenous immunoprecipitation experiments, control (non-targeting shRNA) and UBE2D3-depleted HEK 293T cells were transfected with pcDNA3.1(+)-HA-ubiquitin and the same procedure as above was followed, with the exception that the pulldown was done with RPS20 antibody-conjugated Protein G Dynabeads.

Immunoblotting

Whole-cell lysates were prepared by scraping cells in SDS sample buffer (125 mM Tris pH 6.8, 20% glycerol, 4% SDS). Lysates were sheared with a 25G needle and boiled for 10 minutes at 95°C. Protein concentration was determined by standard BCA protein assay (Pierce). Equal amounts of protein were separated on precast polyacrylamide gels (Invitrogen). Immunoblotting was done according to standard methods using IRDye800CW- and IRDye680-labelled secondary antibodies for detection on the Odyssey Infrared imager (LI-COR) or using horseradish peroxidase (HRP)-conjugated secondary antibodies for detection by enhanced chemiluminescence (Supersignal and Supersignal West Pico Plus, Thermo Scientific). Primary antibodies used were against UBE2D (A615, Boston Biochem, 1:2000 or 4330S, CST, 1:500), V5 (R960-25, Invitrogen, 1:1,000), HA (MMS-101R, Covance, 1:1,000), FLAG (F3165, Sigma, 1:1,000), RPS20 (ab133776, Abcam, 1:1,000), Histone H3 (ab1791, Abcam, 1:10,000), CRABP1 (HPA017203, Sigma, 1:1,000), TSPAN8 (A06997-1, Sanbio, 1:1,000), GAPDH (PA1-987, ThermoFisher, 1:5,000), HSP90 α/β (sc-7947, Santa Cruz, 1:1,000), □-actin (A5316, Sigma, 1:10,000) and γ-tubulin (T6557, Sigma, 1:10,000).

Experimental Design and Statistical Rationale

The number of samples in the SILAC-based diGly proteomics was two, since one experimental condition (*Ube2d3* shRNA1) was studied with respect to its appropriate control (control shRNA). Six biological replicates of both sample and control were performed in a 'forward' SILAC fashion, resulting in a total number of 12 samples. For every sample the appropriate control was analysed

in the same experiment, *i.e.* sample in the SILAC heavy channel and control in the SILAC light channel. Statistical analysis was performed using a two-sided one-sample t test, which is appropriate for SILAC H:L ratio analysis based on multiple replicates.

For all three experimental conditions in the LFQ diGly proteomics (control shRNA, *Ube2d3* shRNA1 and *Ube2d3* shRNA2) three biological replicates were used for comparison of protein or ubiquitinated peptide LFQ abundances, as this number of replicates is the minimum number of replicates required in order to be able to reliably use a two-sided student's t-test for comparison of each shRNA-cell line versus the wild-type control condition.

The total number of samples in the TULIP2 assays were 24, corresponding to the TULIP2 UBE2D-WT, TULIP2ΔGG UBE2D3-WT and TULIP2 UBE2D3-C85A in four biological replicates, either treated or not treated with the proteasome inhibitor MG132. TULIP2ΔGG UBE2D3-WT and TULIP2 UBE2D3-C85A both serve as negative controls. Four replicates were chosen to have enough statistical power, while keeping the size of the project on a reasonable size. Statistical analysis was performed using two-sided t-tests.

DATA AVAILABILITY

SILAC-based and LFQ diGly proteomics: MS raw data and data for protein identification and quantification were submitted as supplementary tables to the ProteomeXchange Consortium via the PRIDE partner repository[92] with the data identifier PXD035045. For reviewing purposes, the following credentials can be used:

- Username: reviewer_pxd035045@ebi.ac.uk
- Password: IRIBP4Jh

TULIP2: the mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD026054. For reviewing purposes, the following credentials can be used:

- Username: reviewer_pxd026054@ebi.ac.uk
- Password: GMBhTQpr

For LFQ proteomics annotated spectra have been deposited to MS-Viewer: For proteome, the link is:

https://msviewer.ucsf.edu/prospector/cgi-bin/mssearch.cgi?report_title=MS-Viewer&search_key=3opswum57r&search_name=msviewer (key: 3opswum57r)

For the diGly data set the link is:

https://msviewer.ucsf.edu/cgi-bin/mssearch.cgi?report_title=MS-Viewer&search_key=kapdokzmqp&search_name=msviewer (key: kapdokzmqp).

For TULIP2 mass spectrometry annotated spectra have been deposited to MS-Viewer:

https://msviewer.ucsf.edu/cgi-bin/mssearch.cgi?report_title=MS-Viewer&search_key=lqf88o5m68&search_name=msviewer (key: lqf88o5m68).

AUTHOR CONTRIBUTIONS

J.J.L.J. and Z.Y. conceptualization; Z.Y., K.B., D.S-L., O.B.B., D.K. and M.F. methodology; Z.Y., D.K. and M.F. validation; Z.Y., K.B., D.S-L., O.B.B. and D.K. formal analysis; Z.Y., K.B., D.S-L., O.B.B. and V.B. investigation; Z.Y. writing-original draft; Z.Y. and J.J.L.J. writing-review & editing; Z.Y. and D.S-L. visualization; J.J.L.J., J.A.A.D., M.A. and R.G-P supervision.

ACKNOWLEDGEMENTS

We thank A. Sapmaz and H. Ovaa for the HA-ubiquitin plasmid and for help with the ubiquitin-immunoprecipitation protocol. This work was supported by an institutional grant of the Dutch Cancer Society and of the Dutch Ministry of Health, Welfare and Sport to the Netherlands Cancer Institute, by a Research Project Grant from the Dutch Cancer Society (KWF-NKI2012-5305) to J.J.L.J., by a Young Investigator Grant from the Dutch Cancer society (11369/2017-2) to R.G-P and by the X-omics Initiative, part of the NWO National Roadmap for Large-Scale Research Infrastructures, to M.A.

SUPPLEMENTARY FIGURES AND DATASET LEGENDS

Table S1: List of identified proteins in SILAC-based global proteomics

Table S2: Proteins increased or decreased in abundance in SILAC-based proteomics

Table S3: Proteins in Venn diagrams

Table S4: List of identified proteins in LFQ *Ube2d3* sh1 global proteomics

Table S5: List of identified proteins in LFQ *Ube2d3* sh2 global proteomics

Table S6: Proteins increased or decreased in abundance LFQ *Ube2d3* sh1 and LFQ *Ube2d3* sh2

Table S7: List of non-modified and diGly modified peptides in SILAC-based diGly proteomics with *Ube2d3* sh1

Table S8: List of diGly modified peptides including sites of modifications within each peptide for SILAC-based diGly proteomics with *Ube2d3* sh1

Table S9: List of diGly modified peptides increased or decreased in ubiquitination in SILAC-based diGly proteomics with *Ube2d3* sh1

Table S10: List of diGly modification sites in label-free quantification with *Ube2d3* sh2

Table S11: List of diGly modification sites increased or decreased in ubiquitination in label-free quantification with *Ube2d3* sh2

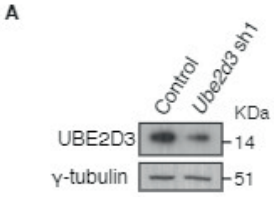
Table S12: IPA analysis of SILAC-based diGly proteomics hits increased in ubiquitination with *Ube2d3* sh1

Table S13: IPA analysis of LFQ diGly proteomics hits increased in ubiquitination with *Ube2d3* sh2

Table S14: Overlapping pathways IPA analysis of SILAC-based & label-free diGly proteomics hits increased in ubiquitination with *Ube2d3* sh1 (SILAC) and *Ube2d3* sh2 (LFQ)

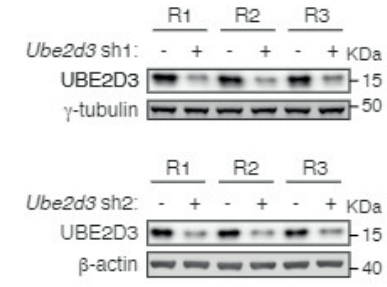
Table S15: IPA analysis diGly proteomics hits overlapping between SILAC *Ube2d3* sh1 and LFQ *Ube2d3* sh2

Table S16: List of identified proteins in TULIP2 assays



C

Numbers of identified proteins SILAC Ube2d3 sh1			
Experiment	Total # proteins	Total Up	Total Down
1	6377	116	209
2	6975	119	152
3	6337	138	206
4	6503	127	232
5	7136	132	199
6	6877	165	176



D

Identified diGly peptides and the effect of *Ube2d3* depletion (SILAC sh1) on their ubiquitination status	
	# diGly peptides
Total number of peptides	19599
- with 1 diGly site	18794
- with 2 diGly sites	788
- with 3 diGly sites	16
- with 4 diGly sites	1
Total number with increased ubiquitination*	482
Total number with decreased ubiquitination*	3573

Figure S1. Validation of UBE2D3 depletion in proteomics experiments and numbers of identified proteins and peptides in SILAC Ube2d3 sh 1 experiments. **A**, Representative immunoblots for Ube2d3 depletion in the SILAC-based (diGly) proteomics experiments. **B**, Immunoblots for Ube2d3 depletion in all three replicates of the LFQ (diGly) proteomics. **C**, Table with numbers of identified proteins in all six SILAC Ube2d3 sh1 replicates, including numbers of proteins that are ≥ 1.5 fold increased or decreased in abundance upon UBE2D3 depletion. **D**, Table with identified diGly peptides and the effect of UBE2D3 depletion on the ubiquitination status of diGly modified peptides in all SILAC Ube2d3 sh1 experiments. Asterisk indicates that these numbers are based on a cut-off of ≥ 1.5-fold increase or decrease in ubiquitination (increase: log2 ≥ 0.585; decrease: log2 ≤ -0.585).

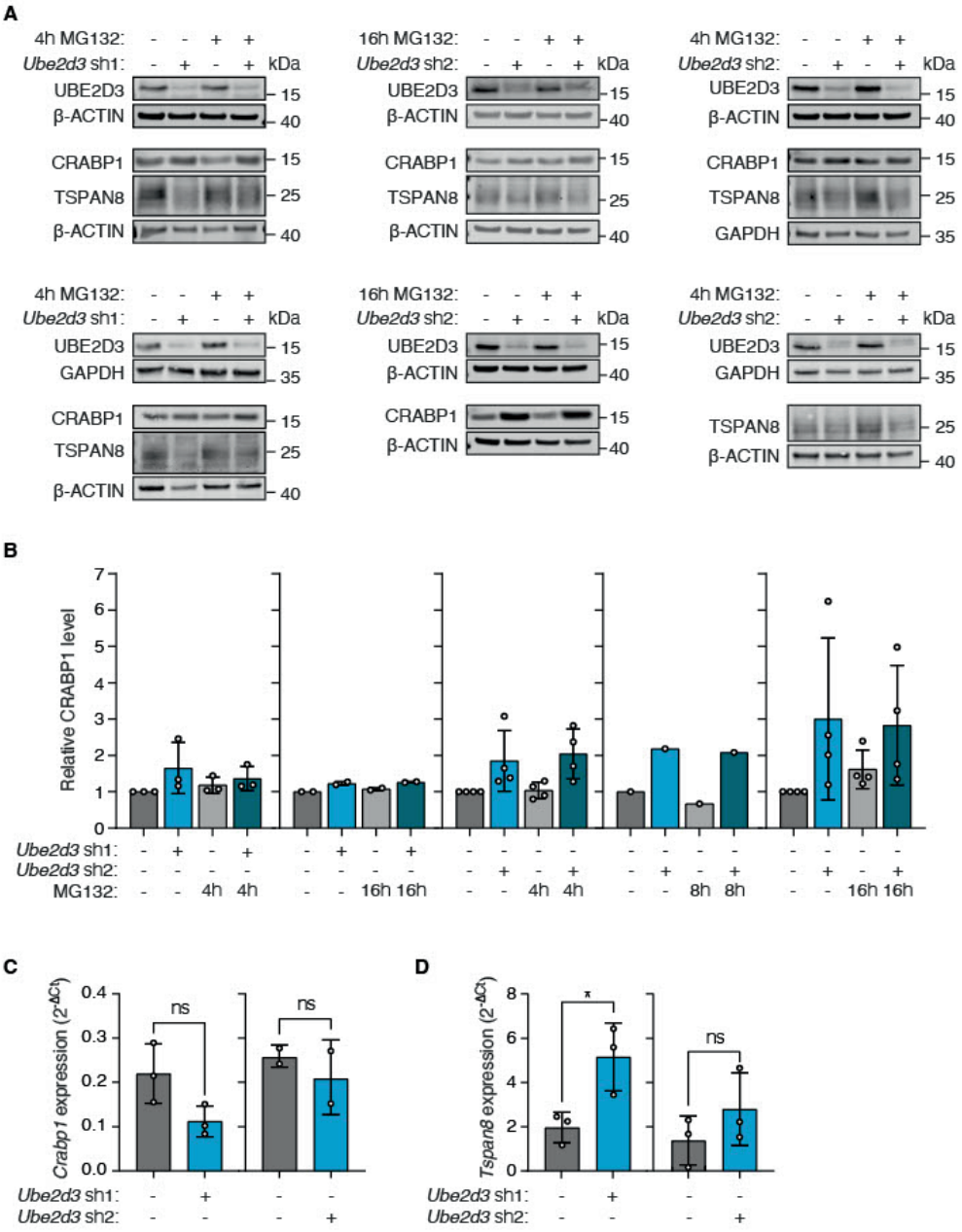
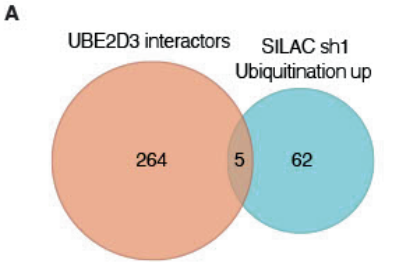
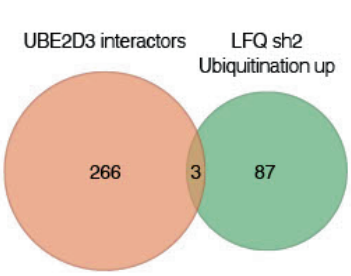


Figure S2. Validation of CRABP1 and TSPAN8 as targets of UBE2D3. **A**, Two additional biological replicates for CRABP1 and TSPAN8 protein levels in Ube2d3 sh1 cells and three additional biological replicates for CRABP1 and TSPAN8 protein levels in Ube2d3 sh2 cells are shown for the immunoblot experiments in Figure 2G. **B**, Quantification of CRABP1 protein levels relative to control (scrambled) shRNA cells in DMSO. Quantified from immunoblots shown in Figure 2G and Figure S2A. The mean ± SD is shown and each dot

represents an independent experiment. For Ube2d3 sh1: n=3 at 4h and n=2 at 16h of DMSO or MG132 treatment. For Ube2d3 sh2: n=4 at 4h, n=1 at 8h and n=4 at 16h of DMSO or MG132 treatment. Statistical significance was calculated using a two-tailed Student t-test. Differences in CRABP1 protein levels between control shRNA and Ube2d3 shRNA samples did not reach statistical significance. **C-D**, q-RT PCR analysis of Crabp1 (C) and Tspan8 (D) mRNA levels in MEFs. The mean ± SD for 3 independent experiments is shown, except for Crabp1 expression in Ube2d3 sh2, which represents an n=2. Each dot represents an independent. Statistical significance was calculated using a two-tailed Student t-test (ns, non-significant ($p \geq 0.05$), *$p = 0.0303$).



Overlapping hits	
Name	**Biological function**
ARIH1	E3 ligase
DZIP3	E3 ligase
FBXW11	F-box protein
UBE2S	E2 ubiquitin-conjugating enzyme
UHRF1	E3 ligase



Overlapping hits	
Name	**Biological function**
HERC2	E3 ligase
UBE4B	E3 ligase
UBTD1	Ubiquitin domain-containing protein

Figure S3. Venn diagrams of proteins increased in ubiquitination upon UBE2D3 depletion. **A**, Venn diagrams illustrating the overlap between known UBE2D3 interactors (BioGRID, IntAct, MINT, STRING and UniProt databases) and proteins that go up in ubiquitination upon UBE2D3 depletion in the SILAC diGly proteomics experiments. The table below shows the overlapping hits and their biological function. **B**, Venn diagrams illustrating the overlap between known UBE2D3 interactors and proteins that go up in ubiquitination upon UBE2D3 depletion in the LFQ sh2 experiments. The table below shows the overlapping hits and their biological function.

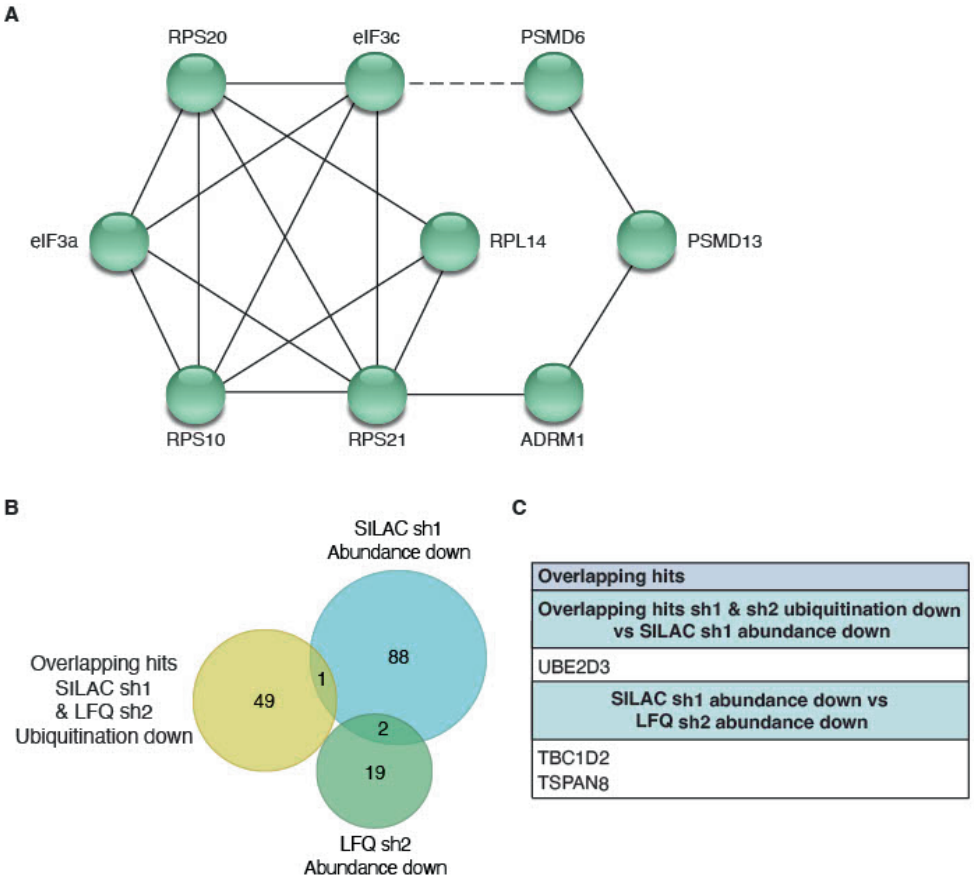


Figure S4. Decreased protein abundance and ubiquitination upon UBE2D3 depletion. **A**, STRING analysis of proteins that are decreased in their ubiquitination upon Ube2d3 depletion with sh1 and sh2, zoomed in on a network of proteins in volved in mRNA translation. Lines between proteins represent known interactions, dotted line represent a putative interaction. **B**, Venn diagrams illustrating the overlap between the 50x overlapping hits from the SILAC Ube2d3 sh1 and LFQ Ube2d3 sh2 that are significantly downregulated in their ubiquitination and the proteins significantly downregulated in their abundance in the SILAC Ube2d3 sh1 and LFQ Ube2d3 sh2 proteomics. **C**, Table showing the overlapping hits in B.

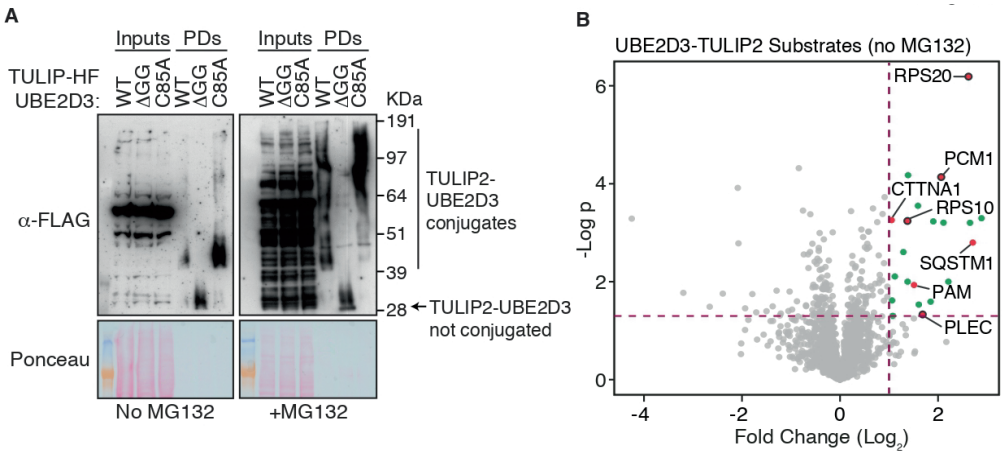


Figure S5. UBE2D3-TULIP2 Western blot validation and volcano plot of results without proteasome inhibition. **A**, Expression of UBE2D3-TULIP2, UBE2D3-TULIP2ΔGG and UBE2D3-C85A-TULIP2 in HeLa cells was induced o/n with doxycycline, not treated or treated with proteasome inhibitor (no MG132 or +MG132) and lysed. TULIP2 conjugates were purified following TULIP2 methodology. 0.1% of whole cell extract serves as input and 5% of the purified proteins serves as pull downs (PDs). Ponceau S is provided as a loading control. **B**, Volcano plot showing UBE2D3 substrates from TULIP2 experiments (n=4) without proteasome inhibition (no MG132). Green dots indicate statistically enriched proteins in the UBE2D3-TULIP2 samples compared to UBE2D3-TULIP2ΔGG samples for p-value = 0.05 and S0 = 0.1. Red dots indicate top hits significantly decreased in their ubiquitination in SILAC Ube2d3 sh1 experiments. Red dots with a black stroke around them also overlap with top hits significantly decreased in their ubiquitination in LFQ Ube2d3 sh2 experiments.

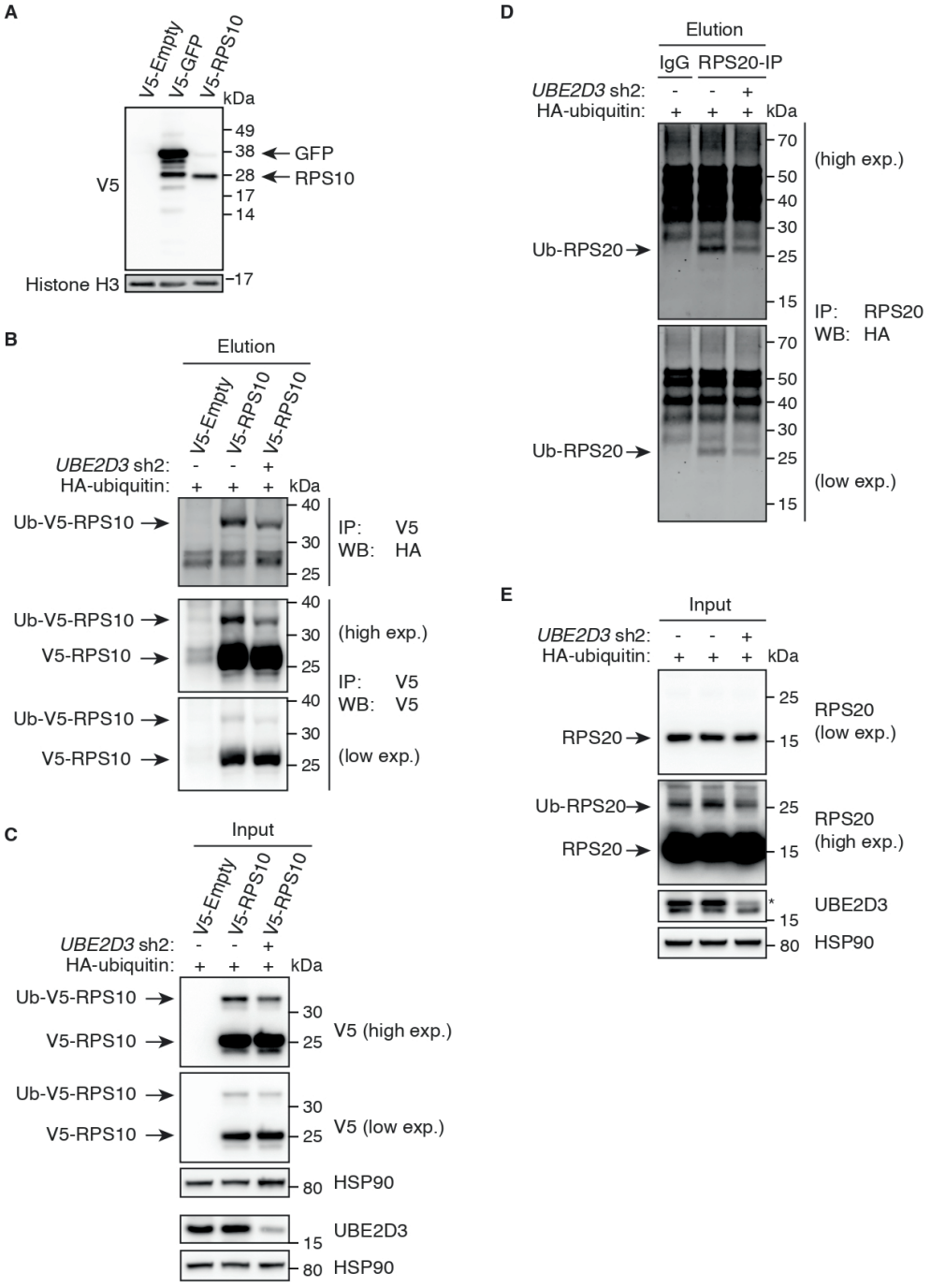

A
V5-Empty
V5-GFP
V5-RPS10
kDa
49
38
28
17
14
GFP
RPS10
V5
Histone H3
17
B
Elution
V5-Empty
V5-RPS10
V5-RPS10
UBE2D3 sh2:
HA-ubiquitin:
kDa
Ub-V5-RPS10
IP: V5
WB: HA
Ub-V5-RPS10
V5-RPS10
(high exp.)
IP: V5
WB: V5
Ub-V5-RPS10
V5-RPS10
(low exp.)
C
Input
V5-Empty
V5-RPS10
V5-RPS10
UBE2D3 sh2:
HA-ubiquitin:
kDa
Ub-V5-RPS10
V5-RPS10
V5 (high exp.)
Ub-V5-RPS10
V5-RPS10
V5 (low exp.)
HSP90
UBE2D3
HSP90
D
Elution
IgG
RPS20-IP
UBE2D3 sh2:
HA-ubiquitin:
kDa
Ub-RPS20
(high exp.)
IP: RPS20
WB: HA
Ub-RPS20
(low exp.)
E
Input
UBE2D3 sh2:
HA-ubiquitin:
kDa
RPS20
RPS20
(low exp.)
Ub-RPS20
RPS20
RPS20
(high exp.)
UBE2D3
HSP90

Figure S6. RPS10 and RPS20 diGly modified peptides in LFQ sh2, V5-RPS10 expression and IPs for RPS10 and RPS20. **A**, Immunoblotting to verify expression of V5-RPS10 in HEK 293T cells. **B-C**, IP assay in 293T cells +/- UBE2D3 sh2 transfected with V5-tagged RPS10 and HA-tagged ubiquitin. Immunoblots representing eluates of the V5-IP, showing decreased ubiquitination of RPS10 in UBE2D3-depleted cells (B) and immunoblots of input samples (C). Representative blots of n = 2 are shown. **D-E**, Endogenous RPS20 IP assay in 293T cells +/- UBE2D3 sh2, transfected with HA-tagged ubiquitin. Immunoblots representing eluates of the IgG negative control and RPS20 IPs, showing decreased ubiquitination of RPS20 in UBE2D3-depleted cells (D) and immunoblots of input samples (E). Representative blots of n = 2 are shown.

REFERENCES

1 Kwon, Y. T. & Ciechanover, A. The Ubiquitin Code in the Ubiquitin-Proteasome System and Autophagy. *Trends Biochem Sci* **42**, 873-886, doi:10.1016/j.tibs.2017.09.002 (2017).

2 Oh, E., Akopian, D. & Rape, M. Principles of Ubiquitin-Dependent Signaling. *Annu Rev Cell Dev Biol* **34**, 137-162, doi:10.1146/annurev-cellbio-100617-062802 (2018).

3 Song, L. & Luo, Z. Q. Post-translational regulation of ubiquitin signaling. *J Cell Biol* **218**, 1776-1786, doi:10.1083/jcb.201902074 (2019).

4 Jensen, J. P., Bates, P. W., Yang, M., Vierstra, R. D. & Weissman, A. M. Identification of a family of closely related human ubiquitin conjugating enzymes. *J Biol Chem* **270**, 30408-30414, doi:10.1074/jbc.270.51.30408 (1995).

5 Buchwald, G. *et al.* Structure and E3-ligase activity of the Ring-Ring complex of polycomb proteins Bmi1 and Ring1b. *EMBO J* **25**, 2465-2474, doi:10.1038/sj.emboj.7601144 (2006).

6 Hashizume, R. *et al.* The RING heterodimer BRCA1-BARD1 is a ubiquitin ligase inactivated by a breast cancer-derived mutation. *J Biol Chem* **276**, 14537-14540, doi:10.1074/jbc.C000881200 (2001).

7 Mattiroli, F. *et al.* RNF168 ubiquitinates K13-15 on H2A/H2AX to drive DNA damage signaling. *Cell* **150**, 1182-1195, doi:10.1016/j.cell.2012.08.005 (2012).

8 Polanowska, J., Martin, J. S., Garcia-Muse, T., Petalcorin, M. I. & Boulton, S. J. A conserved pathway to activate BRCA1-dependent ubiquitylation at DNA damage sites. *EMBO J* **25**, 2178-2188, doi:10.1038/sj.emboj.7601102 (2006).

9 Saville, M. K. *et al.* Regulation of p53 by the ubiquitin-conjugating enzymes UbcH5B/C in vivo. *J Biol Chem* **279**, 42169-42181, doi:10.1074/jbc.M403362200 (2004).

10 Schmidt, C. K. *et al.* Systematic E2 screening reveals a UBE2D-RNF138-CtIP axis promoting DNA repair. *Nat Cell Biol* **17**, 1458-1470, doi:10.1038/ncb3260 (2015).

11 Hattori, H. *et al.* RNAi screen identifies UBE2D3 as a mediator of all-trans retinoic acid-induced cell growth arrest in human acute promyelocytic NB4 cells. *Blood* **110**, 640-650, doi:10.1182/blood-2006-11-059048 (2007).

12 Gao, X. *et al.* UBE2D3 gene overexpression increases radiosensitivity of EC109 esophageal cancer cells in vitro and in vivo. *Oncotarget* **7**, 32543-32553, doi:10.18632/oncotarget.8869 (2016).

13 Lee, K. A. *et al.* Ubiquitin ligase substrate identification through quantitative proteomics at both the protein and peptide levels. *J Biol Chem* **286**, 41530-41538, doi:10.1074/jbc.M111.248856 (2011).

14 Sarraf, S. A. *et al.* Landscape of the PARKIN-dependent ubiquitylome in response to mitochondrial depolarization. *Nature* **496**, 372-376, doi:10.1038/nature12043 (2013).

15 Thompson, J. W. *et al.* Quantitative Lys--Gly-Gly (diGly) proteomics coupled with inducible RNAi reveals ubiquitin-mediated proteolysis of DNA damage-inducible transcript 4 (DDIT4) by the E3 ligase HUWE1. *J Biol Chem* **289**, 28942-28955, doi:10.1074/jbc.M114.573352 (2014).

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

17 Xu, G., Paige, J. S. & Jaffrey, S. R. Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nat Biotechnol* **28**, 868-873, doi:10.1038/nbt.1654 (2010).

18 Sap, K. A., Bezstarosti, K., Dekkers, D. H. W., Voets, O. & Demmers, J. A. A. Quantitative Proteomics Reveals Extensive Changes in the Ubiquitinome after Perturbation of the Proteasome by Targeted dsRNA-Mediated Subunit Knockdown in Drosophila. *J Proteome Res* **16**, 2848-2862, doi:10.1021/acs.jproteome.7b00156 (2017).

19 Wagner, S. A. *et al.* A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. *Mol Cell Proteomics* **10**, M111 013284, doi:10.1074/mcp.M111.013284 (2011).

20 Ong, S. E. *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **1**, 376-386, doi:10.1074/mcp.m200025-mcp200 (2002).

21 Salas-Lloret, D., Agabitini, G. & Gonzalez-Prieto, R. TULIP2: An Improved Method for the Identification of Ubiquitin E3-Specific Targets. *Front Chem* **7**, 802, doi:10.3389/fchem.2019.00802 (2019).

22 Garzia, A. *et al.* The E3 ubiquitin ligase and RNA-binding protein ZNF598 orchestrates ribosome quality control of premature polyadenylated mRNAs. *Nat Commun* **8**, 16056, doi:10.1038/ncomms16056 (2017).

23 Juszkiewicz, S. & Hegde, R. S. Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. *Mol Cell* **65**, 743-750 e744, doi:10.1016/j.molcel.2016.11.039 (2017).

24 Sundaramoorthy, E. *et al.* ZNF598 and RACK1 Regulate Mammalian Ribosome-Associated Quality Control Function by Mediating Regulatory 40S Ribosomal Ubiquitylation. *Mol Cell* **65**, 751-760 e754, doi:10.1016/j.molcel.2016.12.026 (2017).

25 Peuscher, M. H. & Jacobs, J. J. DNA-damage response and repair activities at uncapped telomeres depend on RNF8. *Nat Cell Biol* **13**, 1139-1145, doi:10.1038/ncb2326 (2011).

26 Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **11**, 2301-2319, doi:10.1038/nprot.2016.136 (2016).

27 Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731-740, doi:10.1038/nmeth.3901 (2016).

28 Kramer, A., Green, J., Pollard, J., Jr. & Tugendreich, S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* **30**, 523-530, doi:10.1093/bioinformatics/btt703 (2014).

29 Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* **47**, D607-D613, doi:10.1093/nar/gky1131 (2019).

30 Sap, K. A. *et al.* Global quantitative proteomics reveals novel factors in the ecdysone signaling pathway in Drosophila melanogaster. *Proteomics* **15**, 725-738, doi:10.1002/pmic.201400308 (2015).

31 Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **13**, 2513-2526, doi:10.1074/mcp.M113.031591 (2014).

32 Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* **2**, 1896-1906, doi:10.1038/nprot.2007.261 (2007).

33 Goedhart, J. & Luijsterburg, M. S. VolcaNoseR is a web app for creating, exploring, labeling and sharing volcano plots. *Sci Rep* **10**, 20560, doi:10.1038/s41598-020-76603-3 (2020).

34 Berlin, I., Schwartz, H. & Nash, P. D. Regulation of epidermal growth factor receptor ubiquitination and trafficking by the USP8.STAM complex. *J Biol Chem* **285**, 34909-34921, doi:10.1074/jbc.M109.016287 (2010).

35 Yang, X. *et al.* A public genome-scale lentiviral expression library of human ORFs. *Nat Methods* **8**, 659-661, doi:10.1038/nmeth.1638 (2011).

36 Liu, R. Z. *et al.* CRABP1 is associated with a poor prognosis in breast cancer: adding to the complexity of breast cancer cell response to retinoic acid. *Mol Cancer* **14**, 129, doi:10.1186/s12943-015-0380-7 (2015).

37 Nagpal, I. & Wei, L. N. All-trans Retinoic Acid as a Versatile Cytosolic Signal Modulator Mediated by CRABP1. *Int J Mol Sci* **20**, doi:10.3390/ijms20153610 (2019).

38 Napoli, J. L. Functions of Intracellular Retinoid Binding-Proteins. *Subcell Biochem* **81**, 21-76, doi:10.1007/978-94-024-0945-1_2 (2016).

39 Deng, Y., Cai, S., Shen, J. & Peng, H. Tetraspanins: Novel Molecular Regulators of Gastric Cancer. *Front Oncol* **11**, 702510, doi:10.3389/fonc.2021.702510 (2021).

40 Heo, K. & Lee, S. TSPAN8 as a Novel Emerging Therapeutic Target in Cancer for Monoclonal Antibody Therapy. *Biomolecules* **10**, doi:10.3390/biom10030388 (2020).

41 Maisonial-Besset, A. *et al.* Tetraspanin 8 (TSPAN 8) as a potential target for radio-immunotherapy of colorectal cancer. *Oncotarget* **8**, 22034-22047, doi:10.18632/oncotarget.15787 (2017).

42 Zhu, R. *et al.* TSPAN8 promotes cancer cell stemness via activation of sonic Hedgehog signaling. *Nat Commun* **10**, 2863, doi:10.1038/s41467-019-10739-3 (2019).

43 Gullberg, D. Shift happens--a paradigm shift for the role of integrins in fibrosis. *Matrix Biol* **28**, 383, doi:10.1016/j.matbio.2009.09.008 (2009).

44 Garnett, M. J. *et al.* UBE2S elongates ubiquitin chains on APC/C substrates to promote mitotic exit. *Nat Cell Biol* **11**, 1363-1369, doi:10.1038/ncb1983 (2009).

45 Baek, K. *et al.* NEDD8 nucleates a multivalent cullin-RING-UBE2D ubiquitin ligation assembly. *Nature* **578**, 461-466, doi:10.1038/s41586-020-2000-y (2020).

46 Hill, S. *et al.* Robust cullin-RING ligase function is established by a multiplicity of poly-ubiquitylation pathways. *Elife* **8**, doi:10.7554/eLife.51163 (2019).

47 Kelsall, I. R. *et al.* Coupled monoubiquitylation of the co-E3 ligase DCNL1 by Ariadne-RBR E3 ubiquitin ligases promotes cullin-RING ligase complex remodeling. *J Biol Chem* **294**, 2651-2664, doi:10.1074/jbc.RA118.005861 (2019).

48 Popov, N., Schulein, C., Jaenicke, L. A. & Eilers, M. Ubiquitylation of the amino terminus of Myc by SCF(beta-TrCP) antagonizes SCF(Fbw7)-mediated turnover. *Nat Cell Biol* **12**, 973-981, doi:10.1038/ncb2104 (2010).

49 Wu, K., Kovacev, J. & Pan, Z. Q. Priming and extending: a UbcH5/Cdc34 E2 handoff mechanism for polyubiquitination on a SCF substrate. *Mol Cell* **37**, 784-796, doi:10.1016/j.molcel.2010.02.025 (2010).

50 Wu, W. *et al.* HERC2 is an E3 ligase that targets BRCA1 for degradation. *Cancer Res* **70**, 6384-6392, doi:10.1158/0008-5472.CAN-10-1304 (2010).

51 Mohanty, S. *et al.* The E3/E4 ubiquitin conjugation factor UBE4B interacts with and ubiquitinates the HTLV-1 Tax oncoprotein to promote NF-kappaB activation. *PLoS Pathog* **16**, e1008504, doi:10.1371/journal.ppat.1008504 (2020).

52 Okumura, F., Hatakeyama, S., Matsumoto, M., Kamura, T. & Nakayama, K. I. Functional regulation of FEZ1 by the U-box-type ubiquitin ligase E4B contributes to neuritogenesis. *J Biol Chem* **279**, 53533-53543, doi:10.1074/jbc.M402916200 (2004).

53 Torrino, S. *et al.* UBTD1 is a mechano-regulator controlling cancer aggressiveness. *EMBO Rep* **20**, doi:10.15252/embr.201846570 (2019).

54 Friedl, P. & Mayor, R. Tuning Collective Cell Migration by Cell-Cell Junction Regulation. *Cold Spring Harb Perspect Biol* **9**, doi:10.1101/cshperspect.a029199 (2017).

55 Sumigray, K. D. & Lechler, T. Cell adhesion in epidermal development and barrier formation. *Curr Top Dev Biol* **112**, 383-414, doi:10.1016/bs.ctdb.2014.11.027 (2015).

56 Janiszewska, M., Primi, M. C. & Izard, T. Cell adhesion in cancer: Beyond the migration of single cells. *J Biol Chem* **295**, 2495-2505, doi:10.1074/jbc.REV119.007759 (2020).

57 Mulder, J. *et al.* p116Rip is a novel filamentous actin-binding protein. *J Biol Chem* **278**, 27216-27223, doi:10.1074/jbc.M302399200 (2003).

58 Fan, X. *et al.* 14-3-3 Proteins Are on the Crossroads of Cancer, Aging, and Age-Related Neurodegenerative Disease. *Int J Mol Sci* **20**, doi:10.3390/ijms20143518 (2019).

59 Obsilova, V. & Obsil, T. Structural insights into the functional roles of 14-3-3 proteins. *Front Mol Biosci* **9**, 1016071, doi:10.3389/fmolb.2022.1016071 (2022).

60 Liyasova, M. S. *et al.* Cbl interacts with multiple E2s in vitro and in cells. *PLoS One* **14**, e0216967, doi:10.1371/journal.pone.0216967 (2019).

61 Levkowitz, G. *et al.* Ubiquitin ligase activity and tyrosine phosphorylation underlie suppression of growth factor signaling by c-Cbl/Sli-1. *Mol Cell* **4**, 1029-1040, doi:10.1016/s1097-2765(00)80231-2 (1999).

62 Umebayashi, K., Stenmark, H. & Yoshimori, T. Ubc4/5 and c-Cbl continue to ubiquitinate EGF receptor after internalization to facilitate polyubiquitination and degradation. *Mol Biol Cell* **19**, 3454-3462, doi:10.1091/mbc.E07-10-0988 (2008).

63 Peng, H. *et al.* Ubiquitylation of p62/sequestosome1 activates its autophagy receptor function and controls selective autophagy upon ubiquitin stress. *Cell Res* **27**, 657-674, doi:10.1038/cr.2017.40 (2017).

64 Cha-Molstad, H. *et al.* Amino-terminal arginylation targets endoplasmic reticulum chaperone BiP for autophagy through p62 binding. *Nat Cell Biol* **17**, 917-929, doi:10.1038/ncb3177 (2015).

65 Bertrand, M. J. *et al.* cIAP1/2 are direct E3 ligases conjugating diverse types of ubiquitin chains to receptor interacting proteins kinases 1 to 4 (RIP1-4). *PLoS One* **6**, e22356, doi:10.1371/journal.pone.0022356 (2011).

66 Journo, C. *et al.* NRP/Optineurin Cooperates with TAX1BP1 to potentiate the activation of NF-kappaB by human T-lymphotropic virus type 1 tax protein. *PLoS Pathog* **5**, e1000521, doi:10.1371/journal.ppat.1000521 (2009).

67 Shembade, N., Ma, A. & Harhaj, E. W. Inhibition of NF-kappaB signaling by A20 through disruption of ubiquitin enzyme complexes. *Science* **327**, 1135-1139, doi:10.1126/science.1182364 (2010).

68 Zhou, A. Y. *et al.* IKKepsilon-mediated tumorigenesis requires K63-linked polyubiquitination by a cIAP1/cIAP2/TRAF2 E3 ubiquitin ligase complex. *Cell Rep* **3**, 724-733, doi:10.1016/j.celrep.2013.01.031 (2013).

69 Sarraf, S. A. *et al.* Loss of TAX1BP1-Directed Autophagy Results in Protein Aggregate Accumulation in the Brain. *Mol Cell* **80**, 779-795 e710, doi:10.1016/j.molcel.2020.10.041 (2020).

70 Turco, E. *et al.* Reconstitution defines the roles of p62, NBR1 and TAX1BP1 in ubiquitin condensate formation and autophagy initiation. *Nat Commun* **12**, 5212, doi:10.1038/s41467-021-25572-w (2021).

71 Zhang, S. *et al.* PCNA is ubiquitinated by RNF8. *Cell Cycle* **7**, 3399-3404, doi:10.4161/cc.7.21.6949 (2008).

72 Joazeiro, C. A. P. Mechanisms and functions of ribosome-associated protein quality control. *Nat Rev Mol Cell Biol* **20**, 368-383, doi:10.1038/s41580-019-0118-2 (2019).

73 Filbeck, S., Cerullo, F., Pfeffer, S. & Joazeiro, C. A. P. Ribosome-associated quality-control mechanisms from bacteria to humans. *Mol Cell* **82**, 1451-1466, doi:10.1016/j.molcel.2022.03.038 (2022).

74 Cano, F. *et al.* The RNA-binding E3 ubiquitin ligase MEX-3C links ubiquitination with MHC-I mRNA degradation. *EMBO J* **31**, 3596-3606, doi:10.1038/emboj.2012.218 (2012).

75 Mark, K. G., Simonetta, M., Maiolica, A., Seller, C. A. & Toczyski, D. P. Ubiquitin ligase trapping identifies an SCF(Saf1) pathway targeting unprocessed vacuolar/lysosomal proteins. *Mol Cell* **53**, 148-161, doi:10.1016/j.molcel.2013.12.003 (2014).

76 Yen, H. C. & Elledge, S. J. Identification of SCF ubiquitin ligase substrates by global protein stability profiling. *Science* **322**, 923-929, doi:10.1126/science.1160462 (2008).

77 Sanchez-Quiles, V. *et al.* Cylindromatosis Tumor Suppressor Protein (CYLD) Deubiquitinase is Necessary for Proper Ubiquitination and Degradation of the Epidermal Growth Factor Receptor. *Mol Cell Proteomics* **16**, 1433-1446, doi:10.1074/mcp.M116.066423 (2017).

78 Sapmaz, A. *et al.* USP32 regulates late endosomal transport and recycling through deubiquitylation of Rab7. *Nat Commun* **10**, 1454, doi:10.1038/s41467-019-09437-x (2019).

79 Labrecque, J., Dumas, F., Lacroix, A. & Bhat, P. V. A novel isoenzyme of aldehyde dehydrogenase specifically involved in the biosynthesis of 9-cis and all-trans retinoic acid. *Biochem J* **305 (Pt 2)**, 681-684, doi:10.1042/bj3050681 (1995).

80 Molotkov, A. & Duester, G. Genetic evidence that retinaldehyde dehydrogenase Raldh1 (Aldh1a1) functions downstream of alcohol dehydrogenase Adh1 in metabolism of retinol to retinoic acid. *J Biol Chem* **278**, 36085-36090, doi:10.1074/jbc.M303709200 (2003).

81 Stevers, L. M. *et al.* Modulators of 14-3-3 Protein-Protein Interactions. *J Med Chem* **61**, 3755-3778, doi:10.1021/acs.jmedchem.7b00574 (2018).

82 Huang, Y. *et al.* Nuclear translocation of the 4-pass transmembrane protein Tspan8. *Cell Res*, doi:10.1038/s41422-021-00522-9 (2021).

83 Hildebrandt, A. *et al.* The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A) translation. *Genome Biol* **20**, 216, doi:10.1186/s13059-019-1814-0 (2019).

84 Juszkiewicz, S. *et al.* ZNF598 Is a Quality Control Sensor of Collided Ribosomes. *Mol Cell* **72**, 469-481 e467, doi:10.1016/j.molcel.2018.08.037 (2018).

85 Garshott, D. M., Sundaramoorthy, E., Leonard, M. & Bennett, E. J. Distinct regulatory ribosomal ubiquitylation events are reversible and hierarchically organized. *Elife* **9**, doi:10.7554/eLife.54023 (2020).

86 Choe, Y. J. *et al.* Failure of RQC machinery causes protein aggregation and proteotoxic stress. *Nature* **531**, 191-195, doi:10.1038/nature16973 (2016).

87 Klaips, C. L., Jayaraj, G. G. & Hartl, F. U. Pathways of cellular proteostasis in aging and disease. *J Cell Biol* **217**, 51-63, doi:10.1083/jcb.201709072 (2018).

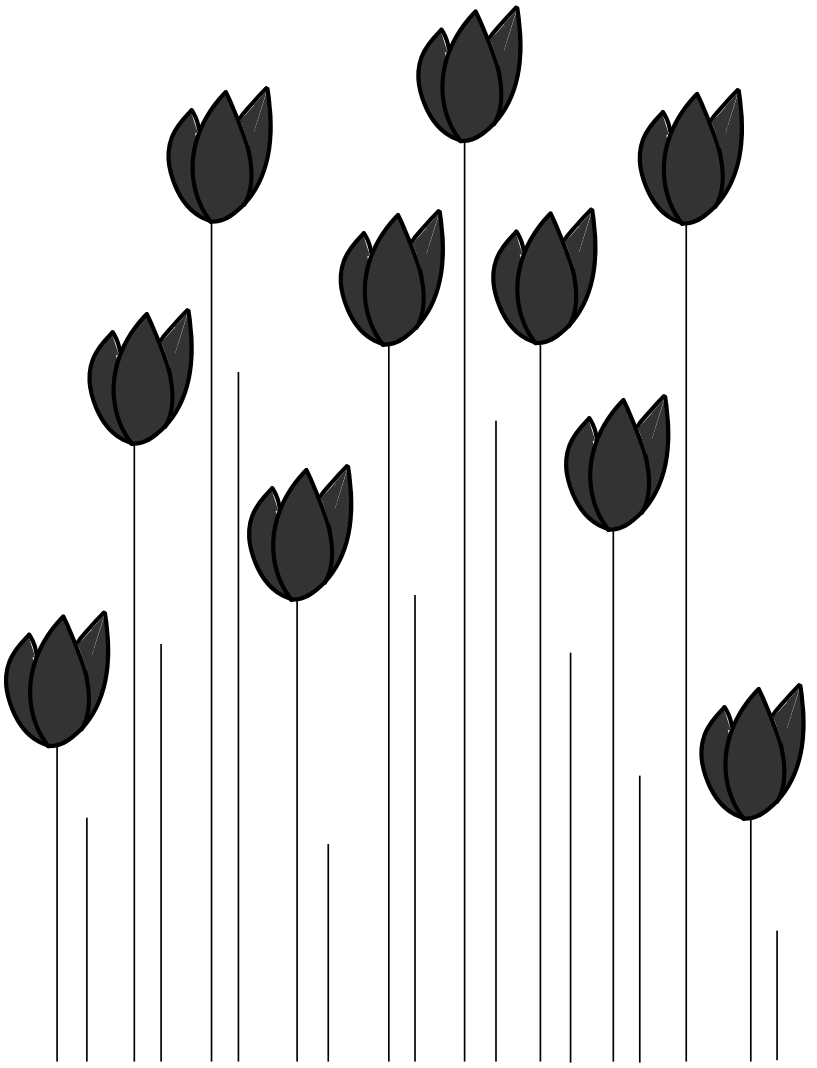
88 Pegoraro, G. *et al.* Identification of mammalian protein quality control factors by high-throughput cellular imaging. *PLoS One* **7**, e31684, doi:10.1371/journal.pone.0031684 (2012).

89 Ma, S., Attarwala, I. Y. & Xie, X. Q. SQSTM1/p62: A Potential Target for Neurodegenerative Disease. *ACS Chem Neurosci* **10**, 2094-2114, doi:10.1021/acschemneuro.8b00516 (2019).

90 Moreno, J. A. & Tiffany-Castiglioni, E. The chaperone Grp78 in protein folding disorders of the nervous system. *Neurochem Res* **40**, 329-335, doi:10.1007/s11064-014-1405-0 (2015).

91 Filatova, E. V. *et al.* Expression analysis of genes of ubiquitin-proteasome protein degradation system in MPTP-induced mice models of early stages of Parkinson's disease. *Dokl Biochem Biophys* **456**, 116-118, doi:10.1134/S1607672914030107 (2014).

92 Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* **47**, D442-D450, doi:10.1093/nar/gky1106 (2019).



7

GENERAL DISCUSSION AND FUTURE PERSPECTIVES

Daniel Salas-Lloret[1]

[1] Cell and Chemical Biology, Leiden University Medical Centre, Leiden, The Netherlands

Protein dynamics are able to change cell's fate in a very fast and efficient way. When stressed, e.g. via DNA damage or cell division, instead of producing newly synthesized proteins, the cell relies on the modification of already existing proteins to execute vital functions. These modifications can be the conjugation of small molecules such as ubiquitin (Ub) or Small Ubiquitin-like Modifiers (SUMOs). This can lead to protein degradation, conformational changes or relocation within the cell of critical proteins (1-3).

One way to detect these proteins modifications, or the fluctuation of protein levels during particular cellular conditions, is the use of Mass-Spectrometry (MS)-based approaches. As it was described in Chapter 1, in the last decade there has been a remarkable advance in the MS-based proteomics field. Nowadays, what you can achieve with the use of MS approaches depends on the amount of samples you can prepare and your imagination for developing new sample-preparation strategies.

From the total proteome to specific proteins

There are several MS approaches for studying cell proteomes. Cell lysis and subsequent protein digestion enable the study of the cell proteome in different conditions. However, one of the main drawbacks of bottom-up or shotgun proteomics of the whole lysate is the enormous complexity of the sample (**Figure 1A**). Although the development of more sophisticated equipment and fractionation techniques can help to reduce sample complexity, the variability between samples can still be very high, affecting the identification of critical proteins. Nowadays, notable research on the MS-based proteomics field comprises two main approaches to generate MS data: Data-Dependent Acquisition (DDA) and Data-Independent Acquisition (DIA), which are going to be critical in global proteomics. Tandem MS (MS/MS or MS^2) experiments are broadly used for amino acid sequencing of peptides and consist of several sequential stages. Firstly, the masses of the sample peptide ions are determined in a first MS round (MS1). Secondly, the selected peptide ion precursors from MS1 are isolated via their mass-to-charge ratio (*m/z*) values, whilst other peptide ion precursors with different *m/z* values are just filtered away. Thirdly, the resulting peptide ion precursors are activated with an inert gas such as Argon that induces their fragmentation. Finally, the *m/z* values of the fragmented peptide ion precursors are determined by a second MS round (MS2). The DDA only puts forward the most intense peptide ion precursors generated during MS1, while the DIA sends all generated peptide ions. However, DIA requires the pre-definition of isolation windows to cover the whole MS1 *m/z* range and requires study-specific spectral libraries, commonly designed by DDA (4). Generally, whole lysate shotgun proteomics can benefit from DIA analysis to gain a better coverage of the whole proteome. Nevertheless, resources and time required to create a library for a particular study might be a limitation to consider. Therefore, new advances are emerging for DIA optimization such as the use of artificial intelligence, single-cell proteomics (4, 5) or the combination of these two analysis into Data Dependent-Independent Acquisition (DDIA) (6, 7).

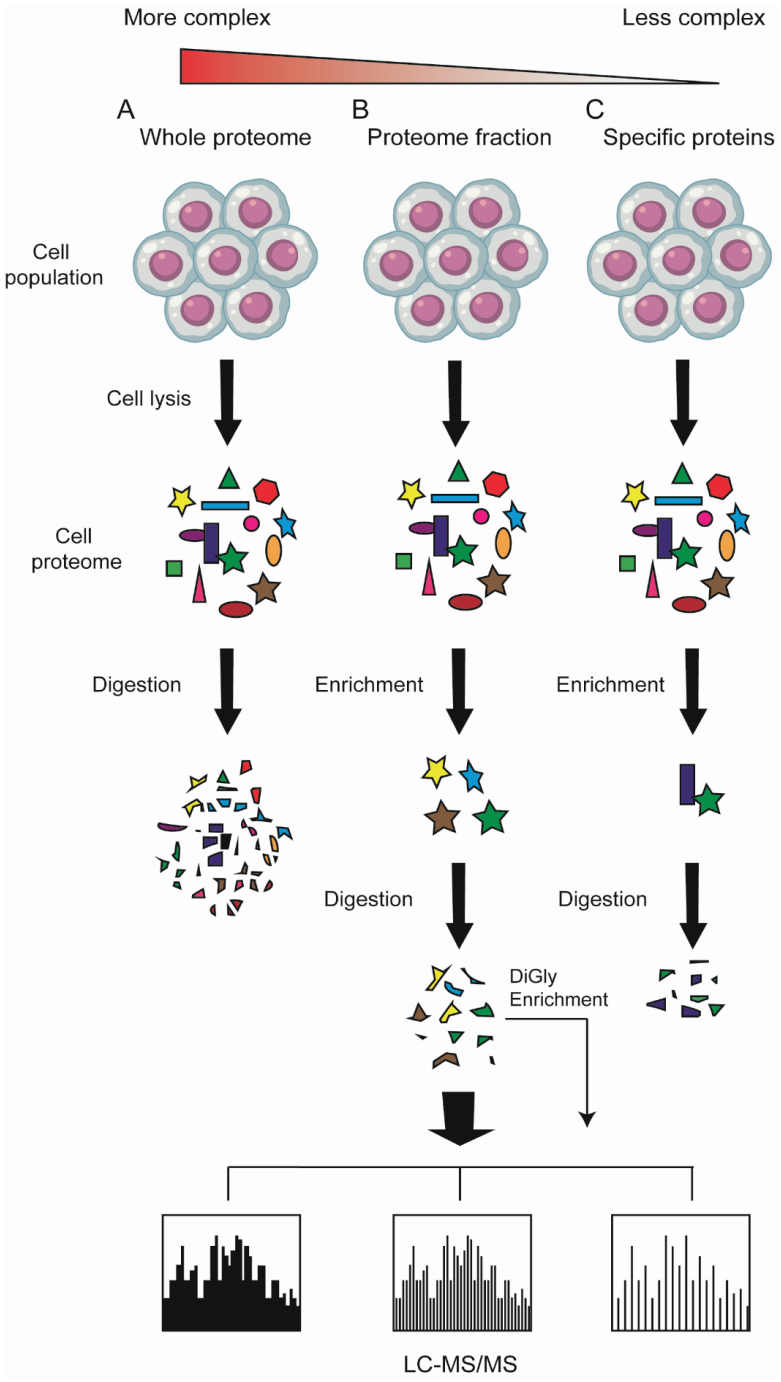


Figure 1. Sample complexity and different approaches depending on the research goal. **A**. Whole proteome of cell lysate followed by tryptic digestion and subsequent LC-MS/MS. **B**. Proteome fraction enrichment equivalent to His-Ub or His-SUMO pull downs followed by tryptic digestion and subsequent LC-MS/MS. **C**.

Specific proteins enrichment equivalent to TULIP2, SATTs or GFP trap approaches followed by tryptic digestion and subsequent LC-MS/MS.

Instead of dealing with all proteins within the cell, in chapter 4 we used His-SUMO1 and His-SUMO2/3 to enrich the SUMO proteome and link SUMO substrates to particular SUMO E3s, employing SUMO Activated Target Traps (SATTs). This enrichment reduces the sample complexity and facilitates the identification of specific proteins (**Figure 1B**). For the quantification of these proteins, we relied on Label-Free Quantification (LFQ) techniques, which require several biological repeats for an accurate quantification. Other laboratories have used different MS strategies to link SUMO substrates to E3 ligases. For example, instead of LFQ, it is possible to combine SUMO immunoaffinity with Stable Isotope Labeling of Amino acid in Cell culture (SILAC) to study the global changes in SUMOylated proteins upon E3 ligase overexpression (8). The SUMO proteome is already a complex sample and the use of SILAC may introduce even more complexity due to isotope labeling. Nowadays, there are other labeling approaches meant to reduce sample complexity and quantify the amount of proteins in a very reliable way (9). Another approach that could be employed to find the SUMO E3 ligase for a substrate of interest, with reduce sample complexity, is the employment of SUMO-ID (10). One fragment of the split Turbo-ID is fused to SUMO while the complementary fragment is fused to a SUMO substrate. The rationale is that upon SUMO-SIM interaction or SUMOylation of the target, both fragments of the Turbo-ID enzyme are close enough to allow refolding of the enzyme and labelling of proximal complexes, which can be purified and identified by MS. Although this approach can further reduce the sample complexity compared to global SUMO proteome screenings, however the remaining fraction of biotinylated proteins is still very large and complex. In all these studies DDA analysis was performed, although the implementation of DIA could still improve the data completeness across samples (11).

Among the global SUMO or Ubiquitin proteomes, we can reduce sample complexity even more by enriching only a small fraction of the modified proteins (**Figure 1C**). In case of the SUMO proteome, we employed SATTs to enrich SUMOylated targets of particular SUMO E3 ligases. In chapter 5 and 6 we used TULIP2 methodology to enrich BRCA1-BARD1 and UBE2D3 specific ubiquitination substrates. Using these SUMO or Ubiquitin traps we can specifically enrich proteins of interest within proteomes. Although the sample complexity is reduced, there are other drawbacks to take into account. For instance, differentiation between background binders and truly specific targets can be challenging. Therefore, for the identification of specific targets, it is necessary to employ accurate controls and to stabilize optimal conditions. Additionally, there are other ways to enrich a very small fraction of the proteome and reduce sample complexity. One broadly used method is the co-immunoprecipitation (co-IP) of a protein of interest and subsequent MS analysis. This co-IP can be performed with antibodies that recognize the protein of interest, or by employing different tags such as GFP (12). This approach can work well for identifying stable complexes, but it has drawbacks in capturing transient states like modification of proteins. If the identification of modified proteins is the main goal, there are very efficient ways to reduce complexity of both SUMO and Ub proteomes and identify the modification sites. After tryptic digestion, peptides can also be enriched by antibodies that recognize the desired modification such as the UbiSite antibody for Ub and the 8A2 antibody for SUMO2/3 (13, 14). Enrichment of the modified peptides not only increases the fidelity of the modified proteins identified, but also allows the identification of the modification site. Other substrate-trapping strategies have already been combined with these antibodies to identify modification sites (15). Overall, there is not a single technique for identifying your proteins of interest. However, the MS

community has put a lot of effort to offer a broad set of methodologies, from which you can choose the one that allows you to reach your goals.

Challenges in identifying E3 substrates

Due to the hierarchical organization of the ubiquitination cascade, there are more than 600 E3s and more than 10.000 ubiquitination substrates. Some of these substrates are redundant and can be ubiquitinated by several E3s via different molecular mechanisms that we described in chapter 1. Therefore, linking a particular ubiquitination target to an E3 ligase can be very challenging. Historically, early approaches to this relied on protein-protein interactions such as yeast two-hybrid assays, co-immunoprecipitation and protein-protein arrays (16, 17). However, another challenge to identify E3 substrates is the transient nature of the interaction between the E3 and the target. The development of MS approaches greatly helped tackling this challenge together with the design of MS tools such as the NEDDylator approach and the Ubait Traps previously reviewed in chapter 1. Unfortunately, with the use of MS new challenges arose, for example some ubiquitination targets are highly ubiquitinated while others are ubiquitinated at very low levels within the cell making the identification of the last ones challenging too. MS approaches using DDA can be very powerful identifying a mayor ubiquitination event, but low abundant peptides can be masked among the background and most abundant targets. Trying to overcome this challenge, it is possible to enhance the interaction of low affinity targets using Tandem Ubiquitin-Binding Entities (TUBEs) (15) and the use of DIA for broader identification of substrate proteins (18). Nevertheless, the E3 specificity problem would still be difficult to overcome. Ideally, when using trapping MS-based approaches, the E3 ligase would be purified together with its targets, employing negative controls helping to identify specific interactions. However, this could be occluded by a large number of background binder, which would make it hard to distinguish between background binders and specific proteins. Additionally, if the use of negative controls includes cell lines deficient for the E3 ligase of interest, there might be compensation effects by other E3 ligases complicating the identification of the target proteins for your E3.

TULIP2 as a versatile tool

Ubiquitin proteome screenings often requires fractionation assays due to the complexity of a sample. However, it is possible to make use of the ubiquitination cascade and the denaturing conditions to reduce sample complexity and trap ubiquitination targets. In 2015, O'Connor et al, published the first ubiquitin-activated trap, where it was possible to find ubiquitination targets for an E3 ligase of interest (19). In chapter 3, this methodology was improved for finding specific targets of a particular E3 employing high denaturing conditions and optimized MS analysis. We showed in chapter 4 that this methodology can be modified for SUMO E3 ligases allowing to get the first comprehensive SUMO proteome in an E3-specific manner.

In chapter 5, we employed TULIP2 methodology not only for identifying E3 specific targets but also to address functional analysis. We showed for the first time that TULIP2 with BRCA1 can be used to rescue phenotypes as it does not alter its physiological function. However, this can be different for other proteins, thus the functional analysis of TULIP2 constructs should be checked before MS experiments. In our study, we demonstrated the functionality of BRCA1-TULIP2 by analyzing RAD51 foci formation, a known marker for Homologous Recombination (HR), and by using a survival assay against olaparib treatment. BRCA1 deficient cells expressing BRCA1-TULIP2 were able to restore the formation of RAD51, where we observed the colocalization of BRCA1-TULIP2 with RAD51, and the resistance to olaparib treatments. The TULIP2 Gateway cloning

system makes cloning different mutants of the E3 ligase of interest straightforward. Additionally, the DOX inducible promotor allows a close to endogenous expression of the TULIP2 constructs. These features enable to produce different mutant versions of the E3 (i.e. catalytic dead mutant, nuclear localization mutant or interaction mutant), which would be crucial for identifying specific targets. To gain deeper knowledge of the E3 ligase of interest, combination of fluorescent microscopy and MS experiments could be performed. In addition, TULIP2 can also be used for pulling down interactors of the E3 under study without the generation of additional cell lines. Changing the denaturing buffer conditions to milder or native conditions would allow the identification of protein interactors. These interactor partners can be validated by western blotting or by LC-MS/MS after TULIP2 His-Pull down. Altogether, TULIP2 experiments including MS experiments with high denaturing conditions would allow the identification of ubiquitination targets, changing to non-denaturing conditions would identify interactor proteins, and fluorescent microscopy experiments would reveal the functionality of the TULIP2 constructs.

Furthermore, instead of an E3, it is possible to clone in an E2 conjugating enzyme and find its ubiquitination targets. In chapter 6, we found out the ubiquitination targets of the E2 UBE2D3 employing the TULIP2 methodology. In this chapter, we could observe that using different approaches completely changes the MS data of the screening. The SILAC approach for UBE2D3 and Label-Free Quantification (LFQ) gave rise to very different sets of ubiquitination substrates, where less than 2% were found in common. As mentioned before, variability between samples can already have a big impact in the generated data. Therefore, using different approaches as LFQ and SILAC in complex samples, can give rise to very different outcomes. In order to find consistency within the data, we performed TULIP2. Using UBE2D3-TULIP2, we identified proteins within the 2% of the common targets from the LFQ and SILAC screenings. Then, TULIP2 can also be employed to validate substrates from broader screenings and find out specific targets for E2 enzymes. Although it is possible to find E2 specific targets using TULIP2 methodology, E2s can be promiscuous and the specific targets might be overshadowed by the redundant ones, making its identification more challenging than for E3s.

When using TULIP2 methodology there are some drawbacks to take into account. The fused ubiquitin to the E3 of interest can be used by other endogenous E3s for the ubiquitination of their substrates, leading to the identification of false specific targets after MS/MS analysis. However, this downside can be counteracted by the employment of catalytic dead mutants as controls. By using the diGly mutant version of TULIP2, we can filter out background binders and identify both specific and non-specific ubiquitination substrates of the E3 ligase of interest. With the use of catalytic dead mutants of the E3 ligase of interest, we are able to filter out the non-specific substrates. Therefore, after filtering out the background binders and the non-specific targets, the resulting ubiquitination substrates are specific targets for the E3 of interest. Additionally, we can compare the ubiquitination targets identified by TULIP2 methodology of different E3s to check redundancy/specificity of substrates. In chapter 5, BARD1-TULIP2 performed in normal growth conditions allowed the identification of PCNA, which is a known redundant ubiquitination substrate. Differently, RAD18-TULIP2 only detected PCNA as ubiquitination substrate after UV damage. TULIP2 methodology performed at different cellular conditions and on different E3s deciphered how a redundant ubiquitination target was ubiquitinated by distinctive E3s depending on the signaling pathway.

There are other drawbacks that still should be improved. The percentage of background binders is still very high even whilst using controls such as diGly and catalytic dead mutants, though

this could be improved by obtaining cleaner samples. One way to achieve this is finding other pull downs technologies with higher specificity. Another pitfall is that protein concentration needs to be carefully equalized among samples. Although LFQ normalizes protein concentration, we can observe big differences in background proteins when samples are not properly equalized. Not surprisingly, even samples with similar starting protein concentrations can result into different final protein concentrations after completing TULIP2 experiments. Generally, this happens due to technical manipulation of the samples. In line, after statistical analysis of the MS data, some experiments can cluster together, not only because their biological significance, but because of experimental issues (i.e. day of the experiment or the use of new buffers). This downside can be addressed by normalization methods. For pull downs experiments, it would be interesting to normalize against the pulled protein LFQ values to avoid significant differences coming from alterations in total protein concentration or pull down efficiencies. Overall, TULIP2 is a powerful tool to identify E3-specific ubiquitination targets and it can be easily updated with the emerging of new technologies. For instance, to find ubiquitination sites, TULIP2 can be combined with UbiSite antibodies to identify the ubiquitination sites of the targets. Finally, labelling technologies such as Tandem Mass Tag (TMT) could be applied together with DIA to generate a deeper database of the ubiquitination targets (20). TMT could help in normalization methods and reduce sample variability, while DIA might allow the identification of ubiquitination targets that occur at a very low level in the cell.

Impact of finding E3 targets in the clinic

In chapter 3 we developed the TULIP2 methodology which allows the identification of ubiquitin E3 specific targets. There are numerous ubiquitin E3 ligases where their dysregulation is involved in neurological pathologies. For example FBXL7 in Alzheimer's disease, PARK2 for Parkinson's disease, BTBD9 for Tourette syndrome and HERC2 for Angelman-like syndrome (21). Many of these diseases can be characterized by a mutation in an E3 ligase or proteins that regulate their function, which normally leads to the aggregation of proteins (22). The substrate proteins for some E3s are known and it is therefore possible to intervene to ameliorate their aggregation. However for other disorders, only the mutation of the E3 ligase is known (23, 24). Here, the TULIP2 methodology could be implemented to identify the targets of these E3s in order to find the proteins that cause the disease, ultimately helping develop therapies to combat these diseases.

E3 ligase dysregulation has not only been found in neurological disorders but also in cancer (25). We showed in chapter 2 how the DNA Damage Response (DDR) is orchestrated by different E3s. Additionally, we discussed potent inhibitors that can be used in clinical trials such as the commercial UBA1 inhibitor, TAK-243. In most cases, dampening of an E3 ligase is preferred over blanket E3 ligase inhibition. Making use of the interaction between an E3 and a target protein, it is possible to develop Proteolytic Targeting Chimeras (PROTACs) to induce protein targeted degradation (26). PROTACs are small protein degrader molecules commonly consisting of two ligands fused by a linker. One ligand recruits and binds a protein of interest (POI) while the other recruits and binds an E3 ubiquitin ligase. Ideally, the binding between the POI and the E3 by the PROTAC results in the proteasome degradation of the POI with the recycling of the PROTAC to target new copies of the POI. There are some PROTACs already in clinical trials, for example the Estrogen Receptor (ER) degrader ARV-471, which is at the time of writing at phase II for patients with $ER^{+}/HERC2^{-}$ locally advance or metastatic breast cancer (27). ARV-471 binds both the ER and cereblon (CRBN) E3 ligase to promote proteasomal degradation of the ER. However, there are some limitations. For example, not every E3 ligase can target every POI. Different E3 ligases have

shown different activity and selectivity for protein degradation, mainly because of incompatibility with the active site or because some targets are tissue-specific (28). Additionally, the mechanism to target the POI differs from the E3 family and will affect the PROTAC design. Therefore, finding new E3 ligases and their targets using Mass-Spectrometry approaches such as the TULIP2, would be very beneficial for the development of new and more efficient PROTACs. All newly discovered E3s and their substrates can be included in databases such as the UbiHub (29), which makes it possible to explore E3 ligases and their targets in an user friendly way. Similarly to PROTACs, there are also other types of protein degrader such as the molecular glues (30). Although they are not heterobifunctional like PROTACs, molecular glues enhance protein-protein interactions between a POI and an E3 ligase to promote its proteasomal degradation. However, the degradation of the POI is not their only purpose, there are also intra- and intermolecular glues that act outside of the ubiquitin proteasome system. For example the allosteric molecular glue SHP099 stabilizes a close conformation of SHP2 (intramolecular) for its inactivation (31) and antigenic peptides act as a glue between MHC presenter proteins and T cell receptors enhancing the immune response (intermolecular)(32). Combination of PROTACs and molecular glues may be new trends in the near future.

In chapter 5, we showed how BARD1-TULIP2 methodology was used to find novel ubiquitination substrates and, subsequently, new functions for its E3 activity, which opened new therapeutic opportunities in breast and ovarian cancer. There are several types of breast cancer characterized by hormonal status (ER^+, PR^+ or $HERC^+$), malignant tissue type (luminal or basal) and the genetic background. Attending to the genetic context, families afflicted with BRCA1 mutations carry a life-time probability of 90% and 50% for developing breast cancer and ovarian cancer respectively. The BRCA1 gene is essential and cells undergo apoptosis when BRCA1 is not functional. Cell death introduces selective pressure for mutations that allow cells to survive (first bottle-neck; **Figure 2**). These mutations commonly occur on oncogenes such as P53, which allows cells to keep dividing, resulting in a tumor with high genetic instability and DNA repair deficiencies. Fortunately, we can take advantage of the DNA repair deficiencies by inflicting DNA damage to the tumor. However, cancer cells can adapt further and go through a new selective pressure where cells acquire mutations in genes, such as 53BP1, that resolve these DNA damage deficiencies and become resistant to some therapies (second bottle-neck; **Figure 2**). In addition, BRCA1 mutants commonly lead to basal-like/triple-negative (ER^-, PR^-, $HERC^-$) breast tumors, which are often aggressive and have poor prognosis (33). Consequently, new vulnerabilities in BRCA1 mutant cancers could help to avoid tumor development and progression.

In our research, we found that BRCA1/BARD1 E3 activity is not essential for HR, but has a key role in single-stranded DNA (ssDNA) gap prevention and replication fork protection. Accumulation of ssDNA gaps leads to genetic instability and DSBs (34). Therefore, BRCA1 deficient and mutant cells unable to repair deleterious DSBs undergo apoptosis (**Figure 2**). However, BRCA1 expression can be restored, in some cases, lacking the E3 activity (35) leading to the accumulation of genetic instability because of ssDNA gaps and replication fork de-protection. These tumors are proficient in HR and resistant to DSBs chemotherapies. Thus, we propose that a novel solution would be to invest in agents that perturb the replication fork and homeostasis to treat these tumors.

Finally, although the E3 activity of BRCA1/BARD1 does not seem to be essential to prevent tumor formation (36), it is important for keeping genetic maintenance and avoid genetic instability, which is a hallmark of cancer. Nowadays, considerable research in the field of breast and ovarian cancer is being done on finding synthetic lethality with BRCA1 deficient cells. Here,

we elucidated BRCA1 E3 activity-involving mechanisms that could be used for finding new synthetic lethality combinations. In chapter 5, we suggested how synthetic lethality can work with RAD18. While BRCA1/BARD1 ubiquitinates PCNA during unperturbed conditions to prevent ssDNA accumulation and contributes in replication fork protection, RAD18 ubiquitinates PCNA upon replication fork blockade to load trans-lesion synthesis (TLS) polymerases such as REV1. In BRCA1 deficient cells there is ssDNA accumulation and cells are unable to repair DSBs. Therefore, interfering with RAD18 activity by either RAD18 knock downs or REV1 inhibitors, is synthetic lethal in BRCA1 deficient cells, becoming a good therapeutic strategy for P53-BRCA1 deficient tumors (37). However, other combinations such as PolΘ have been also described and inhibitors against this protein have rapidly been developed (38, 39). Given that breast cancer is a diverse and complex disease, the molecular mechanisms underlaying each variant may be different, as such they need to be understood individually to create selective treatments.

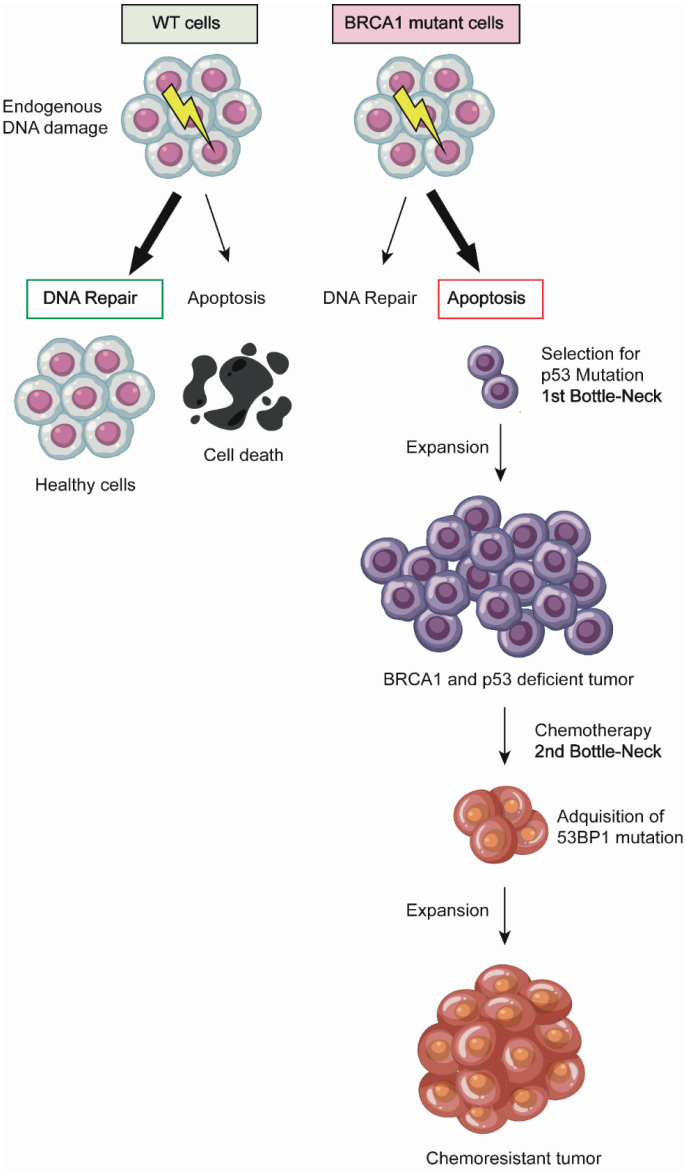


Figure 2. Depiction of breast cancer formation in patients with BRCA1 mutations. Cells undergo endogenous DNA damage. However, cells harboring BRCA1/BARD1 mutations are not able to fully repair some DNA lesions. Upon DNA damage, cells can either fix the lesion or go to apoptosis, this decision is regulated by proteins such as p53. WT cells can fix the damage, but as BRCA1 deficient cells have problems repairing DNA damage, go to apoptosis. After several apoptotic rounds, there is a bottle-neck where cells with p53 mutations can survive and keep dividing. After expansion, a BRCA1 and p53 mutant tumor is formed. In the clinic, DNA damaging agents such as olaparib can be used as chemotherapy for treating this malignancies, which yields a second bottle-neck for

cancer cells. Unfortunately, after chemotherapy, cells containing 53BP1 mutations become resistant to these agents and can survive and proliferate, giving rise to a chemo resistant tumor.

The complexity of SUMOylation

During the 90s, SUMOylation was not conceived as a crucial modification, as SUMO-deficient single substrates usually did not show a strong phenotype in cells. However, in the early 2000s it was shown that SUMOylation was essential for cell viability (40, 41), but the mechanism was still poorly understood. Given the low number of SUMO E3 ligases coupled with mice experiments where depletion of a SUMO E3 ligase was either not essential or compensated by other members of the PIAS family, it was believed that SUMO E3 ligases shared most of their substrates leading to a redundancy of SUMO E3 ligases for cell viability (42, 43). Additionally, although SUMOylation comprises different SUMO family members, namely SUMO1, SUMO2, SUMO3 and SUMO4, considering SUMO5 as a pseudogene (44), it seems that SUMO2 can compensate for the loss of other SUMOs. Mice studies knocking out either SUMO1 or SUMO3 showed mice viability, while knocking out SUMO2 was embryonic lethal (45, 46). Therefore, a post-translational modification with several E3s and five different SUMO family members did not result in a highly complex signaling network.

In chapter 4, we observed that each of 8 different SUMO E3 ligases have specific targets and PIAS1 showed preference for SUMO1 while PIAS4 had preference for SUMO2, indicating that the E3 ligase confers substrate specificity and that SUMO1 and SUMO2 have different targets. In chapter 5, by employing BARD1-TULIP2 we found SUMO1 as modified target during physiological conditions and SUMO2/3 upon stressful conditions using DNA damaging agents such UV. This indicates, not only a cross-talk between SUMOylation and Ubiquitination, but also different roles for SUMO1 and SUMO2/3. It seems that SUMO1 controls physiological conditions and keeps the homeostasis within cells, while SUMO2/3 is involved in pathological conditions such as DNA damage (47). Probably, if SUMO1 is not present, cells will be perturbated and SUMO2/3/4 could compensate to keep cell viability. However, if SUMO2 is not present, embryogenesis cannot continue, leading to embryonic lethality and cell death.

The complexity of SUMOylation increases when different SUMO polymers can be formed. As mentioned above, cellular stresses are characterized by an increase of SUMO chains (48, 49). SUMO polymers preferentially consist of SUMO2/3 chains on lysine K11 which is located in the SUMO consensus motif. However, SUMO1 chains and SUMO2/3 other than K11 chains have been identified by MS experiments (50, 51). Validation of MS approaches have been carried out by different labs. Employing a SUMO2 K11 mutant, RNF4 was still able to recognize and ubiquitinate PML for proteasomal degradation. The proposed RNF4-polySIM recognized polymers were the non-canonical K5 and K7 SUMO2 chains. Additionally, the SUMO2 K11Q mutant showed an increased K5 and K35 chain formation after heat shock stress for a possible compensation of K11 lost (52). Interestingly, although we found in chapter 4 that all SUMO E3s were able to form the canonical K11 SUMO2/3 chains and the non-canonical K5, only NSCME2 was able to form K32/33 SUMO2 non-canonical chains and both NSCME2 and PIAS4 were able to form K7 SUMO3 chains. Other than compensation roles, non-canonical chains might be involved in other pathways that require further investigation.

Nowadays, we consider the possibility of SUMOylation as a very complex signaling network where different E3s not only modify specific targets but also select the SUMO moiety for the modification. We even observed the formation of SUMO1-SUMO2/3 hybrid chains, which were

only detected in PIAS4 and NSCME2 SATTs. The role of hybrid chains is still uncertain, they seem not to be recognize by RNF4 for proteasomal degradation, but they could be potential targets for another STUbL such as Arkadia/RNF111 (53). Overall, the SUMO network might be much more complex that we can imagine as the possibilities increase as new features are discovered for these small-ubiquitin modifiers.

REFERENCES

1. E. R. Watson, N. G. Brown, J. M. Peters, H. Stark, B. A. Schulman, Posing the APC/C E3 Ubiquitin Ligase to Orchestrate Cell Division. *Trends Cell Biol* **29**, 117-134 (2019).
2. E. C. Osmundson *et al.*, The HECT E3 ligase Smurf2 is required for Mad2-dependent spindle assembly checkpoint. *J Cell Biol* **183**, 267-277 (2008).
3. S. Wegmann *et al.*, Linkage reprogramming by tailor-made E3s reveals polyubiquitin chain requirements in DNA-damage bypass. *Mol Cell* 10.1016/j.molcel.2022.02.016 (2022).
4. L. Krasny, P. H. Huang, Data-independent acquisition mass spectrometry (DIA-MS) for proteomic applications in oncology. *Mol Omics* **17**, 29-42 (2021).
5. A. Mund *et al.*, Deep Visual Proteomics defines single-cell identity and heterogeneity. *Nat Biotechnol* **40**, 1231-1240 (2022).
6. S. Guan, P. P. Taylor, Z. Han, M. F. Moran, B. Ma, Data Dependent-Independent Acquisition (DDIA) Proteomics. *J Proteome Res* **19**, 3230-3237 (2020).
7. J. Chen *et al.*, HRPIF data mining based on data-dependent/independent acquisition for Rhei Radix et Rhizoma metabolite screening in rats. *J Chromatogr B Analyt Technol Biomed Life Sci* **1190**, 123095 (2022).
8. C. Li *et al.*, Quantitative SUMO proteomics identifies PIAS1 substrates involved in cell migration and motility. *Nat Commun* **11**, 834 (2020).
9. C. F. Tsai *et al.*, An Improved Boosting to Amplify Signal with Isobaric Labeling (iBASIL) Strategy for Precise Quantitative Single-cell Proteomics. *Mol Cell Proteomics* **19**, 828-838 (2020).
10. O. Barroso-Gomila *et al.*, Identification of proximal SUMO-dependent interactors using SUMO-ID. *Nat Commun* **12**, 6671 (2021).
11. C. D. Kelstrup *et al.*, Performance Evaluation of the Q Exactive HF-X for Shotgun Proteomics. *J Proteome Res* **17**, 727-738 (2018).
12. F. Dossin *et al.*, SPEN integrates transcriptional and epigenetic control of X-inactivation. *Nature* **578**, 455-460 (2020).
13. V. Akimov *et al.*, UbiSite approach for comprehensive mapping of lysine and N-terminal ubiquitination sites. *Nat Struct Mol Biol* **25**, 631-640 (2018).
14. I. A. Hendriks *et al.*, Site-specific characterization of endogenous SUMOylation across species and organs. *Nature Communications* **9** (2018).
15. M. Watanabe *et al.*, A substrate-trapping strategy to find E3 ubiquitin ligase substrates identifies Parkin and TRIM28 targets. *Commun Biol* **3**, 592 (2020).
16. B. Stynen, H. Tournu, J. Tavernier, P. Van Dijck, Diversity in genetic in vivo methods for protein-protein interaction studies: from the yeast two-hybrid system to the mammalian split-luciferase system. *Microbiol Mol Biol Rev* **76**, 331-382 (2012).
17. H. C. Yen, S. J. Elledge, Identification of SCF ubiquitin ligase substrates by global protein stability profiling. *Science* **322**, 923-929 (2008).
18. O. Karayel, A. C. Michaelis, M. Mann, B. A. Schulman, C. R. Langlois, DIA-based systems biology approach unveils E3 ubiquitin ligase-dependent responses to a metabolic shift. *Proc Natl Acad Sci U S A* **117**, 32806-32815 (2020).
19. H. F. O'Connor *et al.*, Ubiquitin-Activated Interaction Traps (UBAITs) identify E3 ligase binding partners. *EMBO Rep* **16**, 1699-1712 (2015).
20. J. Zecha *et al.*, TMT Labeling for the Masses: A Robust and Cost-efficient, In-solution Labeling Approach. *Mol Cell Proteomics* **18**, 1468-1478 (2019).
21. A. J. George, Y. C. Hoffiz, A. J. Charles, Y. Zhu, A. M. Mabb, A Comprehensive Atlas of E3 Ubiquitin Ligase Mutations in Neurological Disorders. *Front Genet* **9**, 29 (2018).
22. B. Levine, G. Kroemer, Biological Functions of Autophagy Genes: A Disease Perspective. *Cell* **176**, 11-42 (2019).

23. L. Lescouzeres, P. Bomont, E3 Ubiquitin Ligases in Neurological Diseases: Focus on Gigaxonin and Autophagy. *Front Physiol* **11**, 1022 (2020).
24. J. Zhu, N. P. Tsai, Ubiquitination and E3 Ubiquitin Ligases in Rare Neurological Diseases with Comorbid Epilepsy. *Neuroscience* **428**, 90-99 (2020).
25. D. Senft, J. Qi, Z. A. Ronai, Ubiquitin ligases in oncogenic transformation and cancer therapy. *Nat Rev Cancer* **18**, 69-88 (2018).
26. M. Bekes, D. R. Langley, C. M. Crews, PROTAC targeted protein degraders: the past is prologue. *Nat Rev Drug Discov* **21**, 181-200 (2022).
27. L. B. Snyder *et al.*, The discovery of ARV-471, an orally bioavailable estrogen receptor degrading PROTAC for the treatment of patients with breast cancer. *Cancer Research* **81** (2021).
28. X. Y. Zhang, V. M. Crowley, T. G. Wucherpfennig, M. M. Dix, B. F. Cravatt, Electrophilic PROTACs that degrade nuclear proteins by engaging DCAF16. *Nature Chemical Biology* **15**, 737-+ (2019).
29. L. Liu *et al.*, UbiHub: a data hub for the explorers of ubiquitination pathways. *Bioinformatics* **35**, 2882-2884 (2019).
30. S. L. Schreiber, The Rise of Molecular Glues. *Cell* **184**, 3-9 (2021).
31. Y. N. Chen *et al.*, Allosteric inhibition of SHP2 phosphatase inhibits cancers driven by receptor tyrosine kinases. *Nature* **535**, 148-152 (2016).
32. S. L. Schreiber, Immunophilin-sensitive protein phosphatase action in cell signaling pathways. *Cell* **70**, 365-368 (1992).
33. M. J. Larsen *et al.*, Classifications within Molecular Subtypes Enables Identification of BRCA1/BRCA2 Mutation Carriers by RNA Tumor Profiling. *Plos One* **8** (2013).
34. T. Abe *et al.*, AND-1 fork protection function prevents fork resection and is essential for proliferation. *Nat Commun* **9**, 3091 (2018).
35. R. Drost *et al.*, BRCA1 RING function is essential for tumor suppression but dispensable for therapy resistance. *Cancer Cell* **20**, 797-809 (2011).
36. R. Shakya *et al.*, BRCA1 tumor suppression depends on BRCT phosphoprotein binding, but not its E3 ligase activity. *Science* **334**, 525-528 (2011).
37. A. Taglialatela *et al.*, REV1-Polzeta maintains the viability of homologous recombination-deficient cancer cells through mutagenic repair of PRIMPOL-dependent ssDNA gaps. *Mol Cell* **81**, 4008-4025 e4007 (2021).
38. P. S. Patel, A. Algouneh, R. Hakem, Exploiting synthetic lethality to target BRCA1/2-deficient tumors: where we stand. *Oncogene* **40**, 3001-3014 (2021).
39. D. Zatreanu *et al.*, Poltheta inhibitors elicit BRCA-gene synthetic lethality and target PARP inhibitor resistance. *Nat Commun* **12**, 3636 (2021).
40. T. Hayashi *et al.*, Ubc9 is essential for viability of higher eukaryotic cells. *Exp Cell Res* **280**, 212-221 (2002).
41. K. Nacerddine *et al.*, The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. *Dev Cell* **9**, 769-779 (2005).
42. W. Roth *et al.*, PIASy-deficient mice display modest defects in IFN and Wnt signaling. *J Immunol* **173**, 6189-6199 (2004).
43. K. A. Wong *et al.*, Protein inhibitor of activated STAT Y (PIASy) and a splice variant lacking exon 6 enhance sumoylation but are not essential for embryogenesis and adult life. *Mol Cell Biol* **24**, 5577-5586 (2004).
44. Y. C. Liang *et al.*, SUMO5, a Novel Poly-SUMO Isoform, Regulates PML Nuclear Bodies. *Sci Rep* **6**, 26509 (2016).
45. E. Evdokimov, P. Sharma, S. J. Lockett, M. Lualdi, M. R. Kuehn, Loss of SUMO1 in mice affects RanGAP1 localization and formation of PML nuclear bodies, but is not lethal as it can be compensated by SUMO2 or SUMO3. *J Cell Sci* **121**, 4106-4113 (2008).
46. L. Wang *et al.*, SUMO2 is essential while SUMO3 is dispensable for mouse embryonic development. *EMBO Rep* **15**, 878-885 (2014).
47. M. Li *et al.*, SUMO2 conjugation of PCNA facilitates chromatin remodeling to resolve transcription-replication conflicts. *Nat Commun* **9**, 2706 (2018).
48. I. A. Hendriks *et al.*, Uncovering global SUMOylation signaling networks in a site-specific manner. *Nat Struct Mol Biol* **21**, 927-936 (2014).

49. H. Saitoh, J. Hinchey, Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. *J Biol Chem* **275**, 6252-6258 (2000).
50. M. H. Tatham *et al.*, Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. *J Biol Chem* **276**, 35368-35374 (2001).
51. I. A. Hendriks, A. C. Vertegaal, A comprehensive compilation of SUMO proteomics. *Nat Rev Mol Cell Biol* **17**, 581-595 (2016).
52. A. Gartner *et al.*, Acetylation of SUMO2 at lysine 11 favors the formation of non-canonical SUMO chains. *EMBO Rep* **19** (2018).
53. A. M. Sriramachandran *et al.*, Arkadia/RNF111 is a SUMO-targeted ubiquitin ligase with preference for substrates marked with SUMO1-capped SUMO2/3 chain. *Nat Commun* **10**, 3678 (2019).

&

APPENDIX

&

SAMENVATTING

In dit proefschrift focussen we op de post-translationele modificatie van eiwitten (PTM) met speciale aandacht voor Ubiquitinatie en SUMOylatie. We beschrijven zowel de Ubiquitine- als de SUMO E3-ligase families en laten niet alleen de verschillende mechanismen zien om respectievelijk Ubiquitine of SUMO aan een substraateiwit te conjugeren, maar ook de signaalroutes waarbij ze betrokken zijn. Naast monomeren kunnen E3's ook polymeren van ubiquitine of SUMO koppelen aan een substraateiwit waardoor polymeren ontstaan. Afhankelijk van het ketentype en de polymeermorfologie zal dit tot verschillende cellulaire uitkomsten leiden. Aan het einde van hoofdstuk 1 analyseren we de voordelen en beperkingen van huidige massaspectrometrie (MS)-benaderingen om zowel Ubiquitine- als SUMO-conjugaten te identificeren.

In hoofdstuk 2 bespreken we recente literatuur over DNA-schadeherstel (DDR) en hoe dit wordt georkestreerd door verschillende E3-ligasen. Van deze E3's richten we ons op BRCA1-BARD1 als een van de belangrijkste coördinatoren van de DDR en de implicaties ervan voor borst- en eierstokkanker. De BRCA1-BARD1 heterodimeer wordt gevormd door de interactie van BRCA1 en BARD1 via hun RING-domeinen (Really Interesting New Gene). De vorming van dit heterodimeer is essentieel voor de enig bekende enzymatische activiteit, met slechts één goed gedefinieerd ubiquitine substraat, het histon H2A. Na meer dan 30 jaar onderzoek naar BRCA1, blijft de biologische relevantie van de E3-ligase-activiteit onduidelijk en is de rol in het repareren van DNA schade via homologe recombinatie (HR) en de rol bij het onderdrukken van tumoren nog steeds niet opgelost. We bespreken de controverse over dit heterodimeer en de huidige richtingen in het veld. Om dit hoofdstuk af te ronden, onthullen we verschillende manieren om de ubiquitine machinerie aan te pakken en de DDR te beheersen om kanker te overwinnen.

In hoofdstuk 3 hebben we een massaspectrometrie-techniek ontwikkeld die we Target of Ubiquitin Ligases Identified by Proteomics 2 (TULIP2) noemen. Er zijn veel ziektes, waaronder neurodegeneratieve aandoeningen en types kanker, die het gevolg zijn van ontregeling van E3-ligasen. TULIP2 maakt de identificatie mogelijk van specifieke substraten van deze E3-ligasen met behulp van MS en biedt daardoor nieuwe mogelijkheden om een bepaalde ziekte aan te pakken. In vergelijking met eerdere methodologieën verhoogt TULIP2 meer dan 50 keer de opbrengst van ubiquitinatie conjugaten en acht keer het signaal van de E3-ligase die wordt bestudeerd. Over het algemeen biedt dit hoofdstuk een duidelijke en gedetailleerde methodologie die kan worden geïmplementeerd in elk laboratorium dat geïnteresseerd is in de identificatie van substraten van een E3-ligase.

De TULIP2-methodologie kan worden aangepast om SUMO-conjugaten te vangen door de gefuseerde ubiquitine te vervangen met een SUMO-groep. In hoofdstuk 4 hebben we SUMO Activated Target Traps (SATTs) ontwikkeld, die volgens de TULIP2-methodologie de identificatie van een E3-specifiek SUMO-proteoom mogelijk maakten. Met behulp van

acht verschillende SUMO E3-ligasen (PIAS1, PIAS2, PIAS3, PIAS4, NSMCE2, ZNF451, LAZSUL (ZNF451-3) en ZMIZ2) identificeerden we 427 potentiële SUMO1 en 961 potentiële SUMO2/3 substraten op een E3-specifieke manier. Hoewel we overlap vonden tussen E3-ligasen, vonden we ook een hoge specificiteit, zelfs op het substraat-isovormniveau. Om op een gebruiksvriendelijke manier door de dataset te kunnen bladeren, hebben we de online webapp Polar Volcano plots ontwikkeld, die vrij toegankelijk is (Polar Volcano plots (shinyapps.io)).

Vervolgens hebben we TULIP2-technologie gebruikt om borst- en eierstokkanker te bestuderen. In hoofdstuk 5 hebben we zowel BRCA1-TULIP2 als BARD1-TULIP2 gegenereerd om BRCA1-BARD1-specifieke ubiquitine substraten te identificeren. BRCA1 en BARD1 erfelijke mutaties worden in verband gebracht met een hoog risico op het ontwikkelen van borst- en eierstokkanker. Het vinden van BRCA1-BARD1-ubiquitine substraten is cruciaal om de E3-activiteit ervan te begrijpen en daarmee nieuwe mogelijkheden voor de behandeling van kanker. Door TULIP2 methodologie te gebruiken, ontdekten we dat BARD1 bij voorkeur substraten ubiquitineert die verschillen van BRCA1. Naast het bekende doelwit H2A, ontdekten we dat de macro H2A (m2HA)-variant een specifiek doelwit is, en PCNA lijkt het belangrijkste ubiquitine substraat te zijn tijdens normale groeiomstandigheden. We rapporteren dat BARD1-gemedieerde PCNA-ubiquitinatie betrokken is bij het voorkomen van accumulatie van enkelstrengs DNA (ssDNA), in tegenstelling tot RAD18-gemedieerde PCNA-ubiquitinatie, die betrokken is bij het repareren van UV-geïnduceerde DNA schade. Bovendien ontdekten we dat de katalytische dode mutant van BRCA1 niet wordt gefosforyleerd op S114, wat leidt tot een gecompromitteerde bescherming van de replicatievork. Over het algemeen concluderen we dat BRCA1-BARD1 E3-activiteit belangrijk is voor genetische stabiliteit door de homeostase van de replicatievork te reguleren in plaats van betrokken te zijn bij het HR-pad, wat nieuwe therapeutische mogelijkheden opent.

In hoofdstuk 6 zochten we specifieke substraateiwitten voor een E2-conjugerend enzym, UBE2D3. Uit eerder onderzoek blijkt dat deze E2 interacteert met BRCA1-BARD1 op plaatsen met DNA-schade en met RNF8 voor PCNA-ubiquitinatie. UBE2D3 wordt doorgaans gebruikt als E2 voor in vitro ubiquitinatie, de rol in vivo is echter nog niet gedefinieerd. In hoofdstuk 6 hebben we op SILAC gebaseerde en labelvrije kwantitatieve ubiquitine diGly proteomics uitgevoerd, samen met TULIP2-technologie, om globale proteoom- en ubiquitinome-veranderingen na UBE2D3 depletie te onderzoeken. Het ubiquitine-proteoom verandert volledig na UBE2D3-depletie, wat een prominent effect heeft op moleculaire routes die verband houden met mRNA-translatie. Met de UBE2D3-TULIP2-methodiek waren we in staat om, naast PCNA als reeds bekend substraat, twee ribosomale eiwitten (RPS10 en RPS20) als specifieke substraten te detecteren, die cruciaal zijn voor ribosoom-geassocieerde eiwitkwaliteitscontrole (RQC). Ubiquitinatie van RPS10 en RPS20 was ook UBE2D3-afhankelijk in SILAC en diGly proteomics, wat een nieuwe in vivo rol voor UBE2D3 onthulde.

In het laatste hoofdstuk geven we enkele inzichten over de keuzemogelijkheden voor MS-experimenten, afhankelijk van het onderzoeksdoel, en bespreken we het TULIP2-technologiepotentieel, dat in dit proefschrift is ontwikkeld. Vervolgens bespreken we verschillende MS-benaderingen en nieuwe trends om E3-ligase doeleiwitten te identificeren. We bespreken onze bevindingen op het gebied van SUMO in de context van de gepubliceerde literatuur om te laten zien hoe een voorheen verondersteld klein post-translationeel modificatienetwerk zeer complex en geavanceerd blijkt te zijn. Daarnaast bespreken en tonen we het potentieel van MS voor het monitoren van borst- en eierstokkanker, waarbij we verschillen in het ubiquitine-proteoom observeren. Ten slotte bespreken we hoe al onze bevindingen van potentieel nut kunnen zijn in de kliniek.

RESUMEN POPULAR

Las células constituyen los tejidos de nuestro cuerpo y son las encargadas de producir cambios ante diferentes situaciones, como por ejemplo producir saliva cuando vemos comida, aumentar la temperatura cuando combatimos patógenos o taponar las heridas que nos hacemos. Sin embargo, hay acciones que en ocasiones no percibimos, como por ejemplo la reparación del ADN cuando se ha dañado. El ADN se encuentra en el núcleo de las células y puede ser transcrito y traducido a proteínas, las cuales van a participar en la mayoría de los procesos celulares. Muchas proteínas se encuentran en las células mientras que otras se sintetizan cuando es conveniente. No obstante, hay situaciones donde se requiere que una proteína desaparezca o que vaya a un lugar determinado. Es entonces cuando estas proteínas se pueden modificar. En el capítulo 1 hablamos de estas modificaciones, más conocidas como modificaciones post-traduccionales de las proteínas. Entre ellas, nos centramos en la adición de pequeñas proteínas como ubiquitina (Ub) o Small Ubiquitin-Like Modifier (SUMO), resultando en ubiquitinación o SUMOylacion respectivamente. Estas proteínas son conjugadas a una proteína diana mediante la acción de una cascada enzimática que involucra tres enzimas: E1, E2 y E3. Esta última se denomina E3 ligasa y es la responsable de dirigir Ub o SUMO a una proteína diana. El papel más conocido de la ubiquitinación es el de hacer desaparecer una proteína mediante su degradación. Una forma de identificar proteínas es el uso del espectrómetro de masas, el cual se ha desarrollado vertiginosamente en los últimos años.

En el capítulo 2, revisamos trabajos recientes donde las E3s juegan un papel importante en la reparación del ADN. Cuando el ADN no se repara de manera adecuada da lugar a mutaciones, que si ocurren en proteínas encargadas de controlar la proliferación celular, pueden originar enfermedades como el cáncer. Entre estas E3s, nos centramos en BRCA1-BARD1 como uno de los protagonistas principales que coordinan la respuesta a daño en el ADN (DDR). BRCA1 es una proteína multifuncional, y forma un heterodímero con BARD1 que confiere actividad E3 de ubiquitina. Se ha demostrado que esta proteína es esencial para reparar los cortes de doble cadena (DSBs) en el ADN por una vía de reparación que se conoce como Recombinación Homologa (HR). Sin embargo, el papel de su actividad E3 para reparar estos DSB es controvertido. Mutaciones en estas proteínas resultan en altas probabilidades de desarrollar cáncer de mama y ovario y, después de más de 30 años investigándolas, su actividad E3 para prevenir estas enfermedades es todavía desconocida.

En el capítulo 3, desarrollamos una tecnología basada en espectrometría de masas con la que podemos identificar las proteínas diana de las E3s de ubiquitina. Esta herramienta la hemos llamado Targets for Ubiquitin Ligases Identified by Proteomics 2 (TULIP2). La desregulación de E3s no está asociada solo con cáncer, sino también con otras patologías como enfermedades neurodegenerativas. Utilizando el TULIP2 podemos identificar las proteínas diana de una E3 de interés para conocer mejor su función y el porqué de una

enfermedad. Conociendo estas proteínas podemos desarrollar fármacos para dar nuevas oportunidades terapéuticas.

La tecnología TULIP2 se puede modificar para identificar las proteínas diana de las E3s de SUMO en vez de las de ubiquitina. En el capítulo 4 presentamos las SATTs (SUMO Activated Target Traps), donde hemos cambiado Ub por SUMO en el constructo TULIP2. La tecnología SATT nos ha permitido identificar proteínas SUMOyladas para ocho E3 ligasas diferentes de manera específica (PIAS1, PIAS2, PIAS3, PIAS4, NSMCE2, ZNF451, LAZSUL (ZNF451-3) y ZMIZ2). En total hemos identificado 427 proteínas diana de SUMO1 y 961 de SUMO2/3. Además, hemos diseñado una herramienta web para explorar qué E3 ligasa modifica una proteína diana de interés (Polar Volcano plots (shinyapps.io)).

En el capítulo 5 hemos empleado la tecnología TULIP2 para identificar las proteínas diana de BRCA1-BARD1 y poder determinar cuál es su función biológica. A día de hoy los tratamientos contra el cáncer de mama que son BRCA1 deficientes se basan en fármacos que causan cortes de doble cadena (DSB) en el ADN, aprovechando que estas células no tienen la proteína BRCA1 o la tienen mutada, y por tanto, no pueden reparar estos DSBs. Desafortunadamente, surgen variantes de cáncer que restauran la capacidad de reparar los DSBs y confieren resistencia contra estos tratamientos. Algunas de estas variantes vienen de versiones de BRCA1 que no poseen actividad E3, pero si pueden reparar los DSBs. Por este motivo averiguar las proteínas diana de BRCA1-BARD1 puede ayudar a abrir nuevas posibilidades contra estas variantes. El TULIP2 de BARD1, además de identificar la histona H2A como la proteína diana más conocida para este heterodímero, ha permitido detectar otra más, PCNA. Las células proliferan y, si son tumorales, lo hacen de manera más rápida y menos controlada. Para ello, las células tienen que replicar el ADN y PCNA se ubiquitina para controlar este proceso. Si hay problemas en la replicación, esta se para y se protege para evitar daños en el ADN. Nosotros hemos observado que la ubiquitinación de PCNA mediada por BRCA1-BARD1 evita que se acumule ADN de cadena sencilla, cuya acumulación resulta en daño en el ADN. Además, hemos observado que la actividad E3 de BRCA1-BARD1 es importante para la protección de la replicación. Cuando utilizamos fármacos que alteran la replicación, como por ejemplo hydroxyurea, las células con mutaciones en BRCA1 que no tienen actividad E3, resultan más sensibles que las células sanas. Con estas observaciones, nuestro trabajo ofrece nuevas oportunidades terapéuticas contra variantes de cáncer de mama que no disponen de la actividad E3 en BRCA1.

La tecnología TULIP2 también se puede emplear para averiguar las proteínas diana de las E2s. En el capítulo 6, hemos empleado diferentes estrategias basadas en espectrometría de masas para averiguar los procesos celulares en los que la E2 UBE2D3 está involucrada. Entre las diferentes estrategias hemos utilizado SILAC, LFQ-DiGly y el TULIP2. Las tres estrategias han permitido identificar proteínas involucradas en diferentes procesos celulares como el procesamiento del ARNm y en el control de la calidad de las proteínas producidas en los ribosomas (RQC). La identificación de las proteínas diana de

esta E2 ha revelado un nuevo papel de UBE2D3 *in vivo* que hasta entonces era desconocido.

En el último capítulo se discute sobre diferentes métodos que se pueden emplear utilizando espectrometría de masas, el potencial que tiene la tecnología TULIP2 que se ha desarrollado en esta tesis y cómo se podrían implementar los avances que hemos originado en la clínica.

&

ABBREVIATIONS

ABC	Ammonium Bicarbonate
AML	Acute Myeloid Leukemia
BARD1	BRCA1-Associated RING Domain 1
BCA	Bicinchoninic Acid
BioID	Proximity-dependent Biotin Identification
BrdU	5’-bromo-2’-deoxyuridine
BRCA1	Breast cancer susceptibility type 1
ChIP	Chromatin Immunoprecipitation
CldU	5-chloro-2’-deoxyuridine
CUL	Cullin
DDA	Data Dependent Acquisition
DDR	DNA Damage Response
DIA	Data Independent Acquisition
DiGly	Di Glycine (Ubiquitin tryptic remnant)
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
DUB	Deubiquitinating Enzyme
Dox	Doxycycline
DSB	Double Strand Break
HR	Homologous Recombination
HU	Hydroxyurea
IdU	5-Iodo-2’-deoxyuridine

iPOND	isolation of Proteins On Nascent DNA
IR	Ionizing Radiation
K	Lysine
LFQ	Label Free Quantification
LUBAC	Linear Ubiquitin Chain Assembly Complex
MoaD	Bacteria protein molybdopterin converting factor subunit 1
MS	Mass Spectrometry
NHEJ	Non-Homologous End Joining
PARylation	ADP-ribosylation
PARPi	PARP inhibitor
PBS	Phosphate-Buffered Saline
PBST	PBS supplemented with 0.05% Tween 20
PCNA	Proliferating Cell Nuclear Antigen
PIAS	Protein Inhibitor of Activated STAT
PINK1	PTEN-induce putative kinase 1
PRC1	Polycomb Repressive Complex 1
PTMs	Post Translational Modifications
RING	Really Interesting New Gene
RT	Room Temperature
SILAC	Stable Isotope Labeling by Amino acids in Cell culture
SIMs	SUMO Interacting Motifs
SSA	Single Strand Annealing
SSB	Single Strand Break
ssDNA	Single-Stranded DNA
SR	Substrate Receptor

STUbLs	SUMO-targeted Ubiquitin Ligases
SUMO	Small Ubiquitin-Like Modifier
ThiS	Thiamine biosynthesis protein S
TLS	Trans-Lesion Synthesis
TS	Temple Switching
TULIP	Targets for Ubiquitin Ligases Identified by Proteomics
Ub	Ubiquitin
UBDs	Ubiquitin Binding Domains
UbLs	Ubiquitin-Like proteins
USP	Ubiquitin Proteasome System
UV	Ultraviolet Light
WB	Western Blot
WT	Wild Type
ZNF451	Zinc finger 451

CURRICULUM VITAE

Daniel Salas Lloret was born on the 29th of June 1994 in Alicante (Spain), although he was raised in Alcázar de San Juan, Ciudad Real (Spain). It was there where he completed the International Bachillerato (High School). In 2012, he started his undergraduate studies in the field of Biochemistry with mention in Biotechnology at University of Castilla La-Mancha (UCLM) in Toledo, Spain. At the end of his undergraduate studies, he performed his bachelor thesis in the laboratory of Professor Christoph Wülfing in the School of Cellular and Molecular Medicine at Bristol University, UK. There, he studied the role of actin regulators at the immune synapse of CD4 and CD8 T-cells. Subsequently, he enrolled in a post-graduate Biotechnology MSc qualification at Autonomous University of Madrid (UAM) in Spain. His master thesis was completed in Dr. Mark J. van Raaij laboratory at National Center of Biotechnology (CNB-CSIC), where Daniel gained expertise in the field of protein origami working on the design of putative self-assembling viral proteins-based building blocks. After completing is MSc education, Daniel was awarded with an European funded grant to work in the laboratory of Professor Andrés Aguilera López in the department of Genetic Instability and Cancer at the Andalusian Molecular Biology and Regenerative Medicine Center (CABIMER). There, he learned how to work with yeast and developed a project focused on topoisomerases and genetic instability. In June 2018, Daniel joined Dr. Román González Prieto as PhD student funded by the Dutch Cancer Society (KWF) in the department of Cell and Chemical Biology at Leiden University Medical Center (LUMC) in The Netherlands. Here, he developed new mass-spectrometry technologies for the identification of E3 ligases substrates for both ubiquitin and small ubiquitin-like modifiers (SUMO). He employed this technology for the development of a comprehensive and interactive E3-specific SUMO proteome and studying the BRCA1-BARD1 E3 ligase for breast cancer vulnerabilities. During his PhD, Daniel attended several conferences and workshops around Europe where he presented his work through posters and oral presentations. In 2023, Daniel joined Professor Dr. Alfred Vertegaal laboratory as post-doctoral fellow.

ACKNOWLEDGEMENTS

As every journey, the PhD also has an end, even if you do not believe it yet. You would agree with me that it is not a smooth journey and that it requires a lot of effort and dedication. Here, I want to make clear that not only in the PhD but in science in general, when there is success, it is never a one person achievement. Do not ever think that only one person can achieve something big in science. It is just not possible. Therefore, it would not be different in my case. I would like to thank everybody who has been involved in this journey and manifest that I could have never finished my PhD without them. There are that many people that I could not fit everyone in this section, thus I apologize to anyone left out.

First of all, I would like to thank Román for giving me the opportunity to do a PhD. Not only for allowing me to do it, but also to put your trust in me.

I also want to thank Alfred for his guidance and support since the very beginning. Your door has always been open for me and that is something I am very grateful for.

Likewise, I would like to thank my guidance committee, David and Martijn, for the guidance and support during the progress meetings. We got really good ideas after our discussions and I always felt safe knowing I could go to you with any matter.

This would have neither been possible without all my co-workers, which now are also friends. You were all amazing, I feel very lucky to have met you all. We have lived lots of moments, shared really good news and real bad ones. I will always keep all the memories with me. For all of you, you already know who you are, thank you so much and we will keep in touch for more adventures.

SUMO wrestlers, all of you are also included in the words above. Additionally, I also wanted to thank you for being more than colleagues. We laughed/cried when our experiments went wrong and we celebrated when things were going right. I felt the support of everyone at every moment. When Fredrik joined my Karate lessons at the very beginning, and when we all gathered for dinner, sailing and partying. We supported each other as a family (you were my family in Leiden). Matty, you are the pillar that keeps everything in one piece, you take care of everyone in the lab including myself, and I am very thankful to have had a figure like you during my PhD. Going to work was much more nicer because of you all.

I also want to thank the people who helped me directly with my projects. Arnoud de Ru, thank you for your support and patience running, literally, thousands of mass-spectrometry samples. All the collaborators, who invested a lot of time and effort. Finally, my master students who are also co-authors in some of the chapters, they all have a great future ahead.

Marta, you made the dark days brighter. You moved with me to The Netherlands and dealt with my stress during all these years, that is something I will always keep in my heart. Thank you for being by my side, supporting me and helping me in each step of my life, you made the PhD journey so much easier.

Next, although they live thousands of kilometers away, they have been my support during my PhD, my family and friends in Spain.

No tiene sentido escribir esta parte en inglés, por lo que iré directo al grano. A mi familia, muchas gracias por apoyarme desde el primer momento en hacer una tesis en el extranjero.

Sabíais que el doctorado iba a suponer verme menos, faltar en cumpleaños, viajes, rutas, actividades, comidas y una larga lista que, espero cambie gracias a la consecución de este título. Aun así, no dudasteis en ningún momento. Cada minuto que nos veíamos, ya fuera aquí en Los Países Bajos o en España, contaba el doble.

A mis amigos, muchas gracias que, a pesar de la distancia, nuestra amistad nunca se ve afectada. Da igual que pasen 2 semanas que 2 años, siempre nos vemos con la misma sonrisa. Muchas gracias por vuestras visitas y, aunque no lo supierais, por estar a mi lado durante estos años.

&

LIST OF PUBLICATIONS

1. **D. Salas-Lloret**, G. Agabitini, R. Gonzalez-Prieto, TULIP2: An Improved Method for the Identification of Ubiquitin E3-Specific Targets. Front Chem 7, 802 (2019).

2. **D. Salas-Lloret**, R. Gonzalez-Prieto, Insights in Post-Translational Modifications: Ubiquitin and SUMO. Int J Mol Sci 23 (2022).

3. **D. Salas-Lloret** et al., SUMO-activated target traps (SATTs) enable the identification of a comprehensive E3-specific SUMO proteome. Sci Adv 9, eadh2073 (2023).

4. **D. Salas-Lloret** et al., BRCA1/BARD1 ubiquitinates PCNA in unperturbed conditions to promote replication fork stability and continuous DNA synthesis. bioRxiv 10.1101/2023.01.12.523782, 2023.2001.2012.523782 (2023). Under revision in Mol. Cell (2023)

5. Z. Yalcin, **D. Salas-Lloret** et al., Ubiquitinome Profiling Reveals in Vivo UBE2D3 Targets and Implicates UBE2D3 in Protein Quality Control. Mol Cell Proteomics 22, 100548 (2023).

6. **D. Salas-Lloret** and R. Gonzalez-Prieto, Unveiling BRCA1-BARD1 ubiquitin ligase heterodimer. DNA repair, Ubiquitin and Cancer. Revised in DNA repair, 2022.

7. Zeliha Yalcin, Shiu Yeung Lam, Marieke Peuscher, Jaco Torr, Prasanna Iyengar, **Daniel Salas-Lloret**, Inge de Krijger, Nathalie Moatti, Aurora Cerutti, Roman Gonzalez-Prieto, Jacqueline Jacobs, UBE2D3 facilitates NHEJ by orchestrating ATM signalling through multi-level control of RNF168. Nat. Struct. Mol. Biol. (2023). Under revision.