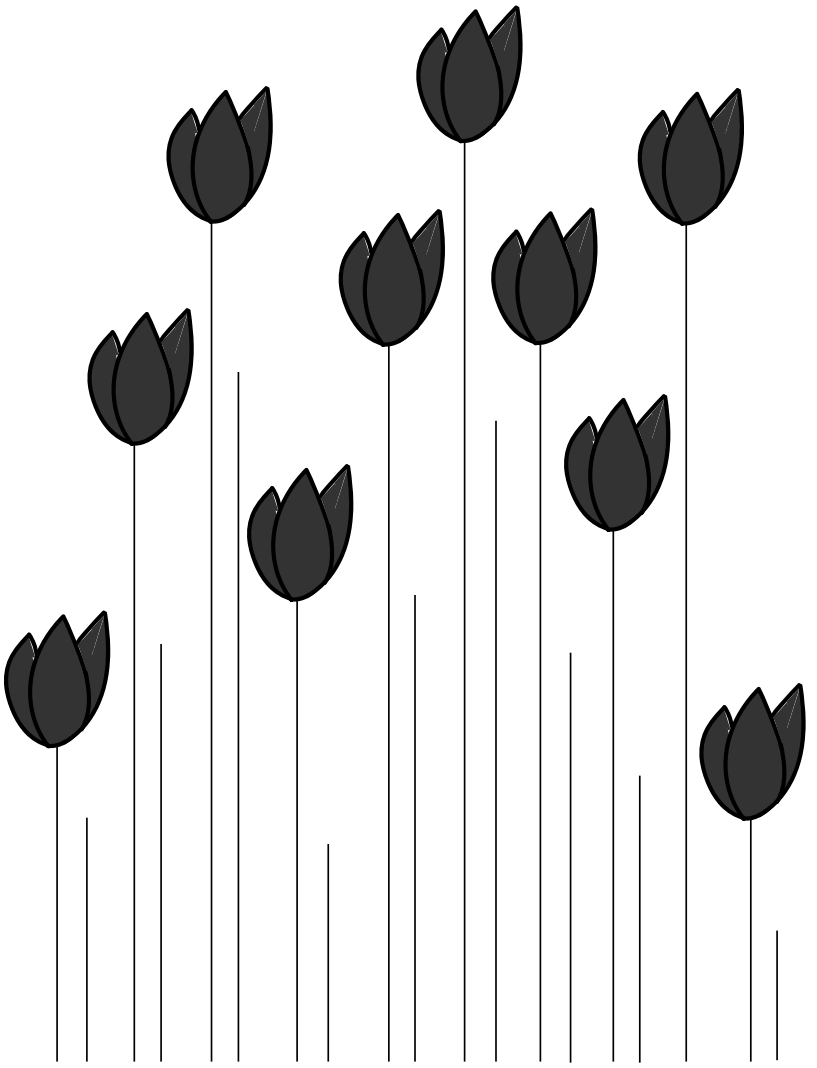
Proteomics and Functional Investigation of SUMO and Ubiquitin E3 ligases

Salas Lloret, D.



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).

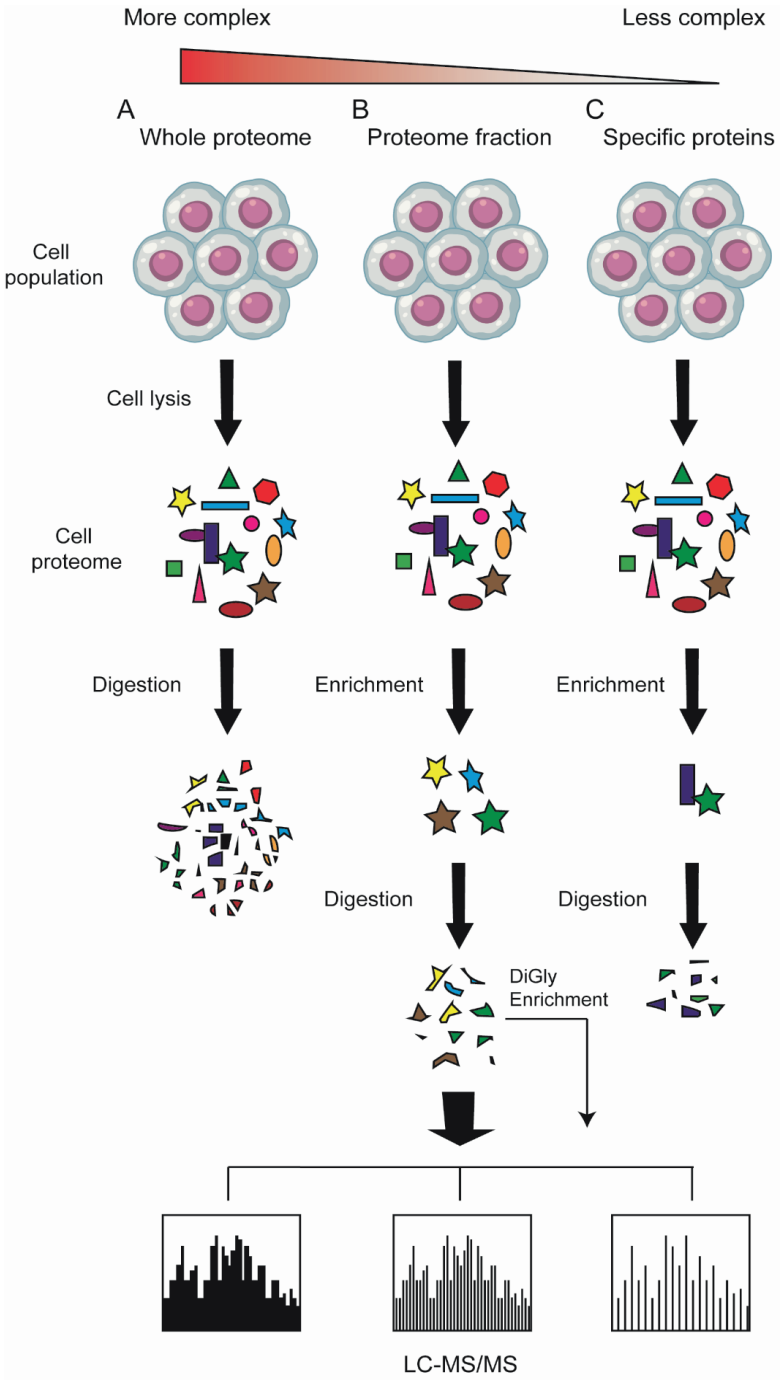


Figure 1. Sample complexity and different approaches depending on the research goal. **A**. Whole proteome of cell lysate followed by tryptic digestion and subsequent LC-MS/MS. **B**. Proteome fraction enrichment equivalent to His-Ub or His-SUMO pull downs followed by tryptic digestion and subsequent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

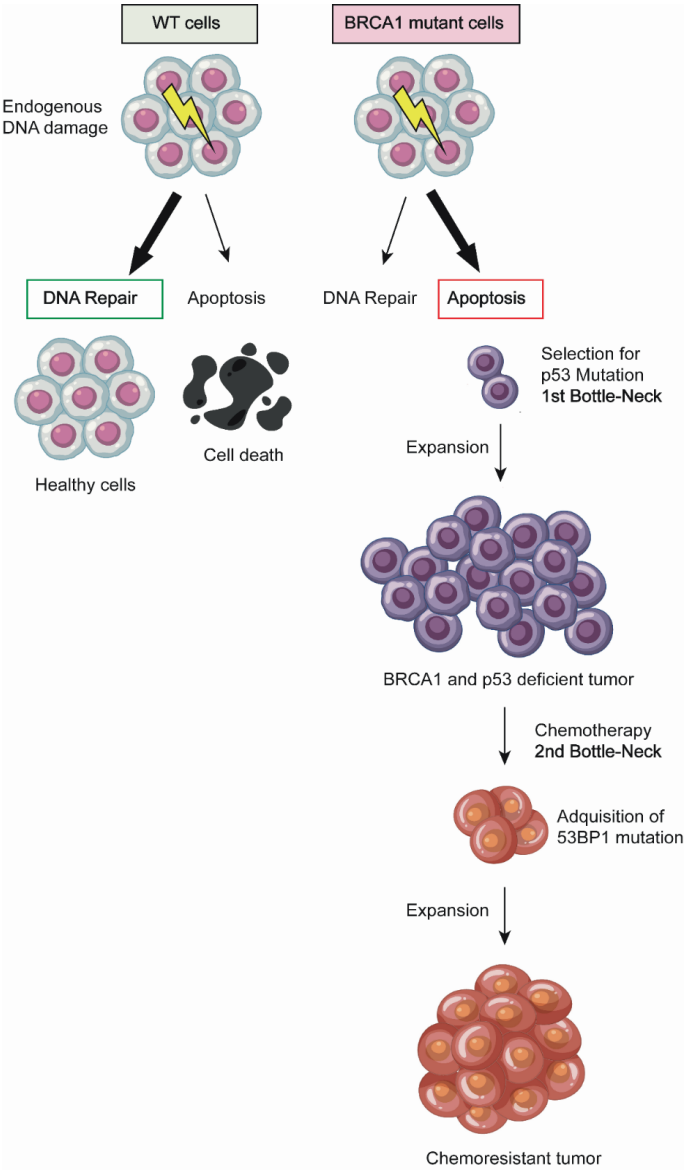


Figure 2. Depiction of breast cancer formation in patients with BRCA1 mutations. Cells undergo endogenous DNA damage. However, cells harboring BRCA1/BARD1 mutations are not able to fully repair some DNA lesions. Upon DNA damage, cells can either fix the lesion or go to apoptosis, this decision is regulated 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).