



Universiteit
Leiden
The Netherlands

Regulation of target proteins by small ubiquitin-like modifiers

Xiao, Z.

Citation

Xiao, Z. (2019, June 5). *Regulation of target proteins by small ubiquitin-like modifiers*. Retrieved from <https://hdl.handle.net/1887/74009>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/74009>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The following handle holds various files of this Leiden University dissertation:

<http://hdl.handle.net/1887/74009>

Author: Xiao, Z.

Title: Regulation of target proteins by small ubiquitin-like modifiers

Issue Date: 2019-06-05

Regulation of Target Proteins by Small Ubiquitin-like Modifiers

Zhenyu Xiao

Zhenyu Xiao

Regulation of Target proteins by Small Ubiquitin-like Modifiers

Thesis, Leiden University, 2019

ISBN/EAN: 978-94-6375-389-0

Printed by: Ridderprint BV

Cover designed by Zhenyu Xiao

Thesis layout by Zhenyu Xiao

The research described in this thesis was financially supported by the Marie Curie Initial Training Network UPStream.

Regulation of Target Proteins by Small Ubiquitin-like Modifiers

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens het besluit van het College voor Promoties
te verdedigen op woensdag 5 Juni 2019
klokke 09.00 uur

door

Zhenyu Xiao

Geboren te Ningxiang, P.R.China

op 24 December 1987

“I have been bent and broken, but - I hope - into a better shape”

Charles Dickens

(February 7, 1812 - June 9, 1870)

Contents

Chapter 1 Introduction	1
Chapter 2 System-wide analysis of SUMOylation dynamics in response to replication stress reveals novel SUMO target proteins and acceptor lysines relevant for genome stability	35
Chapter 3 Proteomics reveals global regulation of protein SUMOylation by ATM and ATR kinases in replication stress.....	69
Chapter 4 The STUbL RNF4 regulates protein group SUMOylation by targeting the SUMO conjugation machinery	111
Chapter 5 UFMylation: an emerging post-translational modification regulating protein synthesis	165
Chapter 6 General Discussion	187
Summary	199
Samenvatting	201
List of Publications	203
Acknowledgements	205
Curriculum Vitae	207

1 Post-translational modifications

1.1 Post-translational modifications, DNA damage repair and cancer

In 1953, James Watson and Francis Crick reported the twisted-ladder structure of deoxyribonucleic acid – DNA¹. Since then, the mechanisms that cells have employed to preserve and transmit the genetic information encoded by DNA have been subject to extensive investigations². DNA carries genetic information used in the development and reproduction of all cellular living organisms. However, the stability of DNA is constantly threatened by all kinds of damages (Figure 1). Unrepaired DNA damage leads to genomic instability, which is a characteristic of most cancers³⁻⁶.

In order to preserve genome integrity, cells are obliged to detect different types of damages and to prevent injury to their genomes by building up an arsenal of repair factors and by employing a number of multi-step and interconnected mechanisms⁷. All those factors and mechanisms are strictly regulated to prevent unessential alterations in DNA structure in a timing controlled, distribution restricted and DNA damage specific manner, to initiate an appropriate repair pathway (Figure 1) and also in the case of irreparable damage, to induce cell cycle arrest or apoptosis⁸.

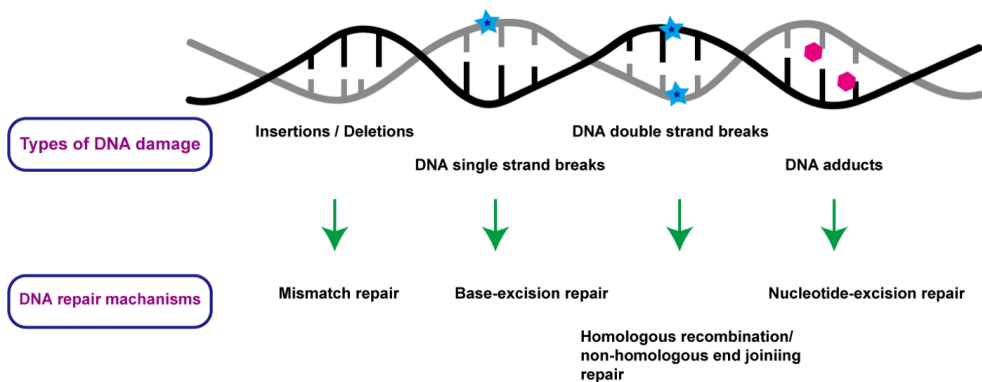


Figure 1. Different types of DNA damages. There are several types of DNA damages due to intrinsic and extrinsic factors and are repaired by damage-specific mechanisms, such as mismatch repair, base/nucleotide-excision repair, or homologous recombination (HR)/non-homologous end-joining (NHEJ).

Among numerous regulatory mechanisms that are available to the cell, rapid, reversible and flexible posttranslational protein modifications (PTMs) that covalently act during or after protein biosynthesis are markedly suited and selectively taken on to fine-tune these regulatory networks⁹⁻¹¹.

PTMs are chemical alterations to amino acids in proteins that reversibly regulate biochemical properties of proteins via specific domains by the addition of a modifying chemical group such as phosphate, acetyl and methyl or by an entire polypeptide including ubiquitin and SUMO (Small Ubiquitin Like Modifier) to one or more of its amino acid residues mainly through enzymatic activities (Figure 2).

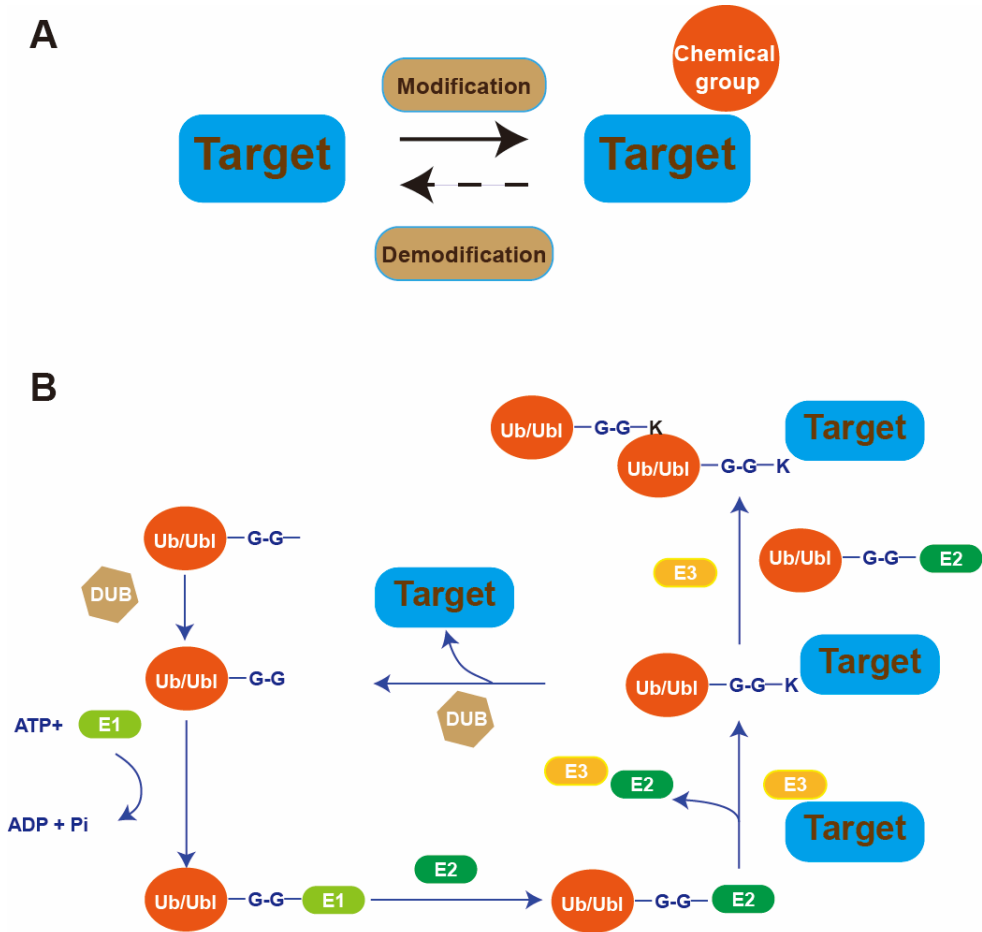


Figure 2. The process of chemical and ubiquitin like (UBL) modifications. A) The reversible process of chemical modifications is regulated via an enzymatic step to modify and de-modify target proteins. B) The reversible process of UBL modification is regulated through a three enzymatic step to modify and de-modify target proteins.

PTMs change the localization, conformation, interactions and stability of their targets proteins and bring functional diversity to them^{11,12}. Some PTMs such as glycosylation are stable modifications of the protein to ensure proper folding and stability of newly synthesized proteins¹³. Others like phosphorylation, ubiquitination and ubiquitin-like modifications such as SUMOylation are more dynamic to quickly alter the function of proteins^{10,14}. Moreover, different PTMs can act individually and/ or in concert to initiate, terminate and fine-tune different signalling pathways¹⁰.

Alterations and disruption in PTMs pathways have been tightly connected to many types of diseases, especially cancer^{12,15-18}. At the same time, regulators of PTMs serve as potential targets for drug development against those diseases. Global identification of protein posttranslational

modifications and intensive research in the biological function of PTMs in cellular processes will open new avenues to advance techniques for diagnosis and therapeutics and therefore are of great importance and an interesting field in basic scientific and drug development research¹⁸⁻²⁰.

1.2 phosphorylation, Ubiquitination, Ubiquitin-like proteins

1.2.1 Phosphorylation

Phosphorylation is one of the most extensively investigated and abundant PTMs^{19,21}. It modifies activities of a large fraction of proteins by adding a phosphoryl group to a target protein in both prokaryotic and eukaryotic cells²². Protein phosphorylation is reversible, it requires protein kinases to deposit phosphates and protein phosphatases to remove phosphates from the target proteins²¹. In eukaryotes, phosphorylation normally take place on serine, threonine, tyrosine and histidine residues whereas in prokaryotes it also happens on the arginine and lysine residues^{23,24}. Protein phosphorylation regulates key cellular functions such as cell growth, differentiation and apoptosis^{25,26}.

1.2.2 Ubiquitin and Ubiquitination

Ubiquitin was first identified in eukaryotic cells in 1975 as a widely expressed protein of unknown function²⁷. It can be attached via the free carboxyl group of the glycine at its C-terminus to the ϵ -amino group of lysine in its target protein^{14,28}. Furthermore, ubiquitin was reported to be able to bind to electron-rich nucleophiles in a protein known as non-canonical ubiquitination²⁹. Furthermore, a N-terminal amino group in a protein could also be conjugated to ubiquitin in a similar manner as lysine. This kind of non-canonical ubiquitination was first identified in MyoD, and later on, characterized in ERK3 and p21 by mass spectrometry analysis³⁰⁻³². N-terminal ubiquitination of proteins was further found to target proteins for degradation³². Cysteine can also be attached to ubiquitin. It was first reported as another non-canonical ubiquitination in the peroxisomal import factor (Pxp5). Beside thioester bonds, hydroxyester linkages can also be formed between ubiquitin and serines, threonines and tyrosines, which are more stable and suggested as part of the apoptotic pathways^{33,34}.

Ubiquitination is an enzymatic process that depends on ATP^{14,33,35,36}. It involves the enzymatic cascade of three enzymes. They are ubiquitin-activating enzymes (E1), ubiquitin-conjugating enzymes (E2) and ubiquitin ligases (E3). Ubiquitin is first activated by the ubiquitin-activating enzyme E1 and is then passed on to the active cysteine site of the E2s. The E2s act with E3 ubiquitin ligases to catalyse the finishing stage of the ubiquitination by adding ubiquitin to its targets. Ubiquitin E3 ligases contain either the Homologous to the E6-AP Carboxyl Terminus (HECT) domain or RING-betweenRING-RING (RBR) domain that both transiently bind ubiquitin or the Really Interesting New Gene (RING) domain that catalyses the straight transfer step to add ubiquitin to the substrates from E2 enzymes^{28,36,37}.

There are three types of ubiquitin modifications; mono-ubiquitination (adding one ubiquitin to one substrate residue), multi-ubiquitination (adding single ubiquitins to multiple substrate residues of one protein) and poly-ubiquitination (forming a ubiquitin chain on a single substrate residue by linking the

C-terminal glycine of a ubiquitin to an inner lysine of ubiquitin) (Figure 3)³⁸. There are seven lysine residues in ubiquitin, which are K6 (lysine 6), K11 (lysine 11), K27 (lysine 27), K29 (lysine 29), K33 (lysine 33), K48 (lysine 48) and K63 (lysine 63). While K48-linked chains are the first identified chains that target proteins for proteolytic degradation, K63-linked chains have been well-characterized as not associated with proteasomal degradation. The function of other lysine chains and, mixed chains remain less clear³⁹. Targeting proteins for ubiquitin-mediated 26S proteasome degradation is an irreversible choice which needs to be tightly regulated in a specific manner.

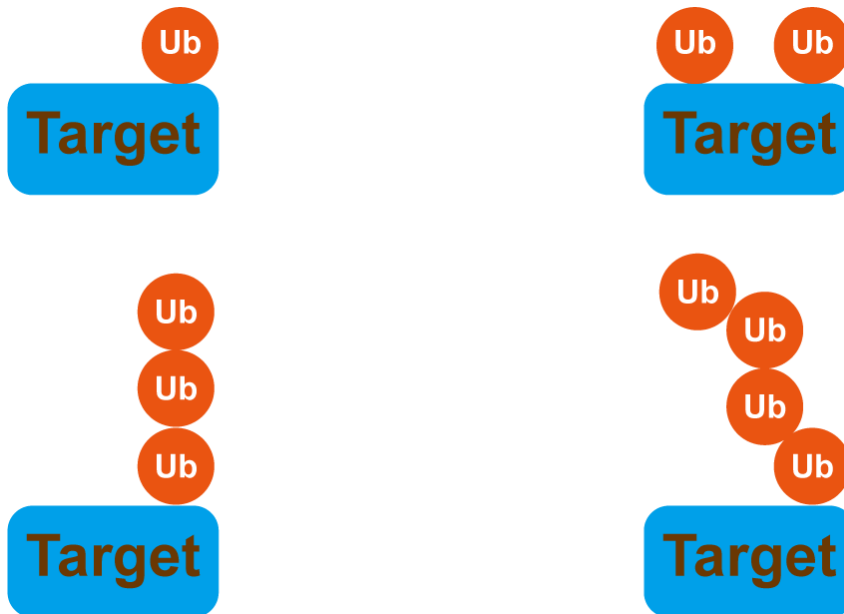


Figure 3. Four types of ubiquitination. mono-ubiquitination is to modify targets via adding one ubiquitin to one substrate residue, multi-ubiquitination is to modify targets via adding single ubiquitin to multiple substrate residues of one protein and poly-ubiquitination is to modify targets via forming a ubiquitin chain on a single substrate residue by linking the C-terminal glycine of a ubiquitin to an inner lysine of ubiquitin, there are two types of poly-ubiquitination, linear ubiquitination or branched ubiquitination.

Moreover, ubiquitin can also be linked linearly (the carboxyl group of glycine of one ubiquitin molecule linked to the amino group of methionine of another ubiquitin) (Figure 3). Linear ubiquitin chains are now known to be catalysed by two RING type E3 ligase HOIL1 and HOIP and are important for NF- κ B activation and cellular responses to inflammatory cytokines^{40,41}.

1.2.3 Ubiquitin-like proteins

It is worth noting that although ubiquitin is the first identified and most famous and well-studied protein post-translational modifier, there are other families of proteins identified that are able to modify proteins in an ubiquitin like enzymatic cascade. They share a common β -grasp fold and are known as ubiquitin-like proteins (UBLs)^{10,14,42-45}.

UBLs identified so far are: Small Ubiquitin-like Modifier (SUMO)^{12,18}, ubiquitin fold-modifier-1 (UFM-1)⁴⁶, autophagy-8 (ATG8) and -12 (ATG12)⁴⁷, ubiquitin-like protein FUBI (FUBI)⁴⁸, ubiquitin cross-reactive protein (UCRP or ISG15)⁴⁹, ubiquitin-related modifier-1 (URM1)⁵⁰, human leukocyte antigen F-associated (FAT10)⁵¹, neuronal-precursor-cell-expressed developmentally downregulated protein-8 (NEDD8)⁴⁴, MUB (membrane-anchored UBL)⁵² and ubiquitin-like protein-5 (UBL5)⁴³. Whilst sequences of those proteins share only moderate homology with ubiquitin, they have similar three-dimensional folds^{10,53,54}.

1.2.4 UFM-1 and UFMylation

Ubiquitin-fold modifier 1 (UFM-1) is the most recently identified new member of the ubiquitin-like protein family. It was first described in 2004 and is a protein of 9.1 kDa⁵⁵. Although UFM-1 shares little sequence identity to ubiquitin, they have a similar (α -helix and β -sheet) tertiary structure. Unlike ubiquitin functions that are extensively studied, the biological function of UFM-1 remains largely unknown⁵⁶.

UFM-1 covalently modifies the lysine residues of its substrates and is attached by an enzymatic cascade analogous to ubiquitination, including E1 (UBA5), E2 (UFC1), and E3 (UFL1) enzymes and this process can be reversed by UFM-1 specific proteases⁵⁶. UFM-1 is translated as an inactive precursor form (pro-UFM1) which has two additional amino acids beyond the single active conserved glycine. In human cells, so far the only active UFM-1 specific protease is UFSP2⁵⁷. UFSP2 cleaves the UFM-1 C-terminal part to expose a C-terminal glycine. UFM-1 only possesses a single active glycine at the C-terminus, which is required for the covalent attachment to its target proteins. UFM-1, UBA5, and UFC1 are all conserved in metazoan and plants but not in yeast, suggesting its potential roles in various multicellular organisms.

Interestingly, UBA5 can also activate SUMO-2 and transfer SUMO-2 to the nucleus. However loss of UBA5 only affects UFM-1 conjugation^{58,59}.

UFM-1 and its system have been demonstrated to play a significant role in regulating protein interaction, localization and function^{56,60}. It is also suggested that UFM-1 is involved in pathological conditions or diseases, like tumorigenesis⁶¹, ischemic heart diseases⁶² and diabetes⁶¹. However, to date, the biological function of UFM-1 remains poorly understood.

Although currently, several groups have reported their identification of UFM1-targets, such as ASC (activating signal co-integrator 1 or TRIP4)¹⁵ and UFBP1⁶³ and recently, 494 UFMylated proteins

got identified, of which 82 passed the cut off⁶⁴. None of them have been independently validated and no biological functional studies have been conducted to elucidate the role of protein UFMylation⁶⁴.

Previously, uS3, uS10, uL16, subunits of the ribosome have also been identified by mass spectrometry to be novel UFM-1 targets⁶⁵, thereby linking UFM-1 to a central biological process.

2 SUMO and SUMOylation

2.1 SUMO

Small Ubiquitin-related Modifier (SUMO) was first found as a Small Ubiquitin like MODifier during the 1990s when several other ubiquitin-like proteins were identified^{66,67}. It is a small protein of 12 kDa and is about 100 amino acids in length (depending on which organism the protein comes from and depending on the identity of the SUMO family member). Although SUMO strongly resembles ubiquitin with a similar structural protein fold and its ability to form chains, they share little sequence similarity with each other at the amino acid level. All SUMO proteins are translated into immature precursors whose C termini need to be removed by SUMO proteases (as described later in more detail) to expose the di-glycine (GG) motif. Mature SUMO is then ready for entering the SUMOylation cycle and conjugation to targets. SUMO often does not direct proteins for proteasomal degradation⁶⁸.

In budding yeast, a single form of SUMO protein exists, encoded by the SMT3 gene, which is essential⁶⁹. In contrast there are three functional genes that can be translated to three SUMO proteins in mammalian cells, which are SUMO-1, 2, 3. SUMO-1 is a highly conserved small ubiquitin like protein that has been first identified as modifier of Ran GTPase-Activating Protein 1 (RanGAP1) at its lysine 526 residue^{66,67}. Later on, two homologous coding sequences have been isolated and sequenced from mouse and human cDNA libraries and they are referred to as SUMO-2 (SMT3A) and SUMO-3 (SMT3B)⁷⁰.

SUMO-1 shares only about 50% amino acid sequence similarity with SUMO-2 and SUMO-3. SUMO-2 and SUMO-3 are almost indistinguishable⁷¹. They share a similarity of about 97% with each other in humans and cannot be distinguished by antibodies. Thus, they are grouped together into a subfamily named SUMO-2/3⁶⁸.

SUMO-1 and SUMO-2/3 were found to be conjugated to substrates under physiological conditions⁷². SUMO-2/3 conjugation can be further induced by cellular stresses⁷³. The overall cellular levels of SUMO-1 is much lower than that of SUMO-2/3. Although mice are viable after SUMO-1 and SUMO-3 knockout, SUMO-2 deficient mice die during the embryonic period^{71,72,74}. There is an internal SUMOylation site at lysine 11 of SUMO-2 and SUMO-3 and therefore they can form poly-SUMO chains, however, there is no internal SUMOylation site in SUMO-1. Although SUMO-1 cannot form poly-SUMO chains, the C-terminus of it can be linked to SUMO-2 or SUMO-3 which ends further chain elongation⁷⁵.

While identifying specific genes that are related to susceptibility to type 1 diabetes (T1D), the SUMO-4 gene was cloned⁷⁶. SUMO-4 shows similarity to SUMO-2/3, but differs in having a proline instead of a glutamine at position 90. As a result, SUMO-4 can only be conjugated under stress-conditions like starvation. SUMO-4 was shown to be conjugated to IκBα and negatively regulated NFκB transcriptional activity⁷⁶. However, no unique endogenous SUMO-4 peptides have been found, thus true evidence for SUMO-4 at the endogenous proteome level by mass spectrometry is missing and further evidence that SUMO-4 can be translated into protein needs to be provided.

SUMO-1, 2, and 3 are mainly localized in the nucleus and were shown to covalently attach to many kinds of proteins to regulate their interaction, stability, localization and function, not only under physiological conditions but also in response to cellular stresses⁶⁸. However, our understanding of the role that SUMOs are playing in cells is quite limited due to the fact that the stoichiometry of SUMOylation is very low, which makes it very challenging to be detected⁷⁷.

SUMOs are mainly attached to nuclear proteins in a redundant manner. Consequently, mutating one single SUMO acceptor lysine in a SUMO target protein often has no clear physiologic effect. Many studies support the concept of protein group modification, thus a pre-enrichment and a proteomic wide analysis of SUMO targets and a group modification view of them are key to understand SUMO signaling transduction^{73,78-83}.

2.2 SUMOylation

Resembling ubiquitination, SUMOs can also be attached to targets through sequential activity of three well conserved enzymes, involving a SUMO E1 activating enzyme, a SUMO E2 conjugating enzyme and SUMO E3 ligases¹⁰. SUMOylation is an energy consuming process that depends on ATP^{68,77,84,85}.

All eukaryotic SUMO proteins are initially synthesized in an immature form that cannot be conjugated to targets and need to be further processed by SUMO proteases^{86,87}. SUMO proteases remove some C-terminal residues to expose two glycine residues that can be subsequently conjugated to a lysine residue on the target protein. Mature SUMO is activated by forming a thioester bond with the catalytic cysteine in the E1 enzyme. Consequently, SUMO is relocated from E1 to the catalytic cysteine of a single E2 conjugating enzyme-UBC9. SUMO E3 ligases catalyse the final step efficiently by transferring SUMO to the acceptor lysine residues of target proteins. The isopeptide bond between SUMO and its target proteins can be cleaved by SUMO proteases. SUMOylation modifies a single or multiple lysines within target proteins (mono-SUMOylation or multi-SUMOylation) and SUMO can also modify itself on its lysines (poly-SUMOylation). SUMO proteases balance the free and conjugated SUMO pool and regulate transient, dynamic and reversible SUMOylation in the cell^{68,88}.

2.3 SUMO machinery

2.3.1 SUMO E1 activating enzymes

The SUMO activating enzyme (E1) was first identified in 1997 as activator of Smt3p in budding yeast to induce its conjugation to other proteins. The E1 is found in the form of a heterodimeric enzyme consisting of Aos1p (activation of Smt2p)/Uba2p (Ubiquitin Activating enzyme E1-like)^{69,89}. Later on, the SUMO-1 activating enzyme was purified from human cells and shown to be a heterodimer consisting of SAE1/SAE2 (SUMO activating enzyme), homologous to Aos1p/Uba2p. SAE1 was able to catalyse the thioester formation between SUMO-1 and SAE2, in an ATP-dependent manner⁹⁰.

2.3.2 SUMO E2 conjugating enzyme

In contrast to many E2 conjugating enzymes that regulate ubiquitination, in human cells, there is one unique SUMO E2 conjugating enzyme characterized – UBC9, that plays an essential role in SUMOylation^{89,91}.

Activated SUMOs can be transferred from a reactive cysteine residue in SAE2 to UBC9 on its active cysteine to form a SUMO-UBC9 thioester complex⁹². It is believed that UBC9 directly selects SUMO target proteins. Many of these targets carry a SUMOylation consensus site (ψ KxE, ψ : hydrophobic amino acid and x: any amino acid) to which UBC9 can bind directly, although with low affinity. In vitro, UBC9 binds to the SUMOylation consensus site and this is adequate to conjugate SUMO to this lysine residue of the substrate. However in vivo, this process might be facilitated by SUMO E3 ligases^{93,94}.

Mouse embryos deficient for UBC9 die at the early post-implantation stage⁹⁵. UBC9 deficient blastocysts are still viable and can be further cultured for at most two days. Without UBC9, inner cell mass (ICM) regression happens due to increased apoptosis, supporting the pivotal role of SUMOylation for the ICM development during the post-implantation phase of embryogenesis⁹⁵.

2.3.3 SUMO E3 ligases

SUMO E3 ligases catalytically transfer SUMOs from the E2 conjugating ligase UBC9 to lysine residues within target proteins and determine the efficiency of SUMOylation and specificity⁹⁶. SUMO E3 ligase activity has been demonstrated for several types of proteins. The most convincing and conserved SUMO E3 ligases are Siz/PIAS (SP) proteins. These SP E3 ligases contain only one catalytically C3HC4 domain that resembles the RING finger domain of ubiquitin E3 ligases and an adjacent SP C-terminus domain that does not coordinate a Zn^{2+} ion⁹⁷⁻⁹⁹. Siz1/2 was first characterized as an E3-like factor in the SUMO pathway by attachment of SUMO to septins in vivo in *S. cerevisiae*. Pli1 is another SP E3 ligase in fission yeast¹⁰⁰⁻¹⁰³.

In mammals, protein inhibitor of activated STAT PIASy (PIAS4) was first reported as a novel SUMO E3 ligase that remarkably stimulates LEF1 SUMOylation⁹⁷. In total, there are four PIAS family members, PIAS1, PIAS2 (PIASx), PIAS3 and PIAS4 (PIASy). They were identified as inhibitors of the

JAK-STAT signalling pathways. In addition, the Mms21 (the methyl methanesulphonate-sensitivity protein) (NSMCE2 in humans and Nse2 in yeast) also bears an SP-RING domain and contributes to SUMOylation of proteins such as Scc1¹⁰⁴.

The Ran binding protein 2 (RanBP2), which is a nuclear pore complex protein is regarded as another type of SUMO E3 ligase. It interacts with SUMOylated RanGAP1 and forms a stable complex that localized at kinetochores and the mitotic spindle. It can also tightly bind UBC9 and mediates SUMOylation of Topoisomerase 2 α (Top2 α) and is required for Top2 α localization at mitotic chromosomes¹⁰⁵⁻¹⁰⁷.

Furthermore, there are some unexplored SUMO E3 ligases such as ZNF451 family members and Polycomb group (PcG) Chromobox Protein Homolog 4 (CBX4) that show high specificity for SUMO, but their function needs to be further studied¹⁰⁸.

2.3.4 SUMO proteases

SUMO proteases precisely cleave between the SUMO C-terminal glycine and the SUMO substrate lysine or depolymerize poly-SUMO chains. Some SUMO proteases also process SUMO precursors by recognizing SUMO precursors and cleaving C-terminal residues to expose the di-glycine motif and therefore affect SUMO conjugation indirectly⁸⁷.

Ulp1 (UBL-specific protease 1) was the first identified SUMO protease in *S. cerevisiae*. Later on through comparing the catalytic domain sequence of Ulp1 to sequence databases, Ulp2 was also identified in *S. cerevisiae* together with many other putative SUMO proteases in other species¹⁰⁹⁻¹¹¹. In human cells, the first confirmed SUMO protease was sentrin-specific protease 1 (SEN1). Through extensive database searches, other putative human SENPs were identified. They are SENP1, 2, 3 and SENP5, 6, 7^{ref. 112}.

Further investigation showed that SUMO proteases possess substrate specificity. In yeast Ulp1 is not only responsible for SUMO precursor maturation, it is also involved in removing SUMOs from protein substrates. Ulp2 cleave isopeptide linkages between SUMO and substrates but it also possesses high activity to cleave SUMO-SUMO linkages in poly-SUMO chains^{110,113}. In human cells, although not well understood, SENP1 shows a preference for deconjugating SUMO-1, SENP2 could efficiently processes SUMO-2 and SUMO-3 precursors, SENP5 has significant activity to maturate SUMO-3 precursor but has no or limited activity towards SUMO-1. SENP6, 7 are not responsible for SUMO precursor maturation, but have a preference to process SUMO chains⁸⁷.

DES11 is another known SUMO protease that has been identified in yeast two-hybrid screening while searching for partners of BZEL. BZEL is a transcriptional repressor that is expressed in effector lymphocytes and binds to the promoter of target genes like blimp-1. DES11 has SUMO-1, 2, and 3 deconjugation activity from BZEL as well as deconjugation of poly-SUMO-2/3 chains^{86,114}.

USPL1 (Ubiquitin-specific protease-like 1) was identified as a protein that can cleave SUMO from its targets. Overexpression of USPL1 caused loss of SUMO-2/3 conjugates but not SUMO-1, which

identifies it as a SUMO protease specifically for SUMO-2/3. Knock down of USPL1 does not cause accumulation of SUMO conjugates, but does show impaired cell growth which can be attenuated by the non-catalytic function of USPL1^{ref. 115}.

3 Cellular roles of SUMOylation in DNA damage repair

Factors that cause alteration of the chemical structure of DNA can be called DNA damaging agents. DNA damage can be produced by exogenous agents like alkylating agents including methyl methanesulfonate (MMS), crosslinking agents such as mitomycin C, cisplatin and deoxyribonucleotide pool depleting agents including hydroxyurea (HU) or physical factors such as ionizing radiation and ultraviolet light¹¹⁶⁻¹¹⁸. Endogenous processes during DNA metabolism can also generate DNA damage, for example DNA base loss during hydrolysis caused by spontaneous DNA depurination, oxidized DNA bases caused by reactive oxygen species (ROS), a natural byproduct derived from normal metabolism of oxygen, dNTP mis-incorporation during DNA replication and replication fork stalling or collapse caused by DNA replication stress. All these endogenous processes let cells experience huge numbers of spontaneous DNA lesions per day¹¹⁹.

To counteract DNA damage, eukaryotic cells have developed well-coordinated defence mechanisms to recruit and activate specific factors in the right place throughout different cell cycle phases¹²⁰. Repair systems for specific types of lesion have been adopted (Figure 1)^{116,121}. DNA mismatch repair is employed to recognize and replace mis-paired DNA bases^{122,123}. Base excision repair is used to remove specific non-bulky chemical alterations of DNA bases¹²⁴. Nucleotide excision repair is introduced to cut a short single-stranded DNA segment that contains a lesion^{125,126}. With the assistance of protein networks involved in Fanconi Anemia protein clusters and many endonucleases, intrastrand crosslink repair pathway is used to excise intrastrand crosslinks, non-homologous end joining (NHEJ) as well as homologous recombination (HR) can work together in different cell phases to promote the repair of the most hazardous damage - DNA double strand breaks (DSBs)¹²⁷⁻¹²⁹.

3.1 DNA Replication stress derived damage and repair

The progress of replication forks is strongly impeded by different types of DNA damage such as DNA lesions, DNA-protein complexes, secondary DNA structures, an RNA-DNA hybrid and DNA-DNA crosslinks which are usually considered as side products produced during DNA replication progression. This kind of DNA damage can be repaired by fine-tuned reactions mentioned above. The response to replication stress derived DNA damage is an interesting area for further research because the instability of DNA replication is one of the hallmarks of cancer and confers genetic diversity during tumorigenesis.

Accurate and complete DNA replication is essential for accurate transmission of genetic information and the maintenance of genomic stability, it requires numerous factors including an ample pool of nucleotides (dNTPs), complete components of the replication machinery, adequate active replication origins, histones and histone chaperones^{127,130}.

As mentioned, faithful completion of DNA replication is under stress caused by spontaneous and exogenous insults¹³¹. A wide variety of obstacles can hamper DNA replication and lead to replication stress. These obstacles include DNA lesions, DNA-protein complexes and secondary DNA structures. Some of these obstacles occur during S phase and will consequently slow down or stop replication fork progression by impeding the capacity of replication polymerases and replication helicases. If those obstacles are not removed, mutations arising from DNA damage will be transmitted during mitosis and consequently can lead to developmental abnormalities and cancer¹³².

Once faced with obstacles that specifically associate with DNA replication, the progression of the DNA polymerase is affected and consequently uncoupled from DNA templates. However DNA replicative helicases sometimes continue to unwind the parental DNA and lead to the formation of single-strand DNA (ssDNA)^{133,134}. In addition, due to cyclin dependent kinase (CDK) controlled resection at S phase and G2 phase, single-stranded DNA (ssDNA) intermediates also form through exonuclease resection of damaged DNA structures such as DSB¹³⁵. The persistence of ssDNA is recognized by the cell cycle checkpoint machinery as a signal of DNA damage that needs to be repaired (Figure 4).

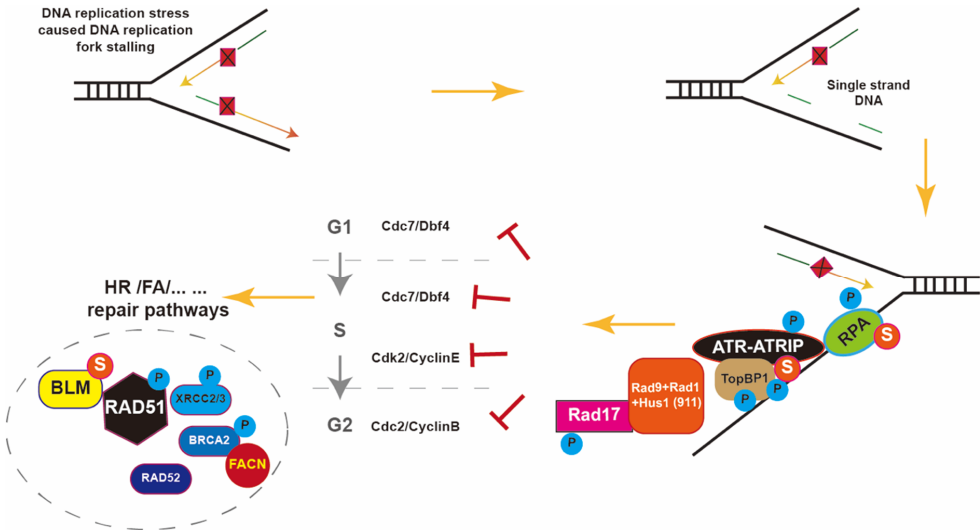


Figure 4. DNA replication fork stalling caused DNA damage and repair. General scheme depicting phosphorylated components that cells have selected to fine-tune regulatory networks during DNA replication fork stalling, to mediate cell-cycle arrest, to slowdown the replication progression and save time for proper DNA damage repair.

Prolonged ssDNA will be coated by replication protein A (RPA) consisting of RPA14, 32 and 70¹³⁶. RPA coated ssDNA not only prevent ssDNA from degradation or DNA secondary structures

formation but serves as a signaling platform to recruit and activate downstream response factors, such as the protein kinase ataxia-telangiectasia mutated and Rad3 related (ATR) and its interaction partner ATRIP (ATR interacting protein). ATR is a phosphatidylinositol 3 kinase like kinase which is defect in ataxia telangiectasia, it acts in many forms of DNA damage including stalled DNA replication forks^{134,137,138}.

Cells will employ a host of mechanisms to mediate cell-cycle arrest, to slowdown the replication progression, to prevent cells from going into mitosis and to save time to repair damaged DNA. Once damage has been repaired, arrested cells re-enter into cell-cycle progression, otherwise cells are stuck in a permanent cell-cycle arrest or undergo apoptosis.

3.1.1 Checkpoint signaling pathways

Checkpoints are cellular surveillance mechanisms. Upon DNA replication stalling, RPA coated ssDNA activates the check point kinase ATR. Activated ATR serves as a central replication stress response kinase to phosphorylate a large number of downstream targets such as the effector kinase CHK1 (checkpoint kinase 1), with the help of several replication checkpoint proteins such as RAD9, RAD17, TOPBP1, BRCA1 and claspin^{139,140}.

Activated CHK1 phosphorylates downstream cell cycle phosphatases, including CDC25-A, -B and -C. Phosphorylated CDC25-A triggers the signal for its degradation by the proteasome and results in inactivated cyclin E/A-CDK2 and cyclin B-CDK1 and arrest of cell cycle progression. Chk1 activation in response to DNA replication fork stalling is thought to be sufficient for Cdc25A degradation, to suppress late origin firing, to save time for the DNA damage repair and consequently to ensure the completion of DNA replication under stress or induce apoptosis¹⁴¹⁻¹⁴³.

Although the ATM checkpoint kinase (ataxia-telangiectasia mutated) is regarded as less imperative for the response to replication blocks than ATR, it can still slow DNA replication progression in response to DNA damage caused DNA replication fork stalling¹⁴⁴. The MRN (MRE11 (meiotic recombination 11)–RAD50–NBS1) complex together with 53BP1, MDC1 and other mediator proteins could sufficiently activate ATM. Activated ATM phosphorylates its effector kinase CHK2. CHK2 can also phosphorylates serine 123 in Cdc25A. Phosphorylation of Cdc25A serves as a signal for the ubiquitin machinery and targets Cdc25A for degradation, inactivating cyclin A-CDK2 and inhibiting further firing of early origins of replication during S phase^{145,146}.

3.1.2 Action of the ATR pathway

ATR plays a vital role in protecting replication forks from stalling. ATR activation needs interactions between the ATR/ATRIP complex and other proteins that contain AADs (ATR-activating domains).

So far in humans, TopBP1 is the only AAD containing protein that is found to be responsible for ATR/ATRIP activation¹⁴⁷. The 9-1-1 (Rad9-Rad1-Hus1) complex is independently recruited to ssDNA. TopBP1 binds the C-term of Rad9 and RHINO and stimulates ATR kinase activity through its AAD.

Additionally, Rad17, RAD9 and the MRN (MRE11-RAD50-NBS1) complex also act to recruit TopBP1 to RPA-ssDNA^{148,149}.

Upon ATR auto-phosphorylation and activation, it phosphorylates a variety of substrates. One substrate of ATR that is being phosphorylated upon DNA damage is SMARCAL1. With the help of SLX4 associated nucleases, phosphorylated SMARCAL1 prevents aberrant fork structures to result in double strand break formation. Other ATR substrates include Chk1, FANCI and Polη, RPA, MCMs, WRN and some others¹⁵⁰⁻¹⁵².

ATR phosphorylates Chk1 at S317 and S345. Chk1 not only controls cell cycle arrest as described before but also regulates replication fork progression and replication related DNA damage. ATR phosphorylation of FANCI plays a vital role in the FANCD2-FANCI complex location to the DNA damage sites and restarting DNA replication forks at intrastrand cross-links¹⁵³. Polη is also an important ATR substrate that plays its role in translesion synthesis repair¹⁵⁴.

3.1.3 Homologous Recombination repair during DNA replication.

HR (Homologous Recombination) not only mediates genetic recombination in meiosis, HR can also repair DNA double strand breaks, interstrand crosslinks and stalled or damaged replication forks¹⁵⁵. Its activity is strictly controlled by CDKs during S and G2 phase of the cell cycle and uses sister chromatids in the genome produced by DNA replication as repair templates.

Homologous recombination is initiated by the binding of the MRN (MRE11-RAD50-NBS1) complex in humans in the case of DNA double strand breaks, subsequently initiating the nucleolytic processing at the DNA break site, with the help of other endonuclease such as BLM, human EXO1 (exonuclease 1) resects DNA ends and generates ssDNA^{156,157}.

Once DNA ends are resected, RPA binds ssDNA efficiently to generated RPA coated ssDNA. This melted DNA secondary structure facilitates the loading of recombinase RAD51, which is mediated by a large number of RAD51 interacting proteins like XRCC2, XRCC3, BRCA2 and RAD52 (Figure 4)^{155,158}.

BRCA2 interacts with the FA protein FANCD1 and promote RAD51-mediated D-loop formation and the second DNA end is used to generate a Holliday junction. During the last step of DNA repair synthesis, the remaining HJ structures are dissolved by a complex of BLM (Bloom syndrome, RecQ helicase-like), TOPIIIa and hRMI1 to form a non-crossover^{155,159}.

3.2 Cellular roles of SUMO in DNA damage response pathways.

SUMOylation regulates almost all kinds of fundamental nuclear functions, ranging from protein degradation, DNA damage response, DNA replication, transcription, cell cycle checkpoint control, signal transduction to chromatin organization, ribosome biogenesis, nuclear trafficking, and pre-mRNA splicing. This thesis focusses on the role of SUMO in DNA damage response pathways.

3.2.1 SUMOylation of RPA

Replication protein A (RPA) is a protein that binds to single-stranded DNA (ssDNA). It was shown to be regulated by SUMO in DNA damage response pathways both in yeast and in human cells. In yeast through SUMO pull down under denaturing conditions, SUMOylation of Rfa1 (RPA1 homolog) can be detected upon treatment with the alkylating agent methyl-methanesulfonate (MMS)⁷⁸.

In human cells, RPA consists of three subunits, RPA1, 2, 3. The largest subunit RPA1 was shown to stably associate with SENP6. After Hydroxyurea treatment or UV irradiation, which efficiently caused replication fork stoppage, RPA1 did not dissociate from SENP6¹⁶⁰. However, SENP6 depletion caused the accumulation of SUMOylated RPA and SUMOylation of RPA increased its interaction with RAD51, promoted RAD51 foci formation and promoted homologous repair. Further study revealed that there are two SUMO sites within RPA, lysine 449 and lysine 577. Lysine 449 was modified by a poly-SUMO chain and lysine 577 was only mono-SUMOylated¹⁶¹.

3.2.2 SUMOylation of ATRIP

Ataxia telangiectasia and Rad3 related (ATR) - interacting protein (ATRIP) is a protein involved in checkpoint signalling after DNA damage. ATRIP is SUMOylated at lysine 234 and lysine 289¹⁶². In UBC9 depleted cells, ATRIP SUMOylation was drastically reduced and failed to be recruited to the DNA damage site. Furthermore, SUMOylation of ATRIP serves as a glue for the interaction of multiple functional partners such as ATR, TopBP1, RPA and MRN in the ATR pathway and it is indispensable for activation of ATR signalling. Moreover, it is reported that the maintenance of ATR basal kinase activity upon DNA damage requires intact activity of the SUMO E3 ligase PIAS3¹⁶³.

3.2.3 SUMOylation of BLM

BLM is an ATP-dependent helicase that acts on DNA replication stress to ensure homologous recombination^{164,165}. It belongs to the RecQ family. It is reported that two SUMO sites exist in BLM which are lysine 317 and lysine 331. SUMOylated BLM can bind RAD51 and regulates the recruitment of RAD51 at stalled replication forks. BLM can also regulate the accumulation of RPA at stalled replication forks and limits the generation of ssDNA¹⁶⁶.

3.2.4 SUMOylation of SLX4

The SLX4 Fanconi anemia protein is a tumor suppressor and acts as a scaffold for nucleases, it plays an important role in the maintenance of genomic stability¹⁶⁷. The increased SUMOylation of SLX4 can be visualized during S and G2 phases and is decreased when the cell cycle is completed. It contains three SIMs that are needed for both SLX4 SUMOylation and its binding to SUMO-2. The SIM domains were also required for the localization of SLX4 to PML nuclear bodies and for proper interstrand DNA crosslink repair. Furthermore, SLX4 has been proposed as an E3 ligase that can SUMOylate itself and the XPF-ERCC1 endonuclease^{168,169}.

3.2.5 SUMOylation of BRCA1 and 53BP1

BRCA1 is a well-known ubiquitin E3 ligase and breast cancer tumor suppressor. It plays a vital role in the maintenances of genomic stability. It is poorly accumulated at the DNA damage foci in the absence of the SUMO E3 ligases PIAS1 and PIAS4, which means that BRCA1 is a SUMO regulated ubiquitin ligase. Moreover, BRCA1 is SUMOylated by PIAS1 and PIAS4 by both SUMO-1 and SUMO-2/3 in a DNA damage dependent way¹⁷⁰.

53BP1 is a DNA damage checkpoint protein that is SUMOylated upon DNA damage. 53BP1 co-localized with SUMO-1 at nuclear foci after four hours of ionizing radiation treatment. 53BP1 depletion only impaired SUMO-1 accumulation in DNA damaged laser tracks. Both PIAS4 and UBC9 depletion impaired 53BP1 accumulation at DNA damage sites⁹⁸.

3.2.6 SUMOylation of MDC1

Mediator of DNA damage checkpoint 1 (MDC1) is a central protein involved in the regulation of checkpoint activation and subsequent DNA repair following DNA damage. It mediates cell cycle arrest in response to DNA damage during S phase and G2/M phases of the cell cycle. Once MDC1 is recruited to DNA damage sites, it is SUMOylated by SUMO-1 or SUMO-2/3 at lysine 1840. PIAS4 is the major E3 ligase for MDC1 SUMOylation. To facilitate the DNA damage response, SUMOylated MDC1 can be further ubiquitinated and directed to 26S proteasome by the STUbL RNF4¹⁷¹.

4 Cross talk among PTMs with focus on SUMO

In addition to modification by SUMO, other post translational modifications such as phosphorylation and ubiquitination can also collaborate and influence each other to fully control protein activity (Figure 5).



Figure 5. Crosstalk between post-translational modification (PTM). PTMs are able to regulated the function of targets not only by individual PTM types and sites but also by cooperatively action.

4.1 Cross talk between SUMOylation and Phosphorylation

More than ten years ago, a highly conserved PDSM (phosphorylation dependent SUMOylation motif) was discovered. The PDSM contains a SUMO consensus site ψ KxE and a downstream proline directed phosphorylation site (ψ KxE ψ SP), which regulates the SUMOylation of a substrate. Within this consensus motif, phosphorylation close to SUMOylation sites could positively regulate SUMOylation of several substrates¹⁷². Of the 46 human proteins that were found containing a PDSM, 71% are transcriptional regulators, such as heat-shock factors (HSFs) and the estrogen-related receptor nuclear receptors.

Later on, the negatively charged amino acid - dependent SUMOylation motif (NDSM) was also proposed to promote substrate SUMOylation. The acidic residues adjacent to the core SUMO motif and the negative charge character of the downstream amino acid residues are required to maintain SUMOylation levels of several target proteins¹⁷³.

4.2 Cross talk between SUMOylation and Ubiquitination

Published studies indicated that SUMO and ubiquitin can collaborate or counteract each other. Mixed SUMO and ubiquitin chains have been identified^{84,174-176}. Recently, a combined immune-affinity enrichment approach was used to determine the crosstalk between SUMOylation and ubiquitination^{177,178}.

4.2.1 STUBLs

Research on SUMO-targeted ubiquitin ligases (STUbLs) highlights connections between SUMOylation and ubiquitination. STUbLs are a subset of ubiquitin E3 ligases that contain SUMO interaction motifs (SIMs) to interact with SUMO, and a RING finger domain to catalyse direct transfer of ubiquitin from the E2 conjugating enzyme to the substrates (Figure 6). STUbLs recognize and specifically ubiquitinate SUMO conjugates¹⁷⁹⁻¹⁸¹.

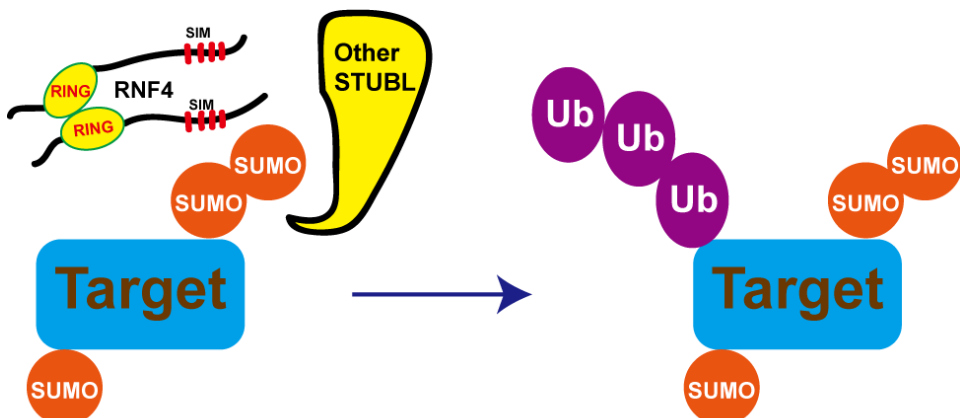


Figure 6. SUMO-targeted ubiquitin ligases (STUbLs). STUbLs recognize and ubiquitinate SUMOylated proteins to regulate their functions and therefore link SUMO modification to the ubiquitin/proteasome system.

4.2.2 SUMO-interacting motifs (SIMs)

SUMO-interacting motifs (SIMs) were first identified in many SUMO-1 interacting proteins in yeast through yeast two-hybrid screening^{182,183}. Later on, this motif has been found in a wide range of proteins, including SUMO-binding proteins, SUMO substrates, SUMO enzymes and STUbLs. SIMs are generally characterized as a motif composed of hydrophobic amino acids ((V/I)X(V/I)(V/I)) and flanking acidic residues. This motif could bind SUMOs and mediate non-covalent interactions between SUMOs and SIM-containing proteins¹⁸⁴.

4.2.3 STUbLs in yeast

The first STUbL identified in yeast via two hybrid interaction screening is Uls1. Uls1 contains four predicted SIMs in its N-terminus and a RING finger domain in its C-terminus. Mutant strains lacking Uls1 show accumulation of SUMO conjugates but efficient ubiquitination of SUMOylated conjugates by Uls1 has not been reported yet^{185,186}. Uls2 was identified as a second STUbL in yeast via two hybrid interaction with SUMO. It is a heterodimer consisting of two RING finger proteins Slx5 and Slx8. Mutations in either of these proteins were lethal and affected the levels of SUMO conjugates. Genetic data showed the connection of Uls1 and Uls2 with DNA damage repair and genome stability¹⁸⁷.

In budding yeast, Rad18 was also identified as a STUbL that can ubiquitinate SUMOylated proliferating cell nuclear antigen (PCNA)¹⁸⁸. PCNA works as a sliding clamp on DNA and attracts and tethers replicative polymerases during DNA replication. It also plays its role in the DNA damage response. Rad18 contains only one SIM with a specificity for mono-SUMOylated conjugates, which is absent from human Rad18.

4.2.4 STUbLs in human

4.2.4.1 RNF4 and RNF4 substrates

The human RNF4 protein was the first identified STUbL in mammalian cells. It resides predominantly in the nucleus and contains only 190 amino acids with four potential SIMs, among which the C-terminal three are functional SIMs that can recognize poly SUMO chains (Figure 6)¹⁸⁹. In the presence of SUMO chains, RNF4 is activated through dimerization. Activated RNF4 binds the UBC4/5 family of E2 enzymes and directs the ubiquitination machinery to SUMOylated proteins and thereby promotes the ubiquitination of SUMO conjugates. *In vitro* experiments showed that RNF4 can also cooperate with the ubiquitin E2 UBC13-UEV1 and synthesize K63 linked chains on its substrate. Furthermore, RNF4 has the ability to rescue yeast cells deficient for Slx5/8, demonstrating the functional conservation in STUbLs among eukaryotes¹⁹⁰.

The promyelocytic leukemia protein (PML) together with its oncogenic fusion product PML-RAR α were identified as the first substrates of RNF4¹⁹¹. In the presence of arsenic trioxide, PML is degraded in a SUMO-dependent manner by the 26S proteasome. RNF4 was shown to promote K48 linked ubiquitin chain formation on PML, leading to the recognition and degradation of PML by the 26S proteasome. In cell culture RNF4 was also reported to disrupt PML nuclear bodies (PML-NBs) in cells treated with arsenic trioxide¹⁹².

RNF4 was furthermore proposed to target SUMOylated RPA for proteasomal degradation based on the fact that in RNF4 depleted cells, RPA persists at the DNA damage lesion¹⁹³. It was proposed that RNF4 mediates RPA degradation to promote the exchange of RPA for RAD51 on ssDNA. However, direct ubiquitination of SUMOylated RPA1 by RNF4 needs to be demonstrated¹⁹⁴.

Kinetochores protein CENP-I has also been reported as a target for RNF4. In SENP6 depleted cells, the SUMOylated levels of CENP-I were increased when RNF4 was co-depleted. In RNF4 depleted cells as well as in cells where 26S proteasome activity was blocked, SUMOylated CENP-I also accumulated¹⁹⁵.

The tumor suppressor proteins FANCI and FANCD2 (ID complex) are the central components of the Fanconi Anemia (FA) pathway. They are mono-ubiquitinated to enable nuclease recruitment and this mono-ubiquitination will promote their loading onto the chromatin. PIAS1 and PIAS4 can SUMOylate the chromatin-loaded ID complex which is then recognized and polyubiquitinated by RNF4¹⁹⁶.

MDC1 and BRCA1 have also been identified as SUMOylated RNF4 targets relevant for genome stability¹⁹⁴. SUMOylation of MDC1 and BRCA1 were increased upon exposure of cells to ionizing radiation and knocking down RNF4 increased the amount of SUMOylated MDC1 and BRCA1¹⁹⁷. Additionally, RNF4 regulates the degradation of the histone demethylase JARID1B/KDM5B in response to MMS to mediate transcriptional repression¹⁹⁸.

Although there is an increase in RNF4 targets, we are still restricted in our understanding of the role of RNF4 because of limited insight into the RNF4-regulated SUMO target proteins. The task to uncover a comprehensive network of RNF4 regulated SUMO targets is therefore evident.

4.2.4.2 RNF111 and RNF111 substrates

A second mammalian STUbL that has been identified is RNF111/Arkadia¹⁹⁹. RNF111 contains three adjacent SIMs and shows a similar activity of RNF4 to ubiquitinate PML upon arsenic trioxide treatment. RNF111 and RNF4 cannot form heterodimers and so far they were only known to work independently on PML.

RNF111/Arkadia uses UBC13-MMS2 as its E2 conjugating enzyme and promotes non-proteolytic, K63-linked poly-ubiquitin chains on SUMOylated xeroderma pigmentosum group C (XPC), which is the generic initiator of global genomic nucleotide excision repair (GG-NER)²⁰⁰.

Ubiquitination by Arkadia regulates the recruitment of XPC and locates XPC to UV induced DNA damage sites, highlighting a fundamental non-proteolytic function of a STUbL, coupling ubiquitination and SUMOylation in response to DNA damage²⁰¹.

5 Technical developments in SUMO proteomics

5.1 SUMO proteomic challenges

The process of SUMOylation is essential in nearly all eukaryotic cells, it has been involved in regulating many cellular functions such as DNA damage repair and cell cycle. In order to understand SUMOylation, it is important to decipher networks of SUMO targets in the cell and identify their SUMO acceptor sites. That is why over the last few years, mass spectrometry analyses of SUMO proteomes has attracted much attention and has been advanced greatly^{81-83,202-207}. However, due to PTMs including SUMOylation, proteomes become highly complex, which means that sensitive mass spectrometers are required for their analysis.

Moreover, SUMO proteomics analysis is challenging for many reasons. First of all, because of the low stoichiometry of SUMOylation and low abundance of SUMOylated proteins. Although, several SUMOylated proteins such as PML and RanGAP1 are quite stable in their SUMOylated form and the steady state SUMOylation level is large in relation to the total level of protein^{105,191,192}. However for the majority of other proteins, only a small proportion of them is SUMOylated and thus their SUMOylated forms are hard to detected⁷⁷.

A second challenge is the high activity of SUMO proteases that can cleave SUMO from its target proteins^{208,209}. In the cell, SUMO proteases are well controlled. These proteases are quite active in many standard buffers used to lyse cells. Furthermore, there are no specific and effective inhibitors for those SUMO proteases available yet. As a result, when processing cells or tissues under these conditions, SUMOylated proteins are largely de-SUMOylated by active SUMO proteases.

Thirdly, a long C-terminal remnant of SUMO after trypsin digestion remains. In contrast to ubiquitin, which after trypsin digestion leaves only a di-glycine remnant, the SUMO-2/3 remnant is 32 amino acids long and cannot be efficiently analysed by current mass spectrometry^{73,210,211}.

To address these issues, several approaches have been established and optimized within the last decade to inactivate SUMO proteases and to enrich SUMOylated proteins from cells or tissue extracts (Figure 7). These cells were grown under normal conditions as well as treated with DNA damaging agents or treated otherwise to study SUMOylation dynamics. At the same time, mass spectrometers have improved significantly during recent years. By combining these two factors, this has led to the identification of thousands of SUMO sites within thousands of SUMO conjugates^{73,80,82,198,205,211,212}.

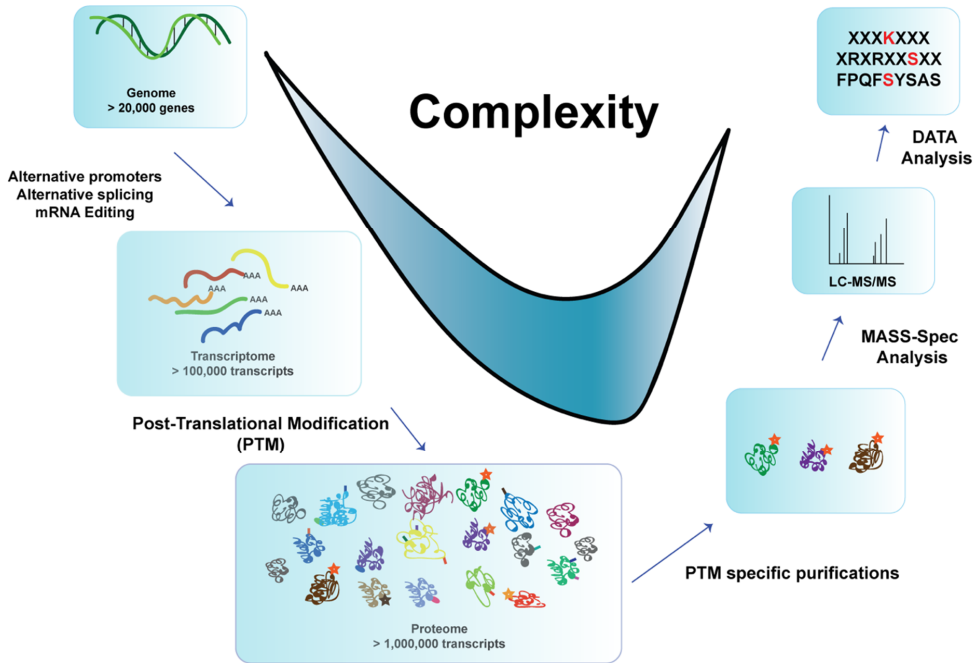


Figure 7. PTM specific purification and identification. Proteomes are much more complex than genomes and transcriptomes due to PTMs. PTM specific approaches have been established and optimized within the last decade to enable their analysis.

5.2 Purifying endogenous SUMO target proteins

The first affinity purification method for the isolation of endogenous poly SUMO conjugates was developed using an RNF4 fragment containing four SIMs. 339 putative endogenous poly SUMO conjugates were identified from HeLa cells by this method after heat shock. The method is not efficient for purification of mono-SUMOylated proteins and has a relatively high background due to the non-specific binding of the SIMs²¹³.

To identify a number of endogenous SUMOylated proteins from more complex organs and tissues, monoclonal antibodies to SUMO-1 and SUMO-2 have been used to highly enrich endogenous SUMO conjugates from HeLa cells and from mouse liver. 584 endogenous SUMO target protein candidates were identified in this manner. However, this method requires very large amounts of starting material and very large amounts of antibodies, which serves as major disadvantages of this method²¹⁰.

5.3 Purifying SUMO target proteins using tagged SUMO

To counteract these disadvantages and to get more efficient purification of SUMO targets, SUMO affinity purification was developed by exogenously expressing N-terminally tagged SUMO family

members to purify epitope-tagged SUMO conjugates. These tags including his6, HA, Myc, His6-FLAG and so on²⁰⁵. Some tags such as histidine or biotin work well under denaturing conditions where proteases are largely blocked. Other epitope tags such as HA and FLAG can be recognized by specific antibodies with relatively higher affinity compared with SUMO antibodies and work well in more stringent buffers. However, exogenously overexpressing N-terminally tagged SUMO usually increases the number of false positives and can only be performed in cultured cells. Thus, SUMOylated candidates characterized by this methodology should be strictly verified through testing SUMOylation of endogenous proteins identified^{79,80,205,211 214}.

5.4 Purifying and identifying SUMO acceptor lysines

Even though plentiful SUMO acceptor lysines in target proteins reside within the consensus motif, bioinformatics analysis of protein sequence is far from sufficient to identify SUMO sites²¹⁵. This consensus motif also occurs without evidence for SUMOylation and SUMOylation occurs on non-consensus sequences. Thus, there is a demand for direct identification and characterization of SUMO acceptor lysines by proteomics approaches in SUMO targets purified from cells. However, identification of SUMO acceptor lysines is even more challenging because of the relative low stoichiometry of SUMOylation and the inefficient identification of the large proteolytic C-terminal remnants of SUMOs by mass spectrometry. To overcome these difficulties, epitope tagged SUMOs have been used with an additional proteolytic cleavage site introduced. These SUMO mutants were exogenously expressed in cells. Using this strategy, thousands of SUMO sites have been identified so far. Endogenous identification of SUMO sites is still very challenging^{82,203,204,219}.

5.5 Proteomic techniques that were used in this thesis

5.5.1 Protein level and Site specific level

In order to study SUMOylation at the protein level, a method has been developed by Ivo Hendriks in our group, employing His10 tagged SUMO-2 (Figure 8)⁷³. A major advantage of the His10 tag over the His6 tag is that a higher concentration of competing imidazole can be introduced during the purification procedure as well as rigorous washing procedures to reduce the binding of contaminating proteins.

To be able to identify SUMO conjugates as well as SUMO acceptor lysines, U2OS cell lines that stably expresses His10-tagged SUMO-2 or lysine-deficient His10-SUMO-2 with a conserved mutation at its C-terminal part, Q87R were generated. SUMO-2 conjugates were enriched by Ni-NTA beads. A label free quantitative proteomics approach was used to study SUMOylated proteins in a system-wide manner as well as at the site-specific level.

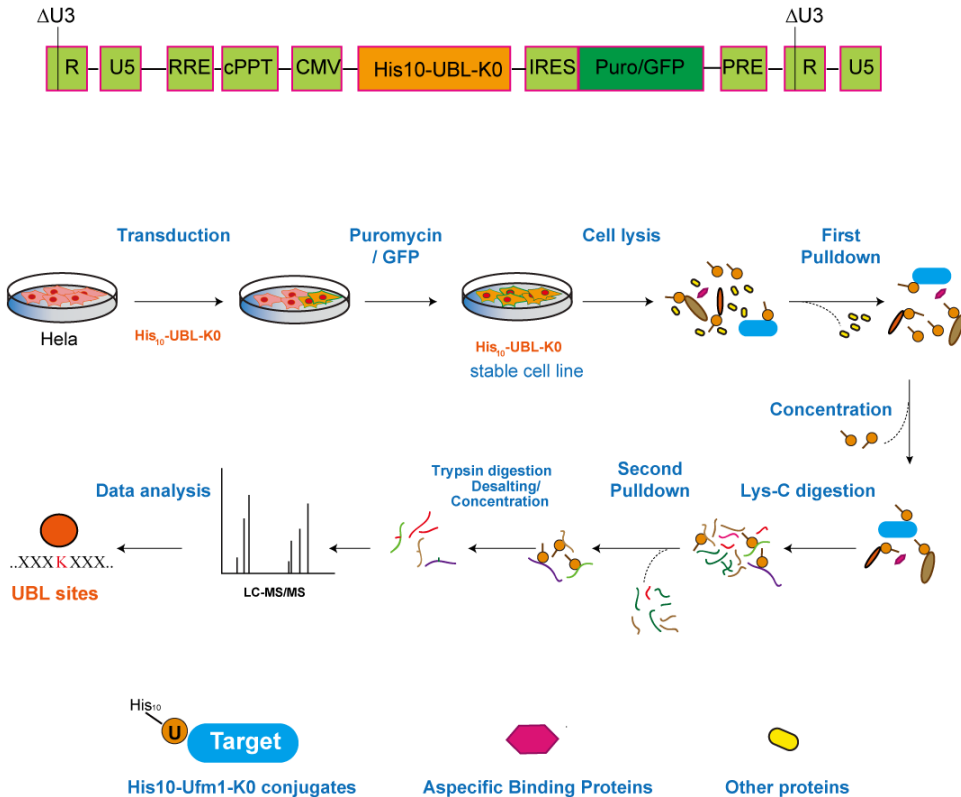


Figure 8. Strategy for identifying UBL modified proteins and acceptor lysines. Cartoon depicting the strategy to identify UBL targets as well as acceptor lysines developed by Alfred Vertegaal lab.

5.5.2 Label-Free Quantitation method

Label-free quantification is a robust strategy in quantitative proteomics that aims to determine the relative amount of proteins in more than one biological sample. Compared with stable isotope labelling, label-free quantification circumvents the use of a stable isotope labeling. Yet, this strategy also has its limitations such as it is relatively high sensitivity to variations in preparing samples and sensitivity to differences in performance of mass spectrometers^{202,214}.

6 Aim of the thesis

In this thesis, I aimed at decoding the role of SUMO in the DNA damage repair pathway. The SUMO system is believed to be involved in this process at several levels^{159,162,168,171,193,196,198}. I focused on the most inevitable DNA obstacle causing DNA replication stress, and the cellular roles of SUMOylation in regulating the recruitment of DNA repair factors to sites of DNA replication damage to influence the choice of employed cellular repair mechanisms for efficient repair.

Post-translational modifications are essential regulators of proteins²¹⁶. PTMs do not only play their roles solo but extensively interact with each other^{217,218}. Our knowledge about proteins modified by a combination of SUMO and ubiquitin, SUMO and phosphate and crosstalk between them is quite limited. In this thesis we also aimed at deciphering the crosstalk between SUMOylation and ubiquitination during the DNA damage response and searching for indirect and direct targets for the human STUbL RNF4, which mediates the ubiquitination of SUMOylated target proteins.

Lastly, we adopted the strategy described for SUMO and introduced His10-tagged UFM-1-K0 to identify UFM-1 acceptor lysines. UFMylation has been found to promote the interactions between proteins⁵⁸. However, limited targets have been found to be regulated by UFM-1^{15,63,65}. We identified and confirmed RPL26 as a key UFM1 target and further confirmed that the UFMylated form of RPL26 can efficiently interact with the Signal Recognition Particle Receptor, implicating that UFMylation could regulate protein transfer to the Endoplasmic Reticulum.

References

- 1 Watson, J. D. & Crick, F. H. Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. 1953. *Rev Invest Clin* **55**, 108-109 (2003).
- 2 Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921, doi:10.1038/35057062 (2001).
- 3 Hoeijmakers, J. H. Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366-374, doi:10.1038/35077232 (2001).
- 4 Wang, H., Zhang, X., Teng, L. & Legerski, R. J. DNA damage checkpoint recovery and cancer development. *Exp Cell Res* **334**, 350-358, doi:10.1016/j.yexcr.2015.03.011 (2015).
- 5 Liakos, A., Lavigne, M. D. & Foustier, M. Nucleotide Excision Repair: From Neurodegeneration to Cancer. *Adv Exp Med Biol* **1007**, 17-39, doi:10.1007/978-3-319-60733-7_2 (2017).
- 6 Zhang, R. *et al.* Genetic variants in nucleotide excision repair pathway predict survival of esophageal squamous cell cancer patients receiving platinum-based chemotherapy. *Mol Carcinog*, doi:10.1002/mc.22877 (2018).
- 7 Chang, H. H. Y., Pannunzio, N. R., Adachi, N. & Lieber, M. R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat Rev Mol Cell Biol* **18**, 495-506, doi:10.1038/nrm.2017.48 (2017).
- 8 Polo, S. E. & Jackson, S. P. Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications. *Genes Dev* **25**, 409-433, doi:10.1101/gad.2021311 (2011).
- 9 Walsh, G. & Jefferis, R. Post-translational modifications in the context of therapeutic proteins. *Nat Biotechnol* **24**, 1241-1252, doi:10.1038/nbt1252 (2006).
- 10 Vertegaal, A. C. Uncovering ubiquitin and ubiquitin-like signaling networks. *Chem Rev* **111**, 7923-7940, doi:10.1021/cr200187e (2011).
- 11 Deribe, Y. L., Pawson, T. & Dikic, I. Post-translational modifications in signal integration. *Nat Struct Mol Biol* **17**, 666-672, doi:10.1038/nsmb.1842 (2010).
- 12 Han, Z. J., Feng, Y. H., Gu, B. H., Li, Y. M. & Chen, H. The post-translational modification, SUMOylation, and cancer (Review). *Int J Oncol* **52**, 1081-1094, doi:10.3892/ijo.2018.4280 (2018).
- 13 Ohtsubo, K. & Marth, J. D. Glycosylation in cellular mechanisms of health and disease. *Cell* **126**, 855-867, doi:10.1016/j.cell.2006.08.019 (2006).
- 14 Kerscher, O., Felberbaum, R. & Hochstrasser, M. Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu Rev Cell Dev Biol* **22**, 159-180, doi:10.1146/annurev.cellbio.22.010605.093503 (2006).
- 15 Yoo, H. M. *et al.* Modification of ASC1 by UFM1 is crucial for ERalpha transactivation and breast cancer development. *Mol Cell* **56**, 261-274, doi:10.1016/j.molcel.2014.08.007 (2014).
- 16 Khalailieh, A. *et al.* Phosphorylation of ribosomal protein S6 attenuates DNA damage and tumor suppression during development of pancreatic cancer. *Cancer Res* **73**, 1811-1820, doi:10.1158/0008-5472.CAN-12-2014 (2013).
- 17 Yoo, H. M., Park, J. H., Jeon, Y. J. & Chung, C. H. Ubiquitin-fold modifier 1 acts as a positive regulator of breast cancer. *Front Endocrinol (Lausanne)* **6**, 36, doi:10.3389/fendo.2015.00036 (2015).
- 18 Chen, Z. & Lu, W. Roles of ubiquitination and SUMOylation on prostate cancer: mechanisms and clinical implications. *Int J Mol Sci* **16**, 4560-4580, doi:10.3390/ijms16034560 (2015).
- 19 Singh, V. *et al.* Phosphorylation: Implications in Cancer. *Protein J* **36**, 1-6, doi:10.1007/s10930-017-9696-z (2017).
- 20 Paska, A. V. & Hudler, P. Aberrant methylation patterns in cancer: a clinical view. *Biochem Med (Zagreb)* **25**, 161-176, doi:10.11613/BM.2015.017 (2015).
- 21 Cohen, P. The origins of protein phosphorylation. *Nat Cell Biol* **4**, E127-130, doi:10.1038/ncb0502-e127 (2002).
- 22 Vlastaridis, P. *et al.* Estimating the total number of phosphoproteins and phosphorylation sites in eukaryotic proteomes. *Gigascience* **6**, 1-11, doi:10.1093/gigascience/giw015 (2017).
- 23 Deutscher, J. & Saier, M. H., Jr. Ser/Thr/Tyr protein phosphorylation in bacteria - for long time neglected, now well established. *J Mol Microbiol Biotechnol* **9**, 125-131, doi:10.1159/000089641 (2005).
- 24 Ciesla, J., Fraczyk, T. & Rode, W. Phosphorylation of basic amino acid residues in proteins: important but easily missed. *Acta Biochim Pol* **58**, 137-148 (2011).
- 25 Rudner, A. D. & Murray, A. W. Phosphorylation by Cdc28 activates the Cdc20-dependent activity of the anaphase-promoting complex. *J Cell Biol* **149**, 1377-1390 (2000).
- 26 Hayakawa, F. & Privalsky, M. L. Phosphorylation of PML by mitogen-activated protein kinases plays a key role in arsenic trioxide-mediated apoptosis. *Cancer Cell* **5**, 389-401 (2004).
- 27 Goldstein, G. *et al.* Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proc Natl Acad Sci U S A* **72**, 11-15 (1975).

- 28 Pickart, C. M. & Eddins, M. J. Ubiquitin: structures, functions, mechanisms. *Biochim Biophys Acta* **1695**, 55-72, doi:10.1016/j.bbamcr.2004.09.019 (2004).
- 29 McDowell, G. S. & Philpott, A. Non-canonical ubiquitylation: mechanisms and consequences. *Int J Biochem Cell Biol* **45**, 1833-1842, doi:10.1016/j.biocel.2013.05.026 (2013).
- 30 Breitschopf, K., Bengal, E., Ziv, T., Admon, A. & Ciechanover, A. A novel site for ubiquitination: the N-terminal residue, and not internal lysines of MyoD, is essential for conjugation and degradation of the protein. *EMBO J* **17**, 5964-5973, doi:10.1093/emboj/17.20.5964 (1998).
- 31 Bloom, J., Amador, V., Bartolini, F., DeMartino, G. & Pagano, M. Proteasome-mediated degradation of p21 via N-terminal ubiquitynylation. *Cell* **115**, 71-82 (2003).
- 32 Coulombe, P., Rodier, G., Bonneil, E., Thibault, P. & Meloche, S. N-Terminal ubiquitination of extracellular signal-regulated kinase 3 and p21 directs their degradation by the proteasome. *Mol Cell Biol* **24**, 6140-6150, doi:10.1128/MCB.24.14.6140-6150.2004 (2004).
- 33 Dikic, I. & Robertson, M. Ubiquitin ligases and beyond. *BMC Biol* **10**, 22, doi:10.1186/1741-7007-10-22 (2012).
- 34 Vosper, J. M. *et al.* Ubiquitylation on canonical and non-canonical sites targets the transcription factor neurogenin for ubiquitin-mediated proteolysis. *J Biol Chem* **284**, 15458-15468, doi:10.1074/jbc.M809366200 (2009).
- 35 Tsuchida, S., Satoh, M., Takiwaki, M. & Nomura, F. Ubiquitination in Periodontal Disease: A Review. *Int J Mol Sci* **18**, doi:10.3390/ijms18071476 (2017).
- 36 Pickart, C. M. Mechanisms underlying ubiquitination. *Annu Rev Biochem* **70**, 503-533, doi:10.1146/annurev.biochem.70.1.503 (2001).
- 37 Metzger, M. B., Hristova, V. A. & Weissman, A. M. HECT and RING finger families of E3 ubiquitin ligases at a glance. *J Cell Sci* **125**, 531-537, doi:10.1242/jcs.091777 (2012).
- 38 Komander, D. The emerging complexity of protein ubiquitination. *Biochem Soc Trans* **37**, 937-953, doi:10.1042/BST0370937 (2009).
- 39 Ikeda, F. & Dikic, I. Atypical ubiquitin chains: new molecular signals. 'Protein Modifications: Beyond the Usual Suspects' review series. *EMBO Rep* **9**, 536-542, doi:10.1038/embor.2008.93 (2008).
- 40 Rittinger, K. & Ikeda, F. Linear ubiquitin chains: enzymes, mechanisms and biology. *Open Biol* **7**, doi:10.1098/rsob.170026 (2017).
- 41 Klein, T. *et al.* The paracaspase MALT1 cleaves HOIL1 reducing linear ubiquitination by LUBAC to dampen lymphocyte NF-kappaB signalling. *Nat Commun* **6**, 8777, doi:10.1038/ncomms9777 (2015).
- 42 Hochstrasser, M. Origin and function of ubiquitin-like proteins. *Nature* **458**, 422-429, doi:10.1038/nature07958 (2009).
- 43 McNally, T. *et al.* Structural analysis of UBL5, a novel ubiquitin-like modifier. *Protein Sci* **12**, 1562-1566, doi:10.1110/ps.0382803 (2003).
- 44 Kamitani, T., Kito, K., Nguyen, H. P. & Yeh, E. T. Characterization of NEDD8, a developmentally down-regulated ubiquitin-like protein. *J Biol Chem* **272**, 28557-28562 (1997).
- 45 Merbl, Y., Refour, P., Patel, H., Springer, M. & Kirschner, M. W. Profiling of ubiquitin-like modifications reveals features of mitotic control. *Cell* **152**, 1160-1172, doi:10.1016/j.cell.2013.02.007 (2013).
- 46 Sasakawa, H. *et al.* Solution structure and dynamics of Ufm1, a ubiquitin-fold modifier 1. *Biochem Biophys Res Commun* **343**, 21-26, doi:10.1016/j.bbrc.2006.02.107 (2006).
- 47 Thompson, A. R., Doelling, J. H., Suttangkakul, A. & Vierstra, R. D. Autophagic nutrient recycling in Arabidopsis directed by the ATG8 and ATG12 conjugation pathways. *Plant Physiol* **138**, 2097-2110, doi:10.1104/pp.105.060673 (2005).
- 48 Rossman, T. G., Visalli, M. A. & Komissarova, E. V. *fau* and its ubiquitin-like domain (FUBI) transforms human osteogenic sarcoma (HOS) cells to anchorage-independence. *Oncogene* **22**, 1817-1821, doi:10.1038/sj.onc.1206283 (2003).
- 49 Loeb, K. R. & Haas, A. L. The interferon-inducible 15-kDa ubiquitin homolog conjugates to intracellular proteins. *J Biol Chem* **267**, 7806-7813 (1992).
- 50 Goehring, A. S., Rivers, D. M. & Sprague, G. F., Jr. Urmylelation: a ubiquitin-like pathway that functions during invasive growth and budding in yeast. *Mol Biol Cell* **14**, 4329-4341, doi:10.1091/mbc.e03-02-0079 (2003).
- 51 Liu, Y. C. *et al.* A MHC-encoded ubiquitin-like protein (FAT10) binds noncovalently to the spindle assembly checkpoint protein MAD2. *Proc Natl Acad Sci USA* **96**, 4313-4318 (1999).
- 52 Downes, B. P., Saracco, S. A., Lee, S. S., Crowell, D. N. & Vierstra, R. D. MUBs, a family of ubiquitin-fold proteins that are plasma membrane-anchored by prenylation. *J Biol Chem* **281**, 27145-27157, doi:10.1074/jbc.M602283200 (2006).
- 53 Welchman, R. L., Gordon, C. & Mayer, R. J. Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Nat Rev Mol Cell Biol* **6**, 599-609, doi:10.1038/nrm1700 (2005).

- 54 Grabbe, C. & Dikic, I. Functional roles of ubiquitin-like domain (ULD) and ubiquitin-binding domain (UBD) containing proteins. *Chem Rev* **109**, 1481-1494, doi:10.1021/cr800413p (2009).
- 55 Cort, J. R., Chiang, Y., Zheng, D., Montelione, G. T. & Kennedy, M. A. NMR structure of conserved eukaryotic protein ZK652.3 from *C. elegans*: a ubiquitin-like fold. *Proteins* **48**, 733-736, doi:10.1002/prot.10197 (2002).
- 56 Komatsu, M. *et al.* A novel protein-conjugating system for Ufm1, a ubiquitin-fold modifier. *EMBO J* **23**, 1977-1986, doi:10.1038/sj.emboj.7600205 (2004).
- 57 Kang, S. H. *et al.* Two novel ubiquitin-fold modifier 1 (Ufm1)-specific proteases, UfSP1 and UfSP2. *J Biol Chem* **282**, 5256-5262, doi:10.1074/jbc.M610590200 (2007).
- 58 Tatsumi, K. *et al.* A novel type of E3 ligase for the Ufm1 conjugation system. *J Biol Chem* **285**, 5417-5427, doi:10.1074/jbc.M109.036814 (2010).
- 59 Zheng, M. *et al.* UBE1DC1, an ubiquitin-activating enzyme, activates two different ubiquitin-like proteins. *J Cell Biochem* **104**, 2324-2334, doi:10.1002/jcb.21791 (2008).
- 60 Daniel, J. & Liebau, E. The ufm1 cascade. *Cells* **3**, 627-638, doi:10.3390/cells3020627 (2014).
- 61 Kim, C. H. *et al.* Overexpression of a novel regulator of p120 catenin, NLBP, promotes lung adenocarcinoma proliferation. *Cell Cycle* **12**, 2443-2453, doi:10.4161/cc.25451 (2013).
- 62 Azfer, A., Niu, J., Rogers, L. M., Adamski, F. M. & Kolattukudy, P. E. Activation of endoplasmic reticulum stress response during the development of ischemic heart disease. *Am J Physiol Heart Circ Physiol* **291**, H1411-1420, doi:10.1152/ajpheart.01378.2005 (2006).
- 63 Lemaire, K. *et al.* Ubiquitin fold modifier 1 (UFM1) and its target UFBP1 protect pancreatic beta cells from ER stress-induced apoptosis. *PLoS One* **6**, e18517, doi:10.1371/journal.pone.0018517 (2011).
- 64 Pirone, L. *et al.* A comprehensive platform for the analysis of ubiquitin-like protein modifications using in vivo biotinylation. *Sci Rep* **7**, 40756, doi: 10.1038/srep40756 (2017).
- 65 Simsek, D. *et al.* The Mammalian Ribo-interactome Reveals Ribosome Functional Diversity and Heterogeneity. *Cell* **169**, 1051-1065 e1018, doi:10.1016/j.cell.2017.05.022 (2017).
- 66 Matunis, M. J., Coutavas, E. & Blobel, G. A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *J Cell Biol* **135**, 1457-1470 (1996).
- 67 Mahajan, R., Delphin, C., Guan, T., Gerace, L. & Melchior, F. A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell* **88**, 97-107 (1997).
- 68 Pichler, A., Fatouros, C., Lee, H. & Eisenhardt, N. SUMO conjugation - a mechanistic view. *Biomol Concepts* **8**, 13-36, doi:10.1515/bmc-2016-0030 (2017).
- 69 Johnson, E. S., Schwenhorst, I., Dohmen, R. J. & Blobel, G. The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer. *EMBO J* **16**, 5509-5519, doi:10.1093/emboj/16.18.5509 (1997).
- 70 Chen, A., Mannen, H. & Li, S. S. Characterization of mouse ubiquitin-like SMT3A and SMT3B cDNAs and gene/pseudogenes. *Biochem Mol Biol Int* **46**, 1161-1174 (1998).
- 71 Saitoh, H. & Hinchey, J. Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. *J Biol Chem* **275**, 6252-6258 (2000).
- 72 Zhang, F. P. *et al.* Sumo-1 function is dispensable in normal mouse development. *Mol Cell Biol* **28**, 5381-5390, doi:10.1128/MCB.00651-08 (2008).
- 73 Hendriks, I. A. *et al.* Uncovering global SUMOylation signaling networks in a site-specific manner. *Nat Struct Mol Biol* **21**, 927-936, doi:10.1038/nsmb.2890 (2014).
- 74 Wang, L. *et al.* SUMO2 is essential while SUMO3 is dispensable for mouse embryonic development. *EMBO Rep* **15**, 878-885, doi:10.15252/embr.201438534 (2014).
- 75 Su, H. L. & Li, S. S. Molecular features of human ubiquitin-like SUMO genes and their encoded proteins. *Gene* **296**, 65-73 (2002).
- 76 Guo, D. *et al.* A functional variant of SUMO4, a new I kappa B alpha modifier, is associated with type 1 diabetes. *Nat Genet* **36**, 837-841, doi:10.1038/ng1391 (2004).
- 77 Hay, R. T. SUMO: a history of modification. *Mol Cell* **18**, 1-12, doi:10.1016/j.molcel.2005.03.012 (2005).
- 78 Psakhye, I. & Jentsch, S. Protein group modification and synergy in the SUMO pathway as exemplified in DNA repair. *Cell* **151**, 807-820, doi:10.1016/j.cell.2012.10.021 (2012).
- 79 Schimmel, J. *et al.* Uncovering SUMOylation dynamics during cell-cycle progression reveals FoxM1 as a key mitotic SUMO target protein. *Mol Cell* **53**, 1053-1066, doi:10.1016/j.molcel.2014.02.001 (2014).
- 80 Xiao, Z. *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. *Mol Cell Proteomics* **14**, 1419-1434, doi:10.1074/mcp.O114.044792 (2015).
- 81 Munk, S. *et al.* Proteomics Reveals Global Regulation of Protein SUMOylation by ATM and ATR Kinases during Replication Stress. *Cell Rep* **21**, 546-558, doi:10.1016/j.celrep.2017.09.059 (2017).

- 82 Hendriks, I. A. & Vertegaal, A. C. A high-yield double-purification proteomics strategy for the identification of SUMO sites. *Nat Protoc* **11**, 1630-1649, doi:10.1038/nprot.2016.082 (2016).
- 83 Hendriks, I. A. *et al.* Site-specific mapping of the human SUMO proteome reveals co-modification with phosphorylation. *Nat Struct Mol Biol* **24**, 325-336, doi:10.1038/nsmb.3366 (2017).
- 84 Creton, S. & Jentsch, S. SnapShot: The SUMO system. *Cell* **143**, 848-848 e841, doi:10.1016/j.cell.2010.11.026 (2010).
- 85 Geiss-Friedlander, R. & Melchior, F. Concepts in sumoylation: a decade on. *Nat Rev Mol Cell Biol* **8**, 947-956, doi:10.1038/nrm2293 (2007).
- 86 Shin, E. J. *et al.* DeSUMOylating isopeptidase: a second class of SUMO protease. *EMBO Rep* **13**, 339-346, doi:10.1038/embor.2012.3 (2012).
- 87 Hickey, C. M., Wilson, N. R. & Hochstrasser, M. Function and regulation of SUMO proteases. *Nat Rev Mol Cell Biol* **13**, 755-766, doi:10.1038/nrm3478 (2012).
- 88 Barry, J. & Lock, R. B. Small ubiquitin-related modifier-1: Wrestling with protein regulation. *Int J Biochem Cell Biol* **43**, 37-40, doi:10.1016/j.biocel.2010.09.022 (2011).
- 89 Johnson, E. S. & Blobel, G. Ubc9p is the conjugating enzyme for the ubiquitin-like protein Smt3p. *J Biol Chem* **272**, 26799-26802 (1997).
- 90 Desterro, J. M., Rodriguez, M. S., Kemp, G. D. & Hay, R. T. Identification of the enzyme required for activation of the small ubiquitin-like protein SUMO-1. *J Biol Chem* **274**, 10618-10624 (1999).
- 91 Desterro, J. M., Thomson, J. & Hay, R. T. Ubc9 conjugates SUMO but not ubiquitin. *FEBS Lett* **417**, 297-300 (1997).
- 92 Wang, J. *et al.* Crystal structure of UBA2(ufd)-Ubc9: insights into E1-E2 interactions in Sumo pathways. *PLoS One* **5**, e15805, doi:10.1371/journal.pone.0015805 (2010).
- 93 Tatham, M. H. *et al.* Unique binding interactions among Ubc9, SUMO and RanBP2 reveal a mechanism for SUMO paralog selection. *Nat Struct Mol Biol* **12**, 67-74, doi:10.1038/nsmb878 (2005).
- 94 Kim, E. T., Kim, K. K., Matunis, M. J. & Ahn, J. H. Enhanced SUMOylation of proteins containing a SUMO-interacting motif by SUMO-Ubc9 fusion. *Biochem Biophys Res Commun* **388**, 41-45, doi:10.1016/j.bbrc.2009.07.103 (2009).
- 95 Nacerddine, K. *et al.* The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. *Dev Cell* **9**, 769-779, doi:10.1016/j.devcel.2005.10.007 (2005).
- 96 Tozluoglu, M., Karaca, E., Nussinov, R. & Haliloglu, T. A mechanistic view of the role of E3 in sumoylation. *PLoS Comput Biol* **6**, doi:10.1371/journal.pcbi.1000913 (2010).
- 97 Sachdev, S. *et al.* PIASy, a nuclear matrix-associated SUMO E3 ligase, represses LEF1 activity by sequestration into nuclear bodies. *Genes Dev* **15**, 3088-3103 (2001).
- 98 Galanty, Y. *et al.* Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks. *Nature* **462**, 935-939, doi:10.1038/nature08657 (2009).
- 99 Bischof, O. *et al.* The E3 SUMO ligase PIASy is a regulator of cellular senescence and apoptosis. *Mol Cell* **22**, 783-794, doi:10.1016/j.molcel.2006.05.016 (2006).
- 100 Takahashi, Y. & Kikuchi, Y. Yeast PIAS-type Uli1/Siz1 is composed of SUMO ligase and regulatory domains. *J Biol Chem* **280**, 35822-35828, doi:10.1074/jbc.M506794200 (2005).
- 101 Pasupala, N., Easwaran, S., Hannan, A., Shore, D. & Mishra, K. The SUMO E3 ligase Siz2 exerts a locus-dependent effect on gene silencing in *Saccharomyces cerevisiae*. *Eukaryot Cell* **11**, 452-462, doi:10.1128/EC.05243-11 (2012).
- 102 Xiao, H. & Zhao, H. Genome-wide RNAi screen reveals the E3 SUMO-protein ligase gene SIZ1 as a novel determinant of furfural tolerance in *Saccharomyces cerevisiae*. *Biotechnol Biofuels* **7**, 78, doi:10.1186/1754-6834-7-78 (2014).
- 103 Takahashi, Y., Iwase, M., Strunnikov, A. V. & Kikuchi, Y. Cytoplasmic sumoylation by PIAS-type Siz1-SUMO ligase. *Cell Cycle* **7**, 1738-1744, doi:10.4161/cc.7.12.6156 (2008).
- 104 Wu, N. *et al.* Scc1 sumoylation by Mms21 promotes sister chromatid recombination through counteracting Wapl. *Genes Dev* **26**, 1473-1485, doi:10.1101/gad.193615.112 (2012).
- 105 Dawlaty, M. M. *et al.* Resolution of sister centromeres requires RanBP2-mediated SUMOylation of topoisomerase IIalpha. *Cell* **133**, 103-115, doi:10.1016/j.cell.2008.01.045 (2008).
- 106 Horio, Y. *et al.* Relationship of mRNA expressions of RanBP2 and topoisomerase II isoforms to cytotoxicity of amrubicin in human lung cancer cell lines. *Cancer Chemother Pharmacol* **66**, 237-243, doi:10.1007/s00280-009-1151-1 (2010).
- 107 Navarro, M. S. & Bachant, J. RanBP2: a tumor suppressor with a new twist on TopoII, SUMO, and centromeres. *Cancer Cell* **13**, 293-295, doi:10.1016/j.ccr.2008.03.011 (2008).
- 108 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* **382**, 585-600, doi:10.1016/j.jmb.2008.07.016 (2008).

- 109 Li, S. J. & Hochstrasser, M. The yeast ULP2 (SMT4) gene encodes a novel protease specific for the ubiquitin-like Smt3 protein. *Mol Cell Biol* **20**, 2367-2377 (2000).
- 110 Bylebyl, G. R., Belichenko, I. & Johnson, E. S. The SUMO isopeptidase Ulp2 prevents accumulation of SUMO chains in yeast. *J Biol Chem* **278**, 44113-44120, doi:10.1074/jbc.M308357200 (2003).
- 111 Takahashi, Y., Mizoi, J., Toh, E. A. & Kikuchi, Y. Yeast Ulp1, an Smt3-specific protease, associates with nucleoporins. *J Biochem* **128**, 723-725 (2000).
- 112 Drag, M. & Salvesen, G. S. DeSUMOylating enzymes--SENPs. *IUBMB Life* **60**, 734-742, doi:10.1002/iub.113 (2008).
- 113 Li, S. J. & Hochstrasser, M. The Ulp1 SUMO isopeptidase: distinct domains required for viability, nuclear envelope localization, and substrate specificity. *J Cell Biol* **160**, 1069-1081, doi:10.1083/jcb.200212052 (2003).
- 114 Kembball-Cook, G., Johnson, D. J., Tuddenham, E. G. & Harlos, K. Crystal structure of active site-inhibited human coagulation factor VIIa (des-Gla). *J Struct Biol* **127**, 213-223, doi:10.1006/jbsi.1999.4158 (1999).
- 115 Schulz, S. *et al.* Ubiquitin-specific protease-like 1 (USPL1) is a SUMO isopeptidase with essential, non-catalytic functions. *EMBO Rep* **13**, 930-938, doi:10.1038/embor.2012.125 (2012).
- 116 Pfeiffer, P. *et al.* DNA lesions and repair. *Mutat Res* **366**, 69-80 (1996).
- 117 Blanpain, C., Mohrin, M., Sotiropoulou, P. A. & Passegue, E. DNA-damage response in tissue-specific and cancer stem cells. *Cell Stem Cell* **8**, 16-29, doi:10.1016/j.stem.2010.12.012 (2011).
- 118 Jackson, S. P. & Bartek, J. The DNA-damage response in human biology and disease. *Nature* **461**, 1071-1078, doi:10.1038/nature08467 (2009).
- 119 Sancar, A., Lindsey-Boltz, L. A., Unsal-Kacmaz, K. & Linn, S. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu Rev Biochem* **73**, 39-85, doi:10.1146/annurev.biochem.73.011303.073723 (2004).
- 120 Murray, J. M. & Carr, A. M. Integrating DNA damage repair with the cell cycle. *Curr Opin Cell Biol* **52**, 120-125, doi:10.1016/j.ceb.2018.03.006 (2018).
- 121 Tuteja, N. & Tuteja, R. Unraveling DNA repair in human: molecular mechanisms and consequences of repair defect. *Crit Rev Biochem Mol Biol* **36**, 261-290, doi:10.1080/20014091074192 (2001).
- 122 Lee, D. F., Drouin, R., Pitsikas, P. & Rainbow, A. J. Detection of an involvement of the human mismatch repair genes hMLH1 and hMSH2 in nucleotide excision repair is dependent on UVC fluence to cells. *Cancer Res* **64**, 3865-3870, doi:10.1158/0008-5472.CAN-03-3193 (2004).
- 123 Kobayashi, K., Karran, P., Oda, S. & Yanaga, K. Involvement of mismatch repair in transcription-coupled nucleotide excision repair. *Hum Cell* **18**, 103-115 (2005).
- 124 Fitzgerald, M. E. & Drohat, A. C. Coordinating the initial steps of base excision repair. Apurinic/apyrimidinic endonuclease 1 actively stimulates thymine DNA glycosylase by disrupting the product complex. *J Biol Chem* **283**, 32680-32690, doi:10.1074/jbc.M805504200 (2008).
- 125 Fleck, O., Lehmann, E., Schar, P. & Kohli, J. Involvement of nucleotide-excision repair in msh2 pms1-independent mismatch repair. *Nat Genet* **21**, 314-317, doi:10.1038/6838 (1999).
- 126 Mullenders, L. H. Transcription response and nucleotide excision repair. *Mutat Res* **409**, 59-64 (1998).
- 127 Auerbach, A. D. & Verlander, P. C. Disorders of DNA replication and repair. *Curr Opin Pediatr* **9**, 600-616 (1997).
- 128 Tanori, M. *et al.* Opposite modifying effects of HR and NHEJ deficiency on cancer risk in Ptc1 heterozygous mouse cerebellum. *Oncogene* **30**, 4740-4749, doi:10.1038/onc.2011.178 (2011).
- 129 Rosidi, B. *et al.* Histone H1 functions as a stimulatory factor in backup pathways of NHEJ. *Nucleic Acids Res* **36**, 1610-1623, doi:10.1093/nar/gkn013 (2008).
- 130 Garcia-Rodriguez, N., Wong, R. P. & Ulrich, H. D. Functions of Ubiquitin and SUMO in DNA Replication and Replication Stress. *Front Genet* **7**, 87, doi:10.3389/fgene.2016.00087 (2016).
- 131 Mazouzi, A., Velimezi, G. & Loizou, J. I. DNA replication stress: causes, resolution and disease. *Exp Cell Res* **329**, 85-93, doi:10.1016/j.yexcr.2014.09.030 (2014).
- 132 Macheret, M. & Halazonetis, T. D. DNA replication stress as a hallmark of cancer. *Annu Rev Pathol* **10**, 425-448, doi:10.1146/annurev-pathol-012414-040424 (2015).
- 133 Byun, T. S., Pacek, M., Yee, M. C., Walter, J. C. & Cimprich, K. A. Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. *Genes Dev* **19**, 1040-1052, doi:10.1101/gad.1301205 (2005).
- 134 Jones, R. M. & Petermann, E. Replication fork dynamics and the DNA damage response. *Biochem J* **443**, 13-26, doi:10.1042/BJ20112100 (2012).
- 135 Aylon, Y., Liefshitz, B. & Kupiec, M. The CDK regulates repair of double-strand breaks by homologous recombination during the cell cycle. *EMBO J* **23**, 4868-4875, doi:10.1038/sj.emboj.7600469 (2004).
- 136 Kim, A. & Park, J. S. Involvement of subcomplexes of 32 and 14 kDa subunits in RPA's DNA binding activity through redox change. *Mol Cells* **13**, 493-497 (2002).

- 137 Sorensen, C. S. & Syljuasen, R. G. Safeguarding genome integrity: the checkpoint kinases ATR, CHK1 and WEE1 restrain CDK activity during normal DNA replication. *Nucleic Acids Res* **40**, 477-486, doi:10.1093/nar/gkr697 (2012).
- 138 Barr, S. M., Leung, C. G., Chang, E. E. & Cimprich, K. A. ATR kinase activity regulates the intranuclear translocation of ATR and RPA following ionizing radiation. *Curr Biol* **13**, 1047-1051 (2003).
- 139 Zou, L. & Elledge, S. J. Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science* **300**, 1542-1548, doi:10.1126/science.1083430 (2003).
- 140 Kumagai, A. & Dunphy, W. G. Claspin, a novel protein required for the activation of Chk1 during a DNA replication checkpoint response in *Xenopus* egg extracts. *Mol Cell* **6**, 839-849 (2000).
- 141 Sorensen, C. S. *et al.* Chk1 regulates the S phase checkpoint by coupling the physiological turnover and ionizing radiation-induced accelerated proteolysis of Cdc25A. *Cancer Cell* **3**, 247-258 (2003).
- 142 Ira, G. *et al.* DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1. *Nature* **431**, 1011-1017, doi:10.1038/nature02964 (2004).
- 143 Sanchez, Y. *et al.* Conservation of the Chk1 checkpoint pathway in mammals: linkage of DNA damage to Cdk regulation through Cdc25. *Science* **277**, 1497-1501 (1997).
- 144 Stiff, T. *et al.* ATR-dependent phosphorylation and activation of ATM in response to UV treatment or replication fork stalling. *EMBO J* **25**, 5775-5782, doi:10.1038/sj.emboj.7601446 (2006).
- 145 Falck, J., Mailand, N., Syljuasen, R. G., Bartek, J. & Lukas, J. The ATM-Chk2-Cdc25A checkpoint pathway guards against radioresistant DNA synthesis. *Nature* **410**, 842-847, doi:10.1038/35071124 (2001).
- 146 Jazayeri, A. *et al.* ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. *Nat Cell Biol* **8**, 37-45, doi:10.1038/ncb1337 (2006).
- 147 Lee, J., Kumagai, A. & Dunphy, W. G. The Rad9-Hus1-Rad1 checkpoint clamp regulates interaction of TopBP1 with ATR. *J Biol Chem* **282**, 28036-28044, doi:10.1074/jbc.M704635200 (2007).
- 148 Cimprich, K. A. & Cortez, D. ATR: an essential regulator of genome integrity. *Nat Rev Mol Cell Biol* **9**, 616-627, doi:10.1038/nrm2450 (2008).
- 149 Duursma, A. M., Driscoll, R., Elias, J. E. & Cimprich, K. A. A role for the MRN complex in ATR activation via TOPBP1 recruitment. *Mol Cell* **50**, 116-122, doi:10.1016/j.molcel.2013.03.006 (2013).
- 150 Couch, F. B. *et al.* ATR phosphorylates SMARCA1 to prevent replication fork collapse. *Genes Dev* **27**, 1610-1623, doi:10.1101/gad.214080.113 (2013).
- 151 Cortez, D., Glick, G. & Elledge, S. J. Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases. *Proc Natl Acad Sci U S A* **101**, 10078-10083, doi:10.1073/pnas.0403410101 (2004).
- 152 Ammazalorso, F., Pirzio, L. M., Bignami, M., Franchitto, A. & Pichierri, P. ATR and ATM differently regulate WRN to prevent DSBs at stalled replication forks and promote replication fork recovery. *EMBO J* **29**, 3156-3169, doi:10.1038/emboj.2010.205 (2010).
- 153 Ishiai, M. *et al.* FANCI phosphorylation functions as a molecular switch to turn on the Fanconi anemia pathway. *Nat Struct Mol Biol* **15**, 1138-1146, doi:10.1038/nsmb.1504 (2008).
- 154 Gohler, T., Sabbioneda, S., Green, C. M. & Lehmann, A. R. ATR-mediated phosphorylation of DNA polymerase η is needed for efficient recovery from UV damage. *J Cell Biol* **192**, 219-227, doi:10.1083/jcb.201008076 (2011).
- 155 Moynahan, M. E. & Jasin, M. Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nat Rev Mol Cell Biol* **11**, 196-207, doi:10.1038/nrm2851 (2010).
- 156 Nimonkar, A. V., Ozsoy, A. Z., Genschel, J., Modrich, P. & Kowalczykowski, S. C. Human exonuclease 1 and BLM helicase interact to resect DNA and initiate DNA repair. *Proc Natl Acad Sci U S A* **105**, 16906-16911, doi:10.1073/pnas.0809380105 (2008).
- 157 Chen, L., Nievera, C. J., Lee, A. Y. & Wu, X. Cell cycle-dependent complex formation of BRCA1.CtIP.MRN is important for DNA double-strand break repair. *J Biol Chem* **283**, 7713-7720, doi:10.1074/jbc.M710245200 (2008).
- 158 Sharan, S. K. *et al.* Embryonic lethality and radiation hypersensitivity mediated by Rad51 in mice lacking Brca2. *Nature* **386**, 804-810, doi:10.1038/386804a0 (1997).
- 159 Shima, H. *et al.* Activation of the SUMO modification system is required for the accumulation of RAD51 at sites of DNA damage. *J Cell Sci* **126**, 5284-5292, doi:10.1242/jcs.133744 (2013).
- 160 Dou, H., Huang, C., Singh, M., Carpenter, P. B. & Yeh, E. T. Regulation of DNA repair through deSUMOylation and SUMOylation of replication protein A complex. *Mol Cell* **39**, 333-345, doi:10.1016/j.molcel.2010.07.021 (2010).
- 161 Marechal, A. & Zou, L. RPA-coated single-stranded DNA as a platform for post-translational modifications in the DNA damage response. *Cell Res* **25**, 9-23, doi:10.1038/cr.2014.147 (2015).
- 162 Wu, C. S. *et al.* SUMOylation of ATRIP potentiates DNA damage signaling by boosting multiple protein interactions in the ATR pathway. *Genes Dev* **28**, 1472-1484, doi:10.1101/gad.238535.114 (2014).

- 163 Wu, C. S. & Zou, L. The SUMO (Small Ubiquitin-like Modifier) Ligase PIAS3 Primes ATR for Checkpoint Activation. *J Biol Chem* **291**, 279-290, doi:10.1074/jbc.M115.691170 (2016).
- 164 Eladad, S. *et al.* Intra-nuclear trafficking of the BLM helicase to DNA damage-induced foci is regulated by SUMO modification. *Hum Mol Genet* **14**, 1351-1365, doi:10.1093/hmg/ddi145 (2005).
- 165 Ouyang, K. J., Yagle, M. K., Matunis, M. J. & Ellis, N. A. BLM SUMOylation regulates ssDNA accumulation at stalled replication forks. *Front Genet* **4**, 167, doi:10.3389/fgene.2013.00167 (2013).
- 166 Bohm, S. & Bernstein, K. A. The role of post-translational modifications in fine-tuning BLM helicase function during DNA repair. *DNA Repair (Amst)* **22**, 123-132, doi:10.1016/j.dnarep.2014.07.007 (2014).
- 167 Ouyang, J. *et al.* Noncovalent interactions with SUMO and ubiquitin orchestrate distinct functions of the SLX4 complex in genome maintenance. *Mol Cell* **57**, 108-122, doi:10.1016/j.molcel.2014.11.015 (2015).
- 168 Gonzalez-Prieto, R., Cuijpers, S. A., Luijsterburg, M. S., van Attikum, H. & Vertegaal, A. C. SUMOylation and PARylation cooperate to recruit and stabilize SLX4 at DNA damage sites. *EMBO Rep* **16**, 512-519, doi:10.15252/embr.201440017 (2015).
- 169 Guervilly, J. H. *et al.* The SLX4 complex is a SUMO E3 ligase that impacts on replication stress outcome and genome stability. *Mol Cell* **57**, 123-137, doi:10.1016/j.molcel.2014.11.014 (2015).
- 170 Morris, J. R. *et al.* The SUMO modification pathway is involved in the BRCA1 response to genotoxic stress. *Nature* **462**, 886-890, doi:10.1038/nature08593 (2009).
- 171 Luo, K., Zhang, H., Wang, L., Yuan, J. & Lou, Z. Sumoylation of MDC1 is important for proper DNA damage response. *EMBO J* **31**, 3008-3019, doi:10.1038/emboj.2012.158 (2012).
- 172 Hietakangas, V. *et al.* PDSM, a motif for phosphorylation-dependent SUMO modification. *Proc Natl Acad Sci U S A* **103**, 45-50, doi:10.1073/pnas.0503698102 (2006).
- 173 Yang, S. H., Galanis, A., Witty, J. & Sharrocks, A. D. An extended consensus motif enhances the specificity of substrate modification by SUMO. *EMBO J* **25**, 5083-5093, doi:10.1038/sj.emboj.7601383 (2006).
- 174 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).
- 175 Uzunova, K. *et al.* Ubiquitin-dependent proteolytic control of SUMO conjugates. *J Biol Chem* **282**, 34167-34175, doi:10.1074/jbc.M706505200 (2007).
- 176 Geoffroy, M. C. & Hay, R. T. An additional role for SUMO in ubiquitin-mediated proteolysis. *Nat Rev Mol Cell Biol* **10**, 564-568, doi:10.1038/nrm2707 (2009).
- 177 Cuijpers, S. A. G., Willemstein, E. & Vertegaal, A. C. O. Converging Small Ubiquitin-like Modifier (SUMO) and Ubiquitin Signaling: Improved Methodology Identifies Co-modified Target Proteins. *Mol Cell Proteomics* **16**, 2281-2295, doi:10.1074/mcp.TIR117.000152 (2017).
- 178 Cuijpers, S. A. G. & Vertegaal, A. C. O. Guiding Mitotic Progression by Crosstalk between Post-translational Modifications. *Trends Biochem Sci* **43**, 251-268, doi:10.1016/j.tibs.2018.02.004 (2018).
- 179 Kohler, J. B. *et al.* Targeting of SUMO substrates to a Cdc48-Ufd1-Npl4 segregase and STUbL pathway in fission yeast. *Nat Commun* **6**, 8827, doi:10.1038/ncomms9827 (2015).
- 180 Garza, R. & Pillus, L. STUbLs in chromatin and genome stability. *Biopolymers* **99**, 146-154, doi:10.1002/bip.22125 (2013).
- 181 Barry, K. C. *et al.* The Drosophila STUbL protein Degringolade limits HES functions during embryogenesis. *Development* **138**, 1759-1769, doi:10.1242/dev.058420 (2011).
- 182 Kroetz, M. B. & Hochstrasser, M. Identification of SUMO-interacting proteins by yeast two-hybrid analysis. *Methods Mol Biol* **497**, 107-120, doi:10.1007/978-1-59745-566-4_7 (2009).
- 183 Minty, A., Dumont, X., Kaghad, M. & Caput, D. Covalent modification of p73alpha by SUMO-1. Two-hybrid screening with p73 identifies novel SUMO-1-interacting proteins and a SUMO-1 interaction motif. *J Biol Chem* **275**, 36316-36323, doi:10.1074/jbc.M004293200 (2000).
- 184 Hecker, C. M., Rabiller, M., Haglund, K., Bayer, P. & Dikic, I. Specification of SUMO1- and SUMO2-interacting motifs. *J Biol Chem* **281**, 16117-16127, doi:10.1074/jbc.M512757200 (2006).
- 185 Tan, W., Wang, Z. & Prelich, G. Physical and Genetic Interactions Between Uls1 and the Slx5-Slx8 SUMO-Targeted Ubiquitin Ligase. *G3 (Bethesda)* **3**, 771-780, doi:10.1534/g3.113.005827 (2013).
- 186 Lescasse, R., Pobiega, S., Callebaut, I. & Marcand, S. End-joining inhibition at telomeres requires the translocase and polySUMO-dependent ubiquitin ligase Uls1. *EMBO J* **32**, 805-815, doi:10.1038/emboj.2013.24 (2013).
- 187 Burgess, R. C., Rahman, S., Lisby, M., Rothstein, R. & Zhao, X. The Slx5-Slx8 complex affects sumoylation of DNA repair proteins and negatively regulates recombination. *Mol Cell Biol* **27**, 6153-6162, doi:10.1128/MCB.00787-07 (2007).
- 188 Parker, J. L. & Ulrich, H. D. A SUMO-interacting motif activates budding yeast ubiquitin ligase Rad18 towards SUMO-modified PCNA. *Nucleic Acids Res* **40**, 11380-11388, doi:10.1093/nar/gks892 (2012).
- 189 Sun, H., Leversson, J. D. & Hunter, T. Conserved function of RNF4 family proteins in eukaryotes: targeting a ubiquitin ligase to SUMOylated proteins. *EMBO J* **26**, 4102-4112, doi:10.1038/sj.emboj.7601839 (2007).

- 190 Bohm, S., Mihalevic, M. J., Casal, M. A. & Bernstein, K. A. Disruption of SUMO-targeted ubiquitin ligases Slx5-Slx8/RNF4 alters RecQ-like helicase Sgs1/BLM localization in yeast and human cells. *DNA Repair (Amst)* **26**, 1-14, doi:10.1016/j.dnarep.2014.12.004 (2015).
- 191 Gartner, A. & Muller, S. PML, SUMO, and RNF4: guardians of nuclear protein quality. *Mol Cell* **55**, 1-3, doi:10.1016/j.molcel.2014.06.022 (2014).
- 192 Weisshaar, S. R. *et al.* Arsenic trioxide stimulates SUMO-2/3 modification leading to RNF4-dependent proteolytic targeting of PML. *FEBS Lett* **582**, 3174-3178, doi:10.1016/j.febslet.2008.08.008 (2008).
- 193 Yin, Y. *et al.* SUMO-targeted ubiquitin E3 ligase RNF4 is required for the response of human cells to DNA damage. *Genes Dev* **26**, 1196-1208, doi:10.1101/gad.189274.112 (2012).
- 194 Galanty, Y., Belotserkovskaya, R., Coates, J. & Jackson, S. P. RNF4, a SUMO-targeted ubiquitin E3 ligase, promotes DNA double-strand break repair. *Genes Dev* **26**, 1179-1195, doi:10.1101/gad.188284.112 (2012).
- 195 Mukhopadhyay, D., Arnaoutov, A. & Dasso, M. The SUMO protease SENP6 is essential for inner kinetochore assembly. *J Cell Biol* **188**, 681-692, doi:10.1083/jcb.200909008 (2010).
- 196 Gibbs-Seymour, I. *et al.* Ubiquitin-SUMO circuitry controls activated fanconi anemia ID complex dosage in response to DNA damage. *Mol Cell* **57**, 150-164, doi:10.1016/j.molcel.2014.12.001 (2015).
- 197 Vyas, R. *et al.* RNF4 is required for DNA double-strand break repair in vivo. *Cell Death Differ* **20**, 490-502, doi:10.1038/cdd.2012.145 (2013).
- 198 Hendriks, I. A., Treffers, L. W., Verlaan-de Vries, M., Olsen, J. V. & Vertegaal, A. C. SUMO-2 Orchestrates Chromatin Modifiers in Response to DNA Damage. *Cell Rep*, doi:10.1016/j.celrep.2015.02.033 (2015).
- 199 Sun, H., Liu, Y. & Hunter, T. Multiple Arkadia/RNF111 structures coordinate its Polycomb body association and transcriptional control. *Mol Cell Biol* **34**, 2981-2995, doi:10.1128/MCB.00036-14 (2014).
- 200 Poulsen, S. L. *et al.* RNF111/Arkadia is a SUMO-targeted ubiquitin ligase that facilitates the DNA damage response. *J Cell Biol* **201**, 797-807, doi:10.1083/jcb.201212075 (2013).
- 201 Ma, T. *et al.* RNF111-dependent neddylation activates DNA damage-induced ubiquitination. *Mol Cell* **49**, 897-907, doi:10.1016/j.molcel.2013.01.006 (2013).
- 202 Liu, N. Q. *et al.* Quantitative proteomic analysis of microdissected breast cancer tissues: comparison of label-free and SILAC-based quantification with shotgun, directed, and targeted MS approaches. *J Proteome Res* **12**, 4627-4641, doi:10.1021/pr4005794 (2013).
- 203 Tammsalu, T. *et al.* Proteome-wide identification of SUMO2 modification sites. *Sci Signal* **7**, rs2, doi:10.1126/scisignal.2005146 (2014).
- 204 Tammsalu, T. *et al.* Proteome-wide identification of SUMO modification sites by mass spectrometry. *Nat Protoc* **10**, 1374-1388, doi:10.1038/nprot.2015.095 (2015).
- 205 Hendriks, I. A. & Vertegaal, A. C. A comprehensive compilation of SUMO proteomics. *Nat Rev Mol Cell Biol* **17**, 581-595, doi:10.1038/nrm.2016.81 (2016).
- 206 Schratzenholz, A., Klemm, M. & Cahill, M. Potential of comprehensive toxico-proteomics: quantitative and differential mining of functional proteomes from native samples. *Altern Lab Anim* **32 Suppl 1A**, 123-131 (2004).
- 207 Choudhary, C. & Mann, M. Decoding signalling networks by mass spectrometry-based proteomics. *Nat Rev Mol Cell Biol* **11**, 427-439, doi:10.1038/nrm2900 (2010).
- 208 Mukhopadhyay, D. & Dasso, M. Modification in reverse: the SUMO proteases. *Trends Biochem Sci* **32**, 286-295, doi:10.1016/j.tibs.2007.05.002 (2007).
- 209 Tatham, M. H. & Hay, R. T. FRET-based in vitro assays for the analysis of SUMO protease activities. *Methods Mol Biol* **497**, 253-268, doi:10.1007/978-1-59745-566-4_17 (2009).
- 210 Becker, J. *et al.* Detecting endogenous SUMO targets in mammalian cells and tissues. *Nat Struct Mol Biol* **20**, 525-531, doi:10.1038/nsmb.2526 (2013).
- 211 Matic, I. *et al.* Site-specific identification of SUMO-2 targets in cells reveals an inverted SUMOylation motif and a hydrophobic cluster SUMOylation motif. *Mol Cell* **39**, 641-652, doi:10.1016/j.molcel.2010.07.026 (2010).
- 212 Hendriks, I. A., D'Souza, R. C., Chang, J. G., Mann, M. & Vertegaal, A. C. System-wide identification of wild-type SUMO-2 conjugation sites. *Nat Commun* **6**, 7289, doi:10.1038/ncomms8289 (2015).
- 213 Bruderer, R. *et al.* Purification and identification of endogenous polySUMO conjugates. *EMBO Rep* **12**, 142-148, doi:10.1038/embor.2010.206 (2011).
- 214 Hendriks, I. A. & Vertegaal, A. C. Label-Free Identification and Quantification of SUMO Target Proteins. *Methods Mol Biol* **1475**, 171-193, doi:10.1007/978-1-4939-6358-4_13 (2016).
- 215 Wilkinson, K. A., Nishimune, A. & Henley, J. M. Analysis of SUMO-1 modification of neuronal proteins containing consensus SUMOylation motifs. *Neurosci Lett* **436**, 239-244, doi:10.1016/j.neulet.2008.03.029 (2008).
- 216 Uy, R. & Wold, F. Posttranslational covalent modification of proteins. *Science* **198**, 890-896 (1977).
- 217 Hunter, T. The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Mol Cell* **28**, 730-738, doi:10.1016/j.molcel.2007.11.019 (2007).

- 218 Ghosh, S. & Dass, J. F. Study of pathway cross-talk interactions with NF-kappaB leading to its activation via ubiquitination or phosphorylation: A brief review. *Gene* **584**, 97-109, doi:10.1016/j.gene.2016.03.008 (2016).
- 219 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* **9**(1), 2456. doi: 10.1038/s41467-018-04957-4 (2018).

Chapter 2

System-wide analysis of SUMOylation dynamics in response to replication stress reveals novel SUMO target proteins and acceptor lysines relevant for genome stability

Zhenyu Xiao¹, Jer-Gung Chang¹, Ivo A. Hendriks¹, Jón Otti Sigurðsson², Jesper V. Olsen² and Alfred C.O. Vertegaal^{1*}

¹Department of Molecular Cell Biology, Leiden University Medical Center, 2300 RC Leiden, the Netherlands

²Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen, Denmark

* To whom correspondence should be addressed: E-mail: vertegaal@lumc.nl

Chapter 2 has been published in **Molecular cell proteomics**

Mol Cell Proteomics. 2015 May;14(5):1419-34. doi: 10.1074/mcp.O114.044792. Epub 2015 Mar 9.

Supplementary Tables are available online.

Running title

SUMOylation and replication stress

Abbreviations

53BP1	Tumor suppressor p53-binding protein 1
Atg12	Ubiquitin-like protein ATG12
Atg8	Ubiquitin-like protein ATG8
BHLHE40/41	Class E basic helix-loop-helix protein 40/41
BLM	Bloom syndrome protein, RecQ helicase-like
BRCA1	Breast cancer type 1 susceptibility protein
CENPC1	CENPC, Centromere protein C
CENPH	Centromere protein H
CHAF1A	Chromatin assembly factor 1, subunit A
DDR	DNA damage response
dNTPs	deoxynucleotide triphosphates
DSBs	Double strand breaks
EME1	Crossover junction endonuclease EME1
FAT10	UBD or ubiquitin D
FOXM1	Forkhead box protein M1
FUBI	Ubiquitin-like protein FUBI
GO	Gene Ontology
HERC2	HECT domain and RCC1-like domain-containing protein 2, E3 ligase
His10-S2	His10-SUMO-2-IRES-GFP
His10-S2-K0-Q87R	His10-SUMO-2-K0-Q87R-IRES-GFP
HU	Hydroxyurea
HUB1	Ubiquitin-like modifier HUB1
IAA	Iodoacetamide
IR	Ionizing radiation
IRES	Internal ribosome entry site
ISG15	Ubiquitin-like modifier ISG15
K0	Lysine-deficient
LFQ	Label free quantification
MAFF	Transcription factor MafF
MCM4	Minichromosome maintenance complex component 4
MDC1	Mediator of DNA-damage checkpoint 1
MIS18A	Protein Mis18-alpha
MMS	Methyl methanesulfonate
MYBL2	Myb-related protein B

NACC1	Nucleus accumbens-associated protein 1
Nedd8	Ubiquitin-like protein NEDD8
PCNA	Proliferating Cell Nuclear Antigen
PIAS	E3 SUMO-protein ligase
PTM	Post-translational modification
RANBP2	RAN binding protein 2
RMI1	RecQ-mediated genome instability protein 1
RNF168	E3 ubiquitin-protein ligase RNF168
SAE1	SUMO activating enzyme subunit 1
SAE2	SUMO-activating enzyme subunit 2, UBA2
Srs2	ATP-dependent DNA helicase SRS2
STRING	Search Tool for the Retrieval of Interacting Genes and Proteins
SUMO	Small Ubiquitin-like Modifier
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
TOP2A	DNA topoisomerase II α
UBC9	SUMO-conjugating enzyme UBC9
Ubl	Ubiquitin like
UFM1	Ubiquitin-fold modifier 1
URM1	Ubiquitin related modifier 1
UV	Ultraviolet

Summary

Genotoxic agents can cause replication fork stalling in dividing cells due to DNA lesions, eventually leading to replication fork collapse when the damage is not repaired. Small Ubiquitin-like Modifiers (SUMOs) are known to counteract replication stress, nevertheless, only a small number of relevant SUMO target proteins are known. To address this, we have purified and identified SUMO-2 target proteins regulated by replication stress in human cells. The developed methodology enabled single step purification of His10-SUMO-2 conjugates under denaturing conditions with high yield and high purity. Following statistical analysis on five biological replicates, a total of 566 SUMO-2 targets were identified. After 2 hours of Hydroxyurea treatment, 10 proteins were up-regulated for SUMOylation and 2 proteins were down-regulated for SUMOylation, whereas after 24 hours, 35 proteins were up-regulated for SUMOylation and 13 proteins were down-regulated for SUMOylation. A site-specific approach was used to map over 1,000 SUMO-2 acceptor lysines in target proteins. The methodology is generic and is widely applicable in the ubiquitin field. A large subset of these identified proteins function in one network that consists of interacting replication factors, transcriptional regulators, DNA damage response factors including MDC1, ATR-interacting protein ATRIP, the Bloom syndrome protein and the BLM-binding partner RMI1, the crossover junction endonuclease EME1, BRCA1 and CHAF1A. Furthermore, centromeric proteins and signal transducers were dynamically regulated by SUMOylation upon replication stress. Our results uncover a comprehensive network of SUMO target proteins dealing with replication damage and provide a framework for detailed understanding of the role of SUMOylation to counteract replication stress. Ultimately, our study reveals how a post-translational modification is able to orchestrate a large variety of different proteins to integrate different nuclear processes with the aim of dealing with the induced DNA damage.

1 Introduction

All cellular processes are tightly regulated via post-translational modifications (PTMs) including small chemical modifications like phosphorylation and acetylation and including modifications by small proteins belonging to the ubiquitin family (1). These PTMs frequently regulate protein-protein interactions via specific domains, exemplified by the archetypical phosphor-tyrosine-interacting SH2-protein-interaction module (2). The reversible nature of these modifications enables rapid and transient cellular signal transduction. As a result of these PTMs, functional proteomes are extremely complex (3).

Ubiquitination, the process of ubiquitin conjugation to target proteins is best known for its role in targeting proteins for degradation by the proteasome, but importantly also regulates target proteins in a degradation-independent manner (4). The ubiquitin-like (Ubl) family includes Small Ubiquitin-like Modifiers (SUMOs), FUBI, HUB1, Nedd8, ISG15, FAT10, URM1, UFM1, Atg12 and Atg8 (5, 6). SUMOs are predominantly located in the nucleus, regulating all nuclear processes, including transcription, splicing, genome stability and nuclear transport (7).

Similar to the ubiquitin system, SUMO conjugation is mediated by E1, E2 and E3 enzymes (8). The SUMO E1 is a dimer consisting of SAE1 and SAE2. A single E2 enzyme, Ubc9, mediates conjugation of SUMO to all target proteins. SUMO E3 enzymes include PIAS protein family members and the nucleoporin RanBP2. SUMO proteases remove SUMOs from target proteins and mediate the maturation of SUMO precursors to enable SUMO conjugation to the epsilon amino group of lysines situated in target proteins (9). A significant set of SUMO-2 acceptor lysines are situated in the SUMO consensus motif Ψ K Ψ E (8, 10). This motif is directly recognized by Ubc9, with coordinated binding of the lysine and the acidic residue of the motif to the catalytic core of the E2 enzyme (11).

The essential role of SUMO to maintain genome stability is particularly well studied (12-14). Organisms deficient for SUMOylation display increased sensitivity for different types of DNA damaging agents including double strand breaks (IR), intra-strand crosslinks (UV), alkylation (MMS) and replication fork blockage (HU) (12-14). Mice deficient for Ubc9 die at the early post-implantation stage showing DNA hypo-condensation and chromosomal aberrancies (15). The trimeric replication clamp PCNA is one of the best studied SUMO target proteins in yeast (16, 17), where SUMOylation enables the interaction with the helicase Srs2 to prevent recombination (18-20). Multiple SUMO target proteins relevant for the DNA Damage Response have been identified in mammalian systems, including DNA topoisomerase I (21), DNA topoisomerase II α and β (22, 23), the BLM helicase (24), 53BP1 (25), BRCA1 (26), HERC2, RNF168 (27) and MDC1 (28-31). In yeast, significant numbers of SUMO target proteins have been identified upon MMS and UV treatment using proteomics approaches (32, 33).

Currently, we are limited in our understanding of the role of SUMOylation in mammalian cells during the DDR since only a limited number of SUMO target proteins are known to be specifically SUMOylated in response to DNA damage. In this study, we are focusing on the role of SUMOylation

with respect to replication, since SUMOylation-impaired organisms are particularly sensitive to replication stress (34-36). We have used a proteomics approach to purify and identify SUMO-2 target proteins and acceptor lysines from cells exposed to replication stress. We have uncovered sets of SUMO target proteins specifically up-regulated or down-regulated in response to replication stress, revealing a highly interactive network of proteins that is coordinated by SUMOylation to cope with replication damage. Our results shed light on the target protein network that is coordinated by SUMOylation to maintain genome stability during replication stress.

2 Results

A quantitative proteomics approach to identify SUMO-2 target proteins and acceptor lysines that are dynamically SUMOylated in response to replication stress.

2.1 Strategy to enrich SUMO-2 conjugates

In order to purify SUMO-2 target proteins from cells under replication stress conditions, we first generated a U2OS cell line stably expressing His10-tagged SUMO-2 (abbreviated as His10-S2). To facilitate flow cytometry sorting of cells expressing a low level of His10-S2, GFP was linked to the His10-S2 cDNA via an internal ribosome entry site (IRES) (Fig. 1A). A homogeneous population of low level GFP expressing cells was selected by flow cytometry to avoid overexpression artifacts. Immunoblotting analysis confirmed the relatively low expression level of His10-SUMO-2 compared to endogenous SUMO-2 levels in U2OS cells, and the efficient enrichment of SUMO-2 conjugates by Ni-NTA purification. Ponceau S staining is shown as a loading control. Coomassie staining showed the low level of non-specific binding of non-SUMOylated proteins (Fig. 1B). Analysis of the cells by confocal microscopy revealed that His10-SUMO-2 was predominantly located in the nucleus (Fig. 1C).

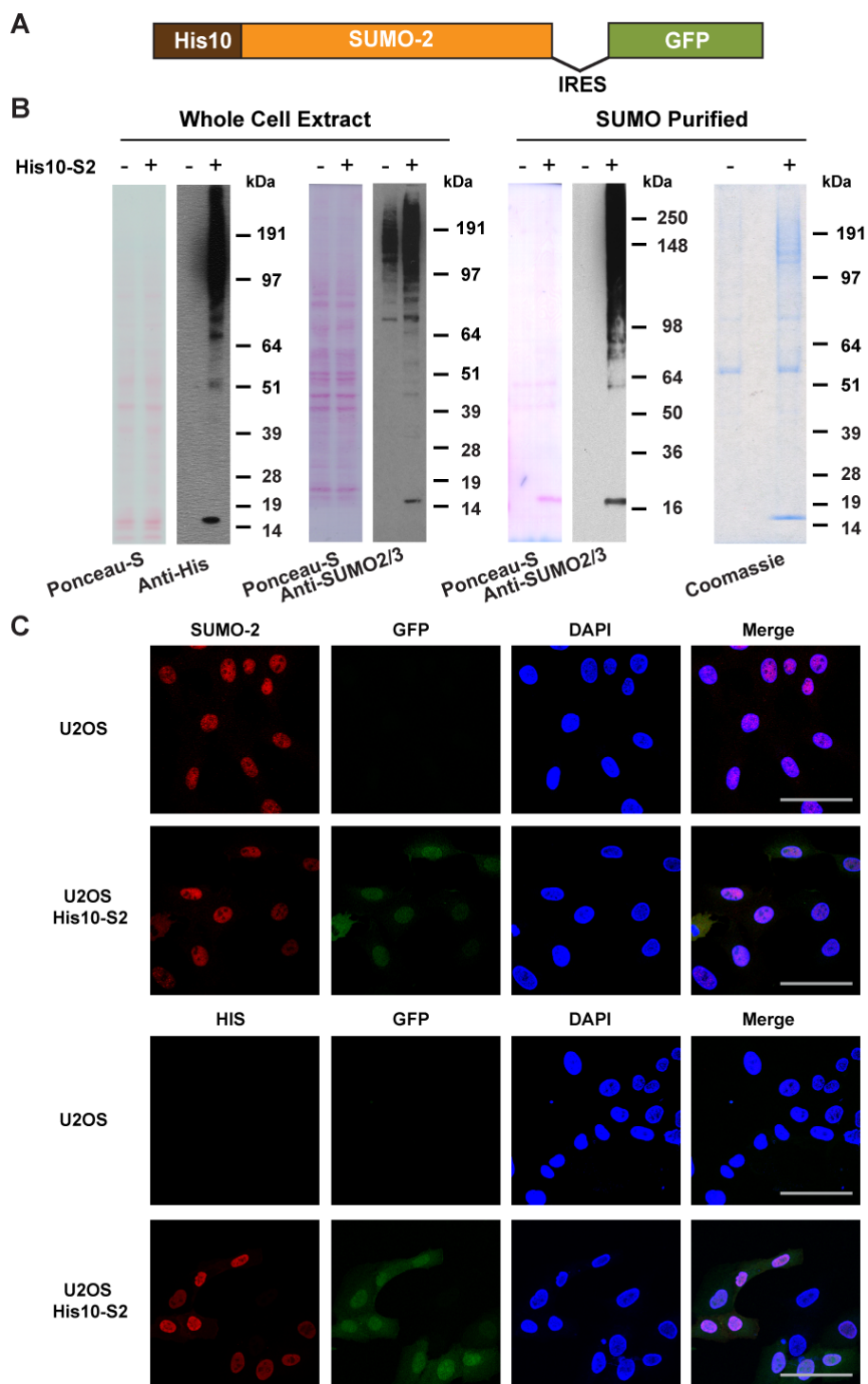


Figure 1. Generation and validation of U2OS cells stably expressing His10-SUMO-2. *A*, Schematic representation of the His10-SUMO-2-IRES-GFP construct used in this project. U2OS cells were infected with a lentivirus encoding His10-SUMO-2 (His10-S2) and GFP separated by an Internal Ribosome Entry Site (IRES) and cells stably expressing low levels of GFP were sorted by flow cytometry. *B*, Expression levels of SUMO-2 in U2OS cells and His10-SUMO-2 (His10-S2) expressing stable cells were compared by immunoblotting. Whole cell extracts were analyzed by immunoblotting using anti-polyHistidine and anti-SUMO-2/3 antibody to confirm the expression of SUMO-2 in U2OS cells and His10-SUMO-2 (His10-S2) expressing stable cells. Ponceau-S staining is shown as a loading control. Additionally, a His-pulldown was performed using Ni-NTA agarose beads to enrich SUMOylated proteins, and purification of His10-SUMO-2 conjugates was confirmed by immunoblotting using anti-SUMO-2/3 antibody. Ponceau-S staining and Coomassie staining were performed to confirm the purity of the final fraction. The experiment was performed in three biological replicates. *C*, The predominant nuclear localization of His10-SUMO-2 was visualized via confocal fluorescence microscopy after immunostaining with the indicated antibodies. DAPI staining was used to visualize the nuclei. Scale bars represent 75 μ m.

2.2 Replication stress induction upon HU treatment

To investigate global changes of proteins that are dynamically SUMOylated during early and late replication damage events, we employed a label free quantitative proteomics approach (Fig. 2A). It was reported before that after short replication blocks, replication forks can stay viable and are able to restart after release from the replication block. In contrast, prolonged stalling of replication forks is known to result in the generation of DNA double strand breaks (DSBs) in S phase and requires HR-mediated restart (43). We cultured U2OS and U2OS cells which stably expressed His10-tagged SUMO-2 (His10-S2) in regular medium and then treated these cells with 2 mM of the replication inhibiting agent hydroxyurea (HU) for 2 hours or for 24 hours in order to induce replication fork stalling.

We first used Ni-NTA purification to enrich His10-SUMO-2 conjugates from U2OS cells stably expressing His10-SUMO-2, after treatment with HU for either 2 hours, 24 hours, or after mock treatment. Parental U2OS cells were included as a negative control. Immunoblotting analysis was employed to assess global purified SUMO-2 conjugates. The total level of SUMO-2 conjugation appeared to be equal and the immunoblotting analysis also confirmed our highly efficient enrichment for SUMO-2 conjugates (Fig. 2B).

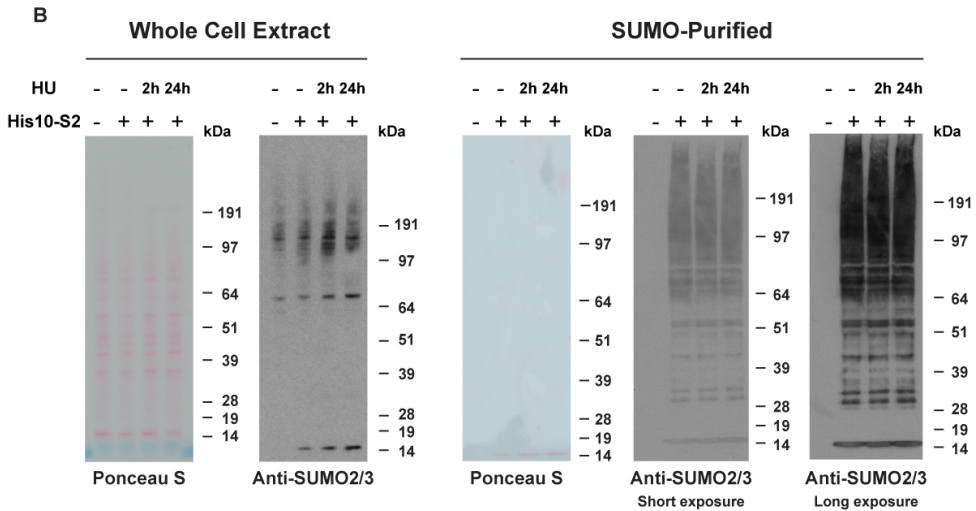
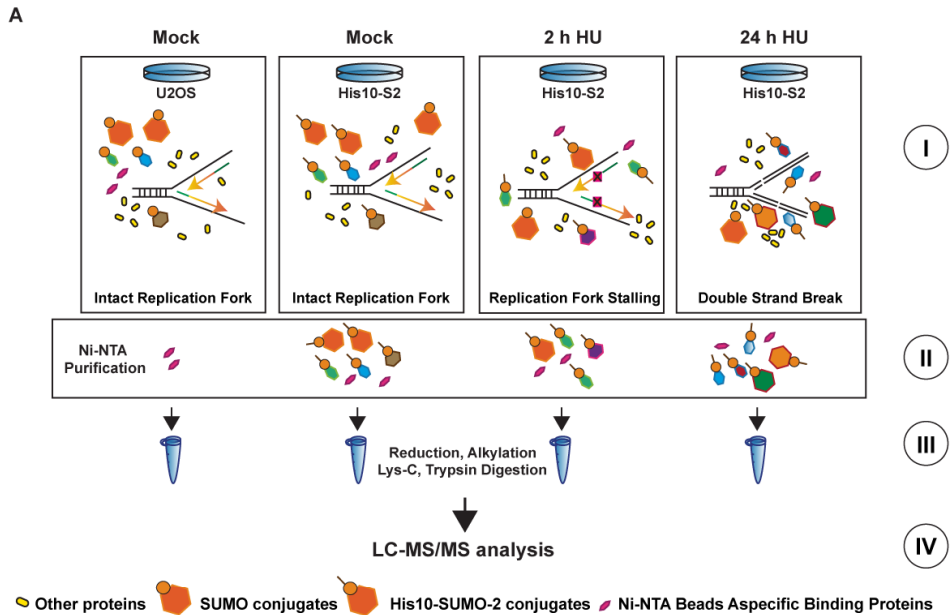


Figure 2. A strategy for discerning SUMOylation dynamics during replication stress. *A*, Cartoon depicting the strategy to study SUMOylation dynamics during replication stress. U2OS cells expressing His10-SUMO-2 were treated with 2 mM Hydroxyurea (HU) for 2 hours or 24 hours to induce DNA replication fork stalling and double strand breaks, respectively. Parental U2OS cells and U2OS cells expressing His10-SUMO-2 were mock treated as negative controls. SUMO-2 target proteins were purified by Ni-NTA purification. To study SUMO-2 targets that dynamically respond to replication stress, 5 biological replicates were performed. *B*, Purification of His10-SUMO-2 conjugates via NTA purification was confirmed by immunoblotting. Whole cell extracts and SUMO-2 purified proteins of the differently treated cells were run on 4-12% Bis-Tris polyacrylamide gels and levels of His10-SUMO-2 conjugates were compared by immunoblotting using anti-SUMO-2/3 antibody.

As indicated in Figure 3A and 3B, flow cytometry analysis from three independent experiments confirmed the enrichment of cells in the G1 phase and a decreased number of G2/M cells after 2 hours HU treatment, and a further decrease of G2/M phase cells upon 24 hours HU treatment, which confirmed stalling of the replication forks.

To further ratify that HU treatments induced the anticipated DNA damage response, we measured the formation of the phosphorylated histone variant H2AX (γ H2AX) foci after 2 hours and 24 hours HU treatment and mock treatment was used as negative control (Fig. 3C). As reported before (44), γ H2AX accumulated during 2 hours HU treatment and further increased numbers of foci were observed after 24 hours HU treatment. Furthermore, we checked the formation of Double Strand Break (DSB)-associated 53BP1 foci, and confirmed a large increase in foci upon 24 hours HU treatment (Fig. 3C).

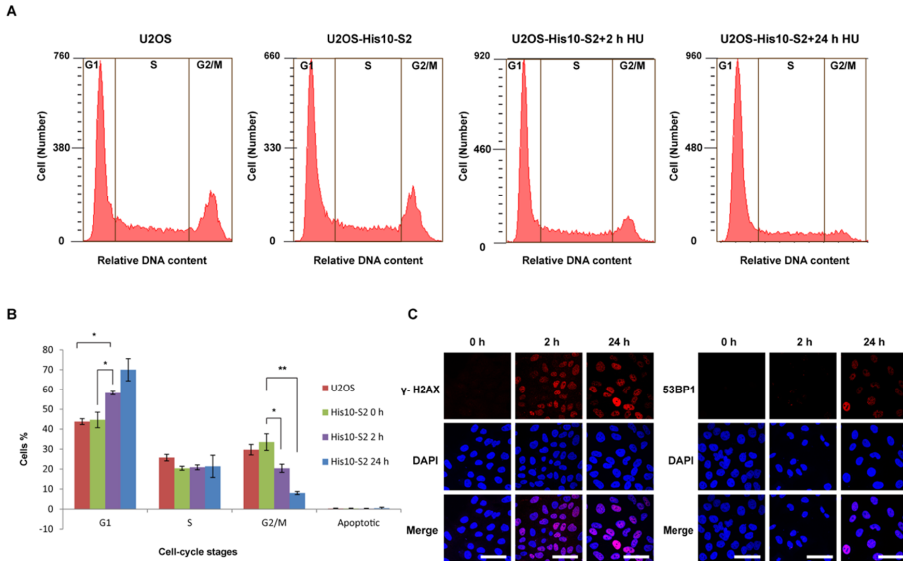


Figure 3. HU-induced DNA damage in U2OS stably expressing His10-SUMO-2. *A*, DNA content analysis of Hydroxyurea treated and non-treated cells. Flow cytometry was employed to confirm an increase in G1 phase cells upon HU treatment and a corresponding decrease in G2/M phase cells. *B*, The percentage of cells in each cell cycle phase is depicted. Error bars indicate the standard deviation from three independent replicates. Asterisks indicate significant differences by two-tailed Student's *t* testing. * $P < 0.05$, ** $P < 0.001$. *C*, Localization of γ H2AX and 53BP1 upon HU treatment. Cells were treated with 2 mM HU for 2 hours or 24 hours or left untreated. Cells were then fixed, permeabilized and immunostained for γ H2AX (red) or 53BP1 (red), and DNA was stained with DAPI (blue). Scale bars represent 75 μ m.

2.3 Identification of SUMOylated proteins using Label Free Quantification

His10-SUMO-2 purification coupled to mass spectrometry and label free quantification as described in the experimental procedures were used to study the abundance of SUMOylated proteins in response to replication stress (Fig. 4). In total, 5 biological replicates with three technical repeats for each condition were analyzed in this study. Label free quantification was performed using the MaxQuant software suite which quantifies proteins by MS1 peak intensities (41). Peak intensities measured during individual runs were matched to all other runs.

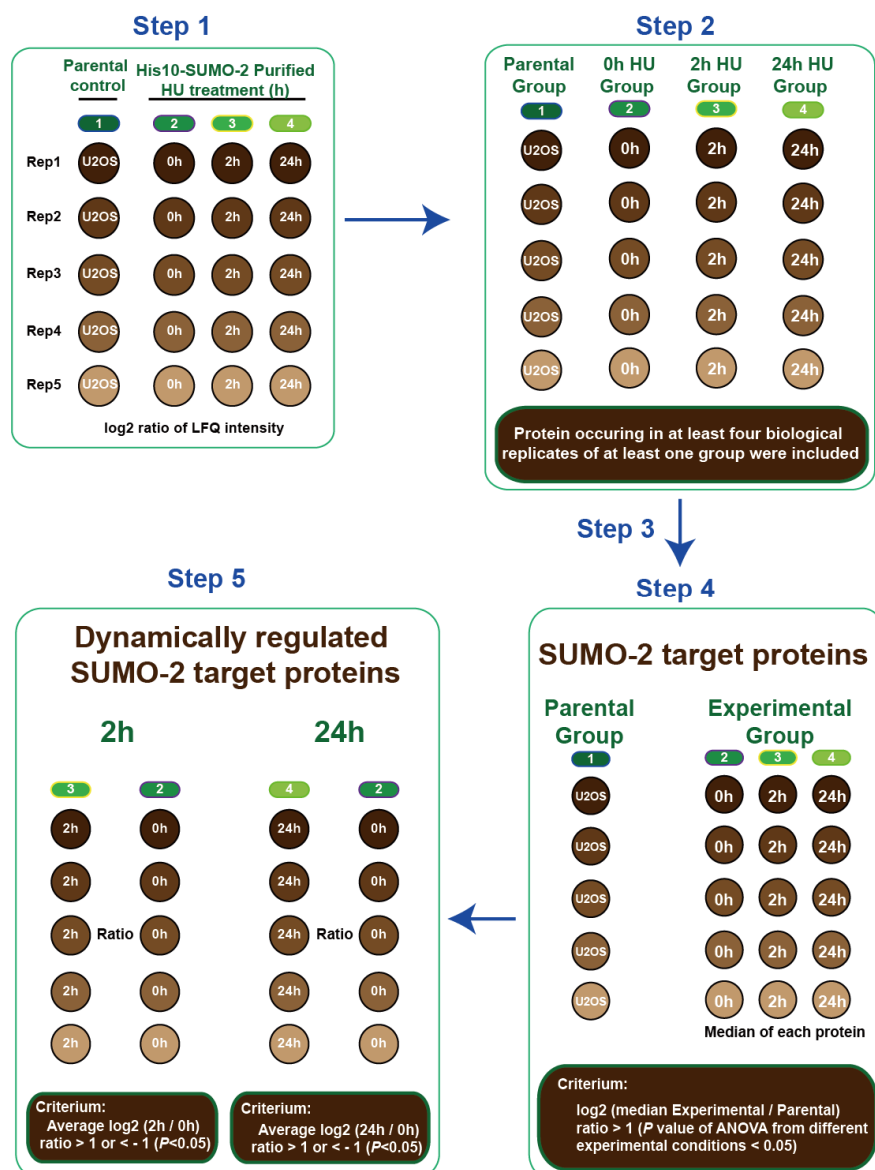


Figure 4. Label Free Quantification Strategy. Cartoon depicting our strategy for Label Free Quantification (LFQ) to select SUMO-2 target proteins and to identify significantly up- or down- regulated SUMO-2 target proteins in response to 2 hours or 24 hours HU treatment.

Step 1: Protein lists generated by MaxQuant were further analyzed by Perseus and LFQ intensities were \log_2 transformed.

Step 2: Different experiments were divided into four groups based on experimental conditions: a parental control group for U2OS control samples and three experimental groups for SUMO-2 samples purified from U2OS cells expressing His10-SUMO-2 treated with HU for 2 hours or 24 hours or mock treated. Inclusion criteria are depicted.

Step 3: Imputation of the missing values by normally distributed values with 1.8 downshift (\log_2) and 0.3 randomized width (\log_2).

Step 4: Proteins were considered as SUMO-2 target proteins using the indicated criteria.

Step 5: Significantly up- or down- regulated SUMO-2 target proteins in response to 2 hours or 24 hours HU treatment were identified as indicated.

2881 proteins were identified from 48821 peptides and 566 of them were considered as SUMO-2 target proteins in response to DNA replication stress (Fig. 5A and Table S1). After filtering out contaminants and putative false positives, all the LFQ intensities were transformed by \log_2 . The multiple scatter plot in Figure 5B shows high correlation within each condition throughout different biological replicates. SUMO-2 purified fractions showed more correlation than parental control due to specific enrichment of SUMOylated proteins by the affinity purification. Next, \log_2 ratios of all the LFQ intensities were used to generate a heat map by hierarchical clustering of all proteins. The heat map also visualized a high correlation between the biological replicates (Fig. 5C).

Subsequently, LFQ ratios corresponding to proteins derived from SUMO-2 enriched fractions purified from either the parental U2OS cell line or U2OS cells stably expressing His10-SUMO-2 were compared, in order to filter out non-specifically binding proteins (Fig. 4). After selecting proteins that were found in at least four biological replicates in at least one experimental condition in SUMO-2 purified samples, missing LFQ ratios were imputed as described in the experimental procedures. Proteins were considered as SUMO-2 target proteins when they were enriched at least two-fold from His10-SUMO-2 expressing cells compared to U2OS parental control cells. The SUMOylated protein list is provided in Table S1.

To assess the biological function of SUMOylated proteins identified in this study, we performed Gene Ontology (GO) term enrichment analysis using Perseus (Fig. 5D, Table S2 and Table S3). For GO Biological Processes, proteins involved in the DNA damage response were found to be significantly enriched. 53 proteins were related to DNA repair, 61 proteins were related to the DNA damage response. For GO Cellular Compartments, 300 proteins were found to be located in the nucleus. For GO Molecular Functions, 260 proteins were involved in DNA binding and 11 proteins were involved in damaged DNA binding.

2.4 Analysis of SUMOylated protein dynamics during replication stress

SUMOylation dynamics in response to replication stress was explored using bioinformatics analysis as described in experimental procedures and in Figure 4. Ten proteins were significantly increased in SUMOylation and 2 proteins were significantly decreased in SUMOylation after 2 hours HU treatment. After 24 hours HU treatment, 35 proteins were significantly increased in SUMOylation and 13 proteins were significantly decreased in SUMOylation (Table S4).

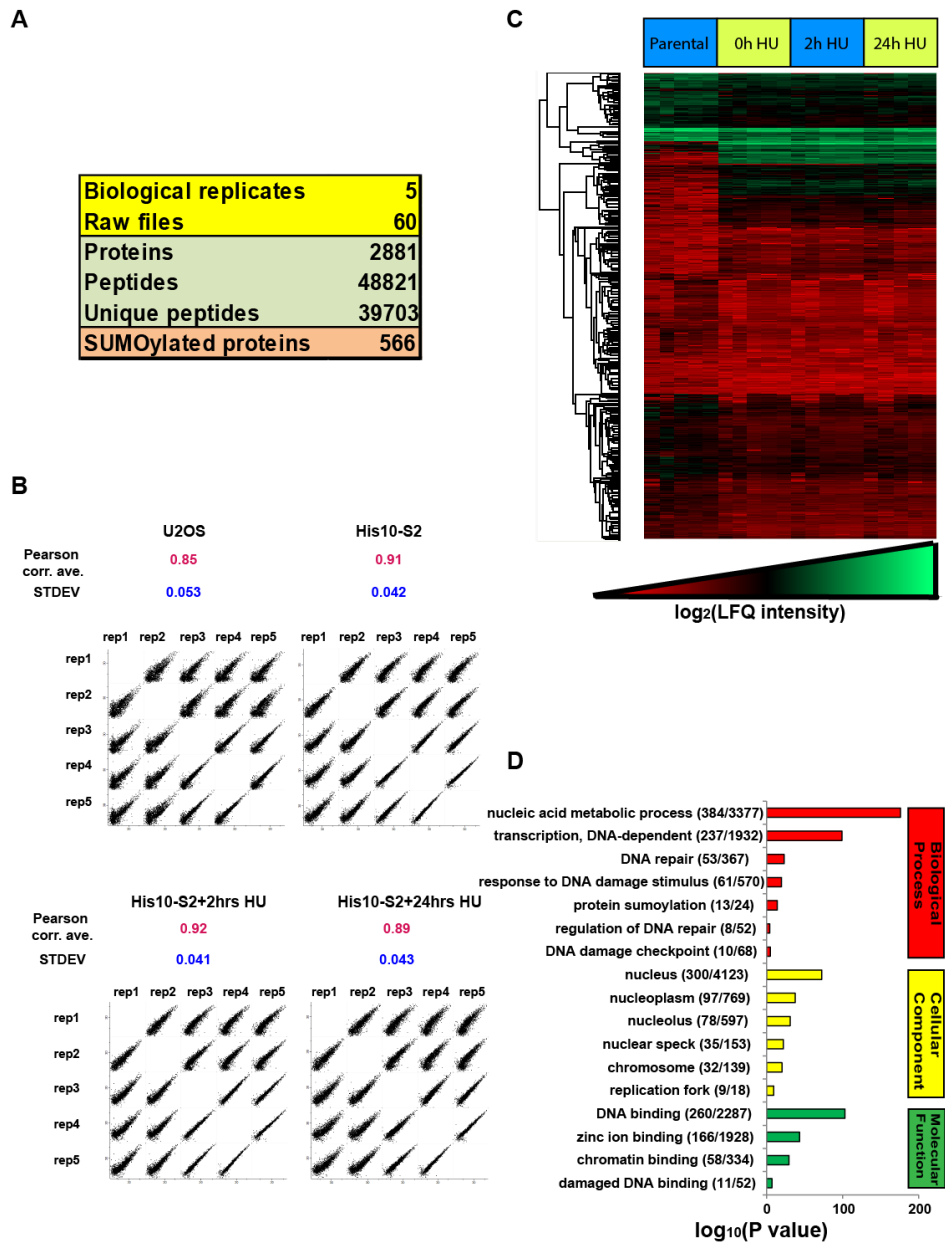


Figure 5. Overview of the SUMO proteomics results. *A*, Overview of the proteomic experiments. Out of 2,881 proteins identified with 48,821 peptides, 566 proteins were considered as SUMO-2 target proteins after filtering by LFQ intensities as described in Figure 4. *B*, LFQ intensity scatter plot. Each condition of each biological replicate was plotted together to visualize the correlation between the experiments. Pearson correlation averages were calculated for each condition and standard deviations (SD) are indicated. *C*, Heat map of log₂ LFQ intensities. Hierarchical clustering was performed for all identified

proteins. Within each biological replicate, the sample order from left to right was U2OS, U2OS His10-SUMO-2 (mock treated), U2OS His10-SUMO-2 (2 hours HU), and U2OS His10-SUMO-2 (24 hours HU). **D**, GO term enrichment analysis of the SUMOylated proteins identified. The bar chart shows GO terms for biological processes, cellular components and molecular functions.

Volcano plots shown in Figure 6A and 6B indicate the significance and magnitude of SUMO-2 target protein changes after 2h and 24h HU treatment. *P* values less than 0.05 were considered significant. STRING analysis of significantly regulated SUMOylated proteins was performed. Figure 6C shows the interaction of proteins significantly increased or decreased in SUMOylation after 2 hours of HU treatment. SUMOylation of CHAF1A and PCNA was significantly decreased. On the other hand, SUMOylation of MCM4, MYBL2 and FOXM1 was found to be increased. Similar to the finding of Li *et al.* (45), PCNA is a hub connecting several other SUMOylated proteins, including CHAF1A, FOXM1, MYBL2, ATRIP, and MCM4. After 24 hours of HU treatment (Fig. 6D), BHLHE41, CHAF1A, and DNMT1 were significantly decreased in SUMOylation. Additionally, BHLHE40, BARD1, MDC1, RMI1, and BRCA1 were greatly increased in SUMOylation. EME1 was found to be modestly down-regulated and additionally high-lighted in Figure 6B. Most of the SUMOylated proteins were significantly interacting to each other throughout the STRING network. In conclusion, DNA replication stress caused by HU treatment changed the SUMOylation of a distinct subset of proteins with key functions in the DNA damage response.

2.5 Site-specific SUMOylation dynamics during replication stress

Previously we have developed methodology to study SUMOylation at the site-specific level (Fig. S1) (38, 39). Here we used similar methodology to map SUMO-2 acceptor lysines in U2OS target proteins. First, we generated U2OS cells stably expressing lysine-deficient His10-SUMO-2 with a conserved mutation at its C-terminal part, Q87R, which mimics the localization of an arginine in yeast SUMO, Smt3. Characterization of the cell line is shown in Figure S2. We used this cell line to enrich SUMOylation sites from mock treated cells, or cells treated with HU for 2 hours or 24 hours to induce replication stress. In total, 1,043 SUMOylation sites were identified in 426 proteins (Tables S5 and S6), at a false discovery rate (FDR) below 1%. Mass accuracy was within 3 p.p.m. for 98.8% of all identified sites and within 6 p.p.m. for all sites, with an average absolute mass error of 0.78 p.p.m. We identified 83 peptides co-modified by SUMOylation and phosphorylation, of which 24 SUMOylated peptides were not found in the non-phosphorylated form (Table S9).

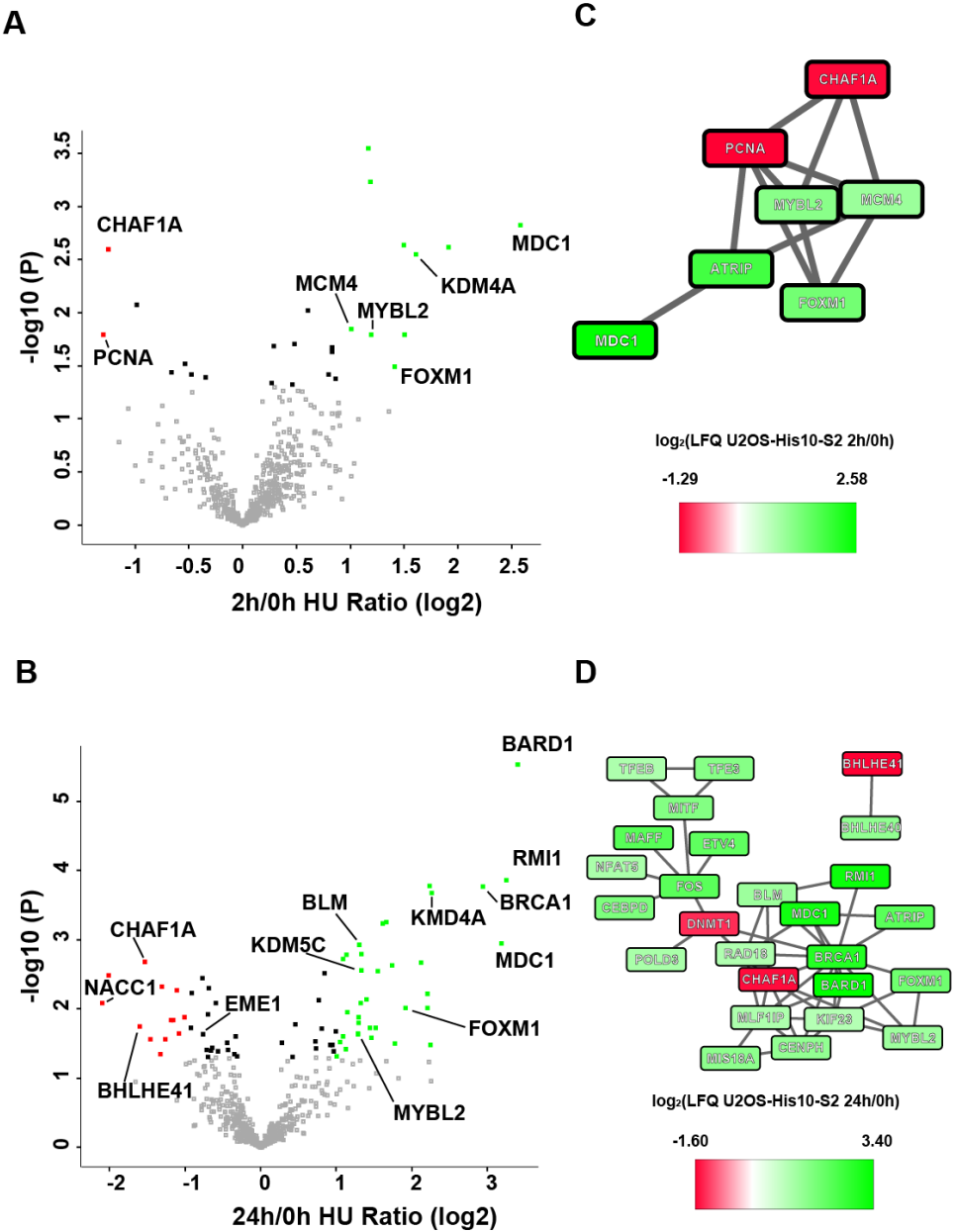


Figure 6. Volcano plots and STRING protein interaction network of dynamically regulated SUMO-2 target proteins. A, B, Volcano plots to show significantly altered SUMO-2 targets in response to 2 hours HU treatment (A) or 24 hours HU treatment (B). The $-\log_{10}(P)$ value of 2h/0h and 24h/0h from pairwise comparisons of SUMO-2 target proteins purified from mock treated cells and HU-treated cells were plotted against the average LFQ ratio 2h/0h (\log_2) and LFQ ratio 24h/0h (\log_2). The red dots represent proteins decreased for SUMOylation in response to HU with an average \log_2 ratio smaller than -1. The green dots represent proteins increased for SUMOylation in response to HU with an average \log_2 ratio greater than 1. C, STRING analysis of dynamically regulated SUMO-2 target proteins after 2 hours Hydroxyurea treatment. P value: 1.42×10^{-7} .

Upregulated SUMOylated proteins are colored in green and downregulated SUMOylated proteins are colored in red. **D**, STRING analysis of dynamically regulated SUMO-2 target proteins after 24 hours Hydroxyurea treatment. *P* value: 6.94×10^{-14} . Upregulated SUMOylated proteins are colored in green and downregulated SUMOylated proteins are colored in red.

In order to study the dynamics of SUMO-2 acceptor sites in response to HU treatment, SUMO-2 site MS1 peak intensities were normalized and compared between samples. Volcano plots were used to visualize SUMOylation site dynamics in response to 2 hours or 24 hours HU treatment. (Fig. 7A and 7B). Sites displaying a significant change of at least two-fold in response to HU treatment are highlighted. After 2 hours of HU treatment, 27 sites were significantly upregulated and 23 sites were significantly downregulated. After 24 hours of HU treatment, 49 sites were upregulated and 38 sites were downregulated (Table S7). Overall, the dynamic SUMO-2 sites correspond well with the results from our site-independent approach (Table S8).

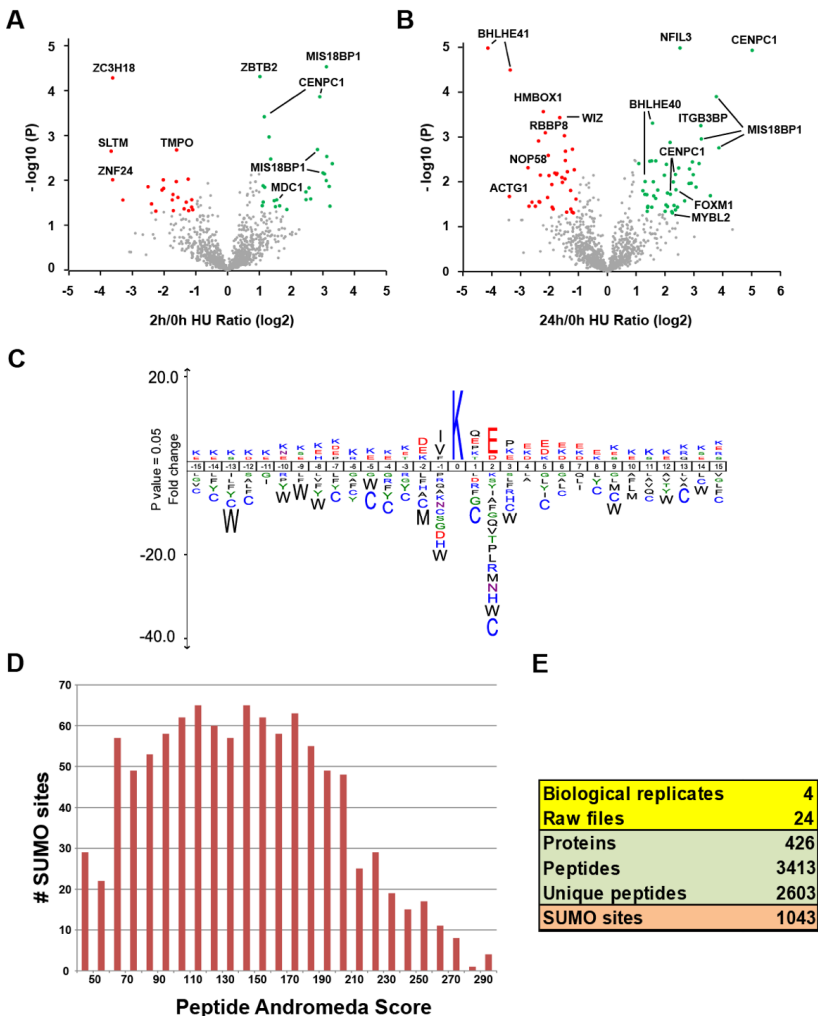


Figure 7. Volcano plots of dynamically regulated SUMOylation sites and SUMOylation motif analysis. **A, B,** Volcano plots showing dynamically regulated SUMO-2 acceptor sites in response to 2 hours HU treatment (A) or 24 hours HU treatment (B). The $-\log_{10}(P)$ value from pairwise comparisons of SUMO-2 acceptor lysines purified from mock treated cells and HU-treated cells were plotted against the average LFQ Ratio 2h /0h (\log_2) and LFQ Ratio 24h /0h (\log_2). The red dots represent sites decreased for SUMOylation in response to HU with an average \log_2 ratio smaller than -1.0 and with $P < 0.05$. The green dots represent sites increased for SUMOylation in response to HU with an average \log_2 ratio greater than 1.0 and with $P < 0.05$. **C,** All SUMO-2 acceptor lysines identified in this study (1,043 sites) were used to generate a SUMOylation motif employing IceLogo software. The height of the amino acid letters represents the fold change as compared to amino acid background frequency. All amino acid changes were significant with $P < 0.05$ by two-tailed Student's t test. **D and E,** Summary of the SUMO-2 acceptor lysines identified (E) with their peptide Andromeda scores (Median = 141.36) (D).

2.6 Verification of dynamic SUMO targets upon DNA replication stress.

We verified SUMOylation dynamics upon DNA replication stress by immunoblotting analysis for a subset of the identified dynamic SUMO-2 targets, including FOXM1, MYBL2, MDC1, and EME1 (Fig. 8). To study whether the HU treatment was efficient, flow cytometry was used to confirm the expected effects on cell cycle progression (Fig. 3A and 3B). All four of these SUMO-2 target proteins were found to be dynamically SUMOylated in accordance with the LFQ data derived from the mass spectrometry analysis, whereas the total amount of SUMO remained stable. As such, we demonstrated the feasibility of our approach, providing a powerful tool for analysis of SUMOylation dynamics in general, as well as a reliable resource of SUMO-2 target proteins dynamically regulated in response to replication stress.

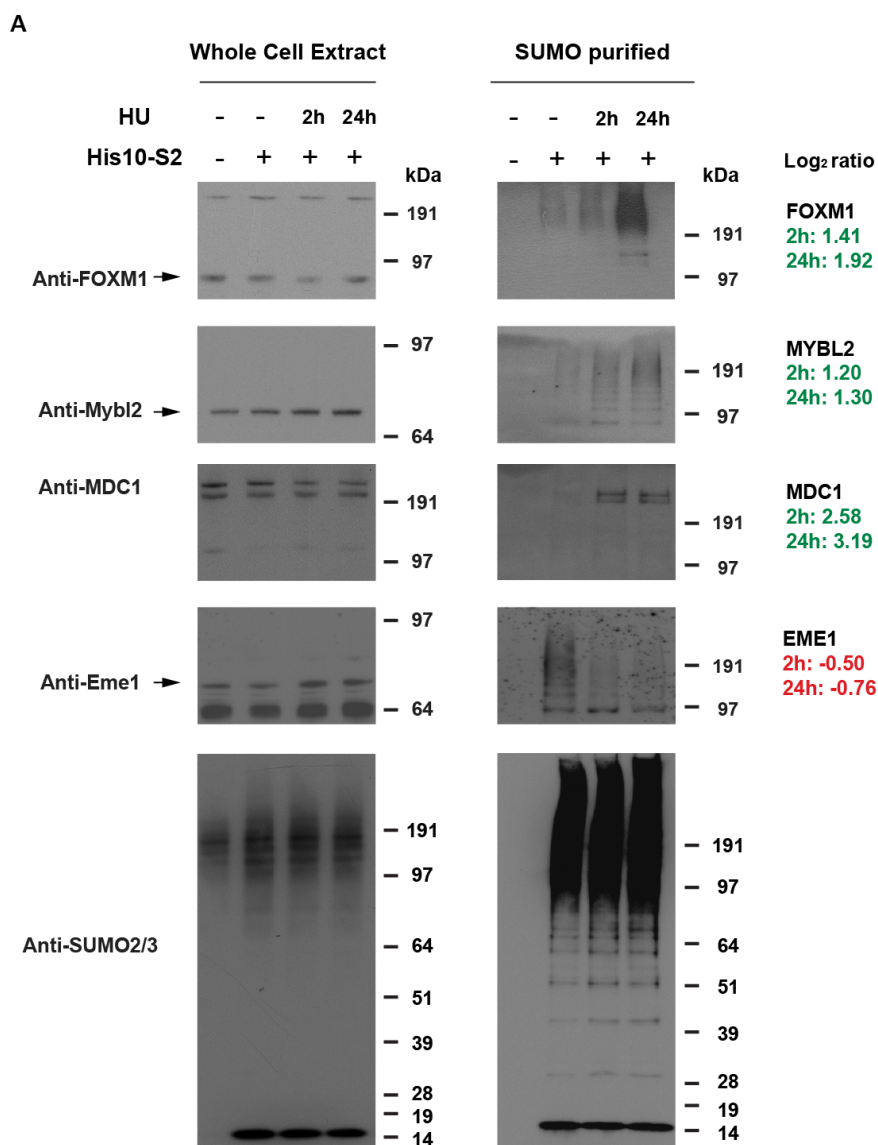


Figure 8. Verification of SUMO targets showing SUMOylation dynamics in response to replication stress. U2OS cells and U2OS cells expressing His10-tagged SUMO-2 were either mock treated or treated with Hydroxyurea (2 mM) for 2 hours or 24 hours as described, and His10-SUMO-2 conjugates were purified by Ni-NTA purification. SUMOylation dynamics induced by DNA replication stress was analyzed for four different SUMO-2 targets identified in the mass spectrometry screen using the indicated antibodies, and equal levels of SUMO conjugates in all samples were verified via immunoblotting using anti-SUMO-2/3 antibody. The fold changes in SUMOylation (log₂) of these proteins as found in our proteomics screen are indicated on the right.

3 Discussion

Our knowledge on the role of SUMOylation to maintain genome stability during replication is limited, due to limited insight into all involved SUMO target proteins. To address this, we have optimized a purification procedure to enrich SUMO target proteins, reaching a depth of 566 proteins. The optimized methodology employs the His10 tag, enabling the use of denaturing buffers to inactivate proteases and combining a high yield with a high purity; a major improvement over the His6 tag as a result of the usage of a much higher concentration of competing imidazole during the purification procedure to reduce the binding of contaminating proteins.

We have used the optimized methodology to study SUMOylation dynamics in response to HU-induced replication stress, resulting from the depletion of dNTPs required for DNA replication. A group of 12 dynamic SUMO-2 targets were identified when cells were treated for 2 hours with HU, including 10 upregulated and 2 downregulated proteins. When treating U2OS cells for 24 hours with HU, 48 dynamic SUMO-2 targets were identified including 35 upregulated and 13 downregulated proteins. As a cautionary note, we cannot exclude the possibility that changes in total levels of some proteins could underlie some of the observed changes in SUMOylation. More than half (2h: 70%, 24h: 52%) of these targets are functionally connected, indicating tight interactions between the SUMO-orchestrated proteins. The identified SUMO-regulated functional groups include key replication factors, DDR-components, a transcription-factor network, centromeric proteins and signal transducers. Identification of sites of modification is the most reliable manner to study post-translational modification of proteins (5). Powerful site-specific methodology is available to study phosphorylation and ubiquitination. In contrast, this has remained a major challenge in the SUMOylation field (5). Several years ago, we have developed a novel approach, enabling the identification of a limited set of SUMO-2 acceptor lysines in endogenous target proteins (38). Recently, we have further optimized this methodology by employing the His10-tag, enabling large-scale identification of SUMOylation sites (39). The use of a lysine-deficient version of SUMO-2 enabled the enrichment of SUMOylated peptides after Lys-C digestion. This is a key step to reduce the complexity of samples prior to mass spectrometry analysis. Using this methodology, we have identified 1,043 SUMOylation sites in this project, including 382 sites not previously identified (39). Similar methodology could be employed to study other ubiquitin-like modifications.

Proliferating Cell-Nuclear Antigen (PCNA) is one of the most strongly downregulated SUMO target proteins after two hours of HU treatment and is also one of the most highly connected components of the SUMO-target protein interaction network as depicted in Figure 6C. PCNA is a trimeric replication clamp that serves as platform for replicating polymerases. SUMOylated PCNA is known to interact with the helicase Srs2 to counteract recombination (18-20). Down-regulation of PCNA SUMOylation at early time points could represent a mechanism to explain increased recombination as a result of HU treatment, to counteract replication fork stalling (46).

MCM4 is a second key replication factor that we identified in our screen. MCM4 is a component of the DNA replication licensing factor Mini Chromosome Maintenance (MCM), which consists of MCM2-7 hexamers. The MCM complex plays an essential role in replication licensing and ensures that the entire genome is replicated exactly once during S-phase, avoiding reduplication or leaving genomic regions unreplicated (47). It acts as a helicase to unwind DNA, enabling access for the replication machinery to duplicate DNA. Interestingly, we previously found that MCM4 SUMOylation preferentially occurs during G1, a time point when MCM complexes are loaded on DNA for replication-licensing (48). Our data indicate a small increase in MCM4 SUMOylation in response to HU-treatment, which potentially could be involved in a cellular attempt to complete replication by firing of dormant origins in response to replication fork stalling (49). It is currently unclear how SUMOylation precisely regulates MCM4.

Furthermore, we identified a set of known SUMOylated DNA damage response factors, including MDC1, BRCA1 and BLM. MDC1 was previously found to be SUMOylated in response to Ionizing Radiation (28-31), and BRCA1 was found to be SUMOylated in response to cisplatin and HU (26). SUMOylation regulates the interaction of BLM and RAD51 at damaged replication forks (50) and the accumulation of ssDNA at stalled replication forks. The identification of SUMOylated BRCA1 and BLM is thus in agreement with the existing literature, and underlines the validity of our approach.

After prolonged exposure of cells to HU, SUMOylation of a cluster of centromeric proteins was induced, including CENPC1, CENPH and MIS18A (51) which is required for regulating CENPA deposition. Our data thus indicate co-regulation of centromeric proteins by SUMOylation in response to replication stress. The functional significance of these findings could further be explored, since other studies have demonstrated that SUMOylation at centromeres plays a key role to regulate cell cycle progression (23, 52-54).

Finally, it is interesting to note that SUMOylation regulates a number of factors involved in other modifications, demonstrating extensive signaling crosstalk. These proteins include the ATR-interacting protein ATRIP, ubiquitin E3 ligases RAD18 and BRCA1, the lysine-specific demethylases KDM5D, KDM5C and KDM4A. SUMOylation of HDAC1 and BRCA1 was previously demonstrated to promote enzymatic activity (26, 55), and it would be interesting to determine the functional relevance of SUMOylation for these other SUMO target proteins. Concerning crosstalk, we also obtained evidence for mixed SUMO-ubiquitin chains in our study, including linkages of SUMO-2 to lysines 11, 48 and 63 of ubiquitin. Furthermore, we found 83 peptides simultaneously modified by both SUMO-2 and phosphorylation (Table S9).

In summary, our study provides a critical framework to understand the role of SUMOylation to maintain genome stability during replication stress. The identified target proteins can be studied in detail at the functional level. Replication problems frequently lead to genome instability, one of the enabling characteristics of cancer cells (56). A detailed understanding of the role of SUMOylation in genome instability could potentially be employed to counteract oncogenesis, similarly to recent successful

efforts to disrupt signal transduction by the ubiquitin-like protein Nedd8 (57, 58). Moreover, the developed methodology could be widely employed to study SUMOylation, but also ubiquitination and signal transduction by other ubiquitin-like proteins.

Ultimately, our study reveals how SUMO regulates a network of target proteins in response to replication stress to coordinate the cellular DNA damage response. This network not only consists of known DNA damage response factors, but also includes replication factors, transcriptional regulators, chromatin modifiers and centromeric proteins, revealing how a post-translational modification is able to orchestrate a large variety of different proteins to integrate different nuclear processes with the aim of dealing with the induced DNA damage.

4 Experimental procedures

4.1 Antibodies

The primary antibodies used were: Mouse monoclonal anti-polyHistidine, clone HIS-1 (Sigma, H1029), mouse monoclonal anti-SUMO-2/3 (Abcam, ab81371), rabbit polyclonal anti-SUMO-2/3 (developed with Eurogentec) (37), rabbit polyclonal anti gamma-H2AX (Bethyl, A300-081A), rabbit polyclonal anti-53BP1 (Bethyl, A300-272A), mouse monoclonal anti EME1 (ImmuQuest, IQ284), rabbit polyclonal anti B-Myb (Mybl2) (Bethyl, A301-654A), rabbit polyclonal anti-FOXM1 (Santa Cruz Biotechnology, sc-502) and rabbit polyclonal anti-MDC1 (Bethyl, A300-052A).

4.2 Electrophoresis and immunoblotting

Whole cell extracts or purified protein samples were separated on Novex Bolt 4-12% Bis-Tris Plus gradient gels (Life Technologies) using MOPS buffer or via regular SDS-PAGE using a Tris-glycine buffer and transferred onto Hybond-C nitrocellulose membranes (GE Healthcare Life Sciences) using a submarine system (Life Technologies). Membranes were stained with Ponceau S (Sigma) to stain total protein and blocked with PBS containing 8% milk powder and 0.05% Tween-20 before incubating with the primary antibodies as indicated. Gels were stained with Coomassie using the Colloidal Blue Staining kit according to the manufacturer's instructions (Life Technologies).

4.3 Flow cytometry

Cells were harvested by a mild trypsin treatment, subsequently washed two times with phosphate-buffered saline (PBS) and resuspended in 1.3 mL of PBS. Afterwards, 3.75 mL of 100% ethanol was added and the cells were fixed at 4 °C overnight. On the day of flow cytometry analysis, the cells were centrifuged at 250 r.c.f. for 2 minutes, the supernatant was removed and the cells were washed with PBS and 2% fetal calf serum. Then, the cells were pelleted again and resuspended in 500 µL of PBS complemented with 2% fetal calf serum, 25 µg/mL propidium iodide (Sigma) and 100 µg/mL RNase A (Sigma) and then incubated for 30 minutes at 37 °C. FACS analysis was performed on a BD LSRII system and all gathered data was analyzed using BD FACS DIVA Software (BD Biosciences Clontech).

4.4 Cell culture, cell line generation and Hydroxyurea treatment

Cells were cultured in Dulbecco's modified Eagle's medium (Life Technologies) supplemented with 10% fetal bovine serum (Life Technologies) and 100 U/mL penicillin and 100 µg/mL streptomycin (Life Technologies). U2OS cells were infected using a bicistronic lentivirus encoding His10-SUMO-2 (His10-S2) and GFP separated by an IRES. Following infection, cells were sorted for low GFP levels using a FACSaria II (BD Biosciences). To induce DNA replication damage, an asynchronously growing cell population was incubated in medium containing 2 mM Hydroxyurea (Sigma) for either 2 hours or 24 hours. In all cases, the cells were then harvested and subjected to flow cytometry analysis, or Ni-NTA purification to enrich SUMO conjugates. For the proteome-wide identification of SUMO-2 acceptor lysines, U2OS cells were infected with bicistronic lentiviruses encoding His10-SUMO-2-K0-Q87R-IRES-GFP, abbreviated as His10-S2-K0-Q87R.

4.5 Immunofluorescence

Primary antibodies used for immunofluorescence were rabbit polyclonal anti-53BP1, rabbit polyclonal anti-phospho-Histone H2AX (Ser139) (gamma H2AX). Secondary antibody was anti-rabbit IgG AlexaFluor 594 (Life Technologies). Cells were cultured on circular glass slides in 24-well plates. After Hydroxyurea treatment for either 2 hours or 24 hours, medium was removed, cells were fixed with 4% paraformaldehyde for 20 minutes at room temperature in PBS and cells were permeabilized with 0.1% Triton X-100 in PBS for 15 minutes. Next, cells were washed twice with PBS and once with PBS with 0.05% Tween-20 (PBS-T). Cells were then blocked for 10 minutes with 0.5% blocking reagent (Roche) in 0.1 M Tris, pH 7.5 and 0.15 M NaCl (TNB) and incubated with primary antibody in TNB for one hour. Coverslips were washed five times with PBS-T and incubated with the secondary antibody in TNB for one hour. Next, coverslips were washed five times with PBS-T and dehydrated by washing once with 70% ethanol, once with 90% ethanol, and once with 100% ethanol. After drying the cells, coverslips were mounted onto a microscopy slide using citifluor/DAPI solution (500 ng/mL) and sealed with nail varnish. Images were recorded on a Leica TCS SP8 confocal microscope system equipped with 405, 488, 552 and 638-nm lasers for excitation, and a 63× lens for magnification, and were analyzed with Leica confocal software.

4.6 Purification of His10-SUMO-2 conjugates

U2OS cells expressing His10-SUMO-2 were washed, scraped and collected in ice-cold PBS. For total lysates, a small aliquot of cells was kept separately and lysed in 2% SDS, 1% N-P40, 50 mM TRIS pH 7.5, 150 mM NaCl. The remaining part of the cell pellets were lysed in 6 M guanidine-HCl pH 8.0 (6 M guanidine-HCl, 0.1 M Na₂HPO₄/NaH₂PO₄, 10 mM TRIS, pH 8.0). The samples were snap frozen using liquid nitrogen, and stored at -80°C.

For His10-SUMO-2 purification, the cell lysates were first thawed at room temperature and sonicated for 5 sec using a sonicator (Misonix Sonicator 3000) at 30 Watts to homogenize the lysate.

Protein concentrations were determined using the bicinchoninic acid (BCA) protein assay reagent (Thermo Scientific) and lysates were equalized. Subsequently, imidazole was added to a final concentration of 50 mM and β -mercaptoethanol was added to a final concentration of 5 mM. His10-SUMO-2 conjugates were enriched on nickel-nitrilotriacetic acid-agarose beads (Ni-NTA) (Qiagen), and subsequently the beads were washed using wash buffers A-D. Wash buffer A: 6 M guanidine-HCl, 0.1 M Na₂HPO₄/NaH₂PO₄ pH 8.0, 0.01 M Tris-HCl pH 8.0, 10 mM imidazole pH 8.0, 5 mM β -mercaptoethanol, 0.1% Triton X-100. Wash buffer B: 8 M urea, 0.1 M Na₂HPO₄/NaH₂PO₄ pH 8.0, 0.01 M Tris-HCl pH 8.0, 10 mM imidazole pH 8.0, 5 mM β -mercaptoethanol, 0.1% Triton X-100. Wash buffer C: 8 M urea, 0.1 M Na₂HPO₄/NaH₂PO₄ pH 6.3, 0.01 M Tris-HCl pH 6.3, 10 mM imidazole pH 7.0, 5 mM β -mercaptoethanol, no Triton X-100. Wash buffer D: 8 M urea, 0.1 M Na₂HPO₄/NaH₂PO₄ pH 6.3, 0.01 M Tris-HCl, pH 6.3, no imidazole, 5 mM β -mercaptoethanol, no Triton X-100. Wash buffers employed for immunoblotting experiments contained 0.2% Triton X-100. Samples were eluted in 7 M urea, 0.1 M NaH₂PO₄/Na₂HPO₄, 0.01 M Tris/HCl, pH 7.0, 500 mM imidazole pH 7.0. For site-specific purification, we used the strategy developed previously by our group (38, 39).

4.7 Sample preparation and mass spectrometry

SUMO-2 enriched samples were supplemented with 1 M Tris-(2-carboxyethyl)-phosphine hydrochloride (TCEP) to a final concentration of 5 mM, and incubated for 20 minutes at room temperature. Iodoacetamide (IAA) was then added to the samples to a 10 mM final concentration, and samples were incubated in the dark for 15 minutes at room temperature. Lys-C and Trypsin digestions were performed according to the manufacturer's specifications. Lys-C was added in a 1:50 enzyme-to-protein ratio, samples were incubated at 37 °C for 4 hours, and subsequently 3 volumes of 100 mM Tris-HCl pH 8.5 were added to dilute urea to 2 M. Trypsin (V5111, Promega) was added in a 1:50 enzyme-to-protein ratio and samples were incubated overnight at 37 °C. For site-specific samples preparation, we used the strategy developed previously by our group (39).

Subsequently, digested samples were desalted and concentrated on STAGE-tips as described previously (40) and eluted with 80% acetonitrile in 0.1% formic acid. Eluted fractions were vacuum dried employing a SpeedVac RC10.10 (Jouan, France) and dissolved in 10 μ L 0.1% formic acid before online nanoflow liquid chromatography-tandem mass spectrometry (nanoLC-MS/MS).

All the experiments were performed on an EASY-nLC 1000 system (Proxeon, Odense, Denmark) connected to a Q-Exactive Orbitrap or a Q-Exactive Plus Orbitrap (Thermo Fisher Scientific, Germany) through a nano-electrospray ion source. For the Q-Exactive, peptides were separated in a 13 cm analytical column with an inner-diameter of 75 μ m, in-house packed with 1.8 μ m C18 beads (Reprospher-DE, Pur, Dr. Manish, Ammerbuch-Entringen, Germany). The Q-Exactive Plus was coupled to 15 cm analytical columns with an inner-diameter of 75 μ m, in-house packed with 1.9 μ m

C18 beads (Reprospher-DE, Pur, Dr. Manish, Ammerbuch-Entringen, Germany), employing a column oven (PRSO-V1, Sonation, Biberach) to heat the column to 50 °C.

The gradient length was 120 minutes from 2% to 95% acetonitrile in 0.1% formic acid at a flow rate of 200 nL/minute. The mass spectrometers were operated in data-dependent acquisition mode with a top 10 method. Full-scan MS spectra were acquired at a target value of 3×10^6 and a resolution of 70,000, and the Higher-Collisional Dissociation (HCD) tandem mass spectra (MS/MS) were recorded at a target value of 1×10^5 and with a resolution of 17,500 with a normalized collision energy (NCE) of 25%. The maximum MS1 and MS2 injection times were 20 ms and 60 ms, respectively. The precursor ion masses of scanned ions were dynamically excluded (DE) from MS/MS analysis for 60 sec. Ions with charge 1, and greater than 6 were excluded from triggering MS2 events.

For samples enriched for identification of SUMO-2 acceptor lysines, a 120 minute gradient was used for chromatography. Data dependent acquisition with a top 5 method was used. Maximum MS1 and MS2 injection times were 20 ms and 250 ms, respectively. Resolutions, normalized collision energy and automatic gain control target were set as mentioned previously. Dynamic exclusion was set to 20 sec.

4.8 Protein level SUMOylation data analysis

For protein-level data analysis, four experimental conditions were performed in biological quintuplicate, and all samples were measured in technical triplicate, resulting in a total of 60 runs. All RAW data was analyzed using MaxQuant (version 1.4.1.2) and its integrated search engine Andromeda. The first search was carried out with 20 ppm, while the main search used 6 ppm for precursor ions. Mass tolerance of MS/MS spectra were set to 20 ppm to search against an *in silico* digested UniProt reference proteome for *Homo sapiens* (13 Nov 2013, 88704 proteins). Additionally, MS/MS data were searched against a list of 262 common mass spectrometry contaminants by Andromeda.

Database searches were performed with Trypsin/P and Lys-C specificity, allowing two missed cleavages. Carbamidomethylation of cysteine residues was considered as a fixed modification. Oxidation of methionine, phosphorylation and acetylation of protein N-termini were considered as variable modifications. Match between runs was performed with 0.7 min match time window and 20 min alignment time window. The minimum peptide length was set to 6. Protein groups considered for quantification required at least 2 peptides, including unique and razor peptides. Peptides and proteins were identified with a false discovery rate of 1% according to Cox *et al.* (41).

Label-free quantification was carried out, as described in Figure 4, using MaxLFQ without fast LFQ. Proteins identified by the same set of peptides were combined to a single protein group by MaxQuant.

Protein lists generated by MaxQuant were further analyzed by Perseus (version 1.5.0.31) (42). Proteins identified as a common contaminant were filtered out, and then all the LFQ intensities were \log_2 transformed as described in Figure 4, step 1. Scatter plots were generated for each experimental

condition to compare the differences between biological replicates and to derive Pearson correlations. Different experiments were annotated into four groups as described in Figure 4, step 2: a control group for the parental U2OS cell line, a 0 hours HU group (untreated), a 2 hours HU group and a 24 hours HU group for the different time points of HU treatment of the U2OS cell line expressing His10-SUMO-2. Proteins identified in at least one treatment condition and found in at least four biological replicates were included for further analysis. For each experimental condition individually, missing values were imputed using Perseus software by normally distributed values with a 1.8 downshift (\log_2) and a randomized 0.3 width (\log_2) as described in Figure 4, step 3. For heat maps, the \log_2 ratios of LFQ intensities were plotted by hierarchical clustering to compare biological replicates. For identification of SUMO-2 target proteins, selection criteria are detailed in Figure 4 step 4. Proteins are considered to be SUMO-2 target proteins when the median \log_2 ratio of the LFQ intensity in the experimental group minus the median \log_2 ratio of the LFQ intensity in the parental control group was greater than 1 and the P value of ANOVA was smaller than 0.05. Term enrichment analysis (Gene Ontology) of SUMOylated proteins by protein annotation was carried out using Perseus. Term enrichment was determined by Fisher Exact testing, and P values were corrected for multiple hypotheses testing using the Benjamini and Hochberg false discovery rate. Final corrected P values were filtered to be less than 0.02.

Subsequently, average \log_2 ratios for 2 hours of HU treatment (LFQ ratio 2h/ 0h HU treatment) and average \log_2 ratios for 24 hours of HU treatment (LFQ ratio 24h/ 0h HU treatment) were calculated, and P values of each protein across all treatment conditions were calculated by ANOVA, using permutation-based FDR. SUMOylated proteins were considered significantly up- or down-regulated when their average \log_2 ratio (2h/0h HU treatment) or average \log_2 ratio (24h/0h HU treatment) was greater than 1 or less than -1, and corresponding ANOVA P values were less than 0.05; selection criteria are detailed in Figure 4 step 5. Volcano plots to demonstrate significant changes in protein SUMOylation upon HU treatment were created by plotting the Student's t -test $-\log_{10}(P)$ value of LFQ 2h/0h against the average \log_2 (LFQ ratio 2h/0h) value, and the Student's t -test $-\log_{10}(P)$ value of 24h/0h against the average \log_2 (LFQ ratio 24h/0h). Significantly regulated SUMOylated proteins after 2 hours or 24 hours of HU treatment were selected to perform functional protein interaction analysis by STRING (string-db.org, version 9.1) using a confidence score of medium or higher ($p > 0.4$). STRING analysis results were visualized using Cytoscape (version 3.1.0).

4.9 Site level SUMOylation data analysis

Site-specific purifications were performed in biological quadruplicate, and all samples were measured in technical duplicate. All 24 RAW files were analyzed by MaxQuant (version 1.5.1.2). The first search was carried out with a mass accuracy of 20 ppm, while the main search used 6 ppm for precursor ions. Database searches were performed with Trypsin/P specificity, allowing two missed cleavages. Carbamidomethylation of cysteine residues was considered as a fixed modification. Mass tolerance of

MS/MS spectra was set to 20 ppm to search against an *in silico* digested UniProt reference proteome for *Homo sapiens* (29 October 2014, 88812 proteins). Additionally, MS/MS data was searched against a list of 245 common mass spectrometry contaminants by Andromeda. Pyro-QQTGG (K), QQTGG (K) and phosphorylation (STY) were set as variable modifications. In order to increase identification certainty, diagnostic peaks were searched within MS/MS spectra corresponding to SUMOylated peptides. The default minimum Andromeda score for accepting MS/MS spectra corresponding to modified peptides was left at 40 and the minimum delta score was left at 6. Additionally identified Pyro-QQTGG and QQTGG sites without diagnostic peaks, matching reversed sequences, or having a localization probability of less than 0.9 were excluded. Match between runs was used with 0.7 min match time window and 20 min alignment time window. Sequence windows of -15 to +15 with respect to the identified SUMO-2 acceptor lysines were used to generate a SUMOylation motif, employing iceLogo software.

Intensities of Pyro-QQTGG and QQTGG were further analyzed by Perseus (1.5.0.31). Pyro-QQTGG and QQTGG site intensities were normalized by the total site-peptide intensities in each sample, and then all the intensities were \log_2 transformed. Experiments were annotated into three groups: untreated control, 2 hours HU treatment and 24 hours HU treatment. Sites occurring in at least one treatment condition and identified in all four biological repeats, were included for further analysis. Missing values were imputed using Perseus software by normally distributed values with a 1.8 downshift (\log_2) and a randomized 0.3 width (\log_2), for each experimental condition individually. Volcano plots to demonstrate significant changes in protein SUMOylation upon HU treatment were created by plotting $-\log_{10}(P)$ values of 2h/0h HU and 24h/0h HU from pairwise comparisons of SUMO-2 target proteins against the average LFQ Ratio 2h/0h (\log_2) and LFQ Ratio 24h/0h (\log_2).

Acknowledgements

This work was supported by the Marie Curie Initial Training Network UPStream of the European Union (A.C.O.V. and J.V.O.). This work was furthermore supported by the European Research Council (A.C.O.V.) and the Netherlands Organization for Scientific Research (NWO) (A.C.O.V.). The NNF Center for Protein Research is supported by a generous donation from the Novo Nordisk Foundation. None of the authors have a financial interest related to this work.

The MS data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository with the dataset identifier PXD001736.

References

1. Deribe, Y. L., Pawson, T., and Dikic, I. (2010) Post-translational modifications in signal integration. *Nat. Struct. Mol. Biol.* 17, 666-672
2. Scott, J. D., and Pawson, T. (2009) Cell Signaling in Space and Time: Where Proteins Come Together and When They're Apart. *Science* 326, 1220-1224
3. Choudhary, C., and Mann, M. (2010) Decoding signalling networks by mass spectrometry-based proteomics. *Nat. Rev. Mol. Cell Biol.* 11, 427-439
4. Komander, D., and Rape, M. (2012) The Ubiquitin Code. *Annu. Rev. Biochem.* 81, 203-229
5. Vertegaal, A. C. O. (2011) Uncovering Ubiquitin and Ubiquitin-like Signaling Networks. *Chem. Rev.* 111, 7923-7940
6. Kerscher, O., Felberbaum, R., and Hochstrasser, M. (2006) Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu. Rev. Cell Dev. Biol.* 22, 159-180
7. Flotho, A., and Melchior, F. (2013) Sumoylation: a regulatory protein modification in health and disease. *Annu. Rev. Biochem.* 82, 357-385
8. Gareau, J. R., and Lima, C. D. (2010) The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nat. Rev. Mol. Cell Biol.* 11, 861-871
9. Hickey, C. M., Wilson, N. R., and Hochstrasser, M. (2012) Function and regulation of SUMO proteases. *Nat. Rev. Mol. Cell Biol.* 13, 755-766
10. Rodriguez, M. S., Dargemont, C., and Hay, R. T. (2001) SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. *J. Biol. Chem.* 276, 12654-12659
11. Bernier-Villamor, V., Sampson, D. A., Matunis, M. J., and Lima, C. D. (2002) Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. *Cell* 108, 345-356
12. Ulrich, H. D. (2012) Ubiquitin and SUMO in DNA repair at a glance. *J. Cell Sci.* 125, 249-254
13. Jackson, S. P., and Durocher, D. (2013) Regulation of DNA damage responses by ubiquitin and SUMO. *Mol. Cell* 49, 795-807
14. Bergink, S., and Jentsch, S. (2009) Principles of ubiquitin and SUMO modifications in DNA repair. *Nature* 458, 461-467
15. Nacerddine, K., Lehenbre, F., Bhaumik, M., Artus, J., Cohen-Tannoudji, M., Babinet, C., Pandolfi, P. P., and Dejean, A. (2005) The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. *Dev. Cell* 9, 769-779
16. Hoege, C., Pfander, B., Moldovan, G. L., Pyrowolakis, G., and Jentsch, S. (2002) RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* 419, 135-141
17. Stelter, P., and Ulrich, H. D. (2003) Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. *Nature* 425, 188-191
18. Pfander, B., Moldovan, G. L., Sacher, M., Hoege, C., and Jentsch, S. (2005) SUMO-modified PCNA recruits Srs2 to prevent recombination during S phase. *Nature* 436, 428-433
19. Armstrong, A. A., Mohideen, F., and Lima, C. D. (2012) Recognition of SUMO-modified PCNA requires tandem receptor motifs in Srs2. *Nature* 483, 59-63
20. Papouli, E., Chen, S., Davies, A. A., Huttner, D., Krejci, L., Sung, P., and Ulrich, H. D. (2005) Crosstalk between SUMO and ubiquitin on PCNA is mediated by recruitment of the helicase Srs2p. *Mol. Cell* 19, 123-133
21. Mao, Y., Sun, M., Desai, S. D., and Liu, L. F. (2000) SUMO-1 conjugation to topoisomerase I: A possible repair response to topoisomerase-mediated DNA damage. *Proc. Natl. Acad. Sci. U. S. A* 97, 4046-4051
22. Mao, Y., Desai, S. D., and Liu, L. F. (2000) SUMO-1 conjugation to human DNA topoisomerase II isozymes. *J. Biol. Chem.* 275, 26066-26073
23. Dawlaty, M. M., Malureanu, L., Jeganathan, K. B., Kao, E., Sustmann, C., Tahk, S., Shuai, K., Grosschedl, R., and van Deursen, J. M. (2008) Resolution of sister centromeres requires RanBP2-mediated SUMOylation of topoisomerase IIalpha. *Cell* 133, 103-115
24. Eladad, S., Ye, T. Z., Hu, P., Leversha, M., Beresten, S., Matunis, M. J., and Ellis, N. A. (2005) Intra-nuclear trafficking of the BLM helicase to DNA damage-induced foci is regulated by SUMO modification. *Hum. Mol. Genet.* 14, 1351-1365
25. Galanty, Y., Belotserkovskaya, R., Coates, J., Polo, S., Miller, K. M., and Jackson, S. P. (2009) Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks. *Nature* 462, 935-939
26. Morris, J. R., Boutell, C., Keppler, M., Densham, R., Weekes, D., Alamshah, A., Butler, L., Galanty, Y., Pangon, L., Kiuchi, T., Ng, T., and Solomon, E. (2009) The SUMO modification pathway is involved in the BRCA1 response to genotoxic stress. *Nature* 462, 886-890

27. Danielsen, J. R., Povlsen, L. K., Villumsen, B. H., Streicher, W., Nilsson, J., Wikstrom, M., Bekker-Jensen, S., and Mailand, N. (2012) DNA damage-inducible SUMOylation of HERC2 promotes RNF8 binding via a novel SUMO-binding Zinc finger. *J. Cell Biol.* 197, 179-187
28. Luo, K., Zhang, H., Wang, L., Yuan, J., and Lou, Z. (2012) Sumoylation of MDC1 is important for proper DNA damage response. *EMBO J.* 31, 3008-3019
29. Galanty, Y., Belotserkovskaya, R., Coates, J., and Jackson, S. P. (2012) RNF4, a SUMO-targeted ubiquitin E3 ligase, promotes DNA double-strand break repair. *Genes Dev.* 26, 1179-1195
30. Yin, Y., Seifert, A., Chua, J. S., Maure, J. F., Golebiowski, F., and Hay, R. T. (2012) SUMO-targeted ubiquitin E3 ligase RNF4 is required for the response of human cells to DNA damage. *Genes Dev.* 26, 1196-1208
31. Vyas, R., Kumar, R., Clermont, F., Helfricht, A., Kalev, P., Sotiropoulou, P., Hendriks, I. A., Radaelli, E., Hocheppied, T., Blanpain, C., Sablina, A., van, A. H., Olsen, J. V., Jochemsen, A. G., Vertegaal, A. C., and Marine, J. C. (2013) RNF4 is required for DNA double-strand break repair in vivo. *Cell Death Differ.* 20, 490-502
32. Psakhye, I., and Jentsch, S. (2012) Protein group modification and synergy in the SUMO pathway as exemplified in DNA repair. *Cell* 151, 807-820
33. Cremona, C. A., Sarangi, P., Yang, Y., Hang, L. E., Rahman, S., and Zhao, X. (2012) Extensive DNA damage-induced sumoylation contributes to replication and repair and acts in addition to the mec1 checkpoint. *Mol. Cell* 45, 422-432
34. Branzei, D., Sollier, J., Liberi, G., Zhao, X., Maeda, D., Seki, M., Enomoto, T., Ohta, K., and Foiani, M. (2006) Ubc9- and mms21-mediated sumoylation counteracts recombinogenic events at damaged replication forks. *Cell* 127, 509-522
35. Branzei, D., Vanoli, F., and Foiani, M. (2008) SUMOylation regulates Rad18-mediated template switch. *Nature* 456, 915-920
36. Branzei, D., and Foiani, M. (2010) Maintaining genome stability at the replication fork. *Nat. Rev. Mol. Cell Biol.* 11, 208-219
37. Vertegaal, A. C., Ogg, S. C., Jaffray, E., Rodriguez, M. S., Hay, R. T., Andersen, J. S., Mann, M., and Lamond, A. I. (2004) A proteomic study of SUMO-2 target proteins. *J. Biol. Chem.* 279, 33791-33798
38. Matic, I., Schimmel, J., Hendriks, I. A., van Santen, M. A., van de Rijke, F., van, D. H., Gnad, F., Mann, M., and Vertegaal, A. C. (2010) Site-specific identification of SUMO-2 targets in cells reveals an inverted SUMOylation motif and a hydrophobic cluster SUMOylation motif. *Mol. Cell* 39, 641-652
39. Hendriks, I. A., D'Souza, R. C., Yang, B., Verlaan-de, V. M., Mann, M., and Vertegaal, A. C. (2014) Uncovering global SUMOylation signaling networks in a site-specific manner. *Nat. Struct. Mol. Biol.* 21, 927-936
40. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* 2, 1896-1906
41. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367-1372
42. Cox, J., and Mann, M. (2012) 1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data. *BMC Bioinformatics* 13 Suppl 16:S12
43. Jones, R. M., and Petermann, E. (2012) Replication fork dynamics and the DNA damage response. *Biochem. J.* 443, 13-26
44. Petermann, E., Orta, M. L., Issaeva, N., Schultz, N., and Helleday, T. (2010) Hydroxyurea-Stalled Replication Forks Become Progressively Inactivated and Require Two Different RAD51-Mediated Pathways for Restart and Repair. *Mol. Cell* 37, 492-502
45. Ma, L., Aslanian, A., Sun, H. Y., Jin, M. J., Shi, Y., Yates, J. R., and Hunter, T. (2014) Identification of Small Ubiquitin-like Modifier Substrates with Diverse Functions Using the Xenopus Egg Extract System. *Mol. Cell. Proteomics* 13, 1659-1675
46. Galli, A., and Schiestl, R. H. (1996) Hydroxyurea induces recombination in dividing but not in G1 or G2 cell cycle arrested yeast cells. *Mutat. Res.* 354, 69-75
47. Blow, J. J., and Dutta, A. (2005) Preventing re-replication of chromosomal DNA. *Nat. Rev. Mol. Cell Biol.* 6, 476-486
48. Schimmel, J., Eifler, K., Sigurdsson, J. O., Cuijpers, S. A. G., Hendriks, I. A., Verlaan-de Vries, M., Kelstrup, C. D., Francavilla, C., Medema, R. H., Olsen, J. V., and Vertegaal, A. C. O. (2014) Uncovering SUMOylation Dynamics during Cell-Cycle Progression Reveals FoxM1 as a Key Mitotic SUMO Target Protein. *Mol. Cell* 53, 1053-1066
49. Ge, X. Q., Jackson, D. A., and Blow, J. J. (2007) Dormant origins licensed by excess Mcm2-7 are required for human cells to survive replicative stress. *Genes Dev.* 21, 3331-3341
50. Ouyang, K. J., Woo, L. L., Zhu, J., Huo, D., Matunis, M. J., and Ellis, N. A. (2009) SUMO modification regulates BLM and RAD51 interaction at damaged replication forks. *PLoS Biol.* 7, e1000252

51. Fujita, Y., Hayashi, T., Kiyomitsu, T., Toyoda, Y., Kokubu, A., Obuse, C., and Yanagida, M. (2007) Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1. *Dev. Cell* 12, 17-30
52. Zhang, X. D., Goeres, J., Zhang, H., Yen, T. J., Porter, A. C., and Matunis, M. J. (2008) SUMO-2/3 modification and binding regulate the association of CENP-E with kinetochores and progression through mitosis. *Mol. Cell* 29, 729-741
53. Bachant, J., Alcasabas, A., Blat, Y., Kleckner, N., and Elledge, S. J. (2002) The SUMO-1 isopeptidase Smt4 is linked to centromeric cohesion through SUMO-1 modification of DNA topoisomerase II. *Mol. Cell* 9, 1169-1182
54. Mukhopadhyay, D., Arnaoutov, A., and Dasso, M. (2010) The SUMO protease SENP6 is essential for inner kinetochore assembly. *J. Cell Biol.* 188, 681-692
55. David, G., Neptune, M. A., and DePinho, R. A. (2002) SUMO-1 modification of histone deacetylase 1 (HDAC1) modulates its biological activities. *J. Biol. Chem.* 277, 23658-23663
56. Hanahan, D., and Weinberg, R. A. (2011) Hallmarks of Cancer: The Next Generation. *Cell* 144, 646-674
57. Soucy, T. A., Smith, P. G., Milhollen, M. A., Berger, A. J., Gavin, J. M., Adhikari, S., Brownell, J. E., Burke, K. E., Cardin, D. P., Critchley, S., Cullis, C. A., Doucette, A., Garnsey, J. J., Gaulin, J. L., Gershman, R. E., Lublinsky, A. R., McDonald, A., Mizutani, H., Narayanan, U., Olhava, E. J., Peluso, S., Rezaei, M., Sitchak, M. D., Talreja, T., Thomas, M. P., Traore, T., Vyskocil, S., Weatherhead, G. S., Yu, J., Zhang, J., Dick, L. R., Claiborne, C. F., Rolfe, M., Bolen, J. B., and Langston, S. P. (2009) An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature* 458, 732-736
58. Nawrocki, S. T., Griffin, P., Kelly, K. R., and Carew, J. S. (2012) MLN4924: a novel first-in-class inhibitor of NEDD8-activating enzyme for cancer therapy. *Expert Opin. Investig. Drugs* 21, 1563-1573

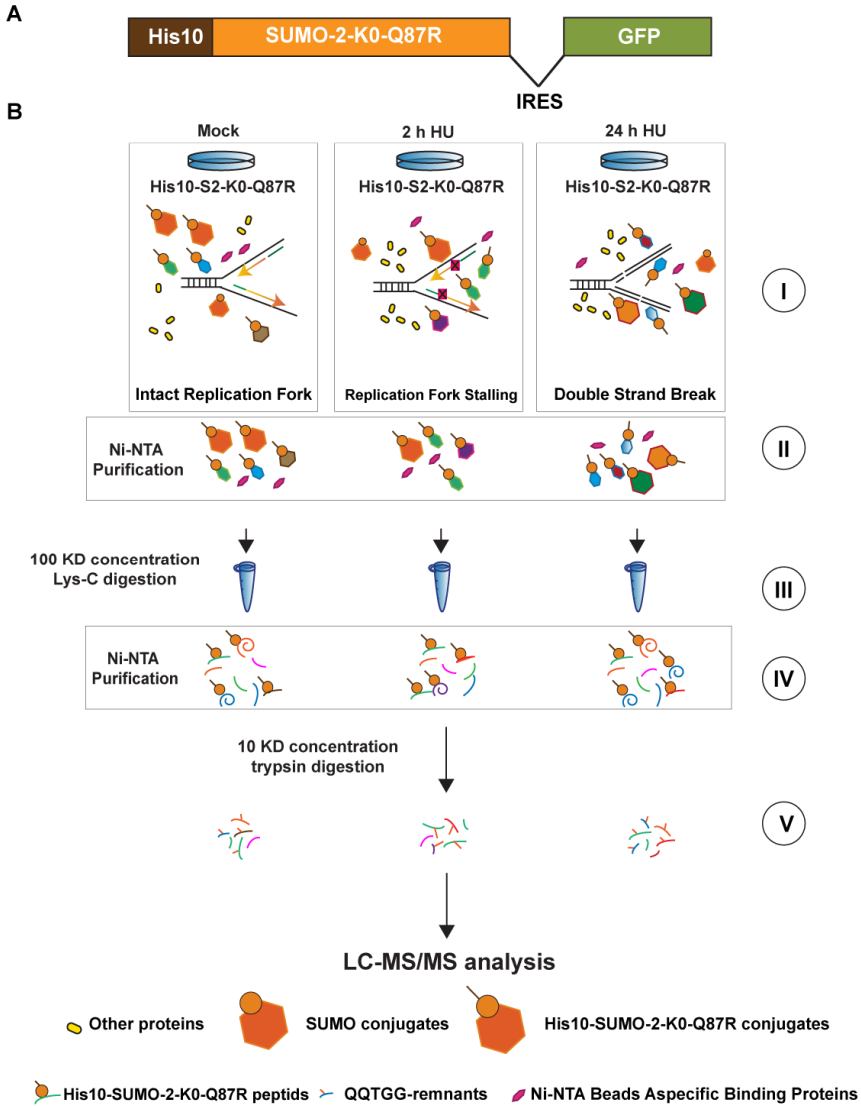


Figure S1. A strategy for identifying SUMO-2 acceptor lysines in endogenous proteins during replication stress.

A. Schematic representation of the His10-SUMO-2-K0-Q87R-IRES-GFP construct used in this project. **B.** Cartoon depicting our strategy to identify SUMO-2 acceptor lysines and their dynamics during replication stress. U2OS cells expressing His10-SUMO-2-K0-Q87R were treated with 2 mM Hydroxyurea (HU) for 2 hours or 24 hours to induce DNA replication fork stalling and double strand breaks, respectively. U2OS cells expressing His10-SUMO-2-K0-Q87R cells were mock treated as negative controls. SUMO-2 target proteins were enriched on Ni-NTA beads. SUMOylated peptides were obtained by Lys-C digestion and Ni-NTA re-purification and subsequently analyzed by mass spectrometry.

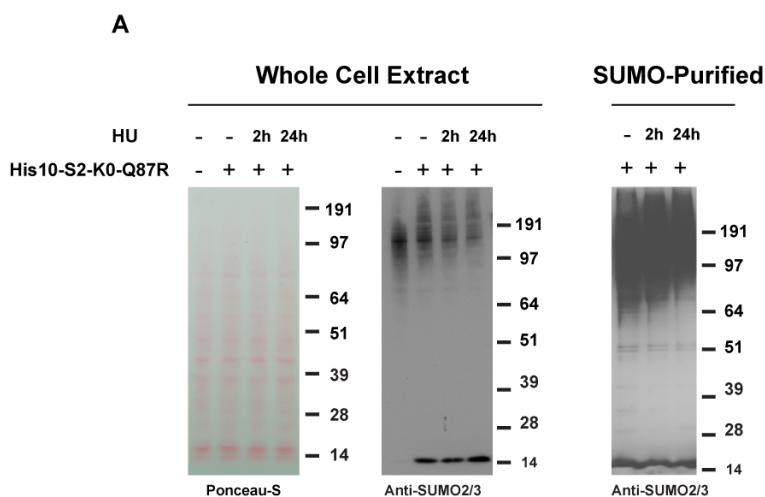


Figure S2. A U2OS cell line stably expressing His10-SUMO-2-K0-Q87R. Immunoblotting analysis was used to verify the expression levels of SUMO-2 in U2OS cells stably expressing His10-SUMO-2-K0-Q87R (His10-K0-S2-Q87R).

Chapter 3

Proteomics reveals global regulation of protein SUMOylation by ATM and ATR kinases in replication stress

Stephanie Munk^{1,2,4}, Jón Otti Sigurðsson^{1,4}, Zhenyu Xiao^{3,4}, Tanveer Singh Batth¹, Giulia Franciosa¹, Louise von Stechow¹, Andres Joaquin Lopez-Contreras², Alfred Cornelis Otto Vertegaal^{3*}, Jesper Velgaard Olsen^{1,5*}

¹Proteomics program, Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen, Denmark.

²Center for Chromosome Stability and Center for Healthy Aging, Institute for Cellular and Molecular Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen, Denmark.

³Department of Molecular Cell Biology, Leiden University Medical Center, 2300 RC Leiden, The Netherlands.

⁴These authors contributed equally

⁵Lead Contact

*Correspondence should be directed to J.V.O (jesper.olsen@cpr.ku.dk) and A.C.O.V (A.C.O.Vertegaal@lumc.nl)

Chapter 3 has been published in **Cell reports**

Cell Rep. 2017 Oct 10;21(2):546-558. doi: 10.1016/j.celrep.2017.09.059.

Supplementary Tables are available online

Summary

The mechanisms that protect eukaryotic DNA during the cumbersome task of replication depend on the precise coordination of several post-translational modifications (PTMs)-based signaling networks. Phosphorylation is a well-known regulator of the replication stress response and recently an essential role for SUMO (small ubiquitin-like modifiers) has also been established. Here we investigate the global interplay between phosphorylation and SUMOylation in response to replication stress. Using the latest SUMO- and phospho-proteomics technologies, we identified thousands of regulated modification sites. We found co-regulation of central DNA damage and replication stress responders of which the ATR activating factor, TOPBP1 was the most highly regulated. Using pharmacological inhibition of the apical DNA damage response kinases, ATR and ATM, we found these to regulate global protein SUMOylation in the protein networks that protect DNA upon replication stress and fork breakage. Combined, we uncovered integration between phosphorylation and SUMOylation in the cellular systems that protect DNA integrity.

1 Introduction

DNA replication is a tremendously challenging, time consuming and vital task for eukaryotic organisms. The maintenance of genomic integrity during this process is challenged by endogenous and exogenous factors that cause replication forks to slow, stall and in extreme cases this leads to DNA breakage (Halazonetis et al., 2008). Cells are equipped with a complex DNA damage response (DDR), consisting of protein networks that enable them to cope with replication stress (RS), and a malfunction in these systems can result in genomic instability and oncogenesis (Jackson and Bartek, 2009). These protective signaling pathways require the precise spatial and temporal coordination of DDR components, which is achieved by dynamic and specific post-translational modifications (PTMs) (Polo and Jackson, 2011). In particular, protein phosphorylation is the well-established driver of the RS response, with the ATR (Ataxia telangiectasia and Rad3-related protein) kinase functioning as the key initiator and orchestrator (López-Contreras and Fernandez-Capetillo, 2010; Shiloh, 2001). Depletion of this central kinase leads to replication fork breakage and genomic instability, thereby instigating a phosphorylation response mounted by the ATM (Ataxia telangiectasia mutated) kinase, which mediates repair and checkpoint activation upon double strand DNA breaks (DSBs) (Murga et al., 2009; Smith et al., 2010). ATM and ATR belong to the same atypical serine/threonine kinase family (the PIKK-related kinases) with similar substrate sequence specificity (Kim et al., 2009), yet they have unique triggers. While ATR responds to the accumulation of single stranded DNA (ssDNA) and regulates replication, ATM is the key mediator of the cellular response to DSBs. DNA-PK is the third member of this kinase family, however its functions are confined to local repair processes (Meek et al., 2008).

Phosphorylation, however, must act in concert with other PTMs, such as ubiquitylation, to elicit efficient responses to genotoxic insults (Ulrich and Walden, 2010). The functions of PTMs in the DNA damage and RS responses have therefore been subject of intense investigations, individually (Beli et al., 2012; Bennetzen et al., 2009; Danielsen et al., 2011; Jungmichel et al., 2013) and in concert (Gibbs-Seymour et al., 2015; González-Prieto et al., 2015; Hunter, 2007). More recently, studies have revealed the significance of protein SUMOylation in the DDR and deregulation of the SUMO system has been shown to confer genomic instability (Bergink and Jentsch, 2009; Bursomanno et al., 2015; Jackson and Durocher, 2013; Xiao et al., 2015). Using various RS inducing agents, these studies have shown that the SUMOylation status of a number of proteins is modulated when DNA replication is perturbed (García-Rodríguez et al., 2016). Furthermore, it has been demonstrated that phosphorylation and SUMOylation intersect at various levels (Gareau and Lima, 2010). A phosphorylation-dependent SUMO modification (PDSM) motif has been suggested to prime SUMOylation (Hietakangas et al., 2006) by enhancing the binding of the SUMO E2 enzyme UBC9 (Mohideen et al., 2009), and phosphorylation was also found to regulate the function of SUMO interacting motifs (SIMs) (Stehmeier and Muller, 2009). However, a potential global coordination of the SUMOylation response and the well-known phosphorylation response to RS remains unexplored.

Quantitative mass spectrometry (MS)-based proteomics and developments in enrichment methodologies have seen tremendous developments in recent years (Hendriks and Vertegaal, 2016). State-of-the-art MS technologies allow for the identification of thousands of SUMOylation sites (Hendriks et al., 2017; Lamoliatte et al., 2014, 2017; Schimmel et al., 2014; Tammsalu et al., 2014), and tens of thousands of phosphorylation sites from cellular systems (Francavilla et al., 2017; Mertins et al., 2016; Olsen et al., 2010). In this study, we utilized complementary proteomics strategies to identify the interplay between the global SUMOylation and phosphorylation responses to replication stressors. We identified regulation of thousands of phosphorylation sites and hundreds of SUMOylation sites in response to treatment with the DNA inter-strand crosslinking (ICL) agent mitomycin C (MMC) and to hydroxyurea (HU), with a number of proteins co-regulated by both PTMs. Our investigations revealed that the well-established apical responders to RS and RS induced DSBs, namely ATR and ATM, both modulate protein SUMOylation at various stages of the RS response. Our findings not only identify an intersection between phosphorylation and SUMOylation in the RS response, but also reveal further levels of signaling regulation in this response by the two most prominent kinases of the DNA damage and RS responses.

2 Results

2.1 Global SUMOylation changes upon MMC treatment

To investigate the interplay between the SUMOylation and phosphorylation responses to RS, we treated U-2-OS osteosarcoma cells with MMC (Figure 1A). MMC, a widely-used chemotherapeutic agent in treatment of various cancers, induces ICLs, thereby impeding normal replication fork progression and causing RS. To study the effects of MMC during DNA replication, cells were synchronized at the G1/S checkpoint by 24 hours of thymidine blocking, and were thereafter released into S-phase with or without MMC for 8 hours (Figure 1B, Figure S1A). After an 8 hours release into MMC, western blotting confirmed increased phosphorylation of checkpoint kinases, CHK1 at S435 and CHK2 at T68, as well as increased levels of phosphorylation of S140 on histone H2A.X (γ H2AX) (Figure S1B). These phosphorylation sites are known targets of ATR and ATM indicating that our experimental conditions generate RS (ATR activation) and DSBs (ATM activation).

For MS based global analysis of SUMOylation we used two previously described SUMO-enrichment approaches to quantify changes in protein SUMOylation and SUMO acceptor sites (Hendriks et al., 2014; Schimmel et al., 2014) on a global scale (Figure 1B). SUMOylated proteins were identified and quantified by immuno-precipitation (IP) of SUMO2-conjugated proteins from U-2-OS cells stably expressing FLAG-SUMO2-Q87R (Figure 1C and Figure S1C). The Q87R mutation allows for identification of SUMO after tryptic digestion due to the resulting remnant (Schimmel et al., 2014). To confidently distinguish SUMOylated from non-SUMOylated proteins, control IPs were also performed from the parental U-2-OS cell line, as non-SUMOylated proteins would be underrepresented

in these compared to FLAG-SUMO-Q87R expressing cells (Figure 1B). Complementarily, we mapped SUMOylation acceptor sites by enrichment of SUMOylated peptides from His10-tagged SUMO2-K0-Q87R expressing U-2-OS cells (Figure 1B) (Xiao et al., 2015). Tryptic peptides from all enriched samples were analyzed by nano-scale liquid chromatography tandem MS (LC-MS/MS) on a Q-Exactive HF instrument (Kelstrup et al., 2014). We used stable isotope labeling by amino acids in cell culture (SILAC) (Ong, 2002) for accurate MS-based quantification and differentially labeled SILAC cells showed comparable cell-cycle distributions upon synchronization (Figure S1A). The SUMO2 expression levels in the two stable cell lines were 3 to 4 fold higher than in the parental cells as observed by MS full scans from proteome measurements and by western blotting (Figures S1D and S1E).

All raw LC-MS/MS files were processed and analyzed together using the MaxQuant software suite (www.maxquant.org) with one percent false discovery rate at peptide, site and protein levels (Cox and Mann, 2008). From this analysis, we confidently identified 3,453 proteins (Table S1). Ratios from proteome measurements of these conditions revealed that the protein abundances in the MMC treated FLAG-SUMO2-Q87R cells were largely unchanged compared to the equivalently treated parental cells. We therefore reasoned that we could determine the proteins significantly SUMOylated in the FLAG-SUMO2-Q87R cells using ratio cutoffs of two standard deviations from the mean (95th percentile) of this ratio distribution (Figure S1F). This analysis resulted in a cutoff of 1.7 fold change, by which 702 proteins were deemed SUMOylated (Figure 1D and Table S1). Using the same strategy for the MMC treated and untreated FLAG-SUMO2-Q87R cells, a resulting ratio cutoff of 1.5 resulted in 187 proteins being having significantly increased SUMOylation upon treatment with MMC (Figures 1D, S1G and Table S1). Additionally, we mapped 311 unique SUMO acceptor sites (Figure 1E). Sequence motif analysis of these showed a strong preference for a glutamate two residues downstream from the modified lysine (Figure 1E), conforming to the previously described SUMOylation consensus motif (ΨKXE) (Sampson et al., 2001). By separately analyzing SUMOylated peptides with or without this motif, we found that indeed the known SUMO consensus motif is the predominant, with the inversed SUMO motif as the second most overrepresented (Figure 1E).

To determine the cellular compartments and biological processes in which the SUMOylated proteins are involved, we performed a Gene Ontology (GO) enrichment analysis. In agreement with previous studies, we found that the majority of SUMOylation occurs on nuclear proteins that are involved in transcription (Figure S1H) (Flotho and Melchior, 2013). Further to this, among the proteins with MMC regulated SUMOylation, we identified 24 transcription factors, for which 24 target genes were found to be co-regulated by at least two of these. Interestingly, these target genes were highly enriched in proteins involved in apoptosis and cancer development (Table S1). GO analysis of the 187 proteins with increased SUMOylation after MMC treatment also revealed this trend, and furthermore these proteins are involved in histone ubiquitylation and DNA repair (Figure 1F).

Many of the identified proteins known to function in DNA repair clustered together in a functional network based on STRING database analysis (Szklarczyk et al., 2015). Fanconi anemia factors, BRCA1

(Breast cancer type 1 susceptibility protein) and TOPBP1 (DNA topoisomerase 2-binding protein 1) were among the regulated SUMOylated proteins after MMC treatment (Figure 1G). These proteins are well-known to play important roles in response to ICL-induced RS and DNA damage. The regulation of SUMOylation levels on these proteins upon MMC treatment indicates that this modification may modulate their function in this response.

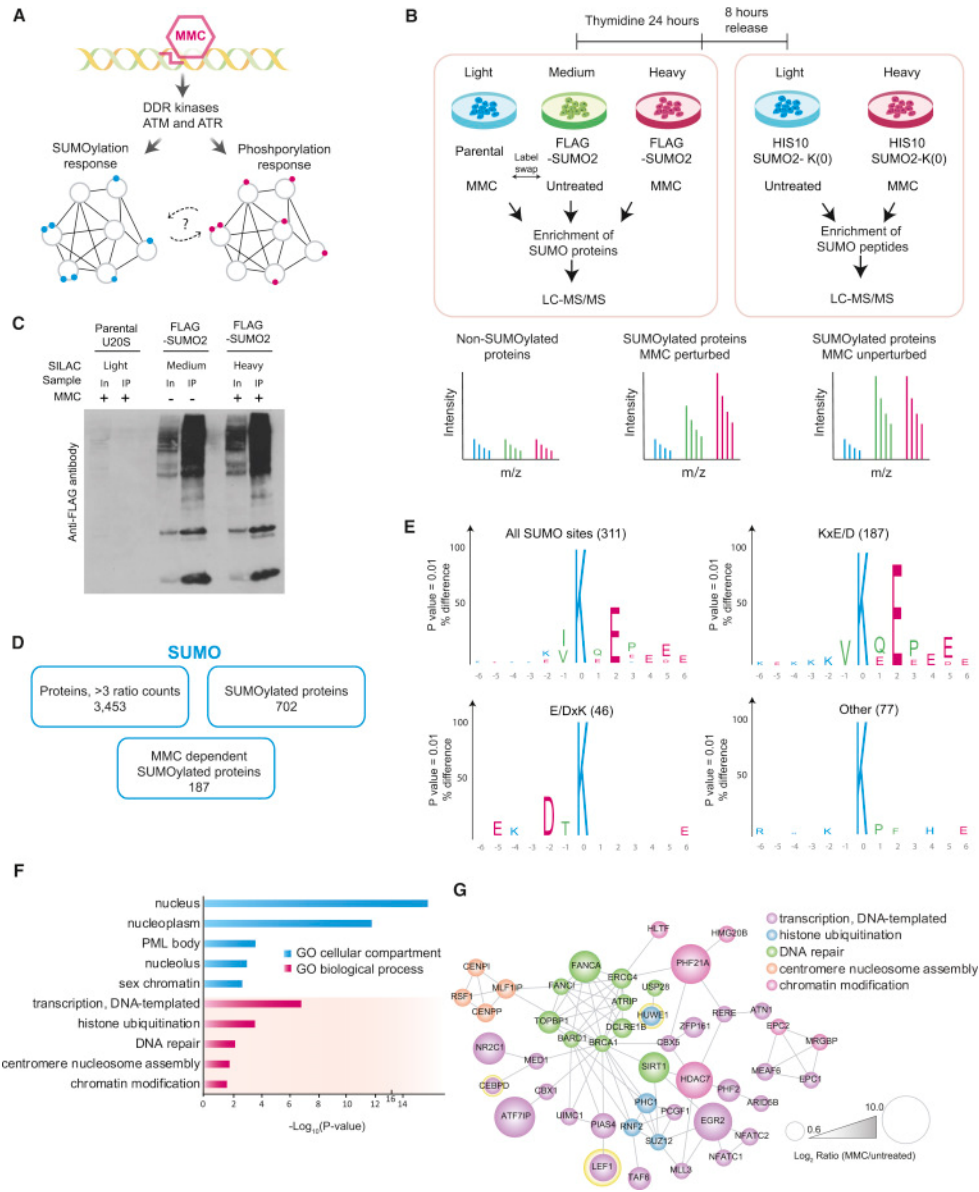


Figure 1 Proteomics analysis of SUMOylation changes upon MMC treatment. **A)** Schematic representation of the aim to study a potential interplay between phosphorylation and SUMOylation in MMC induced RS. **B)** Experimental design for proteomics analysis of SUMOylated proteins from FLAG-SUMO2 and His10-SUMO2 expressing U-2-OS cells to enrich SUMOylated proteins and peptides respectively. **C)** Western blot analysis of SUMO enriched proteins from SILAC labeled U-2-OS cells stably transfected with FLAG-SUMO2. Cells were synchronized and treated as in (A). **D)** Results of proteomics analysis. **E)** Motif analysis of SUMOylation acceptor sites. **F)** Enrichment analysis of GO cellular compartments (GOCC) and biological processes (GOBP) of MMC regulated SUMOylated proteins, using InnateDB. **G)** Functional network analysis of proteins from the GOBP terms enriched in (F). (See also Figure S1 and Table S1).

2.2 Global phosphorylation changes upon MMC treatment

To study the potential interplay between the SUMOylation and phosphorylation responses to MMC, we used a streamlined quantitative phosphoproteomics workflow (Batth et al., 2014) to enrich phosphopeptides from FLAG-SUMO2-Q87R U-2OS cells synchronized and treated with MMC in the same manner as for SUMOylation mapping. Tryptic digests of whole cell lysates were separated by offline high pH reversed-phase fractionation and phospho-peptides were enriched with TiO₂ beads prior to LC-MS/MS (Figure S2A). We quantified 20,900 high confidence phosphorylated sites, of which 650 were induced (SILAC ratio above 1.5) after 8 hours of MMC treatment (Figure 2A and Table S2). Proteins with induced phosphorylation were primarily nuclear and involved in DNA repair as determined by GO analysis, similar to our findings for SUMOylated proteins that were induced by MMC treatment (Figure 2B and S2B).

We performed sequence motif analysis of the 650 up-regulated phosphorylation sites to identify protein kinases that were activated in the response to MMC treatment. A strong overrepresentation of glutamine (Q) at the position directly C-terminal to the phosphorylation sites (P+1) indicated activation of the ATM and ATR kinases, both of which are known to preferentially phosphorylate substrates on serine/threonine residues that are followed by a glutamine (S/T-Q) (Figure 2C). Indeed, we find that 170 (26%) of the phosphorylation sites up-regulated by MMC treatment confer to the S/T-Q motif. Moreover, MS spectra show a clear induction of ATM and ATR target phosphorylation sites on ATM itself and CHK1, respectively (Figure 2D). Conversely, phosphorylation sites on proteins from other signaling pathways, as exemplified by ERK1, remained largely unperturbed (Figure 2D). Functional network analysis of the proteins with increased phosphorylation reveals two highly interconnected clusters of phosphoproteins involved in the DDR, DNA replication and cell cycle (Figure 2E). A number of these proteins were also found to have increased SUMOylation, indicating that phosphorylation and SUMOylation are modulating proteins in the same pathways in the RS response to MMC treatment.

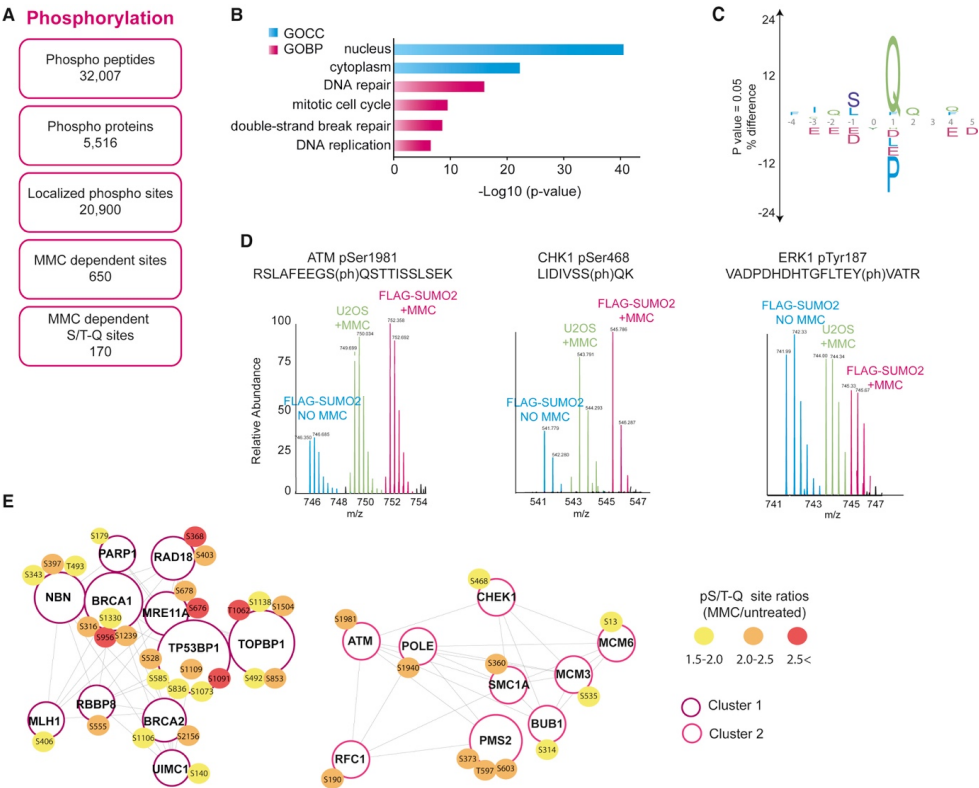


Figure 2 Phosphoproteomics analysis of MMC treated cells. **A)** Overview of number of phosphorylated peptides and proteins from phosphoproteomics analysis of cells treated as shown in Figure S2A. **B)** GOCC and GOBP analysis of proteins with regulated phosphorylation sites after MMC treatment, using InnateDB. **C)** Motif enrichment analysis of 360 MMC dependent phosphorylation sites, done with IceLogo. **D)** Full MS spectra of phosphorylated peptides from ATM, CHK1 and ERK1. **E)** Two highly interconnected MCODE clusters from functional network analysis of all proteins with regulated phosphorylation sites. MCODE was set to determine clusters with the 'Haircut' approach, a minimum node score cutoff of 0.2, K-core was set to 2 and max depth to 100. (See also Figure S2 and Table S2).

2.3 Central DDR proteins are highly phosphorylated and SUMOylated in the response to MMC

To elaborate on this hypothesis and uncover a potential interplay between the SUMOylation and phosphorylation responses to MMC, we integrated our large-scale proteomics datasets of the two modifications. First, we evaluated the datasets for potential biases arising from the MS strategies used for enrichment and detection of proteins with these modifications. The distribution of the relative protein copy numbers (iBAQ values) from the proteome, the phosphorylated proteins and the SUMOylated proteins in these datasets revealed that all three groups of proteins had similar distribution patterns with no apparent abundance biases (Figure S3A). We then assessed the overlap between the

datasets and found that 540 proteins harbored at least one SUMOylation and phosphorylation event (Figure 3A). This comprises two-thirds of the SUMOylated proteins we identified, corresponding to the proportion of the total proteome that is reported to be phosphorylated at any given time (Olsen et al., 2010). While only 17 of these proteins were found to have up-regulation of both modifications upon MMC treatment, this subset included UIMC1 (BRCA1-A complex subunit RAP80), BRCA1, BARD1 (BRCA1-associated RING domain protein 1), and TOPBP1, which are proteins with well-established key functions in the DDR (Figure 3A, 3B and Table S3). We therefore find that quantitative analysis of proteins co-regulated by both PTMs is powerful means to determine and prioritize key players in cellular signaling networks.

To elaborate on the mechanism of regulation of these two PTMs in RS, we further investigated the roles of most prominent DDR and RS activated kinases, namely ATR and ATM, in modulating RS induced SUMOylation (Smith et al., 2010). These kinases are the well-known initiators and key modulators of the global phosphorylation and ubiquitylation responses to DNA damage and RS (Shiloh, 2001). Indeed, ATR is activated upon 8 hours of MMC treatment after thymidine release, as observed by increased phosphorylation of its direct target CHK1 on S345, which can further be attenuated with an ATR inhibitor (ATRi) (Figure S3B). Interestingly, TOPBP1, an important co-activator of ATR, was the highest co-modified protein upon MMC treatment (Figure 3B). By SUMO enrichment from both the FLAG-SUMO2-Q87R and His10-tagged SUMO2-K0-Q87R cells, we were able to confirm that indeed TOPBP1 SUMOylation is increased over time with MMC treatment (Figure 3C). Since the His10-based pull-down procedures involved lysis and enrichment under harsh denaturing conditions, these findings confidently demonstrate that TOPBP1 is indeed differentially SUMOylated by RS and that the observed changes are not due to TOPBP1 interactions with other SUMO-regulated target proteins. Interestingly, TOPBP1 SUMOylation was further induced upon co-treatment of MMC with ATRi, also at earlier time points (Figure 3D). Although TOPBP1 SUMOylation is increased upon treatment with MMC or ATRi only, the combination of the two is required for massive hyper-SUMOylation (Figure 3D). ATM is also activated in these conditions as indicated by increased CHK2 and H2A.X phosphorylation (Figure 3D and S3B), and interestingly the hyper-SUMOylation of TOPBP1 upon MMC and ATRi co-treatment was significantly reduced by ATM inhibition (Figure 3D). Thus, in contrast to well-known phospho-induced SUMOylation, it appears that modulation of phosphorylation networks can also reduce SUMOylation in this context, expanding the repertoire of phospho-SUMO crosstalk.

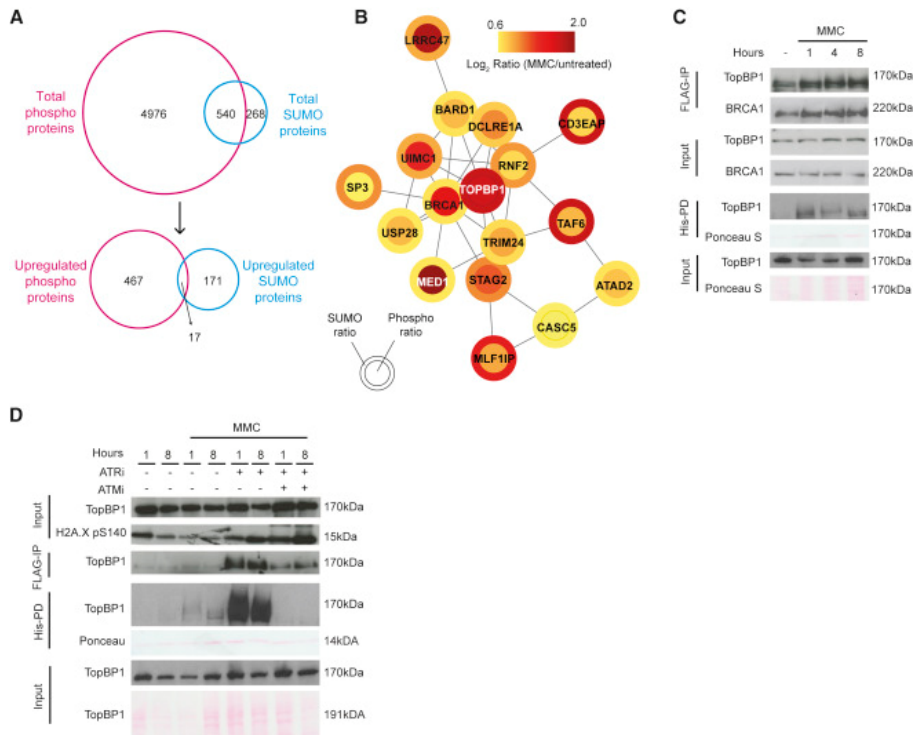


Figure 3 Integrated analyses of SUMOylation and phosphorylation datasets. **A)** Overlap of all identified and regulated SUMOylation and phosphorylation substrates. **B)** Functional network analysis of the 17 proteins with regulated phosphorylation and SUMOylation. **C)** Validation of TOPBP1 and BRCA1 regulated SUMOylation in FLAG-SUMO2 U-2-OS cells blocked for 24 hours with thymidine and then treated for 8 hour with or without MMC. Flag-IP: FLAG-based immuno-precipitation; His-PD: His-based pull down. **D)** Western blot analysis of TOPBP1 SUMOylation upon treatment with 8 hours MMC with and without ATR (ATRi, ATR-45) and ATM inhibitors (ATMi, KU55933). (See also Figure S3 and Table S3).

These observations are in accordance with induction of DNA DSBs and ATM activation that arises upon RS in combination with checkpoint inhibition (Toledo et al., 2013) (Figure 4A). To validate our observations that central DDR kinases modulate hyper-SUMOylation of TOPBP1 upon MMC treatment and determine whether such regulation occurs on other proteins, we performed an additional label-free quantitative proteomics screen. Here we analyzed enriched SUMOylated proteins from MMC treated cells in combination with the ATMi and ATRi (Figure 4B, Figure S4A and Table S4). We confirmed that TOPBP1 is hyper-SUMOylated by co-treatment with MMC and ATRi, and that this was attenuated upon addition of ATMi (Figure 4C and 4D). Remarkably, ATR itself, and its constitutive interactor ATRIP, which localizes ATR to TOPBP1 for activation, both displayed the same hyper-SUMOylation pattern as TOPBP1 (Figure 4C and 4D). While SUMOylation of ATRIP and ATR has previously been reported in response to UV and HU treatments (Wu et al., 2014), we find that hyper-

SUMOylation of ATR, ATRIP, TOPBP1, and XRCC6 (X-ray repair cross-complementing protein 6) arises upon RS in combination with checkpoint inhibition. Importantly, STRING-based functional network analysis of SUMOylation targets significantly regulated upon MMC treatment with and without ATRi and ATMi, reveals that these consist of core ATR activating proteins and DDR responders, showing remarkable orchestration of this functional group (Figure S4B) (Jentsch and Psakhye, 2013).

Together, these proteomics experiments suggest that regulation of phosphorylation and SUMOylation occurs within overlapping networks of RS responders, and that these may be subjected to common control by the same apical DDR kinases.

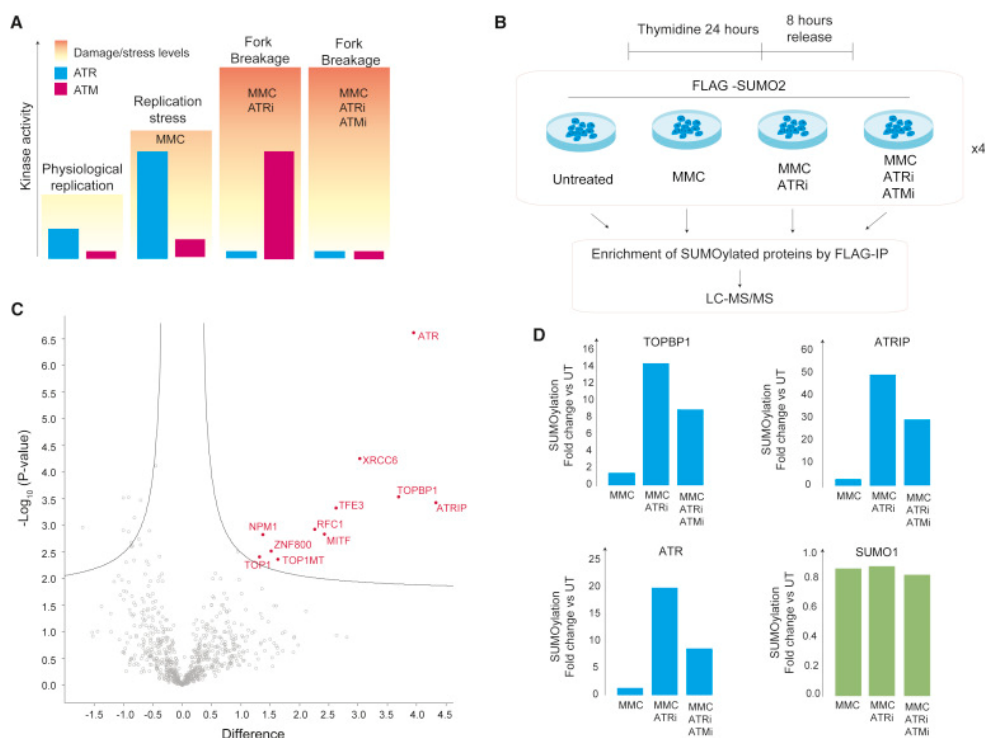


Figure 4 Proteomics analysis of TOPBP1 SUMOylation regulation by ATR and ATM inhibitors. **A)** Schematic representation of kinase activities at progressive stages of RS induced by MMC treatment and in combination with ATR and ATM inhibition. The blue and red bars represent the level of activation of the ATR and ATM kinases respectively. The shaded backgrounds represent the increasing levels of replication stress and damage that can be induced by MMC and ATRi co-treatment, yellow being less and red being extreme RS. **B)** Experimental design for label-free proteomics analysis of TOPBP1 SUMOylation upon MMC treatment with and without ATRi (ATR-45) and ATMi (KU55933), in FLAG-SUMO2 U-2-OS cells. **C)** Volcano plot of all ratios of MMC and ATR treated cells compared to MMC alone from enriched SUMOylated proteins, using t-test to determine significantly modulated (FDR<0.05) targets (indicated in red). **D)** SUMOylation levels for TOPBP1, ATRIP, ATR, and SUMO1.

TOPBP1, ATR and ATRIP from the proteomics analysis, and SUMO as a negative control. UT: untreated (cells that were released into DMSO without MMC or inhibitors) (See also Figure S4 and Table S4).

2.4 ATM and ATR modulate a global SUMOylation response to RS

We next sought to determine whether modulation of protein SUMOylation by ATM and ATR was a general mechanism in other conditions of RS. Using (HU), an inhibitor of dNTP synthesis, which causes DNA replication fork stalling, we could reproduce the pattern of TOPBP1 SUMOylation observed for MMC with and without ATRi and ATMi co-treatment (Figure 5A). TOPBP1 SUMOylation was increased upon 3 hours of HU treatment, further massively enhanced by co-treatment with ATRi, and then attenuated by addition of ATMi (Figure 5A). However, after 30 min HU and ATRi treatment, only modest increase of TOPBP1 SUMOylation was detected. This pattern is in accordance with replication forks breaking after longer treatments with replication stressors and checkpoint inhibition, thereby also inducing ATM signaling (Figure 5A and Figure S5A). Furthermore, treatment with high-dose ionizing radiation (IR), which also induces DSBs and ATM activation, did not induce TOPBP1 SUMOylation, indicating that this regulation is specific to RS associated DNA breaks (Figure 5A). We further validated this pattern of TOPBP1 SUMOylation using two different pharmacological inhibitors for ATM and ATR and with one CHK1 inhibitor (CHK1i) (Figure S5A). Analogous to ATR, inhibition of CHK1, a prominent substrate and mediator of ATR checkpoint signaling, results in replication fork breakage and ATM activation (Figure S5A). Interestingly, TOPBP1 was also hyper-SUMOylated upon CHK1i and HU co-treatment (Figure S5A). Collectively, these observations indicate that modulation of the SUMOylation response to RS by these central DDR kinase could be a general regulatory mechanism, and not only specific to MMC treatment.

To elaborate on the magnitude of this mechanism, we performed a large-scale proteomics experiment to analyze SUMOylation and phosphorylation site regulation under these conditions. Specifically, we enriched SUMOylated and phosphorylated peptides from cells treated with HU in combinations with and without CHK1i and ATMi for analysis by LC-MS/MS (Figure S5B and S5C). CHK1i was used rather than ATRi to permit initiation of the RS response by ATR. Four biological replicates were performed and each sample was analyzed twice by MS for label free quantification (Figure S5D). We identified 3,465 SUMOylated peptides corresponding to 1,590 SUMOylation acceptor sites, of which 2,450 peptides were quantified at least three times in at least one of the three treatment conditions (Figure 5B and Table S5). Using ANOVA significance testing to compare the dynamics of the modifications between treatments, 1,374 SUMOylated peptides, corresponding to 816 SUMO acceptor sites, were deemed regulated in at least one condition (Figure 5B). Similarly, 3,373 high confidence phosphorylation sites were found to be modulated and 127 proteins harbored changes of both PTMs (Figure 5B and 5C). To determine whether there was interdependency between SUMOylation and phosphorylation in our dataset, for example with the PDSM motif (Hietakangas et

al., 2006), we analyzed our raw MS data to identify co-occurring phosphorylation sites on the enriched SUMO peptides. We identified 127 phosphorylation sites in the SUMO-enriched dataset of which 26 were on SUMOylated peptides (Table S5). While the overlap is modest, 64% of these phosphorylation sites harbored a proline in the residue directly C-terminal to the phosphorylated serine/threonine residue, conforming to part of the PDSM motif ($\Psi K x E x x S P$) (Table S5).

We further analyzed our dataset to determine the degree of control that the DDR kinases exert on protein SUMOylation in response to RS. It is evident from the number of significantly perturbed SUMOylation acceptor sites, that regulation of this modification by ATM and ATR is a global mechanism in the response to RS, as more than fifty percent of the quantified sites were significantly regulated (Figure 5B). We performed unsupervised hierarchical clustering of the regulated phosphorylation sites and SUMOylated peptides to determine the dynamics of this regulation (Figures 5D and S5E). For both modifications we identified a cluster that showed the same dependency on CHK1 and ATM as observed for TOPBP1 by western blotting (Figure 5D). In this cluster, protein SUMOylation and phosphorylation sites increased upon co-treatment of HU with CHK1i compared to HU alone, and was attenuated upon further addition of ATMi (Figure 5D and S5E). Interestingly, GO analysis revealed that these clusters were enriched in proteins involved in DNA replication and recombination (Figure 5D and S5E). Among the SUMO-regulated proteins in this cluster were key regulators of DNA replication and homologous recombination such as TOP2A (DNA topoisomerase 2-alpha), BLM (Bloom syndrome protein), and BRCA1 as well as its constitutive interactor BARD1 (Figure 5D). Moreover, the dynamics of the modifications in these specific clusters are in accordance with the expected and observed phosphorylation profiles of targets of ATR and ATM (Figure 5A, 5D, S5A and S5E). Additionally, a cluster of proteins with significantly increased SUMOylation upon HU and CHK1i co-treatment, but unchanged by addition of ATMi, were enriched in proteins involved in DDR and DNA repair (Figure 5D). This included UIMC1, RBBP(CtIP), and interestingly also TOPORS, a SUMO E3 ligase that is known to play a role in the DDR (Lin et al., 2005; Marshall et al., 2010). Noteworthy, a substantial fraction of SUMOylation sites were modulated inversely, being unaffected or only slightly modulated by CHK1 inhibition, yet increasing dramatically upon co-inhibition of ATM (Figure 5D). This further indicates that ATM is a central regulator of protein SUMOylation in the DDR, and possibly more specifically in protein *de*SUMOylation. This subset of SUMO-regulated proteins was enriched for house-keeping biological processes such as RNA metabolism, transcription and chromatin remodeling (Figure 5D). Our findings demonstrate that SUMOylation is regulated globally in response to RS by the chief DDR kinases, ATM and ATR.

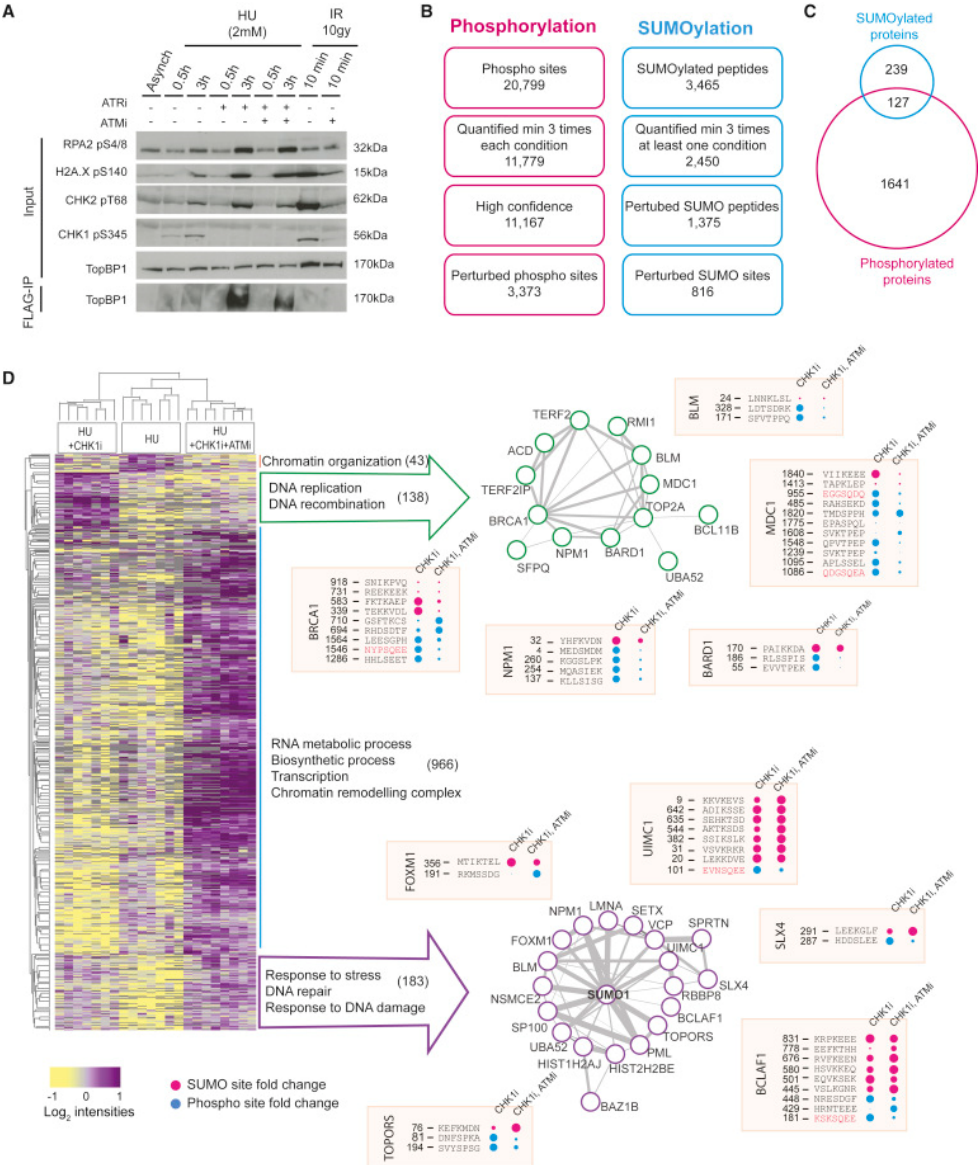


Figure 5 Deep proteomics analysis of phosphorylation and SUMOylation in RS and replication fork breakage. A) Western blot analysis of TOPBP1 SUMOylation and markers of ATR and ATM activity upon treatment with HU with and without ATRi (ATR-45) and/or ATMi (KU55933). **B)** Number of peptides, sites and proteins identified and quantified from the proteomics analysis. Total phosphorylation sites and SUMOylated peptides from all experimental conditions with a 1% FDR rate. Targets quantified at least three times from all biological and technical replicates in at least one condition were used for further analysis. For phosphorylation events, localization probabilities of at least 0.75 (high confidence) was also required. Perturbed SUMOylation peptides and phosphorylation sites that were modulated in any one condition compared to another were determined by ANOVA testing (FDR < 0.05). **C)** Overlap of proteins with regulated SUMOylation and phosphorylated.

D) Unsupervised hierarchical clustering of the 1,375 significantly perturbed SUMOylation. GOBP enrichment analysis of the clusters, with STRING-based functional network analysis of the proteins in the clusters and dotplot representation of SUMOylation and phosphorylation site changes on selected proteins (pink for SUMOylation sites and blue for phosphorylation sites). The modified sequence is with the modification site in the center, and the red phospho-peptide sequences are those that confer to the ATM and ATR sequence motif, S/T-Q. (See also Figure S5 and Table S5).

3 Discussion

Context-specific and dynamic post-translational protein modifications are well-established regulators of the signaling pathways that protect eukaryotic DNA integrity during the tremendous task of replication. Advancements in speed, resolution and sensitivity of MS-based technologies have revolutionized the study of global PTM biology (Olsen and Mann, 2013). With this rise in global PTM data, it has become evident that efficient cellular responses, such as those that safeguard genomic integrity, require the precise and timely coordination of several PTMs and the different enzymes that regulate them (Papouli et al., 2005). Integrated analysis of PTMs is therefore pertinent for our understanding of the molecular mechanisms that respond to DNA damage and RS. Using state-of-the-art proteomics methodologies, we mapped nearly 1,400 regulated SUMOylation acceptor sites and 3,300 regulated phosphorylation sites, in response to the chemotherapeutic agents MMC and HU. Our study reveals that SUMOylation is regulated by the most dominant, apical DDR kinases, ATR and ATM, which are known to initiate and coordinate the phosphorylation responses to RS and replication fork breakage.

In accordance with previous studies, we find that RS elicits increased SUMOylation of the core ATR activating proteins, including TOPBP1 and ATRIP. Interestingly, previous studies have shown the SUMOylation of ATR and its constitutive interactor ATRIP are necessary for efficient ATR dependent checkpoint signaling (Wu and Zou, 2016; Wu et al., 2014). Further to this, here we showed that TOPBP1, a key co-activator of ATR, undergoes increased SUMOylation in response to MMC induced RS. This indicates that SUMOylation of this factor, in addition to that of ATR and ATRIP, may be important for ATR dependent checkpoint signaling. However, further biochemical and molecular biological analysis is required to confirm the precise role of TOPBP1 SUMOylation in ATR activation. In addition, our data suggest that SUMOylation is a common and relevant modification of a number of proteins involved in ATR activation in response to RS.

We aimed to uncover the interplay between phosphorylation and SUMOylation of protein networks in the RS response. Using an integrated proteomics approach, we found that protein SUMOylation was widely modulated by the main regulatory kinases that mediate the phosphorylation response. Parallel proteomics analysis of changes in these two PTMs revealed co-regulation of a number of central RS and DDR responders including BRCA1, BARD1 and TOPBP1. BRCA1 SUMOylation and phosphorylation have individually been found to play a key role in the function of this protein, as SUMOylation has been shown to increase its ubiquitin ligase activity (Morris et al., 2009). It will be

interesting in future analysis to determine whether there is co-dependency or cross regulation of these modifications in the proteins that harbor both phosphorylation and SUMOylation, and in particular the relevance of this for the functions of central DDR proteins.

Our approach, to study co-regulated SUMO- and phospho-modified proteins, proved successful to identify key co-modified targets. TOPBP1 was the most highly co-regulated protein in our dataset upon eight hours of MMC treatment, and we found that TOPBP1 SUMOylation was heavily modulated by ATR inhibition during RS and by ATM upon replication fork breakage. We found this particularly interesting, as these central DDR kinases (particularly ATM) are well known to orchestrate various PTM-based networks upon threats to the DNA (Smith et al., 2010). However, the effect of the apical DDR kinases on global protein SUMOylation in response to DNA damage and replication stress has not yet been shown. We determined that such regulation by kinases not only applies to TOPBP1, but further to nearly 1,400 SUMOylation acceptor sites in response to HU induced RS, demonstrating global regulation of SUMOylation by these kinases in the maintenance of genome stability. Interestingly, we observed decreased SUMOylation of a large subset of proteins upon ATM inhibition under conditions that induce replication fork breakage. This suggests that ATM may be important for global *de*SUMOylation to maintain and control physiological levels of protein SUMOylation.

In our bioinformatics analysis of proteins with increased SUMOylation upon treatment with MMC and HU, we found clusters of co-regulated proteins that are known to function together in the RS response. In addition to the ATR activation proteins, BRCA1 and BARD1, we also found Fanconi Anemia proteins and DSB response proteins like MDC1, NBN and CtIP. This is particularly interesting in light of the recent idea that SUMO functions as a molecular glue to mediate protein complex formation under specific cellular states, and that this modification takes form of a ‘SUMO-spray’ (Jentsch and Psakhye, 2013). A consequence of this hypothesis is that SUMOylation should occur on functionally related proteins, to promote cooperation and interaction in protein networks, and this is precisely what we observe in our dataset. Interestingly, we find that proteins co-modified by SUMOylation and phosphorylation, generally have many regulated sites in response to RS. This poses a challenge for functional studies as site-directed mutagenesis of specific SUMOylation acceptor sites has been shown to result in little effect on overall protein SUMOylation or function (Jentsch and Psakhye, 2013).

Here we present an integrated analysis of global protein phosphorylation and SUMOylation in RS responses, and the largest resource to date of regulated SUMOylation targets in these conditions. We propose that increased SUMOylation occurs on specific and relevant factors in response to distinct DNA lesions, as illustrated by the SUMOylation dynamics upon RS and RS induced DSBs. Our data suggests that these SUMOylation responses are orchestrated by the apical kinases ATR and ATM in parallel with or as part of their phosphorylation signaling. These findings and further investigations of the co-regulation of these two modifications, is currently of great interest, in that the induction of RS provoked DSBs is increasingly used in chemotherapy to induce cancer cell killing (Li and Heyer, 2008).

In light of the essential role of SUMO in maintenance of genomic integrity (Bergink and Jentsch, 2009; Jackson and Durocher, 2013), the increasing interest of this system as a druggable target (Kessler et al., 2012) will require the understanding of how its perturbation affects global signaling networks.

4 Experimental procedures

Further details and an outline of resources used in this work can be found in Supplemental Experimental Procedures.

4.1 Cell culture

Human U-2-OS osteosarcoma cells cultured in complete DMEM, and for SILAC based experiments cells were SILAC labeled as reported previously (Hekmat et al., 2013; Olsen et al., 2006. For further details on cell culture, synchronization and drug treatments see Supplemental experimental procedures.

4.2 Stable cell line generation

To generate stable cell lines for SUMO enrichment U-2-OS cells were infected with lentivirus encoding either FLAG-tagged SUMO-2 (FLAG-SUMO2) or His10-SUMO-2-K0-Q87R (His10-S2-K0) as previously described (Hendriks et al., 2014; Schimmel et al., 2014). Further details are provided in Supplemental experimental procedures.

4.3 SUMO target protein enrichment

Enrichment of SUMOylated proteins was performed as previously described (Schimmel et al., 2014). Briefly cell were harvested in lysis buffer and sonicated, prior to enrichment of SUMOylated protein using Monoclonal ANTI-FLAG M2 beads (Sigma) for 90min at 4°C with rotation. Following washes, the bound proteins were eluted using 1mM FLAG-M2 epitope peptide and thereafter filtered through a Amicon Ultra 10k NMWL spin filter (Millipore). The resulting proteins were processed by in-gel digestion for LC-MS/MS analysis. For detail on enrichment of SUMO target proteins and in-gel digestion, please see Supplemental experimental procedures.

4.4 SUMO-peptide enrichment

SUMOylated peptides were enriched as previously described in (Hendriks et al., 2014). Briefly, thirty 15 cm plates of U-2-OS per condition were harvested in PBS, lysed in a 6M guanidine-HCl lysis buffer and sonicated. SUMOylated proteins were enriched from equal amounts of protein for each condition by Ni-NTA agarose beads overnight at 4°C. Proteins were eluted using 500mM imidazole two times. The eluted proteins were filtered and concentrated in spin filters digested with LysC. SUMOylated peptides were subsequently enriched with Ni-NTA agarose beads at 4°C for 5 hours, and eluted using 500mM imidazole. The enriched peptides were filtered and concentrated prior to digestion with trypsin and analysis by LC-MS/MS. For detailed SUMO-peptide enrichment procedures see Supplemental experimental procedures.

4.5 Mass Spectrometry analysis

Peptide mixtures were analyzed using an EASY-nLC system (Proxeon, Odense, Denmark) connected to a Q-Exactive mass spectrometer (Thermo Fisher Scientific, Bremen, Germany), as described (Kelstrup et al., 2012). Details are provided in Supplemental experimental procedures.

4.6 Raw Data Processing

Raw data was analyzed using MaxQuant v1.4.1. and v1.5.11 against the complete human UniProt database. See Supplemental experimental procedures for detailed descriptions.

4.7 Bioinformatics analysis

All functional network analysis were done using the String database (Szklarczyk et al., 2015) and further processed with Cytoscape (www.cytoscape.org). Hierarchical clustering and ANOVA t-test were performed using Perseus. For ANOVA the FDR threshold was set to 0.05. Sequence motif analysis was performed using IceLogo (Colaert et al., 2009). Details are provided in Supplemental experimental procedures.

Author Contributions

J.O.S. performed experiments described in Figures 1-4. Z.X. performed experiments described in Figure 5. T.S.B provided help and performed part dataset presented in Figure 2. G.F did validation experiments presented in Figure S5. L.v.S. and A.J.L.C. supervised J.O.S and S.M. and gave critical input on manuscript. A.C.O.V. supervised Z.X., J.V.O supervised J.O.S., and S.M. generated and analyzed the data shown in remaining figures. S.M., A.C.O.V. and J.V.O. conceived the study, designed the experiments, critically evaluated the results, and wrote the manuscript.

Data and software availability

All raw mass spectrometric data files for this study have been deposited to the ProteomeXchange Consortium via with the dataset identifier PXD006361.

Acknowledgements

Work at The Novo Nordisk Foundation Center for Protein Research (CPR) is funded in part by a generous donation from the Novo Nordisk Foundation (Grant number NNF14CC0001). The proteomics technology developments applied was part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 686547. We would like to thank the PRO-MS Danish National Mass Spectrometry Platform for Functional Proteomics and the CPR Mass Spectrometry Platform for instrument support and assistance. L.V.S work was supported by the Danish Research Council (research career program FSS Sapere Aude). J.V.O. was supported by the Danish Cancer Society (R90-A5844 KBVU project grant) and Lundbeckfonden (R191-2015-703). A.C.O.V. was supported by the European Research Council (310913) and by the Netherlands Organisation for Scientific Research (NWO 93511037). J.O.S, Z.X., A.C.O.V. and J.V.O. were supported by the Marie Curie Initial Training Networks program of the European Union (290257-UPStream). A.J.L and S. M. were supported by Danish Cancer Society (KBVU-2014) and European Research Council (ERC-2015-STG-294679068) grants. A.J.L. lab was also supported by the Danish National Research Foundation (DNRF115) and Danish Council for Independent Research (Sapere Aude, DFF-Starting Grant 2014).

References

- Batth, T.S., Francavilla, C., and Olsen, J. V. (2014). Off-line high-pH reversed-phase fractionation for in-depth phosphoproteomics. *J. Proteome Res.*
- Beli, P., Lukashchuk, N., Wagner, S.A., Weinert, B.T., Olsen, J. V., Baskcomb, L., Mann, M., Jackson, S.P., and Choudhary, C. (2012). Proteomic investigations reveal a role for RNA processing factor THRAP3 in the DNA damage response. *Mol. Cell* 46, 212–225.
- Bennetzen, M., Larsen, D.H., Bunkenborg, J., Bartek, J., Lukas, J., and Andersen, J.S. (2009). Site-specific phosphorylation dynamics of the nuclear proteome during the DNA damage response.
- Bergink, S., and Jentsch, S. (2009). Principles of ubiquitin and SUMO modifications in DNA repair. *Nature* 458, 461–467.
- Bursomanno, S., Beli, P., Khan, A.M., Minocherhomji, S., Wagner, S.A., Bekker-Jensen, S., Mailand, N., Choudhary, C., Hickson, I.D., and Liu, Y. (2015). Proteome-wide analysis of SUMO2 targets in response to pathological DNA replication stress in human cells. *DNA Repair (Amst)*. 25, 84–96.
- Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., and Gevaert, K. (2009). Improved visualization of protein consensus sequences by iceLogo. *Nat. Methods* 6, 786–787.
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372.
- Danielsen, J.M.R., Sylvestersen, K.B., Bekker-Jensen, S., Szklarczyk, D., Poulsen, J.W., Horn, H., Jensen, L.J., Mailand, N., and Nielsen, M.L. (2011). Mass spectrometric analysis of lysine ubiquitylation reveals promiscuity at site level. *Mol. Cell. Proteomics* 10, M110.003590.
- Flotho, A., and Melchior, F. (2013). Sumoylation: A Regulatory Protein Modification in Health and Disease. *Annu. Rev. Biochem.* 82, 357–385.
- Francavilla, C., Lupia, M., Tsafou, K., Villa, A., Kowalczyk, K., Rakownikow Jersie-Christensen, R., Bertalot, G., Confalonieri, S., Brunak, S., Jensen, L.J., et al. (2017). Phosphoproteomics of Primary Cells Reveals Druggable Kinase Signatures in Ovarian Cancer. *Cell Rep.* 18, 3242–3256.
- García-Rodríguez, N., Wong, R.P., and Ulrich, H.D. (2016). Functions of Ubiquitin and SUMO in DNA Replication and Replication Stress. *Front. Genet.* 7, 87.
- Gareau, J.R., and Lima, C.D. (2010). The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nat. Publ. Gr.* 11.
- Gibbs-Seymour, I., Oka, Y., Rajendra, E., Weinert, B.T., Passmore, L.A., Patel, K.J., Olsen, J. V, Choudhary, C., Bekker-Jensen, S., and Mailand, N. (2015). Ubiquitin-SUMO Circuitry Controls Activated Fanconi Anemia ID Complex Dosage in Response to DNA Damage. *Mol. Cell* 57, 150–164.
- González-Prieto, R., Cuijpers, S.A., Luijsterburg, M.S., van Attikum, H., and Vertegaal, A.C. (2015). SUMOylation and PARylation cooperate to recruit and stabilize SLX4 at DNA damage sites. *EMBO Rep.* e201440017.
- Halazonetis, T.D., Gorgoulis, V.G., and Bartek, J. (2008). An Oncogene-Induced DNA Damage Model for Cancer Development. *Science* (80-.). 319.
- Hendriks, I.A., and Vertegaal, A.C.O. (2016). A comprehensive compilation of SUMO proteomics. *Nat. Rev. Mol. Cell Biol.* 17, 581–595.
- Hendriks, I.A., D'Souza, R.C., Yang, B., Verlaan-de Vries, M., Mann, M., and Vertegaal, A.C. (2014). Uncovering global SUMOylation signaling networks in a site-specific manner. *Nat Struct Mol Biol* 21, 927–936.
- Hendriks, I.A., Lyon, D., Young, C., Jensen, L.J., Vertegaal, A.C.O., and Nielsen, M.L. (2017). Site-specific mapping of the human SUMO proteome reveals co-modification with phosphorylation. *Nat. Struct. Mol. Biol.* 24, 325–336.
- Hietakangas, V., Anckar, J., Blomster, H.A., Fujimoto, M., Palvimo, J.J., Nakai, A., Sistonen, L., and Walter, P. (2006). PDSM, a motif for phosphorylation-dependent SUMO modification.
- Hunter, T. (2007). The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Mol. Cell* 28, 730–738.
- Jackson, S.P., and Bartek, J. (2009). The DNA-damage response in human biology and disease. *Nature* 461.
- Jackson, S.P., and Durocher, D. (2013). Regulation of DNA Damage Responses by Ubiquitin and SUMO. *Mol. Cell* 49, 795–807.
- Jentsch, S., and Psakhye, I. (2013). Control of Nuclear Activities by Substrate-Selective and Protein-Group SUMOylation. *Annu. Rev. Genet* 47, 167–186.
- Jungmichel, S., Rosenthal, F., Altmeyer, M., Lukas, J., Hottiger, M.O., and Nielsen, M.L. (2013). Proteome-wide identification of poly(ADP-Ribosylation) targets in different genotoxic stress responses. *Mol. Cell* 52, 272–285.
- Kelstrup, C.D., Young, C., Lavalley, R., Nielsen, M.L., and Olsen, J. V. (2012). Optimized Fast and Sensitive Acquisition Methods for Shotgun Proteomics on a Quadrupole Orbitrap Mass Spectrometer. *J. Proteome Res.* 11, 3487–3497.

- Kelstrup, C.D., Jersie-Christensen, R.R., Batth, T.S., Arrey, T.N., Kuehn, A., Kellmann, M., and Olsen, J. V (2014). Rapid and deep proteomes by faster sequencing on a benchtop quadrupole ultra-high-field orbitrap mass spectrometer. *J. Proteome Res.* *13*, 6187–6195.
- 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. (2012). A SUMOylation-Dependent Transcriptional Subprogram Is Required for Myc-Driven Tumorigenesis. *Science* (80-.). 335.
- Kim, S.-T., Lim, D.-S., Canman, C.E., and Kastan, M.B. (2009). Substrate Specificities and Identification of Putative Substrates of ATM Kinase Family Members*.
- Lamoliatte, F., Caron, D., Durette, C., Mahrouche, L., Maroui, M.A., Caron-Lizotte, O., Bonneil, E., Chelbi-Alix, M.K., and Thibault, P. (2014). Large-scale analysis of lysine SUMOylation by SUMO remnant immunoaffinity profiling. *Nat. Commun.* *5*, 5409.
- Lamoliatte, F., McManus, F.P., Maarifi, G., Chelbi-Alix, M.K., and Thibault, P. (2017). Uncovering the SUMOylation and ubiquitylation crosstalk in human cells using sequential peptide immunopurification. *Nat. Commun.* *8*, 14109.
- Li, X., and Heyer, W.-D. (2008). Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res.* *181*.
- Lin, L., Ozaki, T., Takada, Y., Kageyama, H., Nakamura, Y., Hata, A., Zhang, J.-H., Simonds, W.F., Nakagawara, A., and Koseki, H. (2005). topors, a p53 and topoisomerase I-binding RING finger protein, is a coactivator of p53 in growth suppression induced by DNA damage. *Oncogene* *24*, 3385–3396.
- López-Contreras, A.J., and Fernandez-Capetillo, O. (2010). The ATR barrier to replication-born DNA damage. *DNA Repair (Amst)*. *9*, 1249–1255.
- Marshall, H., Bhaumik, M., Aviv, H., Moore, D., Yao, M., Dutta, J., Rahim, H., Gounder, M., Ganesan, S., Saleem, A., et al. (2010). Deficiency of the dual ubiquitin/SUMO ligase Topors results in genetic instability and an increased rate of malignancy in mice. *BMC Mol. Biol.* *11*.
- Meek, K., Dang, V., and Lees-Miller, S.P. (2008). Chapter 2 DNA-PK: The Means to Justify the Ends? *Adv. Immunol.* *99*, 33–58.
- Mertins, P., Mani, D.R., Ruggles, K. V, Gillette, M.A., Clauser, K.R., Wang, P., Wang, X., Qiao, J.W., Cao, S., Petralia, F., et al. (2016). Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature* *534*, 55–62.
- Mohideen, F., Capili, A.D., Bilimoria, P.M., Yamada, T., Bonni, A., and Lima, C.D. (2009). A molecular basis for phosphorylation-dependent SUMO conjugation by the E2 UBC9. *Nat. Struct. Mol. Biol.* *16*.
- Morris, J.R., Boutell, C., Keppler, M., Densham, R., Weekes, D., Alamshah, A., Butler, L., Galanty, Y., Pangon, L., Kiuchi, T., et al. (2009). The SUMO modification pathway is involved in the BRCA1 response to genotoxic stress. *Nature* *462*.
- Murga, M., Bunting, S., Montaña, M.F., Soria, R., Mulero, F., Cañamero, M., Lee, Y., Mckinnon, P.J., Nussenzweig, A., and Fernandez-Capetillo, O. (2009). A mouse model of ATR-Seckel shows embryonic replicative stress and accelerated aging. *Nat. Publ. Gr.* *41*.
- Olsen, J. V, and Mann, M. (2013). Status of large-scale analysis of post-translational modifications by mass spectrometry. *Mol. Cell. Proteomics* *12*, 3444–3452.
- Olsen, J. V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M.L., Jensen, L.J., Gnad, F., Cox, J., Jensen, T.S., Nigg, E.A., et al. (2010). Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis. *Sci. Signal.* *3*.
- Ong, S.-E. (2002). 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.
- Papouli, E., Chen, S., Davies, A.A., Huttner, D., Krejci, L., Sung, P., and Ulrich, H.D. (2005). Crosstalk between SUMO and Ubiquitin on PCNA Is Mediated by Recruitment of the Helicase Srs2p monoubiquitination serves alternative functions (Hicke, 2001), and depending on the linkage between the ubiquitin moieties, polyubiquitin chains can also convey. *Mol. Cell* *19*, 123–133.
- Polo, S., and Jackson, S. (2011). Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications. *Genes Dev.* *25*, 409–433.
- Sampson, D. a., Wang, M., and Matunis, M.J. (2001). The Small Ubiquitin-like Modifier-1 (SUMO-1) Consensus Sequence Mediates Ubc9 Binding and is Essential for SUMO-1 Modification. *J. Biol. Chem.* *276*, 21664–21669.
- Schimmel, J., Eifler, K., Otti Sigurðsson, J., Cuijpers, S.A., Hendriks, I.A., Verlaan-de Vries, M., Kelstrup, C.D., Francavilla, C., Medema, R.H., Olsen, J. V, et al. (2014). Molecular Cell Resource Uncovering SUMOylation Dynamics during Cell-Cycle Progression Reveals FoxM1 as a Key Mitotic SUMO Target Protein. *Mol. Cell* *53*, 1053–1066.
- Shiloh, Y. (2001). ATM and ATR: networking cellular responses to DNA damage. *Curr. Opin. Genet. Dev.* *11*, 71–77.
- Smith, J., Tho, L.M., Xu, N., and Gillespie, D.A. (2010). Chapter 3 - The ATM-Chk2 and ATR-Chk1 Pathways

- in DNA Damage Signaling and Cancer. *Adv. Cancer Res.* *108*, 73–112.
- Stehmeier, P., and Muller, S. (2009). Phospho-Regulated SUMO Interaction Modules Connect the SUMO System to CK2 Signaling. *Mol. Cell* *33*, 400–409.
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* *43*, D447–52.
- Tammsalu, T., Matic, I., Jaffray, E.G., Ibrahim, A.F.M., Tatham, M.H., and Hay, R.T. (2014). Proteome-Wide Identification of SUMO2 Modification Sites. *Sci. Signal.* *7*.
- Toledo, L.I., Altmeyer, M., Rask, M.-B., Lukas, C., Larsen, D.H., Povlsen, L.K., Bekker-Jensen, S., Mailand, N., Bartek, J., and Lukas, J. (2013). ATR Prohibits Replication Catastrophe by Preventing Global Exhaustion of RPA. *Cell* *155*, 1088–1103.
- Ulrich, H.D., and Walden, H. (2010). Ubiquitin signalling in DNA replication and repair. *Nat. Publ. Gr.* *11*.
- Wu, C.-S., and Zou, L. (2016). The SUMO (Small Ubiquitin-like Modifier) Ligase PIAS3 Primes ATR for Checkpoint Activation. *J. Biol. Chem.* *291*, 279–290.
- Wu, C.-S., Ouyang, J., Mori, E., Nguyen, H.D., Maréchal, A., Hallet, A., Chen, D.J., and Zou, L. (2014). SUMOylation of ATRIP potentiates DNA damage signaling by boosting multiple protein interactions in the ATR pathway. *Genes Dev.* *28*, 1472–1484.
- Xiao, Z., Chang, J.-G., Hendriks, I.A., Sigurðsson, J.O., Olsen, J. V, and Vertegaal, A.C.O. (2015). 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. *Mol. Cell. Proteomics* *14*, 1419–1434.

Figure S1

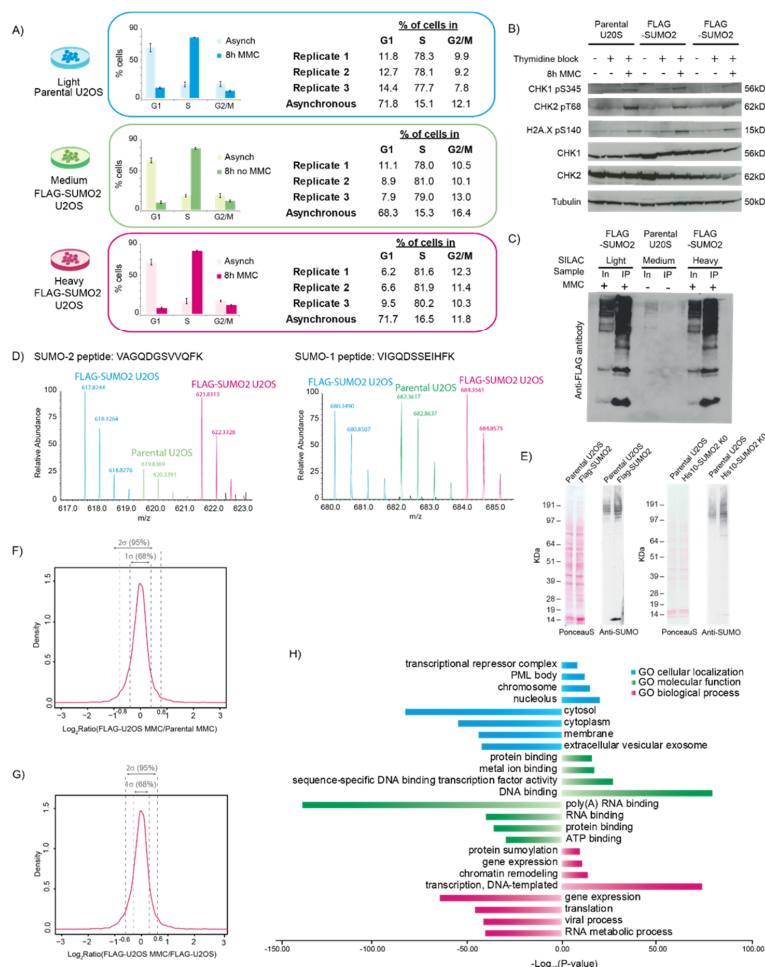


Figure S1 SUMOylation changes upon MMC treatment. Related to Figure 1. A) Flow cytometry-based cell cycle profile of SILAC labeled cells after synchronization for 24 hours with thymidine, and then released with and without 1 μ M MMC for 8 hours. B) Western blot of DNA damage and replication stress markers from parental and FLAG-SUMO2 U-2-OS cells after synchronization and treatment as in (A). C) Enrichment of SUMO from U-2-OS cells stably transfected with FLAG-SUMO2 and parental cells, from label swap SILAC experiment. All SILAC conditions were synchronized as in (A) and MMC treatment was performed as in (A). In: input; IP: immunoprecipitation. D) Full scan MS mass spectra of SUMO2 and SUMO1 derived SILAC peptide triplet MS intensities (relative abundance) from proteome measurements. All cells were synchronized for 24 hours with thymidine, and medium and heavy conditions were released into MMC for 8 hours while light cells were untreated. E) SUMO expression levels in untreated FLAG-SUMO2 U-2-OS (left) and His10-SUMO-2-K0-Q87R cells by western blot analysis of whole cell extracts. F) Distribution of Log_2 ratios of MMC treated FLAG-SUMO2 U-2-OS compared to MMC treated parental U-2-OS, from the unmodified peptides identified in the SUMO-enrichment MS analysis. One and two standard deviations (1σ and 2σ) from the mean are indicated with dashed lines, and used to set cutoffs for determining SUMOylated

Chapter 3

peptides (ratio above or below 2σ). G) Distribution of Log_2 transformed ratios of MMC treated compared to untreated FLAG-SUMO2 U-2-OS of all the unmodified peptides identified in the SUMO-enrichment MS analysis. One and two standard deviations (1σ and 2σ) from the mean are indicated with dashed lines, and used to set cutoffs for determining MMC regulated and non-regulated SUMOylated peptides (ratio above or below 2σ). H) GOCC (Gene Ontology Cellular Compartments), GOMF (Molecular Functions) and GOBP (Biological Processes) enrichment performed with InnateDB from all 702 proteins deemed SUMOylated from the proteomics experiment. Related to Figure 1 and Table S1.

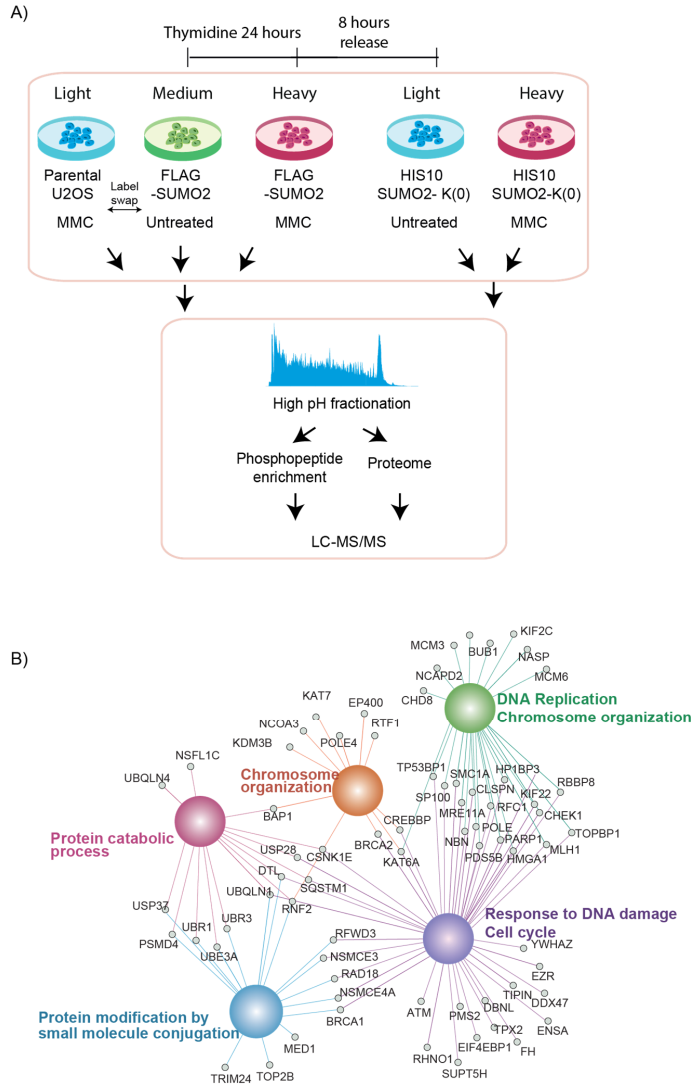
Figure S2

Figure S2 Phosphorylation changes in MMC treated cells. Related to Figure 2. A) Experimental design of quantitative phosphoproteomics analysis. Cells were treated equivalent to the proteomics experiments for SUMO enrichment described in Figures 1 and S1. Cells were lysed and proteins digested with trypsin prior to offline reversed-phase high-pH reversed-phase fractionation of the resulting peptides. An aliquot of each fraction was analyzed directly by LC-MS/MS for proteome measurements, whereas for phosphoproteome measurements phosphopeptides were enriched by TiO₂ prior to LC-MS/MS analysis. B) Phosphoprotein network representation of enriched GOBP terms from all proteins with regulated phosphorylation sites. The colored hubs represent the enriched GOBP terms, while the small nodes indicate the proteins with these terms with significantly up-regulated phosphorylation sites upon MMC treatment. Analysis was performed with Cytoscape using the ClueGO app. Related to Figure 2 and Table S2.

Figure S3

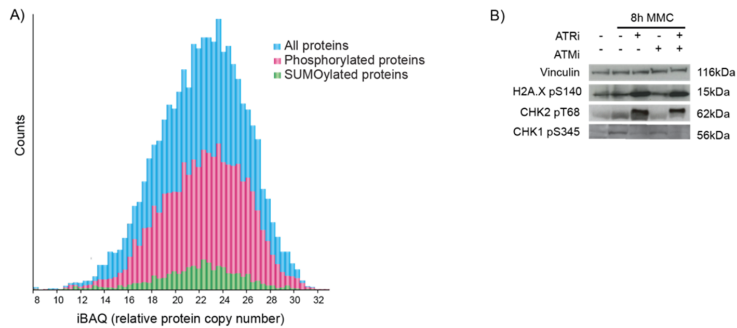


Figure S3 SUMOylation and phosphorylation data integration. Related to Figure 3. A) Distribution of the estimated relative protein copy number (iBAQ) of all proteins, phosphorylated proteins and SUMOylated proteins from datasets acquired in Figures 1 and 2. B) Western blot of markers of activity of DDR kinases ATM and ATR for cells treated with or without MMC in combinations with ATRi (ATR-45) and ATMi (KU55933). Related to Figure 3 and Table S3.

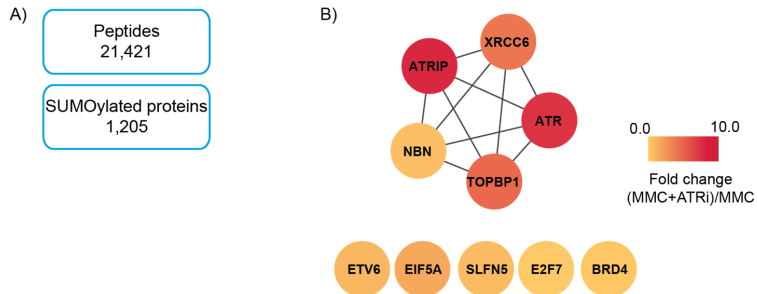
Figure S4

Figure S4 TOPBP1 SUMOylation regulation by ATR and ATM inhibitors. Related to Figure 4. A) Overview of the number of SUMOylated peptides and SUMO target proteins identified from the label free quantitative SUMO-2 proteomics analysis. B) Functional network analysis of proteins with significantly regulated SUMOylation after ATR inhibition. Color gradient represents the absolute fold change of the SUMO enrichment from FLAG-SUMO2 U-2-OS cells treated with MMC in combination with ATRi (ATR-45) compared to MMC only. Analysis was done with the STRING database and Cytoscape. Related to Figure 4 and Table S4.

Figure S5

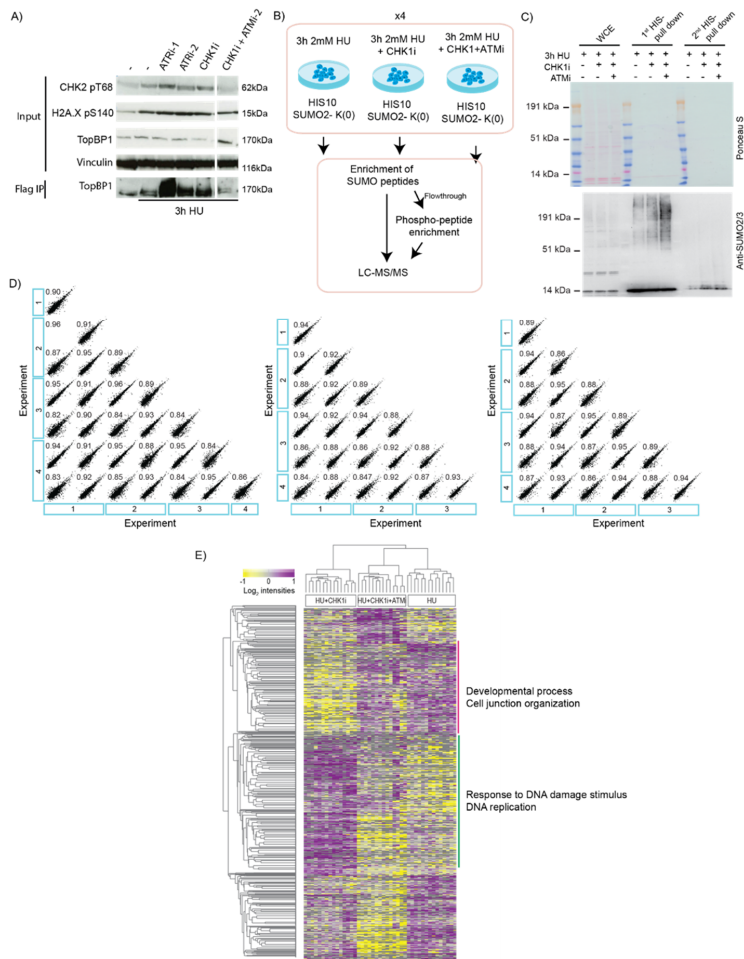


Figure S5 Integrated phosphorylation and SUMOylation analysis in RS and replication fork breakage. Related to Figure 5. A) Western blot analysis of TOPBP1 SUMOylation and markers of ATR and ATM activity upon treatment with HU with and without ATR, CHK1 and ATM inhibitors. ATRi-1: ATR-45; ATRi-2: AZ-20; CHK1i: SCH900776; ATMi-2: KU60019. B) Experimental setup for label free quantitative proteomics analysis of SUMOylation and phosphorylation substrates and sites from HIS10-SUMO2 U-2-OS cells. C) Western blot validation of SUMO enrichment from the experiments performed in (A). WCE: whole cell extract. Two sequential pulldowns were performed, indicated as 1st and 2nd HIS-pull downs. D) Correlation of the quantification of peptides identified from the proteomics analysis, with Pearson correlation coefficients. E) Unsupervised hierarchical clustering of the 3,373 phosphorylation sites deemed to be significantly perturbed by ANOVA testing (FDR <0.05). Two clusters with enriched GOBP terms are indicated by the colored bars.

Supplemental Experimental Procedures

Cells

Human osteosarcoma U-2-OS cells (female) were kindly shared by Roeland Dirks, who acquired them from ATCC. Cells were cultured in DMEM (Gibco, Invitrogen), supplemented with 10% fetal bovine serum (FBS), 100U/mL penicillin (Invitrogen), 100µg/mL streptomycin (Invitrogen) (complete DMEM), at 37°C, in a humidified incubator with 5% CO₂. For SILAC labeling (Ong, 2002), cells were cultured under the same conditions, using SILAC DMEM (Thermo Fisher Scientific) supplemented with 10% dialyzed FBS (Sigma), 100U/mL penicillin (Invitrogen), 100 µg/mL streptomycin (Invitrogen), 2 mM L-glutamine (Gibco) and one of the following three labeling combinations: 1) natural 'light' variants of the amino acids (Lys0, Arg0) (Sigma) 2) medium variants of amino acids {L-[2H4]Lys (+4) and L-[13C6]Arg (+6)} (Lys4, Arg6), and 3) heavy variants of the amino acids {L-[13C6,15N2]Lys (+8) and L-[13C6,15N4]Arg (+10)} (Lys8, Arg10). This is henceforth referred to as complete SILAC DMEM. Medium and heavy variants of amino acids were purchased from Cambridge Isotope Laboratories. We have not performed specific authentication of the cell lines used in this study.

Stable cell line generation

To generate stable cell lines for SUMO enrichment, U-2-OS cells were infected with lentivirus encoding either FLAG-tagged SUMO-2 (FLAG-SUMO2) or His10-SUMO-2-K0-Q87R (His10-SUMO2-K0), both containing a GFP sequence separated by an IRES (Schimmel et al., 2014; Xiao et al., 2015). Two weeks after infection, cells were fluorescence-sorted for a low expression level of GFP using a FACS Aria II (BD Biosciences). The His10-SUMO2-K0-Q87R has the following amino acid sequence:

AHHHHHHHHHHHGGSMSEERPREG

VRTENDHINLRVAGQDGSVVQFRIRRHPTLSRLMRAYCERQGLSMRQIRFRFDGQPINET
DTPAQLEMEDEDTIDVFRQQTGG.

Synchronization and drug treatments of cells

For thymidine blocking cells were grown to approximately 70% confluence on 15cm petri dishes (Greiner bio-one). Cells were treated with 4 mM thymidine (Sigma) in complete DMEM or complete SILAC DMEM for 24 hours and released by washing two times with PBS (Gibco, Life Technologies). For release, the media was replaced with complete DMEM or complete SILAC DMEM with or without drugs as indicated. Mitomycin C (MMC) was used at a concentration of 1 µM (Sigma) and hydroxyurea (HU) at 2 mM. Inhibitors were used at the following concentrations: ATR-45 (Ohio University) at 5 µM, AZ20 at 2µM, KU55933 (Calbiochem) at 10 µM, KU60019 (Selleckchem) at 10 µM and SCH900776 (MedKoo Biosciences) was used at 1 µM. When inhibitors were used in combination with drugs, both were added simultaneously.

Flow cytometry for cell cycle analysis

For cell cycle analysis cells were harvested with trypsin, washed two times with PBS, then fixed and permeabilized overnight at 4°C with ice cold ethanol. Cells were centrifuged at 250xg for 2 minutes. The pellet was resuspended in PBS complemented with 25µg/mL propidium iodide (Sigma) and 100µg/mL RNase A (Sigma) and incubated at 37 °C for 30 minutes. FACS analysis was performed on a BD FACSCalibur (BD Biosciences), and the data was analyzed using the FlowJo software.

SUMO target protein enrichment

Cells were washed two times with PBS and lysed with lysis buffer (1% SDS, 0.5% NP-40, 50 mM sodium fluoride, 1 mM sodium orthovanadate, 5 mM β-glycerol phosphate, 0.5 mM EGTA, protease inhibitor including EDTA (Roche) and 10 mM N-ethylmaleimide (NEM)). Lysates were either used directly for further processing or snap frozen with liquid nitrogen and later thawed at 30°C. Chloroacetamide (CAA) was added to a final concentration of 55 mM. Lysates were sonicated with a microtip sonicator (Sonics Vibra Cell, VCX130) for 3 cycles of 10 seconds on and 15 seconds off at 60% amplitude. Samples were kept cool during sonication but not below 30°C to prevent SDS precipitation. Lysates were incubated for 30 minutes at room temperature then centrifuged at 13,200 rpm at 4°C for 45 minutes, and the supernatant was used for enrichment. An aliquot was kept for testing the input by western blotting. Bradford assay was used to determine protein concentration, and equal amounts of protein were used for each condition. Monoclonal Anti-FLAG M2 beads (Sigma) were washed three times with Wash buffer (50 mM Tris, 150 mM NaCl, 70 mM chloroacetamide, 0.5% NP-40, 5 mM sodium fluoride, 1 mM sodium orthovanadate, 5 mM β-glycerol phosphate, 0.5 mM EGTA, 10 mM NEM, protease inhibitor including EDTA (Roche)). An aliquot of 10 µL bead slurry was used for each milligram of protein in the sample, and samples were incubated with the beads for 90 minutes at 4°C with rotation. Beads were collected at 500xg for 3 minutes, and washed five times in with Wash buffer (without CAA). FLAG-SUMO-2 conjugated proteins were eluted with tenfold bead volumes of 5% SDS in Wash buffer containing 1 mM FLAG-M2 epitope peptide (Sigma), for 10 min at room temperature with shaking. After spinning the supernatant was concentrated by transferring to an Amicon Ultra 10k NMWL spin filter (Millipore) and spinning at 7500xg until 50 µL remained. 10 µL NuPAGE LDS sample buffer was added, the filter was inverted, and the sample collected in a fresh tube by spinning at 2500xg for 2 minutes. The eluted samples were run on an SDS-Page gel for western blotting or in-gel digestion for mass spectrometry analysis.

In-gel digestion

Eluted SUMO-target enriched samples were run on a NuPage 4-12% Bis-Tris protein gel (Novex, Thermo Fisher Scientific), and stained with Colloidal Blue Staining Kit (Invitrogen). Gel pieces were excised and transferred to 1.5 ml tubes where they were destained in 50% ethanol in 25mM ammonium bicarbonate (ABC), then dehydrated in 97% ethanol with shaking. This was followed by incubation

with 10 mM DTT (in 25 mM ABC) for 30 minutes and for an additional 30 minutes with 55 mM CAA (in 25 mM ABC). Samples were washed twice with 25mM ABC with shaking for ten minutes, and again dehydrated with 96% ethanol. Trypsin (Sigma) was added (1.5 µg/mL) for one hour with subsequent addition of an equal volume and incubated overnight with shaking. Trypsin activity was quenched by acidifying the sample with trifluoroacetic acid (TFA) to a final concentration of 1%. The sample was collected, and the remaining gel was sequentially washed with Washing solution A (30% (v/v) acetonitrile (ACN), 3% (v/v) TFA in milliQ), then Washing solution B (80% (v/v) ACN, 0.5% (v/v) acetic acid in milliQ) and finally with 100% ACN. The sample was collected after incubation with each solution for 30 minutes with shaking. Samples were prepared for loading on C-18 stage-tips (Rappsilber et al., 2007) by concentrating and removing ACN with vacuum centrifugation.

SUMO-peptide enrichment

Thirty 15cm petri dishes at 80% confluence with U-2-OS cells expressing His10-SUMO-2-K0-Q87R (approximately 15 million cells per dish) were used for each condition to identify SUMO-2 sites. We employed our previously established SUMO site enrichment method (Hendriks and O Vertegaal, 2016; Hendriks et al., 2014). Cells were washed three times with ice-cold PBS on the plate. After the last wash 2mL ice-cold PBS was added to each plate, cells were scraped and collected in 15 mL tubes. Cells were spun down at 250xg and re-suspended in ice-cold PBS. Small aliquots of cells (~80,000 cells) were kept and lysed in lysis buffer (1% NP-40, 2% SDS, 150 mM NaCl, and 50 mM TRIS, buffered at pH 7.5) for control western blotting. All PBS was aspirated from the main batches of cells. Cell pellets were collected and subsequently lysed in 10 pellet volumes of lysis buffer (6 M guanidine-HCl, 100 mM sodium phosphate, 10 mM TRIS, buffered at pH 8.0). Lysates were sonicated using a microtip sonicator at 30 Watts for ten seconds of sonication time per 10 mL lysate. Sonication steps were repeated once and tubes were inverted to mix in between sonication steps. Protein quantities from the lysates were equalized using the bicinchoninic acid assay. Next, lysates were supplemented with imidazole and β-mercaptoethanol to a final concentration of 50 mM and 5 mM respectively. Ni-NTA agarose beads (Qiagen) at 20 µL (dry volume) per 1 mL lysate were used for SUMO purification. Following overnight incubation at 4°C, beads were centrifuged at 500xg for 2 minutes, and washed using at least 5 bead volumes of wash buffers 1-4 in the following order: wash buffer 1: 6 M guanidine-HCl, 10 mM TRIS, buffered at pH 8.0, 100 mM sodium phosphate, 0.1% Triton X-100, 10 mM imidazole, 5 mM β-mercaptoethanol; wash buffer 2: 8 M urea, 10 mM TRIS, buffered at pH 8.0, 100 mM sodium phosphate, 0.1% Triton X-100, 10 mM imidazole, 5 mM β-mercaptoethanol; wash buffer 3: 8 M urea, 10 mM TRIS, buffered at pH 6.3, 100 mM sodium phosphate, 10 mM imidazole, 5 mM β-mercaptoethanol; wash buffer 4: 8 M urea, 10 mM TRIS, buffered at pH 6.3, 100 mM sodium phosphate, 5 mM β-mercaptoethanol. Subsequently, all wash buffer was aspirated entirely, and proteins were eluted using one bead volume of elution buffer (7 M urea, 10 mM TRIS, buffered at pH 7.0, 100 mM sodium phosphate, 500 mM imidazole) for at least 30 minutes. Elution steps were repeated twice. All

eluted fractions were pooled and filtered using 0.45 micron filters (twice pre-washed with Elution Buffer) to clear any remaining beads from the samples. The flow through from this pull down can be used for enrichment of phosphorylated peptides after trypsin digestion. For mass spectrometric analysis of SUMOylated peptides, samples were concentrated using 100 kDa cut-off spin filters (pre-washed with elution buffer), using a temperature-controlled centrifuge set to 20°C and centrifuging at 8,000xg. After concentration, filters were washed again using 250 µL of elution buffer without imidazole. Concentrated SUMOylated proteins were removed from the filters by centrifuging the filters while placed upside down into an open 1.5 mL LoBind tube. The concentration of purified SUMOylated proteins was determined using the Bradford assay (BioRad). SUMOylated proteins were digested using sequencing grade endoproteinase Lys-C in a 1:25 (w/w) enzyme-to-protein ratio for 4 hours. Subsequently, 10 mM β-mercaptoethanol was freshly added and the samples were further digested overnight by an additional amount of Lys-C equal to the first amount. All digestion steps were performed in the dark at room temperature. Protein samples were subsequently transferred to 15 mL tubes and diluted with half the amounts of guanidine lysis buffer used to lyse the initial cell pellets and were supplemented by adding imidazole to a final concentration of 50 mM and β-mercaptoethanol to a final concentration of 5 mM. Ni-NTA agarose beads at 40 µL (dry volume) per 1 mL sample were used for SUMO enrichment. Following incubation at 4°C for 5 hours, beads were centrifuged at 500xg for 2 minutes and washed again with wash buffer 1- 4 as described before. Proteins were eluted using one bead volume of elution buffer (7 M urea, 100 mM sodium phosphate, 10 mM TRIS, buffered at pH 7.0, 500 mM imidazole). All eluted samples were passed through 0.45 µm filters (twice pre-washed with elution buffer) to remove the remaining beads from the samples. Next, samples were concentrated by passing them through 10 kDa cut-off spin filters (pre-washed with elution buffer) at 14,000xg at 20°C. After concentration, the peptides remaining on the filters were washed twice with 250 µL of elution buffer without imidazole, and re-concentrated. Final concentrated SUMOylated peptides were removed from the filters by centrifuging the filters while placed upside down into a 1.5 mL LoBind tube. The double-purified SUMOylated peptides were trypsin digested, purified on stage-tips and analyzed by mass spectrometry to map specific sites of protein SUMOylation.

Sample preparation for phospho-peptide enrichment

Samples for phospho-peptide enrichment were either prepared freshly as described here or taken from the flow-through after the first HIS-based enrichment in the SUMO-peptide enrichment procedure (and directly trypsin digested as described below). For lysate preparation, cells were washed twice with PBS and subsequently lysed on the plate with boiling (99°C) lysis buffer (6 M guanidinium hydrochloride (GndCl), containing 5 mM tris(2carboxyethyl)phosphine (TCEP) and 10 mM CAA in 100 mM Tris, pH 8.5). The lysate was collected by scraping and subsequently further boiled for 10 minutes at 99°C prior to microtip sonication on a Sonics Vibra Cell (VCX130) at amplitude 50%, for two minutes with 1 second on and 1 second off pulses. Protein concentrations were estimated using Bradford assay (Bio-

Rad). For SILAC samples equal amounts of protein were mixed at this point, and for label free equal amounts of protein were used separately for further processing. Samples were digested with Lys-C in an enzyme:protein ratio of 1:100 (w/w) for 4 hours at 25 °C. GndCl was diluted to 2 M using 25 mM Tris pH 8.5, and the samples were further digested with trypsin (Sigma Aldrich) at 1:100 (w/w) overnight at 37 °C with slow shaking. Trypsin activity was quenched by acidifying the sample with TFA to a final concentration of 1%. The peptide mixtures were desalted and concentrated on a C18-SepPak cartridge (Waters). Sep-Pak columns were washed with 100% ACN and 0.1% TFA buffer before loading the sample, and the column were then washed with 0.1%TFA and stored at 4°C. Peptides were eluted from with 2 mL 40% ACN in 0.1% TFA, and then with 2 mL 60% CAN in 0.1% TFA followed by concentration and removal of ACN by vacuum centrifugation for 40 minutes at 45°C. Peptide concentration in the resulting sample as measured at 280 nM absorbance on a Nanodrop (Thermo Fisher Scientific). The eluted samples were either used directly for phospho-peptide enrichment fractionated as described below prior to further processing.

Offline High pH Reversed-Phase HPLC Fractionation

The eluted and concentrated peptide samples were separated using an Ultimate 3000 high-pressure liquid chromatography (HPLC) system (Dionex) on a C₁₈ Waters Xbridge BEH130 column (3.5 µm 4.6 × 250 mm). Peptides were separated and collected as described previously (Batth et al., 2014). Buffer A (milli-Q water), buffer B (100% ACN), and buffer C (50 mM ammonium hydroxide) were operated simultaneously at a final flow rate of 1 ml/min. Buffer C was operated throughout the separation gradient at 0.1 ml/min. The separation gradient consisted of increasing B from 5% to 25% in 65 minutes. Buffer B was increased to 35% in 10 minutes and 60% in 5 minutes followed by a sharp increase to 90% B in 2 minutes where it was held for additionally 5 minutes. The column was re-equilibrated to starting conditions with a 2 minute ramp down to 5 % buffer B where it was held 6 minutes prior to injection of the next sample. Fractions were collected at 1 minute intervals in a 96 deep well plate using automated fraction collector-3000 coupled to the HPLC system. Fractions were collected until the 80 minute mark in the gradient and manually concatenated to 14 fractions.

Phospho-peptide enrichment

Following fractionation, phospho-peptides were enriched from each fraction, or for enrichment after SUMO-peptide enrichment, samples were directly after Sep-Pak elution as described above.

Titanium dioxide (TiO₂) beads (5 µm Titansphere, GL Sciences) (Larsen, 2005; Olsen et al., 2006) were incubated in 20 mg/ml of 2, 5-dihydroxybenzoic(DHB) acid in 80% ACN and 6% TFA before use. TiO₂ bead slurry was prepared at 20 mg beads per ml DHB solution. Prior to adding the TiO₂ beads, each of the 14 fractions was diluted two-fold with 100% ACN in 12% TFA. 20 µL of TiO₂ beads slurry was added to each of the 14 fractions followed by 15 minutes incubation at room temperature with rotation. A second incubation was done by pooling fractions 1 to 5, 6 to 9 and 10 to 14. These were

incubated with 40 µl of the TiO₂ DBH slurry for another 15 min with rotation. Samples were centrifuged at 4000xg for 5 minutes and supernatant was removed. The beads were loaded on C8 tips (Empore) and washed on-tip with 10% ACN in 6% TFA, followed by 40% ACN in 6% TFA and finally 80% ACN in 6% TFA (Jersie-Christensen et al., 2016). Phospho-peptides were eluted first with 20 µL of 5% NH₄OH followed by 20 µL of 10% NH₄OH with 25% ACN. Eluted peptides were concentrated by vacuum centrifugation and loaded onto C-18 stage-tips.

Electrophoresis and immunoblotting

For Western Blot analysis of SUMOylated proteins enriched from His10-S2-K0 expressing U-2-OS cells, whole cell extracts or purified protein samples (procedures described above) were separated on Novex Bolt 4-12% Bis-Tris Plus gradient gels using MOPS buffer and transferred onto Hybond-C nitrocellulose membranes (GE Healthcare Life Sciences) using a submarine system (Life Technologies). Membranes were stained with Ponceau S (Sigma) to stain total protein and blocked with PBS containing 8% milk powder (w:v) and 0.05% Tween-20 (v:v) before incubating with the primary antibodies.

For analysis of total protein expression, phosphorylated proteins lysates were prepared as described below and SUMO enrichment from FLAG-SUMO2 expressing cells lysates from procedures described above were used. Cells were washed twice in ice-cold PBS, and lysed in complete modified RIPA buffer (50mM Tris, pH 7.5, 150mM NaCl, 1% NP-40, 0.1% sodium deoxycholate, 1 mM EDTA, 5 mM β-glycerolphosphate, 5 mM NaF, 1 mM sodium orthovanadate, complete inhibitor cocktail tablet (Roche)). Lysates were cleared by centrifugation at 13,000xg for 15 minutes. Protein concentration was measured using Bradford assay and lysates were boiled in NuPAGE LDS sample (Invitrogen) buffer with DTT for 15 minutes. Proteins were resolved on NuPAGE Novex 4-12% gradient Bis-Tris gels (Invitrogen) and transferred onto nitrocellulose membranes. After blocking for 1 hour in 5% skimmed milk (w:v) or 5% BSA (w:v) in PBS supplemented with 0.1% Tween-20 (PBST) (v:v), the membranes were incubated with primary antibodies in blocking solution overnight at 4°C. Blots were incubated with horse radish peroxidase coupled secondary antibody (Jackson ImmunoResearch) in 5% skimmed milk in PBST. Detection was performed with Novex ECL chemiluminescent Substrate Reagent Kit (Invitrogen).

Nanoflow LC–MS/MS

For LC-MS/MS analysis, STAGE-tips were eluted twice with 10 µl 40% ACN in 0.5% formic acid. All samples were analyzed on an Easy-nLC 1000 system coupled to the Q Exactive HF instrument (Thermo Fisher Scientific) through a nanoelectrospray ion source. Peptides were separated on an analytical column (inner diameter 75 µm), with 1.9 µm C₁₈ beads (Dr. Maisch, packed in-house) with a flow rate of 250nL/min at 40°C using an integrated column oven (PRSO-V1, Sonation GmbH). The spray voltage was set to 2kV, s-lens RF level to 50 and the heated capillary to 275°C. The mass spectrometer was operated in data-dependent acquisition (DDA) mode, in positive polarity mode with 30 seconds

dynamic exclusion window. Full scan resolution was set to 60,000, scan range to m/z 200-2000 and full-MS ion target value was 3e6. All fragmentation was performed by higher-energy collisional dissociation (HCD) with parallel acquisition. For enriched SUMO-peptides and SUMO-targets, as well as phospho-peptide enriched samples from Figure 1, the peptides were separated on a 15 cm analytical column with a 95 minute gradient from 8% to 75% CAN in 0.1% formic acid. For SUMO-target enriched samples from Figure 4, the gradient was run over 59 minutes. All samples were run with a top 7 method. For enriched SUMO- and phospho-peptides in Figure 5, the samples were run on a 50cm column with a 264 minute gradient from 8% to 75% ACN in 0.1% formic acid. A top 10 MS/MS method was utilized for analysis of these samples.

Quantification and statistical analysis

Raw Data Processing and Analysis

All raw LC-MS/MS data was analyzed by MaxQuant software suite v1.4.1. or v1.5.11. using the Andromeda Search engine. The searches were performed against the complete human UniProt database. The default contaminant protein database was included and identifications from this source were excluded during analysis of the data. We used the “match between runs” option. Cysteine carbamidomethylation was set as a fixed modification while methionine oxidation, protein N-terminal acetylation, and pyro-glutamate formation from glutamine were set as variable modifications. For identification of SUMOylated peptides SUMOylation remnant (QQTGG) with monoisotopic mass of 471.20776 Da and pyroSUMOylation (pyroQQTGG) remnant with monoisotopic mass of 454.18121 Da modification on lysine residues was set as a variable modification. Similarly, phosphorylation of serine, threonine, and tyrosine residues were set as variable modifications for identification of phosphorylation sites. Using the target-decoy strategy the false discovery rate (FDR) was set to 1% for peptide spectral matches (PSM) and protein identification. For each experimental setup the raw MS data was analyzed in a separate MaxQuant analysis. When analyzed together each experiment was processed in separate parameter groups. Mapping of proteins with modifications sites is based on the identified sequence, and protein inference is performed on the global level applying information from peptides identified from all raw MS data. This latter approach was applied for MMC treated FLAG-SUMO2 IPs, HIS10 SUMO2-K(0) IPs, phosphoproteome and proteomes from Figures 1, S2, 2 and S2.

Bioinformatics analysis

Functional protein interaction networks were mapped using the STRING database (Szklarczyk et al., 2015) and further processed with Cytoscape (www.cytoscape.org). For network mapping a minimum confidence score of 0.4 was required. Clusters of highly interconnected nodes within the networks were extracted using the MCODE Cytoscape app. Analysis to identify enrichment of Gene Ontologies (GO) were done using InnateDB (Lynn et al., 2008) and P-values were corrected for multiple testing with by Benjamini Hochberg FDR. Hierarchical clustering, as well as ontology enrichment of clusters, was done

using Perseus (Tyanova et al., 2016). For the clustering illustrated in Figures 5 and S5, at least three quantifications were required in at least one condition. From these we performed a multiple sample t-test using ANOVA with an FDR threshold of 0.05. Normalization was performed by subtracting the median intensity for each condition. Unsupervised hierarchical clustering was performed on the targets that were found significant. Rows were clustered using the Canberra approach and columns by Pearson correlation.

Analysis for enriched sequence motifs for the PTM enriched data was performed using IceLogo (Colaert et al., 2009), and overrepresentation was scored with fold-change with a p-value cutoff of 0.01 or 0.05, using the complete acquired dataset (non-regulated or non-modified) as the reference.

Resource Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
DMEM	Gibco	31966-047
SILAC DMEM	Thermo Fisher Scientific, PAA	PAA-EL15-086
FBS (heat-inactivated)	Gibco	10270-106
Dialyzed FBS	Sigma Aldrich	F0392-500ML
Pen/strep	Gibco	15140-122
Trypsin-EDTA (0.05%)	Gibco	25300-054
L-Glutamine	Gibco	25030-024
L-Arginine monohydrochloride	Sigma Aldrich	A6969
Poly-D-lysine hydrobromide	Sigma Aldrich	P7280
L-Lysine:2HCL (13C, 99%; 15N2, 99%)	Cambridge Isotope Labs	CNLM-290-H
L-Lysine:2HCL (4,4,5,5-D5, 96-98%)	Cambridge Isotope Labs	DLM-2640-0
L-Arginine:HCL (13C6, 99%; 15N4, 99%)	Cambridge Isotope Labs	CNLM-539-H
L-Arginine:HCL (13C6, 99%)	Cambridge Isotope Labs	CNLM-2265-H
PBS	Gibco	20012-068
Thymidine	Sigma Aldrich	T9250
Mitomycin C	Sigma Aldrich	M4287
Hydroxyurea	Sigma Aldrich	H8627
Guanidine hydrochloride	Sigma Aldrich	G3272
ATR-45	Ohio State University	N/A
AZ20	Selleckchem	S7050
KU55933	Calbiochem	118500
KU60019	Selleckchem	S1570
SCH900776	MedKoo Biosciences	118500
Propidium iodide	Sigma Aldrich	P4170
TCEP	Sigma Aldrich	C4706
CAA	Sigma Aldrich	22790
Quick Start Bradford 1X Dye Reagent	BioRad	500-0205
REAGENT or RESOURCE	SOURCE	IDENTIFIER
ANTI-FLAG® M2 Affinity Gel	Sigma Aldrich	A2220
FLAG peptide	Sigma Aldrich	F3290
Trizma base	Sigma Aldrich	T1503
Complete mini EDTA-free protease inhibitor cocktail	Roche	04693124001
Trypsin	Sigma Aldrich	T6567
Lys-C	Wako Chemicals	129-02541

Endoproteinase Lys-C, Sequencing Grade	Promega	V1071
TFA	Sigma Aldrich	T6508
Acetonitrile	Merck	1.00030.2500
Ammonium bicarbonate	Sigma Aldrich	09830
5µM Titansphere	GL Sciences	GS 502075000
2,5-dihydroxybenzoic acid	Sigma Aldrich	85707
Ammonia solution 25%	Merck	1054321011
Amicon Ultra Centrifugal Filter Unit	Millipore	UFC900396
NuPAGE LDS sample buffer 4x	Novex	NP0007
1.0mm x 10 well NuPAGE 4-12% Bis-Tris Gel	Novex	NP032BOX
Amersham Hybond ECL nitrocellulose membrane	GE Healthcare	RPN303D
Novex Bolt 4-12% Bis-Tris Plus gel	Life Teachnologies	BG04125BOX
Colloidal Blue staining kit	Invitrogen	LC-6025
SepPak Classic C18 cartridges	Waters	WAT051910
C18 Column, 130Å, 3.5 µm, 4.6 mm X 250 mm	XBridge PST	186003570
2, 5 Dihydroxybenzoic acid	Sigma Aldrich	85707-1G-F
C18 (Octadecyl) 47mm	Empore	2215
C8 47 mm Extraction Disk	Empore	2214
Monoclonal ANTI-FLAG® M2 antibody	Sigma Aldrich	F1804
ReproSil-Pur 120 C18-AQ, 1.9 µm	Dr. Maisch	r119.aq
Ni-NTA beads	Qiagen	30210
Pierce BCA protein Assay Kit	Thermo Scientific	23227
Vivacon 500 Filter - 100,000MWCO	VN01H41	SartoriusStedim
VIVACON500 Filter - 10,000MWCO	VN01H01	SartoriusStedim
REAGENT or RESOURCE	SOURCE	IDENTIFIER
TopBP1 antibody	Bethyl Laboratories	A300-111A
Anti-Chk1 (phospho S345) antibody	Abcam	ab47318
Chk1 (2G1D5) Mouse mAb	Cell Signaling Technology	2360
Phospho-Chk2 (Thr68) Antibody	Cell Signaling Technology	2661
Chk2 (1C12) Mouse mAb	Cell Signaling Technology	3440
Phospho-Histone H2A.X (Ser139) Antibody	Cell Signaling Technology	2577
Histone H2A.X Antibody	Cell Signaling Technology	2595
BRCA1 Antibody	Cell Signaling Technology	9010
Anti-Vinculin antibody, Mouse monoclonal	Sigma Aldrich	V9264
Anti-RPA32/RPA2 antibody [9A1]	Abcam	ab125681

mouse monoclonal anti-SUMO-2/3	Abcam	Ab81371
Deposited Data		
Raw and analyzed data		PXD006361
Experimental Models: Cell Lines		
U-2-OS	Roeland Dirks (Leiden)	
Software and Algorithms		
MaxQuant 1.5.3.6		http://www.coxdocs.org/doku.php?id=maxquant:start
R software		https://www.r-project.org/
String v10		http://string-db.org/
Cytoscape		www.cytoscape.org
IceLogo		http://iomics.ugent.be/icelogoserver/index.html
Perseus		http://www.perseus-framework.org

Supplemental References

- Bathth, T.S., Francavilla, C., and Olsen, J. V. (2014). Off-line high-pH reversed-phase fractionation for in-depth phosphoproteomics. *J. Proteome Res.*
- Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., and Gevaert, K. (2009). Improved visualization of protein consensus sequences by iceLogo. *Nat. Methods* 6, 786–787.
- Hendriks, I.A., and O Vertegaal, A.C. (2016). A high-yield double-purification proteomics strategy for the identification of SUMO sites. *Nat. Publ. Gr.* 11.
- Hendriks, I.A., D’Souza, R.C., Yang, B., Verlaan-de Vries, M., Mann, M., and Vertegaal, A.C. (2014). Uncovering global SUMOylation signaling networks in a site-specific manner. *Nat Struct Mol Biol* 21, 927–936.
- Jersie-Christensen, R.R., Sultan, A., and Olsen, J. V. (2016). Simple and Reproducible Sample Preparation for Single-Shot Phosphoproteomics with High Sensitivity. (Springer, New York, NY), pp. 251–260.
- Larsen, M.R. (2005). Highly Selective Enrichment of Phosphorylated Peptides from Peptide Mixtures Using Titanium Dioxide Microcolumns. *Mol. Cell. Proteomics* 4, 873–886.
- Lynn, D.J., Winsor, G.L., Chan, C., Richard, N., Laird, M.R., Barsky, A., Gardy, J.L., Roche, F.M., Chan, T.H., Shah, N., et al. (2008). InnateDB: facilitating systems-level analyses of the mammalian innate immune response. *Mol. Syst. Biol.* 2.
- Olsen, J. V., Blagoev, B., Gnäd, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006). Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* 127, 635–648.
- Ong, S.-E. (2002). 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.
- Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips.
- Schimmel, J., Eifler, K., Otti Sigurðsson, J., Cuijpers, S.A., Hendriks, I.A., Verlaan-de Vries, M., Kelstrup, C.D., Francavilla, C., Medema, R.H., Olsen, J. V., et al. (2014). Molecular Cell Resource Uncovering SUMOylation Dynamics during Cell-Cycle Progression Reveals FoxM1 as a Key Mitotic SUMO Target Protein. *Mol. Cell* 53, 1053–1066.
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447–52.
- Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods*.
- Xiao, Z., Chang, J.-G., Hendriks, I.A., Sigurðsson, J.O., Olsen, J. V., and Vertegaal, A.C.O. (2015). 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. *Mol. Cell. Proteomics* 14, 1419–1434

Chapter 4

The STUbL RNF4 regulates protein group SUMOylation by targeting the SUMO conjugation machinery

Ramesh Kumar^{1,4}, Román González-Prieto^{1,4}, Zhenyu Xiao^{1,4},
Matty Verlaan-de Vries¹ and Alfred C.O. Vertegaal^{1,5}

¹Department of Molecular Cell Biology, Leiden University Medical Center, Albinusdreef 2, 2333 ZA, Leiden, the Netherlands.

²current address: Cancer and Stem Cell Biology Program, Duke–NUS Graduate Medical School, 8 College Road, Singapore 169857, Singapore.

³current address: IFOM-p53Lab Joint Research Laboratory, IFOM, FIRC Institute of Molecular Oncology, 20139 Milan, Italy.

⁴ These authors contributed equally

⁵Address correspondence to: A.C.O.V. (e-mail: vertegaal@lumc.nl)

Chapter 4 has been published in **Nature communications**

Nat Commun. 2017 Nov 27;8(1):1809. doi: 10.1038/s41467-017-01900-x.

Supplementary Tables are available online

Summary

SUMO-Targeted Ubiquitin Ligases (STUbLs) mediate the ubiquitylation of SUMOylated proteins to modulate their functions. In search of direct targets for the STUbL RNF4, we have developed TULIP (Targets for Ubiquitin Ligases Identified by Proteomics) to covalently trap targets for ubiquitin E3 ligases. TULIP methodology could be widely employed to delineate E3 substrate wiring. Here we report that the single SUMO E2 Ubc9 and the SUMO E3 ligases PIAS1, PIAS2, PIAS3, ZNF451 and NSMCE2 are direct RNF4 targets. We confirm PIAS1 as a key RNF4 substrate. Furthermore, we establish the ubiquitin E3 ligase BARD1, a tumour suppressor and partner of BRCA1, as an indirect RNF4 target, regulated by PIAS1. Interestingly, accumulation of BARD1 at local sites of DNA damage increases upon knock-down of RNF4. Combined, we provide insight into the role of the STUbL RNF4 to balance the role of SUMO signaling by directly targeting Ubc9 and SUMO E3 ligases.

1 Introduction

Reversible post-translational modifications (PTMs) functionally regulate essentially all proteins¹. These modifications comprise small chemical modifications such as phosphorylation, methylation and acetylation and small proteins that belong to the ubiquitin family². The ubiquitin family includes Small ubiquitin-like modifiers (SUMOs). SUMOylation is essential for viability in eukaryotes with the exception of *S. pombe*³⁻⁵. Mouse embryos deficient for SUMOylation die early after implantation due to chromosomal aberrancies, indicating a key role for SUMO in maintaining genome stability⁴. SUMOylated proteins predominantly localize to the nucleus and are enriched at local sites of DNA damage, consistent with their key role in the DNA damage response (DDR)^{6,7}. SUMO target proteins regulated in response to DNA damage include BRCA1 and 53BP1^{6,7}.

Different types of PTMs functionally cooperate to fine-tune protein activity⁸. Interestingly, SUMO-targeted ubiquitin ligases (STUbLs) regulate the stability of a subset of SUMOylated proteins⁹. These STUbLs were first identified in yeast, and later also found in mammals^{9,10}. Consistent with the important role of SUMOylation and ubiquitylation in genome stability, these STUbLs play key roles in the maintenance of genome integrity¹¹⁻¹⁸. Mice deficient for the STUbL RNF4 die during embryogenesis¹⁵ and RNF4-deficient MEFs showed prolonged DNA damage signaling upon exposure to ionizing radiation¹⁵. Efficient non-homologous end joining and homologous recombination are dependent on RNF4. Similar to SUMO, RNF4 is enriched at local sites of DNA damage, mediated by its SUMO-Interaction Motifs (SIMs)^{13-15,19}.

Currently, we are limited in our understanding of the role of RNF4 because of limited insight into the RNF4-regulated SUMO target proteins. So far, RNF4 has been identified to be involved in the regulation of components from different DNA damage repair pathways. MDC1 and BRCA1 have been found as SUMOylated RNF4 targets relevant for genome stability^{13-15,20}. SUMOylation of MDC1 and BRCA1 was increased upon exposure of cells to ionizing radiation and knocking down RNF4 increased the amount of SUMOylated MDC1 and BRCA1¹⁵. In contrast, SUMOylation of 53BP1 was not upregulated in response to RNF4 knockdown. Recently, the Fanconi Anemia ID (FANCI-FANCD2) complex was found as RNF4 target in the context of DNA crosslink repair²¹. Additionally, RNF4 regulates the degradation of the histone demethylase JARID1B/KDM5B in response to MMS to mediate transcriptional repression²².

Here we set out to identify novel STUbL target proteins and developed TULIP methodology to trap and enrich RNF4 targets. We identify five SUMO E3 ligases and the SUMO E2 enzyme Ubc9 as direct RNF4 targets. Combined, our findings provide novel insight in the efficient downregulation of SUMO signaling by RNF4.

2 Results

2.1 Identification of SUMOylated proteins regulated by RNF4

We used an unbiased proteomics approach to purify and identify SUMOylated proteins regulated by the STUbL RNF4 (Fig. 1a), employing a U2OS cell line expressing low levels of His10-SUMO2²³ and three independent shRNA constructs targeting RNF4. SUMO2 conjugates were purified from cells infected with lentiviruses expressing these shRNA constructs or a non-targeted control shRNA construct, using three biological replicates. SUMO2 conjugates were analyzed by mass spectrometry using three technical replicates (Fig. 1b). Consistent with earlier observations, SUMO2 conjugate levels were significantly increased upon RNF4 knockdown^{22, 24}, indicating that RNF4 is the dominant human STUbL, since no efficient functional compensation occurs for the absence of RNF4 (Fig. 1c and Supplementary Fig. 1a).

Label-free quantification (LFQ) of proteins identified by mass spectrometry indicated that 222 SUMO2 conjugates were consistently upregulated upon RNF4 knockdown (Fig. 1d)(Supplemental Dataset 1). Pearson analysis showed the reproducibility of the experiments (Supplementary Fig.1b). The known RNF4-regulated proteins MDC1, BRCA1, PML and KDM5B/JARID1B were identified in our screen, serving as positive controls^{13-15, 22, 24}.

Network analysis revealed an extensive interaction between the identified proteins, indicating that RNF4 regulates a large protein network (Supplementary Fig.2a). A major set of RNF4-regulated proteins identified in our screen are involved in nucleic acid metabolism with a particular emphasis on SUMOylation, transcription, DNA repair and chromosome segregation (Fig 1e and Supplemental Dataset 2). DDR components identified in our screen include BLM, USP7, RAD18, XRCC5 and PARP1.

Subsequently, we confirmed BARD1 and RAD18 as novel RNF4-regulated proteins with important roles in the DDR (Supplementary Fig.2b). Additionally, we verified the histone-lysine N-methyltransferase SETDB1 as a protein regulated by RNF4 (Supplementary Fig.2b), confirming that the proteins identified in our screen are SUMO2 conjugates regulated by RNF4.

2.2. TULIP methodology to identify direct RNF4 substrates

Our screen yielded PIAS SUMO E3 ligases and a considerable set of other proteins as SUMO conjugates regulated by RNF4 knockdown. Thus, we hypothesized that other identified proteins could be indirectly regulated by RNF4, with the SUMO E3 ligases as primary targets. To address this hypothesis, we developed methodology that would allow us to purify and identify primary ubiquitin E3 ligase substrates. Identifying primary substrates of ubiquitin E3 ligases in cells is notoriously challenging. Clearly, knockdown strategies such as the one employed by us for the first part of our project are helpful, but unable to distinguish between primary and secondary effects. Recently, Ubait methodology was developed to study ubiquitin E3 ligases²⁵. However, O'Conner et al. could not

distinguish between covalent targets and non-covalent interactors because of the mild purification procedure employed.

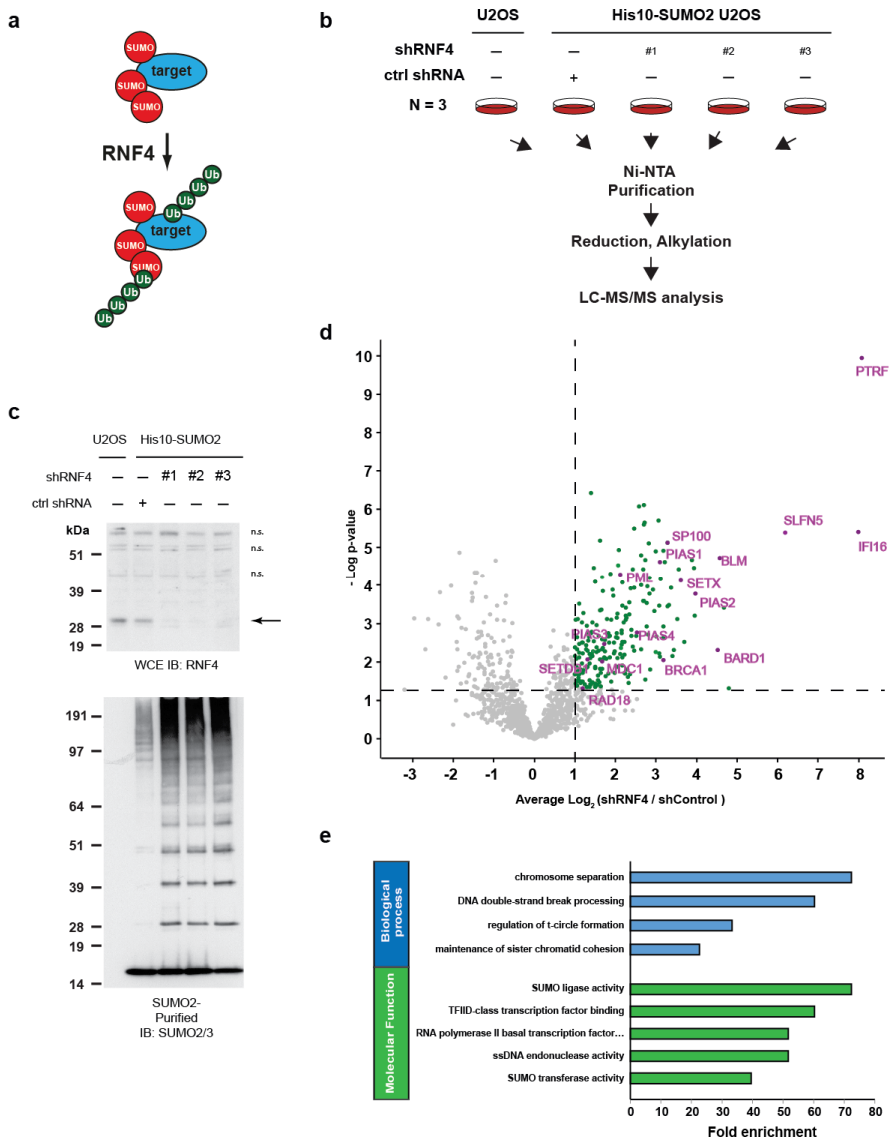


Figure 1. SUMO2 substrates regulated by the SUMO-Targeted Ubiquitin Ligase (STUbL) RNF4. **(a)** The STUbL RNF4 binds and ubiquitinates SUMOylated substrates. **(b)** Experimental approach. **(c)** U2OS cells stably expressing His10-SUMO2 were separately infected with lentiviruses expressing three different shRNAs directed against RNF4 or a control shRNA. Three days post infection, cells were harvested, lysed in a denaturing buffer and His10-SUMO2 conjugates were purified. RNF4 knockdown efficiency and the purification efficiency of His10-SUMO2 conjugates were verified by immunoblotting. Three

biological replicates were performed and the purified proteins were identified by mass spectrometry. n.s. indicates non-specific bands. **(d)** Volcano plot showing RNF4-regulated SUMO2 target proteins. Each dot represents an identified protein. The green dots represent proteins increased for SUMOylation in response to RNF4 knock down with an average log2 ratio greater than 1 when this increase is statistically significant with a $-\log_{10}$ of the p-value higher than 1.3. This corresponds to a two-fold increase with a p-value lower than 0.05 for statistical significance. Selected proteins are highlighted in purple. **(e)** Gene ontology analysis of all SUMO enriched proteins after RNF4 knockdown from **(d)** regarding biological process and molecular function. Full Gene ontology is shown in Supplementary Dataset 2.

To address this pitfall, we designed and employed a complementary strategy termed TULIP: Targets for Ubiquitin Ligases Identified by Proteomics. We constructed three different lentiviral vectors consisting of an inducible linear fusion of our E3 of interest, RNF4, followed by ubiquitin, connected by a linker that contains the 10HIS tag. Two additional constructs were made as negative controls, one lacking Ubiquitin and one with Ubiquitin but lacking the diGly motif. Both negative controls would prevent the covalent binding of our TULIP construct and its target proteins (Fig. 2a). Employing the His10-tag, enabled the use of fully denaturing buffers in the purification procedure, thereby removing non-covalently bound interaction partners (Fig. 2). The TULIP methodology is widely applicable to identify direct substrates for ubiquitin E3 ligases.

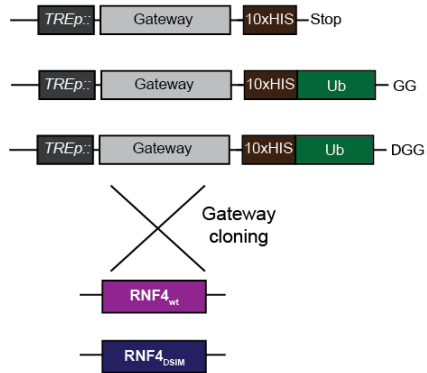
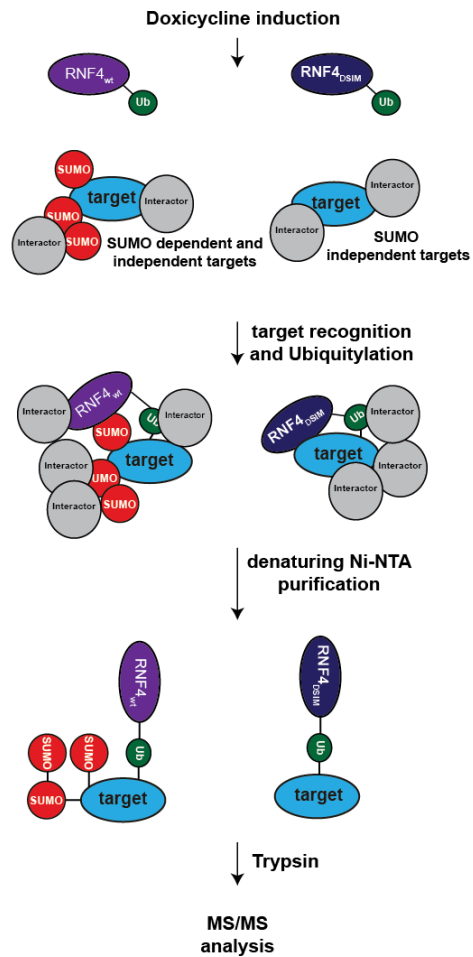
a**b**

Figure 2. RNF4-TULIP constructs and strategy. **(a)** Gateway cassette followed by a linker and either a 10xHIS motif, 10xHIS tagged Ubiquitin or 10xHIS tagged Ubiquitin lacking the C-terminal GG motif, were cloned under the control of TRE promoter in a lentiviral backbone to generate stable cell lines. The Gateway cassette was substituted by RNF4 wild type and SIM deficient constructs using Gateway cloning. **(b)** After stable cell line generation, the expression of the constructs was induced by doxycycline. RNF4 covalently attached to its target proteins using the 10xHIS-Ubiquitin. Target proteins covalently attached to the RNF4-TULIP constructs were purified under denaturing conditions using Ni-NTA beads, avoiding the co-purification of interactors. Subsequently tryptic peptides were analyzed by tandem mass spectrometry to determine the identity of the RNF4 targets.

We induced the expression of the RNF4 wild type and Δ SIM mutant TULIP constructs in U2OS cells, and inhibited the proteasome in order to prevent further degradation of ubiquitylated targets or used DMSO as control. We purified HIS conjugates and analyzed them by immunoblotting (Fig. 3a). We could observe the appearance of RNF4-target conjugates both in wild type and Δ SIM mutant ubiquitin constructs but not in our negative controls. The pattern of the conjugates was different between the wild type and Δ SIM mutant, indicating that we could distinguish between SIM-dependent and independent RNF4 targets.

Next, we identified RNF4-TULIP targets by mass spectrometry (Fig. 3b-e, Supplementary Dataset 3). Gene ontology analysis of RNF4-TULIP targets highlighted the SUMOylation machinery as the most highly enriched category (Supplementary Fig.3, Supplementary Dataset 5). Employing the TULIP strategy allowed us to identify five SUMO E3 ligases PIAS1, PIAS2, PIAS3, ZNF451 and NSMCE2, and the single SUMO E2 ligase UBC9 as RNF4 targets, regulated in a SUMO-Interaction Motif (SIM) and proteasome-dependent manner. These results highlight SUMO E3 ligases, and remarkably also the SUMO E2 ligase, as direct targets for RNF4, explaining the efficient downregulation of SUMO signaling by RNF4.

Proteins identified by the RNF4 knockdown strategy and also by the RNF4-TULIP strategy are depicted in a Venn diagram in Fig. 3f and summarized in Supplementary Dataset 4.

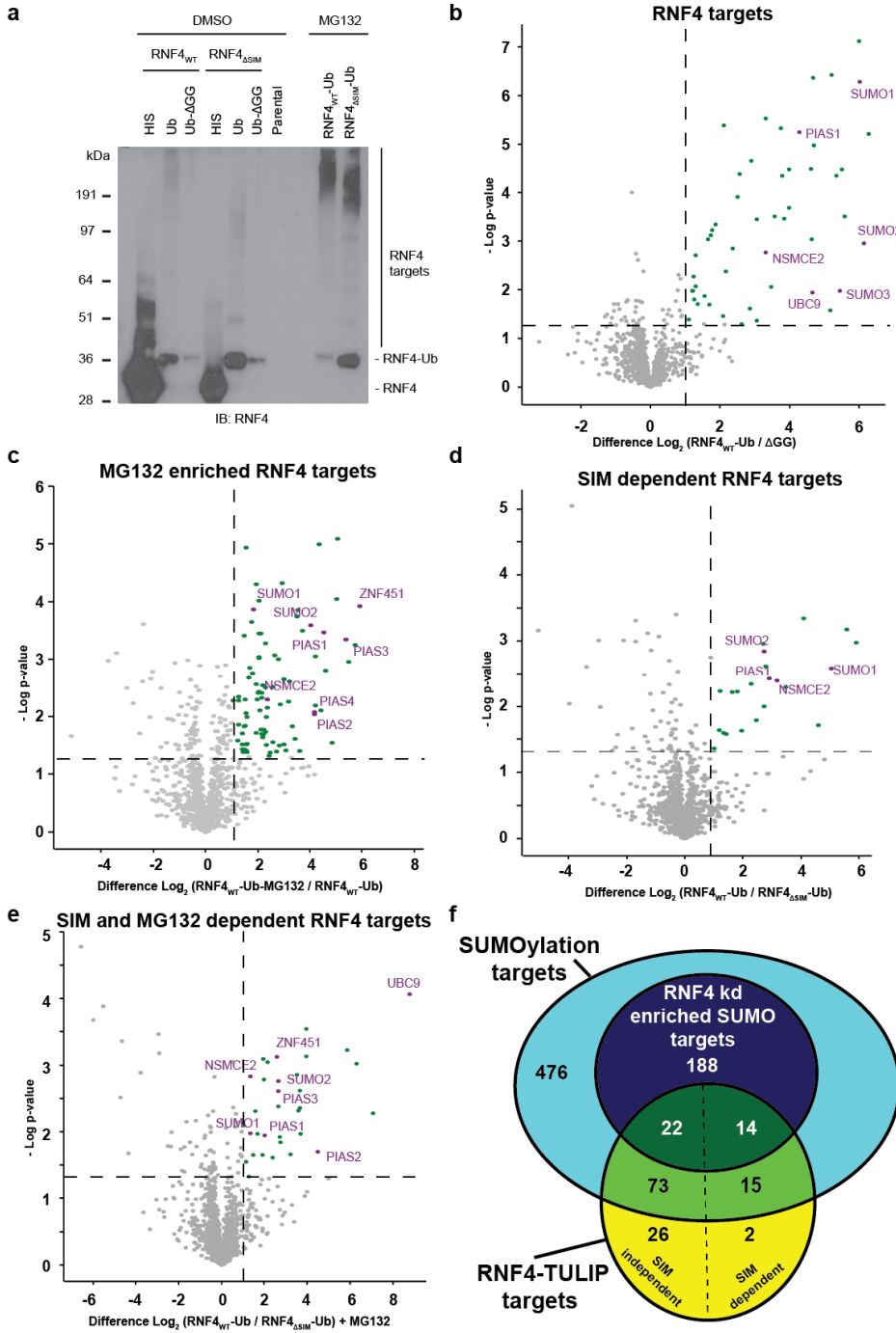


Figure 3. RNF4-TULIP mass spectrometry analysis. (a) Stable U2OS cell lines expressing the different RNF4-TULIP constructs were generated and the expression of the constructs was induced with doxycycline. RNF4-TULIP conjugates were purified and analysed by immunoblotting using RNF4 antibody. Experiments were performed in triplo. (b-e) Volcano plots depicting the statistical differences between three independent sample sets, analysed by mass spectrometry. Dots represent individual proteins. Green dots and purple dots represent proteins that have an statistically significant ($-\log_{10}$ p-value > 1.3) average enrichment higher than 1 (\log_2). This corresponds to a two-fold increase with a p-value lower than 0.05 for statistical significance. Purple dots correspond to components of the SUMOylation machinery. (f) Venn Diagram representing the overlay between the RNF4-TULIP targets (SIM-dependent and independent) and the different SUMOylation targets that are enriched or not upon RNF4 knockdown.

2.3 RNF4 targets the SUMO E3 ligase PIAS1

SUMOylated proteins which are targeted for degradation by RNF4 should be enriched upon RNF4 knockdown and be a direct ubiquitylation target for RNF4-TULIP, enriched after proteasome inhibition in a SIM-dependent manner (Fig. 4a). Proteins matching these conditions are the SUMO E3 ligases, Ubiquitin E3 ligases RNF216 and Rad18, and other SUMOylated targets, which could explain how RNF4 efficiently controls group SUMOylation in cells. PIAS1 was the most highly enriched SUMO E3 ligase using the both RNF4-knockdown and TULIP strategy (Fig. 4a). First, we verified these results for PIAS1 by immunoblotting, showing that PIAS1 is indeed regulated by RNF4 in a SIM-dependent manner and subsequently targeted to the proteasome for degradation (Fig. 4b). Next, we verified that SUMOylated PIAS1 accumulated upon RNF4 knockdown (Fig. 4c).

Subsequently, we investigated the accumulation of SUMO-2/3 levels upon RNF4 knockdown in cells co-depleted for PIAS1 and/or PIAS4 as a negative control (Figs. 4d and Supplementary Fig. 4a). Interestingly, SUMO-2/3 levels did not accumulate in cells co-depleted for PIAS1, but did accumulate upon co-depletion for PIAS4, indicating that PIAS1 is a major target for RNF4.

Next, we tested whether RNF4 mediates the ubiquitylation of PIAS1 in cells. U2OS cells stably expressing His10-ubiquitin were infected with three different knockdown constructs for RNF4, or with a control virus (Fig. 4e and Supplementary Fig. 4b). Cells were harvested and ubiquitin conjugates were purified under denaturing conditions. PIAS1 ubiquitylation was analyzed by immunoblotting, demonstrating that RNF4 is required for efficient PIAS1 ubiquitylation. Our results indicate that RNF4 limits overall SUMOylation levels in cells by targeting SUMO E3 ligases.

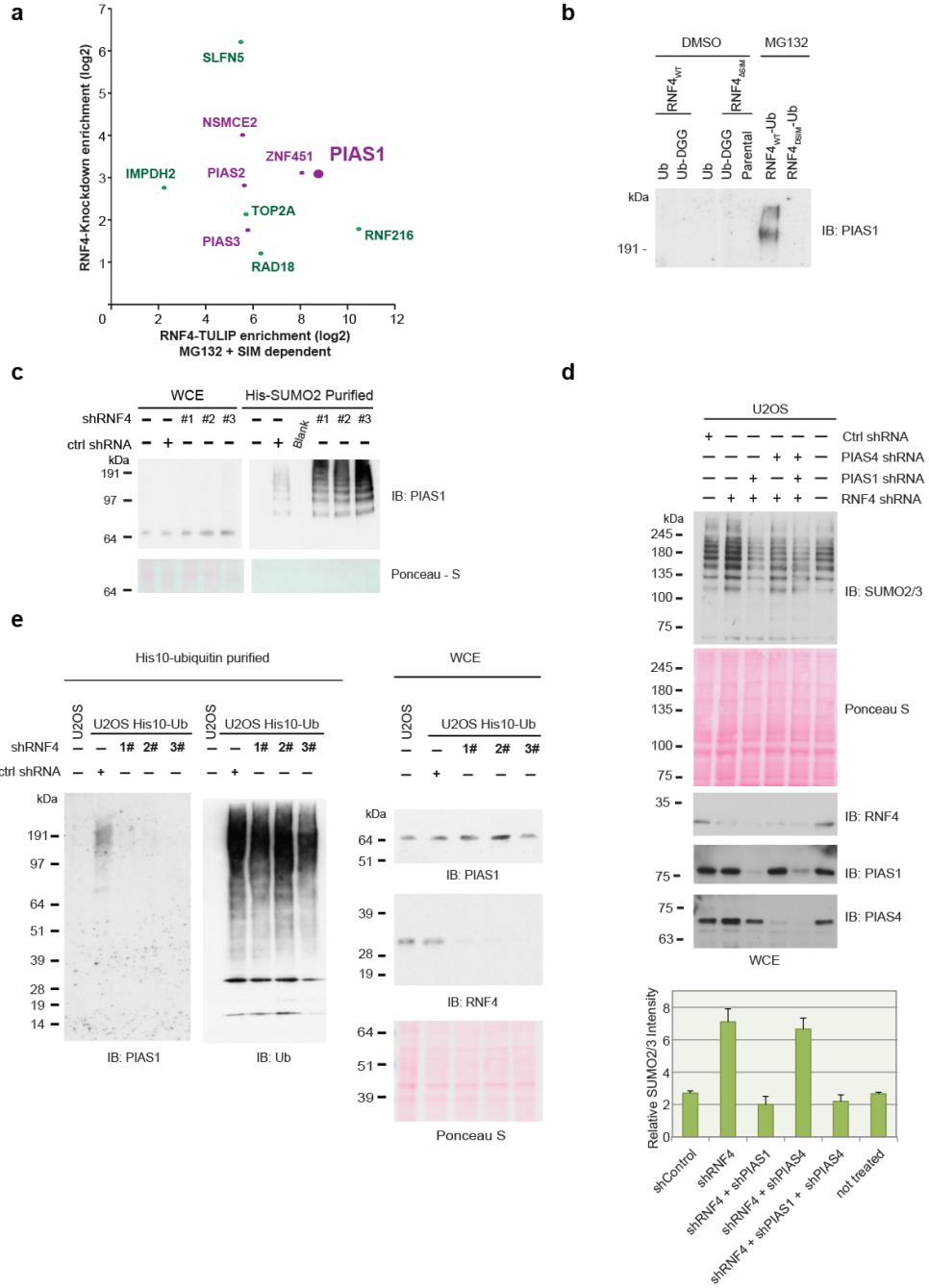


Figure 4. RNF4 limits SUMO signalling by targeting SUMO E3s for degradation. **(a)** Scatter plot showing RNF4-TULIP target proteins that are enriched upon MG132 treatment in a SIM-dependent manner and also found to be SUMO targets enriched upon RNF4 knockdown. Purple dots represent SUMO E3 ligases. **(b)** PIAS1 is a SIM-dependent RNF4 substrate targeted to the proteasome for degradation. Stable U2OS cell lines expressing the different RNF4-TULIP constructs were generated and the expression of the constructs was induced with doxycycline. RNF4-TULIP conjugates were purified from the indicated cell lines after MG132 or DMSO treatment and analysed by immunoblotting using PIAS1 antibody. **(c)** U2OS cells stably expressing His10-SUMO2 were separately infected with lentiviruses expressing three different shRNAs directed against RNF4 or a control shRNA. Three days post infection, cells were harvested and His10-SUMO2 conjugates were purified from denaturing lysates and analysed by immunoblotting against PIAS1. **(d)** The overall increase in protein SUMOylation upon RNF4 knockdown is counteracted by co-knock down of PIAS1. U2OS cells were (co)-infected with lentiviruses expressing shRNAs against RNF4, PIAS1 or PIAS4 or a control shRNA as indicated. Three days after infection, cells were lysed in a denaturing buffer and knockdown efficiencies and overall levels of SUMO2/3 were analysed by immunoblotting. The results were independently confirmed using a second set of shRNAs. The experiment described in was independently performed three times and quantified. Averages and standard deviations of SUMO2/3 conjugate levels are depicted. **(e)** RNF4 regulates PIAS1 ubiquitylation. U2OS cells stably expressing His10-ubiquitin were infected with lentiviruses expressing shRNAs directed against RNF4 or a control shRNA. Three days after infection, cells were lysed in a denaturing buffer and His10-ubiquitin conjugates were purified. The levels of ubiquitylated PIAS1 were verified by immunoblotting. Similarly, RNF4 knockdown efficiency was verified by immunoblotting. The experiment was independently repeated and consistent results were obtained.

2.4 BARD1 is SUMOylated in response to DNA double strand breaks

Subsequently, we studied the regulation of BARD1 by SUMOylation and RNF4 in more detail because BARD1 plays a critical role in the DDR. Consistent with the fate of the first identified RNF4 substrates PML and PML-RAR α ^{24,26}, SUMOylated BARD1 is degraded by the proteasome (Fig. 5a). Additionally, we found that the DNA damaging agents MMS, Bleocin and Ionizing Radiation (IR) stimulate the SUMOylation of BARD1, but did not change overall SUMO levels (Supplementary Fig.5a), indicating that BARD1 is SUMO-modified upon activation of the DDR.

Interestingly, in the absence of exogenous DNA damaging agents, BARD1 SUMOylation could be detected upon RNF4 knockdown (Figs. 5b, Supplementary Figs. 5b and 2b). This could potentially be triggered by replication damage and subsequent replication fork collapse under these conditions²⁷. RNF4 depleted cells exposed to IR resulted in higher levels of BARD1 SUMOylation, which was further increased after inhibition of the proteasome (Fig. 5b and Supplementary Fig. 5b).

To study whether BARD1 and BRCA1 are SUMOylated in replicating cells, we purified SUMO2 conjugates from cells cycle synchronized in different phases of the cell cycle (Fig. 5c, 5d and Supplementary Fig. 5d). Consistently, BARD1 - and BRCA1 SUMOylation could be found in S and S/G2, but not in G1 in the absence of proteasome inhibition. Cells were stained by propidium iodide and analyzed by flow cytometry to verify cell cycle synchronization (Supplementary Fig.5c).

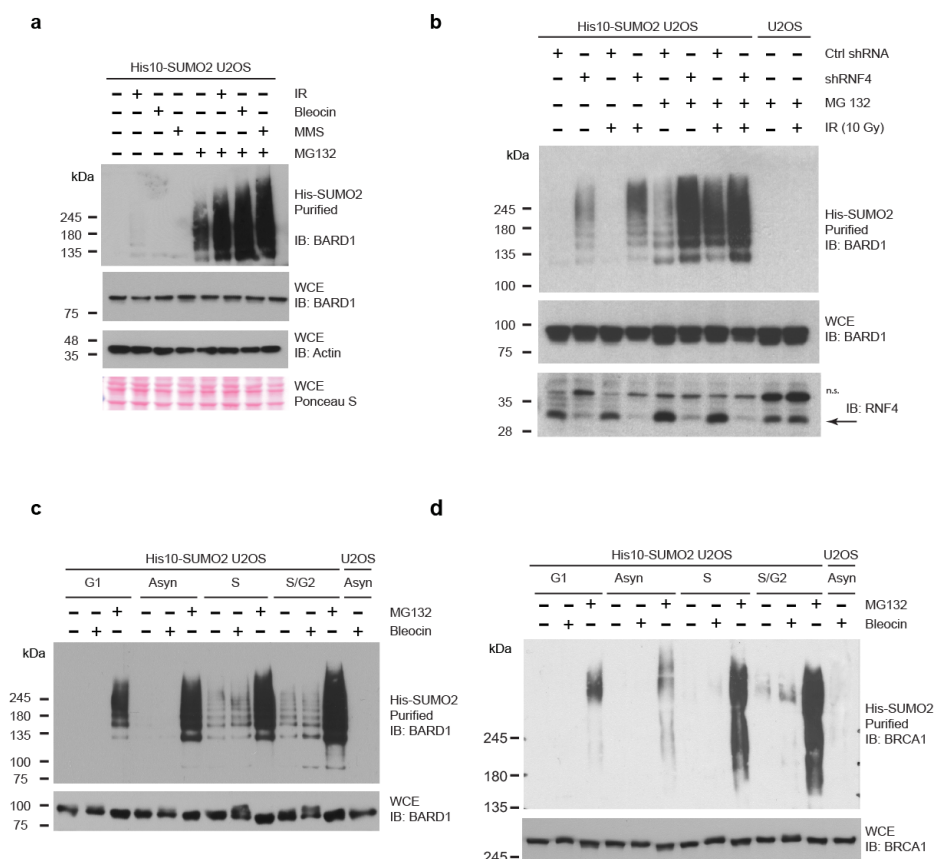


Figure 5. BARD1 SUMOylation is enhanced in response to DNA damage and proteasome inhibition. **(a)** BARD1 is SUMOylated in response to DNA damage and SUMOylated BARD1 is degraded by the proteasome. U2OS cells stably expressing His10-SUMO2 were either mock-treated or treated with MG132 (10 μ M) to inhibit the proteasome. One hour after the start of MG132 treatment, cells were either mock-treated or irradiated with IR (10 Gy), or treated with Bleocin (5 μ g/ml) or MMS (0.001%). The cells were subsequently incubated for 4 h and lysed. His10-SUMO2 conjugates were purified (PD) and analyzed by immunoblotting. Whole cell extracts (WCE) were analyzed by immunoblotting to determine the total levels of BARD1. **(b)** Similar to (a), RNF4 depleted cells were either mock treated or treated with MG132 and/or IR (10 Gy) treated as indicated. His10-SUMO2 purified samples and whole cell extracts were analyzed by immunoblotting using an antibody directed against BARD1. n.s. indicates non-specific bands. **(c)** Cell cycle stage dependent SUMOylation of BARD1. Similar to (a) and (b), U2OS cells stably expressing His10-SUMO2 were arrested in the G1, S or G2/M phase of the cell cycle. These cells were either DMSO treated or treated with Bleocin (5 μ g/ml) or MG132 (10 μ M) as indicated. His10-SUMO2 purified samples and whole cell extracts were analyzed by immunoblotting using an antibody directed against BARD1. **(d)** Cell cycle stage-dependent SUMOylation of BRCA1. Similar to (c), His10-SUMO2 purified samples and whole cell extracts protein samples from G1, S or G2/M were analyzed by immunoblotting using an antibody directed against BRCA1. Knockdown efficiency and the purification of His10-SUMO2 conjugates were verified by immunoblotting, shown in Supplementary Fig. 4. Unprocessed full-size scans of blots are provided in Supplementary Fig.9. Experiments presented in

this section as well as in supplementary figure section was repeated 2-4 times to test the reproducibility of data. The experiments described in Fig. 5a were repeated two times, in Fig. 5b four times and Fig. 5c and 5d three times.

2.5 PIAS1 co-regulates BARD1- BRCA1 SUMOylation upon DNA DSBs

PIAS SUMO E3 ligases, facilitate the transfer of SUMO from Ubc9 to substrates. It has been reported that PIAS4 depletion on its own severely impaired 53BP1 accumulation in laser-induced DNA damage and in Ionizing Radiation Induced Foci (IRIF)⁶. The SUMO E3 ligases PIAS1 and PIAS4 are required for RAD51 accumulation at DNA damage sites²⁸. Here, we have used two different sets of shRNAs to deplete PIAS1 and PIAS4 and tested the SUMOylation of BARD1 and its partner BRCA1. In PIAS1 depleted cells, we noted a significant reduction of BARD1 SUMOylation in cells treated with MG132 alone or in combination with Bleocin. Unlike PIAS1 depletion, DNA damage induced BARD1 SUMOylation was only modestly reduced in PIAS4 depleted cells (Fig. 6a, Supplementary Figs. 6a and 6c). Consistently, BRCA1 SUMOylation was significantly reduced in PIAS1 depleted cells, while PIAS4 depletion did not alter the SUMOylation of BRCA1 (Fig. 6b and Supplementary Fig. 6b). Knockdown efficiencies are shown in Fig 6c and Supplementary Fig. 6c. These observations indicate that PIAS1 is the main SUMO E3 ligase co-regulating the SUMOylation of BARD1 and BRCA1 in response to DNA damage. Our results indicate that BARD1 is an indirect target for RNF4, linked by PIAS1.

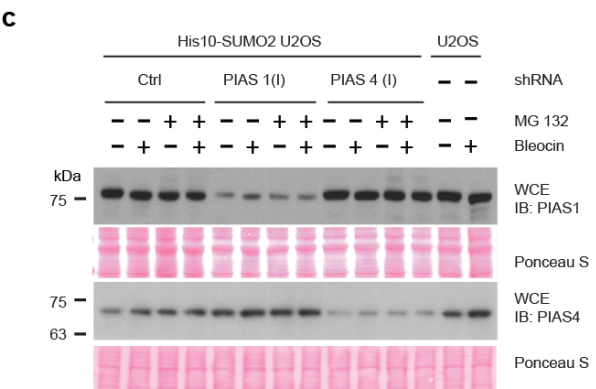
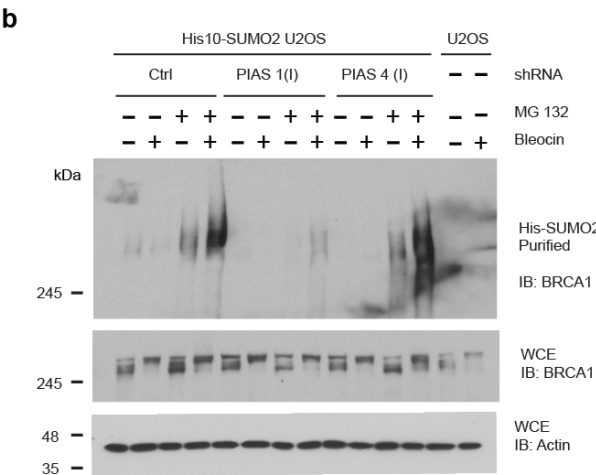
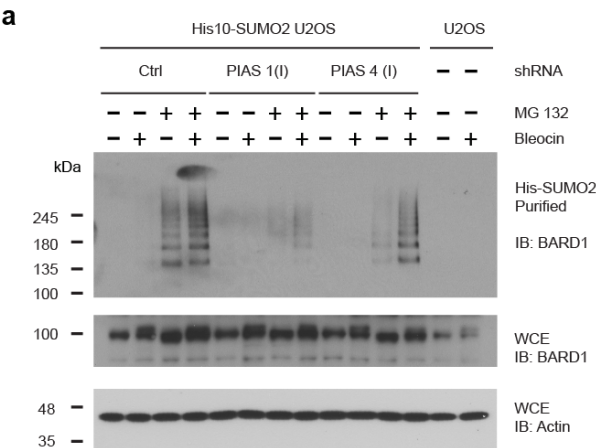


Figure 6. The SUMO E3 ligase PIAS1 is responsible for the SUMOylation of BARD1 and BRCA1. **(a-c)** U2OS cells stably expressing His10-SUMO2 were infected with lentiviruses expressing shRNAs directed against PIAS1 or PIAS4 or a control shRNA. Three days after infection cells were either mock-treated or treated with MG132 (10 μ M) to inhibit the proteasome. One hour after the start of MG132 treatment, cells were either DMSO treated or treated with Bleocin (5 μ g/ml). DNA damaged and undamaged cells were subsequently incubated for 4 hours, lysed in a denaturing buffer and His10-SUMO2 conjugates were purified. **(a)** Levels of SUMOylated BARD1 and total BARD1 were determined by immunoblotting. **(b)** Levels of SUMOylated BRCA1 and total BRCA1 were determined by immunoblotting. **(c)** Knockdown efficiencies of PIAS1 and PIAS4 were determined by immunoblotting. The experiment was independently repeated using other shRNAs directed against PIAS1 and PIAS4. Additionally, the SUMO2 purification efficiency was determined by immunoblotting. These results are provided in Supplementary Fig. 6. Unprocessed full-size scans of blots are provided in Supplementary Fig.9. The experiments were repeated three times.

2.6 BARD1 SUMOylation is dependent on its interaction with BRCA1

In response to genotoxic stress, BARD1 plays a crucial role in DNA repair, both independently and in combination with BRCA1. In our next experiments, we aimed to dissect the role of BRCA1 with regard to the regulation of BARD1. Interestingly, BARD1 total protein levels and SUMOylation were significantly reduced when combined with BRCA1 depletion (Fig. 7a and Supplementary Fig. 7a). Conversely, we studied BRCA1 SUMOylation in BARD1 depleted cells. Similar to BARD1 SUMOylation, BRCA1 SUMOylation was increased after blocking the proteasome in combination with Bleocin treatment (Fig. 7b and Supplementary Fig. 7b). Upon BARD1 depletion, we observed a significant reduction of BRCA1 total protein levels and SUMOylation (Fig. 7b). Our observations strengthen the notion that BRCA1 and BARD1 are mutually dependent on each other for overall protein stability.

Structural studies suggest that leucine 44 of BARD1 is required to mediate the complex formation of BRCA1-BARD1^{29, 30}. To test if RNF4-dependent BARD1 SUMOylation required heterodimer formation with BRCA1, we verified the SUMOylation of the L44R mutant of BARD1. We have used a retroviral expression system to stably express GFP fused to wild type BARD1 or the L44R mutant. We performed RNF4 depletion as well as IR irradiation and purified SUMO2 conjugates. Consistent with our BRCA1 knockdown experiments, the L44R mutant of BARD1 is defective for SUMOylation either in the absence (Fig. 7c and Supplementary Fig. 7d) or in the presence of DNA damage (Fig. 7d and Supplementary Fig. 7e). SUMOylated BARD1 was strongly stabilized by blocking the proteasome (Fig. 7e and Supplementary Fig. 7f). These observations indicate that the BRCA1-BARD1 complex is the substrate for SUMOylation and is subsequently degraded by the proteasome.

BARD1 contains potential consensus sites for SUMOylation. Site directed mutagenesis was performed to generate the point mutants K96R, K127R, K632R and E634A. Interestingly, one of the four BARD1 point mutants, K632R, displayed some reduction in SUMOylation either in the absence or in the presence of IR (Fig. 7c and 7d). However, this site appears not be a classical KxE-type

SUMOylation consensus motif, since no reduction in SUMOylation was observed for the E634A mutant (Fig. 7c and 7d).

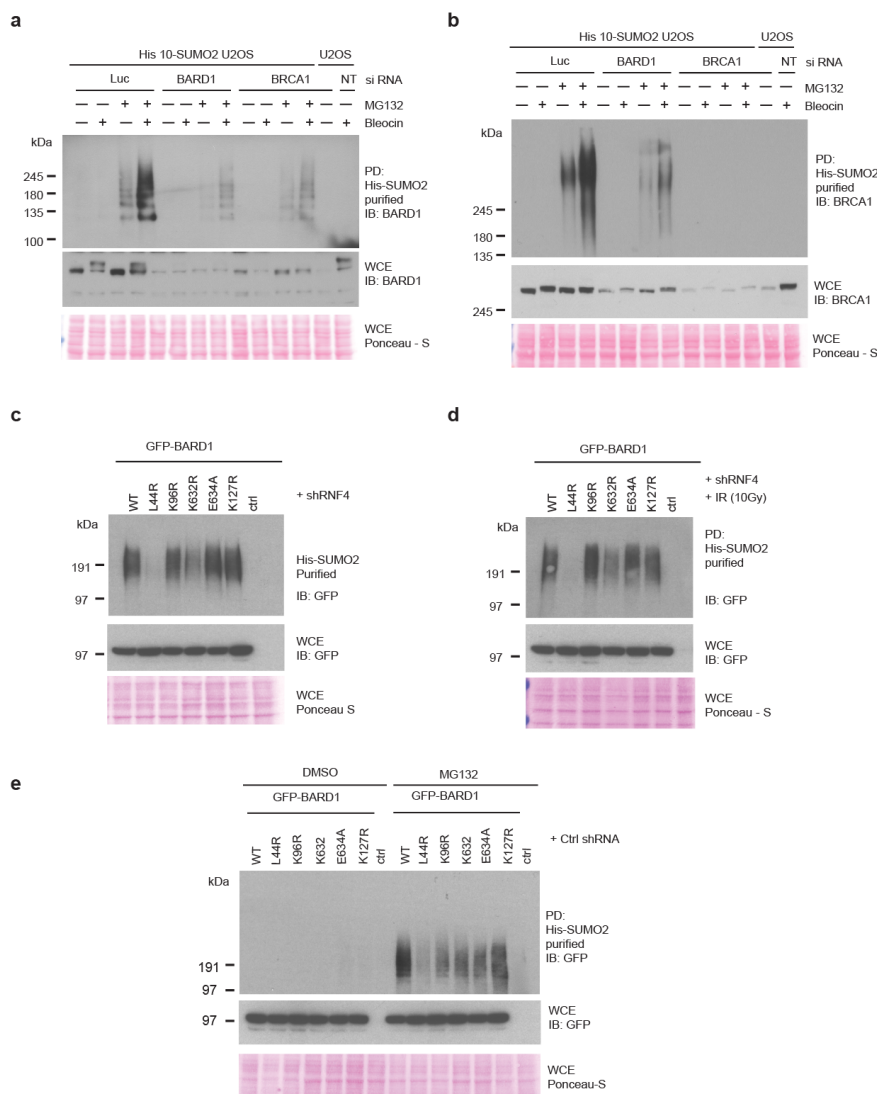


Figure 7. SUMOylation of BARD1 occurs in a BRCA1-dependent manner. **(a and b)** U2OS cells stably expressing His10-SUMO2 were transfected either with siRNAs targeting BRCA1 and BARD1, or a control siRNA as indicated. After three days, cells were treated with MG132 for 5hrs and with Bleocin (5μg/ml) for 4 hours, harvested and His10-SUMO2 conjugates were purified from denaturing lysates. **(a)** Levels of SUMOylated and total BARD1 were determined by immunoblotting. **(b)** Levels of SUMOylated and total BRCA1 were determined by immunoblotting. **(c - e)** U2OS cells stably expressing His10-SUMO2 were infected with retroviruses expressing w.t. BARD1-GFP or the indicated point mutants and selected for puromycin resistance. Four days after antibiotic selection, cells were replated. The next day, cells were either infected with a lentivirus to knockdown RNF4 or with a control lentivirus. Three days after lentiviral infection cells were control treated (c), exposed to

IR (10 Gy) or (d) treated with MG132 (10 μ M) or DMSO for 4 hours (e), harvested, lysed and His10-SUMO2 conjugates were purified from denaturing lysates. Total levels of BARD1-GFP and SUMOylated levels were determined by immunoblotting using antibodies directed against GFP. Experiments were independently repeated. The enrichment of His10-SUMO2 conjugates was verified by immunoblotting as shown in Supplementary Fig.7. Unprocessed full-size scans of blots are provided in Supplementary Fig.9. Experiments presented in Fig. 7a and 7b were performed three times. Experiments described in Fig. 7c, 7d and 7e were performed two times.

2.7 BARD1 and SUMO2/3 colocalizes upon DNA damage

Our findings encouraged us to verify BARD1 co-localization with SUMO2/3 upon DNA damage. To this end, we employed a GFP-BARD1 expression construct. Very little co-localization of BARD1 and SUMO2/3 could be observed in the absence of DNA damage. In line with earlier observations for SUMO1, SUMO2/3 accumulates at nucleoli upon proteasome inhibition³¹, where it colocalizes with GFP-BARD1 (Fig. 8 and Supplementary Fig. 8). A striking IRIF localization of GFP-BARD1 was found upon treatment with the DNA damage inducer Bleocin. The size of these IRIFs was increased after co-treatment of cells with Bleocin and MG132 and a pronounced co-localization between GFP-BARD1 and SUMO2/3 could be observed (Fig. 8 and Supplementary Fig. 8). Combined, this suggests that the SUMOylation of BARD1 occurs at local sites of DNA damage.

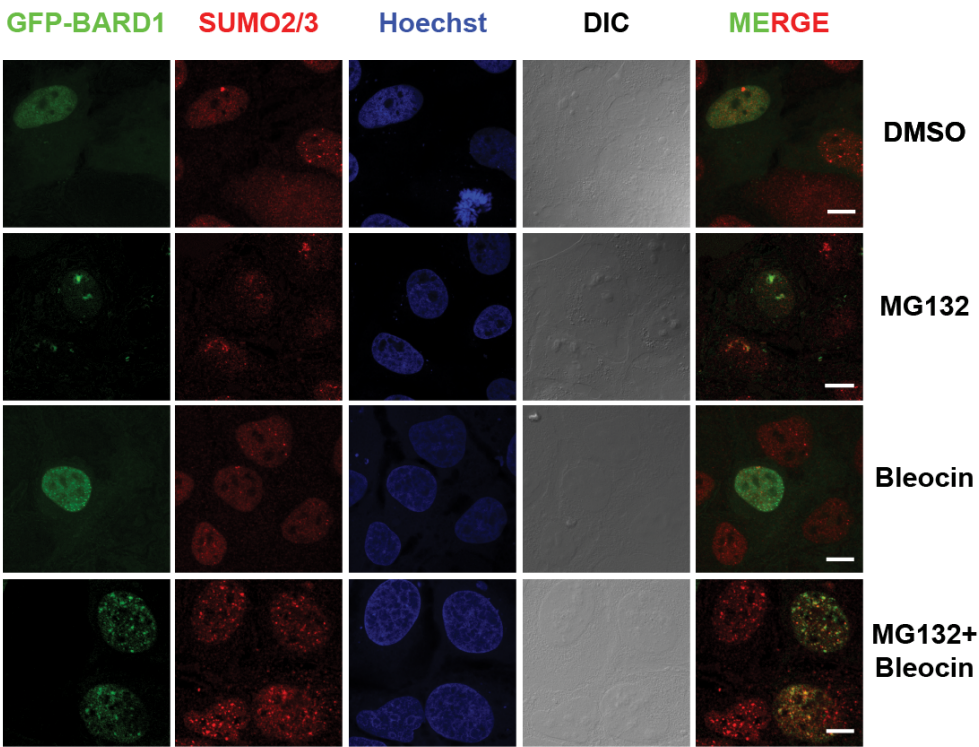


Figure 8. Co-localization of BARD1 and SUMO2/3 in response to DNA damage and proteasome inhibition. U2OS cells were transfected with a GFP-BARD1 (green) expression plasmid. Two days after transfection, cells were treated with DMSO, MG132 (10 μ M) and or Bleocin (5 μ g/ml) for 6 hours as indicated. Cells were fixed and stained with Hoechst (blue) and for endogenous SUMO2/3 (red). Cells were analysed by confocal microscopy. Three biological replicates were performed. Scale bars represent 10 μ M.

2.8 RNF4 regulates BARD1 accumulation at sites of DNA damage

Our results indicate that the BRCA1-BARD1 complex is SUMOylated in response to DNA damage by PIAS1 and subsequently degraded by the proteasome. This would indicate that RNF4 has the ability to balance the accumulation of SUMOylated BRCA1-BARD1 at local sites of DNA damage by targeting PIAS1. To test this hypothesis, we used laser-induced local induction of DNA damage, employing a multi-photon system³². GFP-BARD1 accumulated at laser tracks as expected (Fig. 9a). RNF4 knockdown by RNAi resulted in a significant increase of GFP-BARD1 at these DNA damage tracks, confirming our hypothesis (Fig. 9a-c). Overall, our results indicate that RNF4 primarily targets SUMO E3 ligases and the SUMO E2, to balance SUMO – ubiquitin signaling, affecting downstream proteins including the BRCA1-BARD1 complex (Fig. 10).

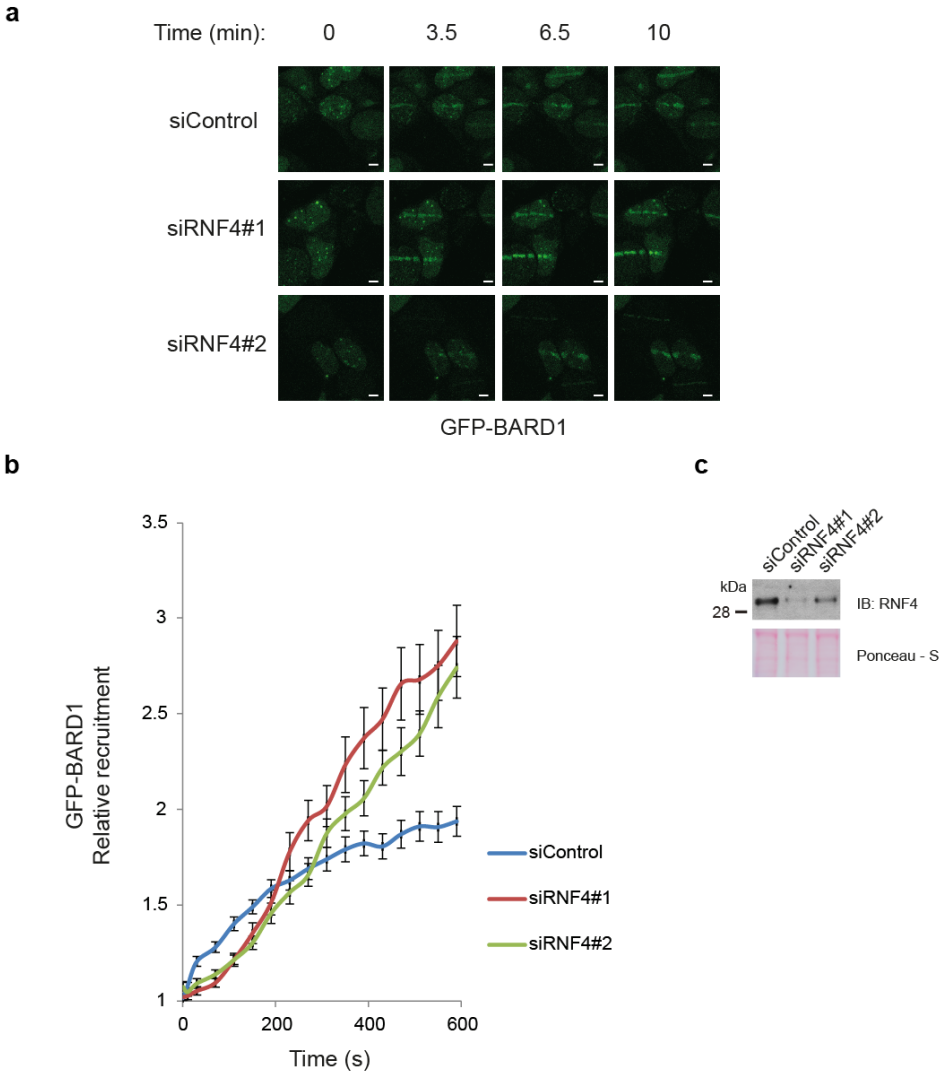


Figure 9. Accumulation of BARD1 at DNA damage tracks is regulated by RNF4. **(a and b)** U2OS cells were co-transfected with a GFP-BARD1 expression plasmid and two siRNAs directed against RNF4, or with a control siRNA. Two days after siRNA transfection, cells expressing low levels of GFP-BARD1 were treated with laser micro-irradiation. Recruitment of GFP-BARD1 to local sites of DNA damage was studied using time lapse microscopy. **(a)** Representative GFP-BARD1 recruitment images from one experiment are shown. Scale bars represent 5 μ M. **(b)** Experiments were performed four times. Relative recruitment of GFP-BARD1 to laser induced DNA damage tracks was quantified. Depicted are average values and SEMs ($n > 40$). Values from 600 sec timepoint were compared using two tailored T-Tests not assuming equal variance (p-values: siControl vs siRNF4#1 = 1.80×10^{-5} ; siControl vs siRNF4#2 = 3.78×10^{-5} ; siRNF4#1 vs siRNF4#2: 0.58) **(c)** RNF4 knockdown was confirmed by immunoblotting.

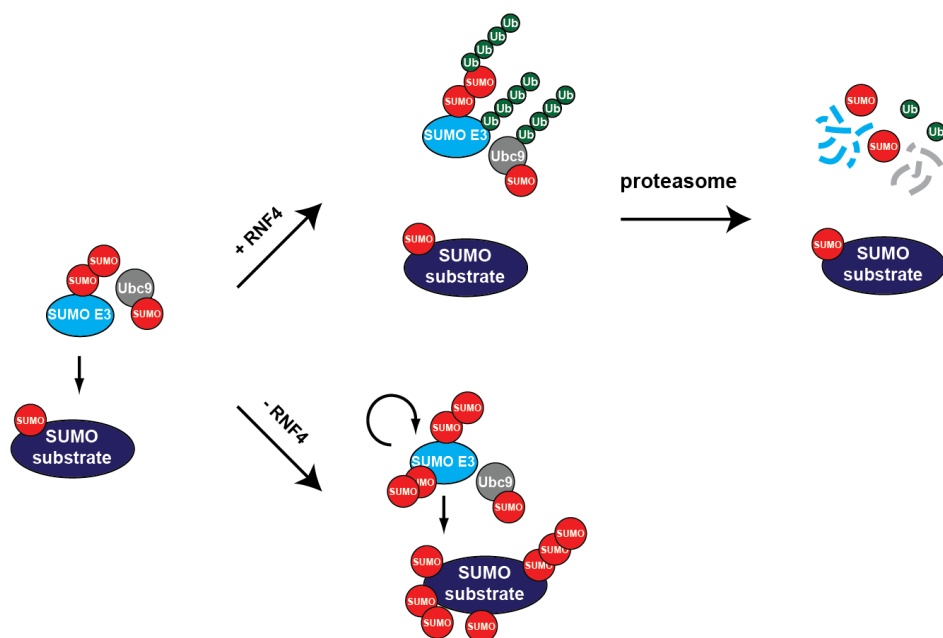


Figure 10. Model explaining the rise and fall of SUMOylation at sites of DNA damage. SUMO E3 ligases are recruited to sites of DNA damage, to modify repair factors including BARD1. Subsequently, the STUbL RNF4 is recruited to ubiquitylate SUMOylated proteins including autoSUMOylated SUMO E3 ligases and autoSUMOylated Ubc9. These proteins are subsequently degraded by the proteasome to resolve the SUMOylation signal at the site of DNA damage, explaining the transient nature of the signal.

3 Discussion

Using two complementary proteomics approaches, we have purified and identified target proteins for the human STUbL RNF4. Interestingly, we found the SUMO E3 ligases PIAS1, PIAS2, PIAS3, ZNF451 and NSMCE2 and the SUMO E2 Ubc9 as targets for RNF4. Subsequent experiments confirmed the regulation of PIAS1 by RNF4 in a SIM- and proteasome-regulated manner. SUMOylated PIAS1 was ubiquitylated by RNF4 and targeted to the proteasome. Knockdown of RNF4 enhanced the accumulation of SUMOylated PIAS1. We are proposing a model where targeting of SUMO E3 ligases and the SUMO E2 by RNF4 is balancing SUMO signal transduction (Fig. 10).

Active SUMO E3 ligases are expected to autoSUMOylate. Given the preference of RNF4 for SUMO chains²⁴, our data indicate that SUMO E3 ligases accumulate SUMO chains by autoSUMOylation, thereby creating binding sites for the STUbL RNF4 (Fig. 10). Evidence for the accumulation of SUMO on the yeast SMT3 E3 ligases Pli1 and Siz1 and subsequent degradation by a yeast STUbL was provided recently^{18, 33}. Alternatively, SUMO E3 ligases could autoSUMOylate at

multiple sites. When these SUMOs are closely spaced, they could also provide efficient binding sites for the closely spaced SIMs in RNF4. We have indeed previously identified such closely spaced SUMOylation sites on SUMO E3 ligases, including PIAS1 lysines 40, 46, 56 and 58, PIAS2 lysines 430, 443, 452, 464 and 489, PIAS3 lysines 46 and 56, ZNF451 lysines 490, 500, 508, 522, 532, 537 and many more, NSMCE2 lysines 30, 41, 47, 65 and 70, but also PIAS4 lysines 59 and 69, 128 and 135 and RanBP2 lysines 1596 and 1605, 2513 and 2531, 2571 and 2592³⁴. Whereas PIAS4 was also found in the TULIP RNF4 samples, its enrichment after MG132 treatment in a SIM-dependent manner was just below the cut off values (Supplementary Dataset 3). The localization of the SUMO E3 ligase RanBP2 at the cytoplasmic side of nuclear pores might explain why it was not identified as RNF4 substrate in our screen, since RNF4 resides predominantly in the nucleus^{24, 26}.

The identification of the SUMO E2 Ubc9 as the most enriched protein from the TULIP screen, in a SIM- and MG132-dependent manner was very striking (Fig. 3e). This finding fits well with the identification of SUMOylated Ubc9 as a key factor in SUMO chain formation³⁵. SUMOylated Ubc9 was severely reduced in its regular activity, but stimulated SUMO polymer formation in cooperation with SUMO thioester charged Ubc9, via noncovalent backside SUMO binding. Targeting SUMOylated Ubc9 is thus an efficient manner to limit SUMO chain formation by RNF4. Ubc9 was also identified in the RNF4 knockdown approach, but statistically it was just below the cut off value.

SUMOylation is frequently a low stoichiometry modification, modifying target proteins only at low levels³⁶. Subsequent ubiquitylation and degradation by RNF4 therefore only affects target proteins at a small percentage, so consequently, no changes in overall protein levels can be observed. This appears to be the case for PIAS1 as shown in Figure 4. Thus, only a small subset of PIAS1 is SUMOylated and ubiquitylated. Nevertheless, this small SUMOylated subfraction of PIAS1 could be functionally very important, since it could represent the functionally active fraction. Targeting the active fractions of SUMO E3 ligases for degradation will have a profound effect on overall SUMOylation levels. Substoichiometric ubiquitylation appears to be a frequent event as noted by Kim et al. in their proteome wide study of ubiquitylation³⁷. They noted that the ubiquitylation of a large set of targets upon proteasomal inhibition did not result in overt changes in total protein levels.

Eight years ago, SUMOylated proteins were shown to accumulate at local sites of DNA damage^{6, 7}. Two SUMO E3 ligases, PIAS1 and PIAS4 were found to be responsible for the accumulation of SUMOylated proteins at these sites^{6, 7}. Two SUMOylated substrates involved were identified, 53BP1 and BRCA1. PIAS1 and PIAS4 activity are required for proper accumulation of ubiquitin adducts, generated by the ubiquitin E3 ligases BRCA1, RNF8 and RNF168. RNF168 and HERC2 were later found to be substrates for SUMOylation as well³⁸. These results highlight the intricate crosstalk between these two major PTMs to build up at local sites of DNA damage.

Subsequently, RNF4 was found to accumulate at sites of DNA damage too¹⁵⁻¹⁷. Potential substrates identified for RNF4 were MDC1 and BRCA1^{15, 19, 20}. Our current project indicates that BRCA1 together with its partner BARD1 is a substrate for SUMOylation by PIAS1, since the L44R mutant of BARD1,

defective for BRCA1 binding, is no longer SUMOylated^{29,30}. Previously it was shown that the BRCA1-BARD1 dimer is activated by SUMOylation, but the authors missed out on BARD1 as SUMO substrate⁷. Our results indicate that the activity of the BRCA1-BARD1 dimer is indirectly regulated by RNF4, since this STUbL targets the BRCA1-BARD1 SUMO E3 ligase PIAS1 for degradation by the proteasome.

BARD1 was originally identified as a binding partner of the key breast cancer susceptibility protein BRCA1, using a yeast two-hybrid approach³⁹. BARD1 is indispensable for embryonic development and mice deficient for BARD1 die between embryonic day 7.5 and 8.5 due to a major reduction in cell proliferation⁴⁰. Similar results were obtained for its partner BRCA1⁴¹. Overall, the protein shares several characteristics with its interaction partner BRCA1, including a RING domain which mediates hetero-dimerization to stabilize the protein since each monomer on its own is unstable^{40,42}. Hetero-dimerization furthermore involves α -helices neighbouring the RING domains in both protein. BARD1 also contains two other domains that function in protein-protein interactions, the BRCT domain and three ankyrin repeats. Together with BRCA1, BARD1 regulates Lys-6 conjugation of ubiquitin⁴³⁻⁴⁷. Substrates ubiquitinated by BRCA1-BARD1 include histones H2A and H2B^{48,49}.

The BRCA1-BARD1 heterodimer functions as a key tumor suppressor. Germline BRCA1 mutations are found in almost half of the breast cancer patients⁵⁰. Many BRCA1 mutations affect its activity as an ubiquitin E3 ligase. However, mutations in BARD1 are less common^{45,51,52}. Defects in the BRCA1-BARD1 dimer result in a strong decrease in genome stability, mechanistically explaining its role in cancer development^{40,53}. BRCA1-BARD1 plays an important role in the repair of double strand DNA breaks via homologous recombination^{54,55}. Interestingly, this fits well with the defects in homologous recombination observed upon knockdown of RNF4, underlining the functional relation between RNF4 and BRCA1-BARD1^{13-15,20}.

The unbiased identification of substrates for ubiquitin E3 ligases in cells is notoriously challenging. We have developed TULIP technology to address this challenge. The ability of TULIP to discriminate between non-covalent binding proteins and covalently bound targets is a key strength. The denaturing buffers used during the purification are furthermore compatible with trypsin digestion of purified samples. Given the gargantuan complexity of the ubiquitin conjugation machinery, the task to delineate substrate – E3 ligase relationships is overwhelming. The TULIP technology we developed in this project could be helpful for this purpose.

4 Experimental procedures

4.1 Plasmid DNA

TULIP plasmids were constructed as follows. pCW57.1, a gift from David Root (Addgene plasmid # 41393) was mutated by site directed mutagenesis using oligos FW-pCW57.1-stop-rem and RV-pCW57.1-stop-rem, to remove the three stop codons in frame with the Gateway cassette before adding

an AgeI restriction site resulting in pCW57.1ns. For –HIS TULIP construction, the AgeI-SpeI fragment from pCW57.1 was amplified by PCR using oligos FW-AgeI-C-term-HIS and RV-SpeI-C-term-HIS, cloned using the Zero-Blunt PCR cloning kit (Thermo-Fisher). This fragment was cut with AgeI-SpeI and cloned into AgeI-SpeI digested pCW57.1ns. For –Ubiquitin and –Ubiquitin-ΔGG TULIP construction, the Ubiquitin cDNA was amplified by PCR using oligos FW-AgeI-10HIS-Ubi and either RV-XmaI-Ubi or RV-XmaI-Ubi-noGG, and cloned using the Zero-Blunt PCR cloning kit (Thermo-Fisher). Inserts were then cut with AgeI-XmaI and cloned into the AgeI site of pCW57.1ns previously de-phosphorylated using Antarctic Phosphatase (New England Biolabs). All plasmids were amplified in the Gateway-compatible *E. coli* strain DB3.1.

RNF4 and RNF4ΔSIM ORFs lacking stop codons were cloned into pDONR207 and transferred to the TULIP plasmids using Gateway technology (Thermo-Fisher).

To generate BARD1 mutants, site directed mutagenesis was performed on the pDONR-BARD1 wild type plasmid with oligos BARD1-L44R_FW and BARD1-L44R_RV to generate pDONR-BARD1-L44R, BARD1-K96R_FW and BARD1-K96R_RV to generate pDONR-BARD1-K96R, BARD1-K632R_FW and BARD1-K632R_RV to generate pDONR-BARD1-K632R, BARD1-E634A_FW and BARD1-E634A_RV to generate pDONR-BARD1-E634A, and, BARD1-K127R_FW and BARD1-K127R_RV to generate pDONR-BARD1-K127R mutant plasmid DNA. The desired mutations were confirmed by DNA sequencing. The Gateway system was used to clone wild-type and mutant plasmid DNA into the pBABE N-terminal GFP retroviral destination vector. All oligo sequences are specified in Supplementary Table 4.

4.2 Retroviral and lentiviral transduction

For retroviral transduction, 1.2 million cells were seeded in a 15-cm dish and the next day these cells were infected with retroviruses at MOI 2. After changing the media, the next day, cells were selected with puromycin for 4 days. Lentiviral transduction was performed essentially as described previously¹⁵. One million cells were seeded in a 15-cm dish and the next day, cells were either infected with shRNA viruses directed against RNF4, PIAS1, PIAS4, BRCA1 and BARD1 or control non-targeting shRNA SHC002 viruses at MOI 2 (Sigma-Aldrich). After changing media on the third day, cells were incubated for another 3 to 4 days as indicated. shRNA constructs are specified in Supplementary Table 3.

4.3 TULIP assays

U2OS cells stably expressing the different TULIP constructs were grown in five 15 cm plates up to 50% confluency. TULIP construct expression was induced adding doxycycline 1μg/mL for 24h. Proteasome inhibitor MG132 10μM or DMSO was added to the cells for 5h and cells were harvested and lysed. HIS conjugates were purified from the denatured lysates.

4.4 Cell culture and cell cycle analysis

U2OS cells (ATCC) and U2OS cells stably expressing His10-SUMO2 were grown in DMEM high glucose medium supplemented with 10% FBS and 100 U/ml penicillin plus 100 µg/ml streptomycin (Thermo-Fisher) at 37°C at 5% CO₂²³. Cells were regularly tested for mycoplasma contamination and found to be negative. To arrest cells at the G1/S boundary, cells were treated with 2 mM thymidine for 19 hrs and then released for 9 hours, followed by a second thymidine (2 mM) block for 17 hrs. To release G1-arrested cells, they were washed two times with PBS and one time with prewarmed cell culture medium. Cells were collected after 4 hours and 8 hours to obtain cell populations enriched for S-phase or G2/M-phase. After washing with PBS, cells were fixed in 70% ethanol and incubated for 30 minutes. Subsequently cells were incubated with Ribonuclease A and stained with propidium iodide (PI) for 15 minutes and analysed by flow cytometry⁵⁶. Drugs used for different treatments are specified in Supplementary Table 2.

4.5 Microscopy

Cells for immunofluorescence microscopy were cultured on glass slides in 24-well plates. After treatment with MG132 (10 µM) and / or Bleocin (5µg/ml) for 6 hours, medium was removed, cells were fixed with 4% paraformaldehyde for 20 minutes at room temperature in PBS, and cells were permeabilized with 0.1% Triton X-100 in PBS for 15 minutes. Next, cells were washed twice with PBS and once with PBS plus 0.05% Tween-20 (PBS-T). Cells were then blocked for 10 minutes with 0.5% blocking reagent (Roche) in 0.1 M Tris, pH 7.5 and 0.15 M NaCl (TNB), and treated with primary antibody as indicated in TNB for one hour. Coverslips were washed five times with PBS-T and incubated with the secondary antibodies as indicated in TNB for one hour. Next, coverslips were washed five times with PBS-T and dehydrated by washing once with 70% ethanol, once with 90% ethanol, and once with 100% ethanol. After drying the cells, coverslips were mounted onto a microscopy slide using citifluor/Hoechst solution (500 ng/mL) and sealed with nail varnish.

4.6 Recruitment of GFP-BARD1 to laser induced DNA damage sites

Approximately 20.000 U2OS cells were seeded in 6 well dishes containing an 18 mm coverslip. The following day 0.5 µg/well of GFP-BARD1 plasmid was transfected using 12µL of PEI (1mg/mL). Transfection was allowed to occur overnight and then cells were washed twice with PBS and siRNA transfections were performed using DharmaFect 1 Transfection Reagent (GE Lifesciences), according to the manufacturer's instructions. Cells were investigated 48 hours after siRNA transfection. siRNA depletion of RNF4 was performed using on-target plus RNF4 siRNAs J-006557-08 and J-006557-07 (GE Lifesciences), and the non-targeted control was performed using siGENOME Non-Targeting siRNA #1 (GE Lifesciences).

Laser track experiments were performed as previously described³². Two days after siRNA transfection, U2OS cells were grown on 18 mm coverslips and transiently transfected with a GFP-

BARD1 construct. Prior to laser micro-irradiation, medium was replaced with CO₂-independent Leibovitz's L15 medium supplemented with 10% FCS and pen/strep. Laser micro-irradiation was carried out on a Leica SP5 confocal microscope equipped with an environmental chamber set to 37°C. DNA damage tracks (1 µm width) were generated with a Mira modelocked titanium-sapphire (Ti:Sapphire) laser (λ = 800 nm, pulse length = 200 fs, repetition rate = 76 MHz, output power = 80 mW) using a UV-transmitting 63× 1.4 NA oil immersion objective (HCX PL APO; Leica). Confocal images were recorded before and after laser irradiation at 20 sec time intervals over a period of 10 min. Images were analysed using Leica LAS X software.

4.7 Purification of His10 conjugates

His10 conjugates were purified essentially as described previously^{15, 57}. U2OS cells expressing His10-SUMO2 were washed, scraped and collected in ice-cold PBS. For total lysates, a small aliquot of cells was kept separately and lysed in 2% SDS, 1% N-P40, 50 mM TRIS pH 7.5, 150 mM NaCl. The remaining part of the cell pellets were lysed in 6 M guanidine-HCl pH 8.0 (6 M guanidine-HCl, 0.1 M Na₂HPO₄/NaH₂PO₄, 10 mM TRIS, pH 8.0). The samples were snap frozen using liquid nitrogen, and stored at -80°C.

For SUMO purification, the cell lysates were first thawed at room temperature and sonicated for 5 sec, using a sonicator (Misonix Sonicator 3000, EW-04711-81) at 30 Watts to homogenize the lysate. Protein concentrations were determined using the bicinchoninic acid (BCA) Protein Assay Reagent (Thermo Scientific) and lysates were equalized. Subsequently, imidazole was added to a final concentration of 50 mM and β-mercaptoethanol was added to a final concentration of 5 mM. His10-SUMO conjugates were enriched on nickel-nitrilotriacetic acid-agarose beads (Ni-NTA) (Qiagen), and the beads were subsequently washed using wash buffers A-D. Wash buffer A: 6 M guanidine-HCl, 0.1 M Na₂HPO₄/NaH₂PO₄ pH 8.0, 0.01 M Tris-HCl pH 8.0, 10 mM imidazole pH 8.0, 5 mM β-mercaptoethanol, 0.1% Triton X-100 (0.2% Triton X-100 for immunoblotting sample preparation). Wash buffer B: 8 M urea, 0.1 M Na₂HPO₄/NaH₂PO₄ pH 8.0, 0.01 M Tris-HCl pH 8.0, 10 mM imidazole pH 8.0, 5 mM β-mercaptoethanol, 0.1% Triton X-100 (0.2% Triton X-100 for immunoblotting sample preparation). Wash buffer C: 8 M urea, 0.1 M Na₂HPO₄/NaH₂PO₄ pH 6.3, 0.01 M Tris-HCl pH 6.3, 10 mM imidazole pH 7.0, 5 mM β-mercaptoethanol, no Triton X-100 (0.2% Triton X-100 for immunoblotting sample preparation). Wash buffer D: 8 M urea, 0.1 M Na₂HPO₄/NaH₂PO₄ pH 6.3, 0.01 M Tris-HCl, pH 6.3, no imidazole, 5 mM β-mercaptoethanol, no Triton X-100 (0.2% Triton X-100 for immunoblotting sample preparation). Samples were eluted in 7 M urea, 0.1 M NaH₂PO₄/Na₂HPO₄, 0.01 M Tris/HCl, pH 7.0, 500 mM imidazole pH 7.0.

4.8 Electrophoresis and immunoblotting

Whole cell extracts or purified protein samples were separated on Novex 4-12% gradient gels (Thermo-Fisher) using MOPS buffer or on Novex 3-8% gradient gels (Thermo-Fisher) using Tris-Acetate buffer

or via regular SDS-PAGE using a Tris-glycine buffer and transferred onto Amersham Protran Premium 0.45 NC Nitrocellulose blotting membrane (GE Healthcare; 10600003) using a submarine system (Thermo-Fisher). The use of Novex 3-8% gradient gels enabled the visualization of phosphorylation shifts. Membranes were stained with Ponceau S (Sigma) to visualize total protein amounts, and blocked with PBS containing 8% milk powder and 0.05% Tween-20 before incubating with the primary antibodies as indicated in Supplementary Table T1.

4.9 Proteomics sample preparation and mass spectrometry

SUMO2 enriched samples were supplemented with 1 M Tris-(2-carboxyethyl)-phosphine hydrochloride (TCEP) to a final concentration of 5 mM, and incubated for 20 minutes at room temperature. Iodoacetamide (IAA) was then added to the samples to a 10 mM final concentration, and samples were incubated in the dark for 15 minutes at room temperature. Lys-C and Trypsin digestions were performed according to the manufacturer's specifications. Lys-C was added in a 1:50 enzyme-to-protein ratio, samples were incubated at 37 °C for 4 hours, and subsequently 3 volumes of 100 mM Tris-HCl pH 8.5 were added to dilute urea to 2 M. Trypsin (V5111, Promega) was added in a 1:50 enzyme-to-protein ratio and samples were incubated overnight at 37 °C.

RNF4-TULIP samples were concentrated using VIVACON 30kDa exclusion filters (Sartorius) to a volume of 50 μ L and Ammonium Bicarbonate (ABC) was added to a final concentration of 50 mM. Samples were reduced with 1mM Dithiothreitol (DTT) for 30 minutes at room temperature, alkylated with 5mM Chloroacetamide (CAA) for 30 minutes at room temperature, reduced once more with 5mM DTT for 30 minutes at room temperature. Next, 200 μ L of 50 mM ABC were added to the samples and 250 ng of Trypsin (V5111, Promega). Samples were incubated overnight at room temperature.

Subsequently, digested samples were desalted and concentrated on STAGE-tips as described previously⁵⁸ and eluted with 80% acetonitrile in 0.1% formic acid. Eluted fractions were vacuum dried, employing a SpeedVac RC10.10 (Jouan, France) and dissolved in 10 μ L 0.1% formic acid before online nanoflow liquid chromatography-tandem mass spectrometry (nanoLC-MS/MS).

All the experiments were performed on an EASY-nLC 1000 system (Proxeon, Odense, Denmark) connected to a Q-Exactive Orbitrap (Thermo Fisher Scientific, Germany) through a nano-electrospray ion source. The Q-Exactive was coupled to a 13 cm analytical column with an inner-diameter of 75 μ m, in-house packed with 1.8 μ m C18 beads (Reprosphere-DE, Pur, Dr. Manish, Ammerbuch-Entringen, Germany) in the case of RNF4 knockdown samples and 1.9 μ m C18-AQ beads in the case of RNF4-TULIP samples.

The gradient length was 120 minutes from 2% to 95% acetonitrile in 0.1% formic acid at a flow rate of 200 nL/minute. The mass spectrometer was operated in data-dependent acquisition mode with a top 10 method. Full-scan MS spectra were acquired at a target value of 3×10^6 and a resolution of 70,000, and the Higher-Collisional Dissociation (HCD) tandem mass spectra (MS/MS) were recorded at a target value of 1×10^5 and with a resolution of 17,500 with a normalized collision energy (NCE)

of 25%. The maximum MS1 and MS2 injection times were 20 ms and 60 ms, respectively. The precursor ion masses of scanned ions were dynamically excluded (DE) from MS/MS analysis for 60 sec. Ions with charge 1, and greater than 6 were excluded from triggering MS2 analysis.

4.10 Data analysis

For the RNF4 knockdown analysis, five experimental conditions were performed in biological triplicate, and all samples were measured in technical triplicate, resulting in a total of 45 runs. For the RNF4-TULIP, nine experimental conditions were measured in biological triplicate, resulting in a total of 27 samples. The raw mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD005425. All RAW data were analysed using MaxQuant (version 1.5.5.1) according to Tyanova et al.⁵⁹. We performed the search against an in silico digested UniProt reference proteome for Homo sapiens (11 Sep 2016).

Database searches were performed with Trypsin/P, allowing four missed cleavages. Oxidation (M) and Acetyl (Protein N-term) were allowed as variable modifications with a maximum number of 5. Match between runs was performed with 0.7 min match time window and 20 min alignment time window. The maximum peptide mass was set to 5000. Label Free Quantification was performed using the MaxLFQ approach, not allowing Fast LFQ⁶⁰. Instrument type was set to Orbitrap.

Protein lists generated by MaxQuant were further analyzed by Perseus (version 1.5.3.3). Proteins identified as common contaminants were filtered out, and then all the LFQ intensities were log2 transformed. Scatter plots were generated for each experimental condition to compare the differences between biological replicates and to derive Pearson correlations. Different biological repeats of the experiment were grouped and only protein groups identified in all three biological replicates in at least one group were included for further analysis. Missing values were imputed using Perseus software by normally distributed values with a 1.8 downshift (log2) and a randomized 0.3 width (log2) considering whole matrix values.

Subsequently, the RNF4 knockdown and RNF4-TULIP results were analysed independently. For the RNF4 knockdown results. Samples were annotated in three different groups: U2OS parental cell line, 10HIS-SUMO2-U2OS and RNF4 knockdown treated cells. Proteins were considered to be SUMO2 target proteins when the median log2 ratio of the LFQ intensity in the experimental group of 10HIS-SUMO2 expressing cells minus the median log2 ratio of the LFQ intensity in the U2OS parental control group was greater than 0 and the P value of ANOVA was smaller than 0.05. Proteins were considered to be enriched after RNF4 knockdown when the average difference (log2) between the 10-HIS-SUMO2 RNF4-knockdown samples and the 10-HIS-SUMO2 U2OS samples was bigger than 1 and the p-value < 0.05 having been identified as a SUMO target protein using ANOVA. For the RNF4-TULIP analysis, different experimental sets were compared with each other. Differences were considered to be significant when the average difference (log2) was larger than 1 with a p-value < 0.05.

Term enrichment analysis (Gene Ontology) was performed using the Gene Ontology Consortium PANTHER Overrepresentation test (release 20160715) using the GO Ontology database (released 2016-10-27)

Volcano plots to demonstrate significant changes in protein enrichments were created by plotting the Student's t-test $-\log_{10}(\text{p-value})$ against the average \log_2 difference value in different comparisons.

Significantly enriched SUMOylated proteins after RNF4 knock down were selected to perform functional protein interaction analysis by STRING (string-db.org, version 10.0) using a high confidence score ($p > 0.7$). STRING analysis results were visualized using Cytoscape (version 3.4.0).

4.11 siRNA transfection

The siRNA duplexes have been previously described²⁹ and were purchased from Dharmacon. BRCA1: 5'- AGG AAA UGC AGA AGA GGA AdTdT -3' and BARD1 5'- GAG UAA AGC UUC AGU GCA AdTdT -3'. 2 million cells were seeded in a 15 cm dish and reverse transfection was performed according to manufacturer's instructions. 18 hours after the transfection, fresh growth medium was added to the plates. 48h after medium refreshment, the indicated drug treatments were performed and cells were harvested.

4.12 Data availability

The datasets generated and analysed during the current study have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD005425.

Acknowledgments

This project was supported by the European Research Council (A.C.O.V.), the Netherlands Organization for Scientific Research (NWO) (A.C.O.V.) and by a postdoctoral fellowship to R.K. from the FIRC Institute of Molecular Oncology Foundation (IFOM). We would like to thank J. Wiegant for microscopy assistance and H.T.M. Timmers and P. de Graaf (University Medical Centre Utrecht, the Netherlands) for the pBABE N-terminal GFP retroviral destination vector.

Author Contributions

A.C.O.V. conceived the project. A.C.O.V., R.K., Z.X. and R.G.P. designed experiments. R.K. performed biochemical experiments on BARD1, BRCA1, PIAS1 and PIAS4. Z.X. performed and analyzed the RNF4 knockdown proteomics experiments. Z.X. verified the identified RNF4 substrates by immunoblotting and performed microscopy experiments. R.G.P. performed and analyzed the TULIP experiments, live-cell microscopy experiments and analyzed all mass spectrometry data. M.V. assisted the project. R.K., Z.X., R.G.P. and A.C.O.V. wrote the manuscript.

Competing financial interests: The authors declare no competing financial interests.

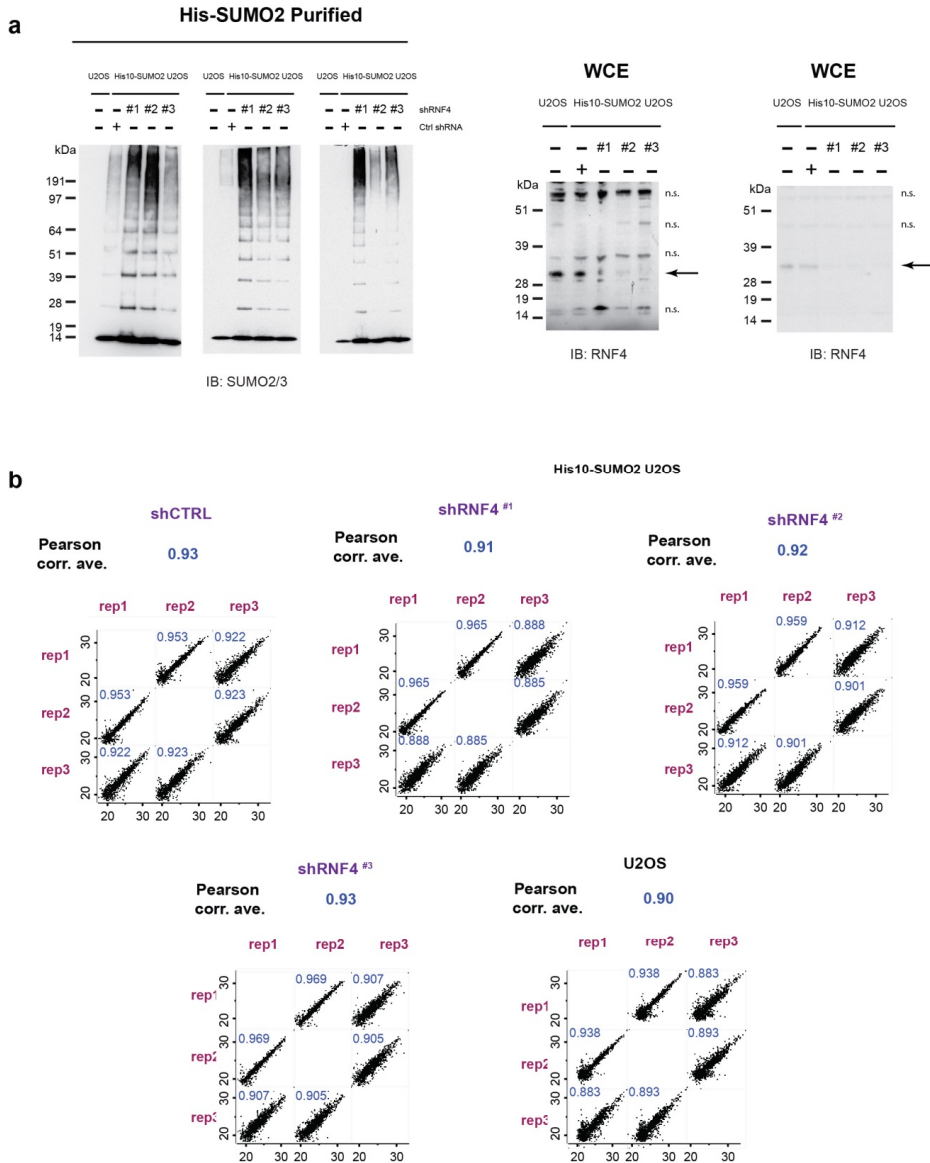
References

1. Hunter,T. & Sun,H. Crosstalk between the SUMO and ubiquitin pathways. *Ernst. Schering. Found. Symp. Proc.* 1-16 (2008).
2. Deribe,Y.L., Pawson,T., & Dikic,I. Post-translational modifications in signal integration. *Nat. Struct. Mol. Biol.* **17**, 666-672 (2010).
3. Wang,L. *et al.* SUMO2 is essential while SUMO3 is dispensable for mouse embryonic development. *EMBO Rep.* **15**, 878-885 (2014).
4. Nacerddine,K. *et al.* The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. *Dev. Cell* **9**, 769-779 (2005).
5. Johnson,E.S., Schvienhorst,I., Dohmen,R.J., & Blobel,G. The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer. *EMBO J.* **16**, 5509-5519 (1997).
6. Galanty,Y. *et al.* Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks. *Nature* **462**, 935-939 (2009).
7. Morris,J.R. *et al.* The SUMO modification pathway is involved in the BRCA1 response to genotoxic stress. *Nature* **462**, 886-890 (2009).
8. Hunter,T. The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Mol. Cell* **28**, 730-738 (2007).
9. Sriramachandran,A.M. & Dohmen,R.J. SUMO-targeted ubiquitin ligases. *Biochim. Biophys. Acta* **1843**, 75-85 (2014).
10. Perry,J.J., Tainer,J.A., & Boddy,M.N. A SIM-ultaneous role for SUMO and ubiquitin. *Trends Biochem. Sci.* **33**, 201-208 (2008).
11. Prudden,J. *et al.* SUMO-targeted ubiquitin ligases in genome stability. *EMBO J.* **26**, 4089-4101 (2007).
12. Sun,H., Leverson,J.D., & Hunter,T. Conserved function of RNF4 family proteins in eukaryotes: targeting a ubiquitin ligase to SUMOylated proteins. *EMBO J.* **26**, 4102-4112 (2007).
13. Galanty,Y., Belotserkovskaya,R., Coates,J., & Jackson,S.P. RNF4, a SUMO-targeted ubiquitin E3 ligase, promotes DNA double-strand break repair. *Genes Dev.* **26**, 1179-1195 (2012).
14. Yin,Y. *et al.* SUMO-targeted ubiquitin E3 ligase RNF4 is required for the response of human cells to DNA damage. *Genes Dev.* **26**, 1196-1208 (2012).
15. Vyas,R. *et al.* RNF4 is required for DNA double-strand break repair in vivo. *Cell Death Differ.* **20**, 490-502 (2013).
16. Poulsen,S.L. *et al.* RNF111/Arkadia is a SUMO-targeted ubiquitin ligase that facilitates the DNA damage response. *J. Cell Biol.* **201**, 797-807 (2013).

17. Grocock,L.M. *et al.* RNF4 interacts with both SUMO and nucleosomes to promote the DNA damage response. *EMBO Rep.* **15**, 601-608 (2014).
18. Westerbeck,J.W. *et al.* A SUMO-targeted ubiquitin ligase is involved in the degradation of the nuclear pool of the SUMO E3 ligase Siz1. *Mol. Biol. Cell* **25**, 1-16 (2014).
19. Guzzo,C.M. *et al.* RNF4-dependent hybrid SUMO-ubiquitin chains are signals for RAP80 and thereby mediate the recruitment of BRCA1 to sites of DNA damage. *Sci. Signal.* **5**, ra88 (2012).
20. Luo,K., Zhang,H., Wang,L., Yuan,J., & Lou,Z. Sumoylation of MDC1 is important for proper DNA damage response. *EMBO J.* **31**, 3008-3019 (2012).
21. Gibbs-Seymour,I. *et al.* Ubiquitin-SUMO Circuitry Controls Activated Fanconi Anemia ID Complex Dosage in Response to DNA Damage. *Mol. Cell* **57**, 150-164 (2015).
22. Hendriks,I.A., Treffers,L.W., Verlaan-de Vries,M., Olsen,J.V., & Vertegaal,A.C. SUMO-2 Orchestrates Chromatin Modifiers in Response to DNA Damage. *Cell Rep.* **10**, 1778-1791 (2015).
23. Xiao,Z. *et al.* System-wide analysis of SUMOylation dynamics in response to replication stress reveals novel SUMO target proteins and acceptor lysines relevant for genome stability. *Mol. Cell Proteomics.* **14**, 1419-1434 (2015).
24. Tatham,M.H. *et al.* RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nat. Cell Biol.* **10**, 538-546 (2008).
25. O'Connor,H.F. *et al.* Ubiquitin-Activated Interaction Traps (UBAITs) identify E3 ligase binding partners. *EMBO Rep.* **16**, 1699-1712 (2015).
26. Lallemand-Breitenbach,V. *et al.* Arsenic degrades PML or PML-RARalpha through a SUMO-triggered RNF4/ubiquitin-mediated pathway. *Nat. Cell Biol.* **10**, 547-555 (2008).
27. Ragland,R.L. *et al.* RNF4 and PLK1 are required for replication fork collapse in ATR-deficient cells. *Genes Dev.* **27**, 2259-2273 (2013).
28. Shima,H. *et al.* Activation of the SUMO modification system is required for the accumulation of RAD51 at sites of DNA damage. *J. Cell Sci.* **126**, 5284-5292 (2013).
29. Li,M. & Yu,X. Function of BRCA1 in the DNA damage response is mediated by ADP-ribosylation. *Cancer Cell* **23**, 693-704 (2013).
30. Morris,J.R., Keep,N.H., & Solomon,E. Identification of residues required for the interaction of BARD1 with BRCA1. *J. Biol. Chem.* **277**, 9382-9386 (2002).
31. Mattsson,K., Pokrovskaja,K., Kiss,C., Klein,G., & Szekely,L. Proteins associated with the promyelocytic leukemia gene product (PML)-containing nuclear body move to the nucleolus upon inhibition of proteasome-dependent protein degradation. *Proc. Natl. Acad. Sci. U. S. A* **98**, 1012-1017 (2001).
32. Gonzalez-Prieto,R., Cuijpers,S.A., Luijsterburg,M.S., van,A.H., & Vertegaal,A.C. SUMOylation and PARylation cooperate to recruit and stabilize SLX4 at DNA damage sites. *EMBO Rep.* **16**, 512-519 (2015).

33. Nie,M. & Boddy,M.N. Pli1(PIAS1) SUMO ligase protected by the nuclear pore-associated SUMO protease Ulp1SENPI/2. *J. Biol. Chem.* **290**, 22678-22685 (2015).
34. Hendriks,I.A. *et al.* Uncovering global SUMOylation signaling networks in a site-specific manner. *Nat. Struct. Mol. Biol.* **21**, 927-936 (2014).
35. Klug,H. *et al.* Ubc9 sumoylation controls SUMO chain formation and meiotic synapsis in *Saccharomyces cerevisiae*. *Mol. Cell* **50**, 625-636 (2013).
36. Hay,R.T. SUMO: a history of modification. *Mol. Cell* **18**, 1-12 (2005).
37. Kim,W. *et al.* Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol. Cell* **44**, 325-340 (2011).
38. Danielsen,J.R. *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).
39. Wu,L.C. *et al.* Identification of a RING protein that can interact in vivo with the BRCA1 gene product. *Nat. Genet.* **14**, 430-440 (1996).
40. McCarthy,E.E., Celebi,J.T., Baer,R., & Ludwig,T. Loss of Bard1, the heterodimeric partner of the Brcal tumor suppressor, results in early embryonic lethality and chromosomal instability. *Mol. Cell Biol.* **23**, 5056-5063 (2003).
41. Hakem,R. *et al.* The tumor suppressor gene Brcal is required for embryonic cellular proliferation in the mouse. *Cell* **85**, 1009-1023 (1996).
42. Yun,M.H. & Hiom,K. Understanding the functions of BRCA1 in the DNA-damage response. *Biochem. Soc. Trans.* **37**, 597-604 (2009).
43. Baer,R. & Ludwig,T. The BRCA1/BARD1 heterodimer, a tumor suppressor complex with ubiquitin E3 ligase activity. *Curr. Opin. Genet. Dev.* **12**, 86-91 (2002).
44. Xia,Y., Pao,G.M., Chen,H.W., Verma,I.M., & Hunter,T. Enhancement of BRCA1 E3 ubiquitin ligase activity through direct interaction with the BARD1 protein. *J. Biol. Chem.* **278**, 5255-5263 (2003).
45. 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 (2001).
46. 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.* **13**, 807-817 (2004).
47. 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.* **278**, 34743-34746 (2003).
48. Thakar,A., Parvin,J., & Zlatanova,J. BRCA1/BARD1 E3 ubiquitin ligase can modify histones H2A and H2B in the nucleosome particle. *J. Biomol. Struct. Dyn.* **27**, 399-406 (2010).
49. Chen,A., Kleiman,F.E., Manley,J.L., Ouchi,T., & Pan,Z.Q. Autoubiquitination of the BRCA1*BARD1 RING ubiquitin ligase. *J. Biol. Chem.* **277**, 22085-22092 (2002).

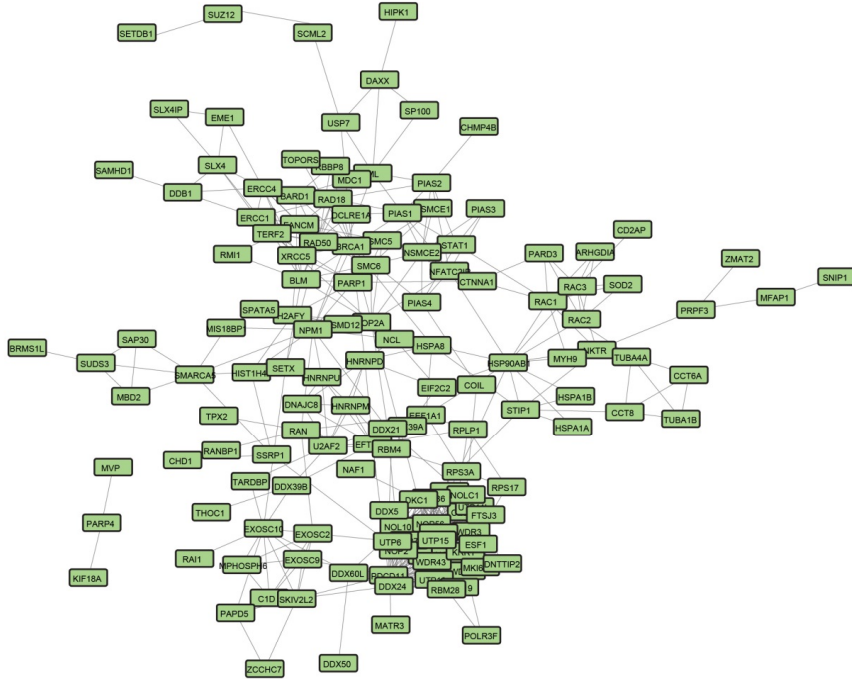
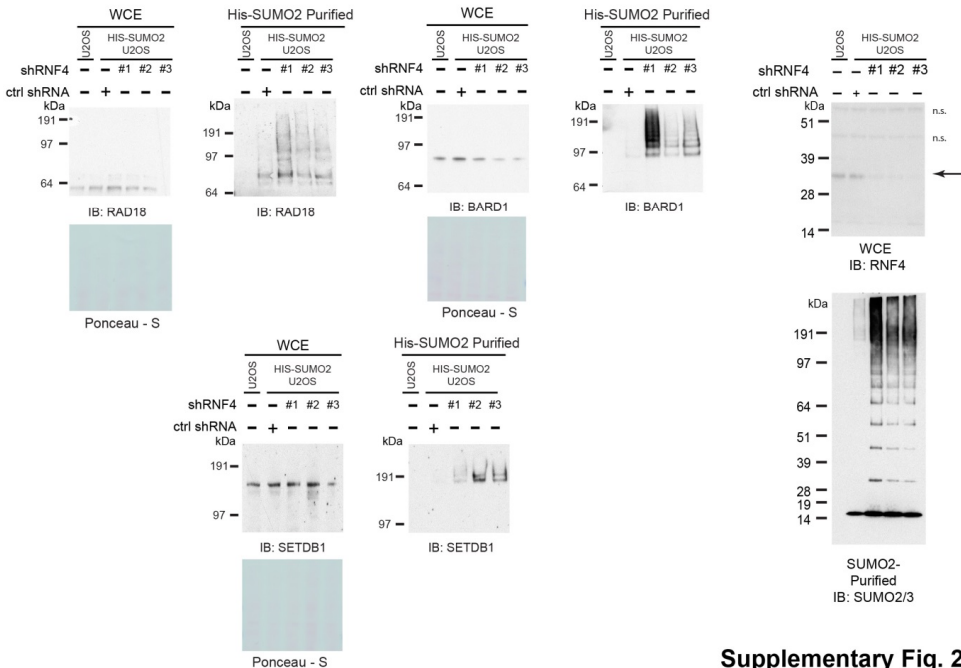
50. Miki,Y. *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* **266**, 66-71 (1994).
51. Brzovic,P.S., Meza,J.E., King,M.C., & Klevit,R.E. BRCA1 RING domain cancer-predisposing mutations. Structural consequences and effects on protein-protein interactions. *J. Biol. Chem.* **276**, 41399-41406 (2001).
52. Ruffner,H., Joazeiro,C.A., Hemmati,D., Hunter,T., & Verma,I.M. Cancer-predisposing mutations within the RING domain of BRCA1: loss of ubiquitin protein ligase activity and protection from radiation hypersensitivity. *Proc. Natl. Acad. Sci. U. S. A* **98**, 5134-5139 (2001).
53. Deng,C.X. & Scott,F. Role of the tumor suppressor gene Brcal in genetic stability and mammary gland tumor formation. *Oncogene* **19**, 1059-1064 (2000).
54. Ohta,T., Sato,K., & Wu,W. The BRCA1 ubiquitin ligase and homologous recombination repair. *FEBS Lett.* **585**, 2836-2844 (2011).
55. Westermarck,U.K. *et al.* BARD1 participates with BRCA1 in homology-directed repair of chromosome breaks. *Mol. Cell Biol.* **23**, 7926-7936 (2003).
56. Schimmel,J. *et al.* Uncovering SUMOylation Dynamics during Cell-Cycle Progression Reveals FoxM1 as a Key Mitotic SUMO Target Protein. *Mol. Cell* **53**, 1053-1066 (2014).
57. Hendriks,I.A. & Vertegaal,A.C. Label-Free Identification and Quantification of SUMO Target Proteins. *Methods Mol. Biol.* **1475**, 171-193 (2016).
58. 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 (2007).
59. Tyanova,S., Temu,T., & Cox,J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301-2319 (2016).
60. 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 (2014).



Supplementary Fig. 1

Supplementary figure 1. SUMO2 substrates regulated by the SUMO-Targeted Ubiquitin Ligase (STUbL) RNF4. **(a)** Additional immunoblots are shown for the biological replicates of the experiment presented in Fig. 1c, using antibodies directed against SUMO2/3 and RNF4. n.s. indicates non-specific bands. **(b)** Pearson correlations between the different replicates analysed by mass spectrometry in Fig. 1d.

a

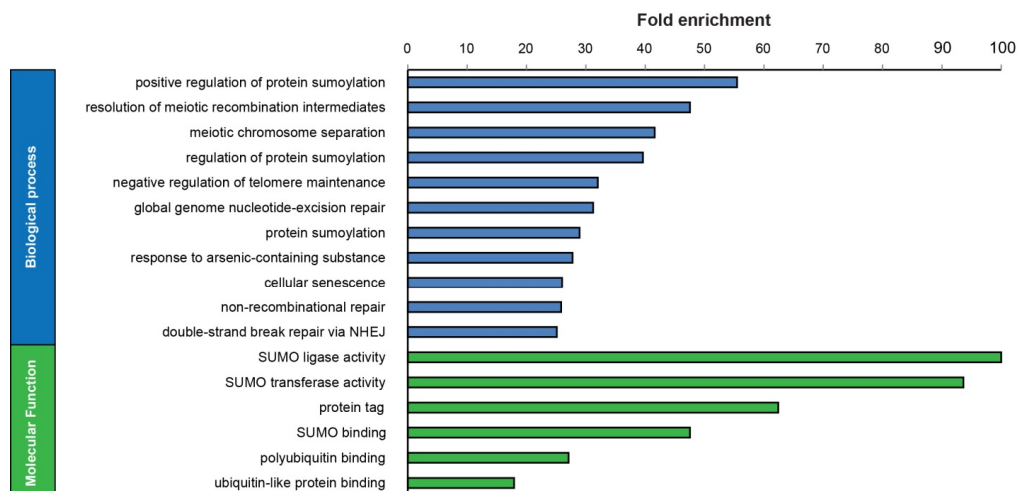
**b**

Supplementary Fig. 2

Supplementary figure 2. (a) Overview of the RNF4-regulated SUMO2 target protein network. STRING protein interaction network of SUMOylation targets enriched after RNF4 knockdown from Fig. 1d. STRING analysis results were visualized using Cytoscape (version 3.4.0). A high confidence score ($p>0.7$) was used. (b) Validation of RNF4 substrates by

Chapter 4

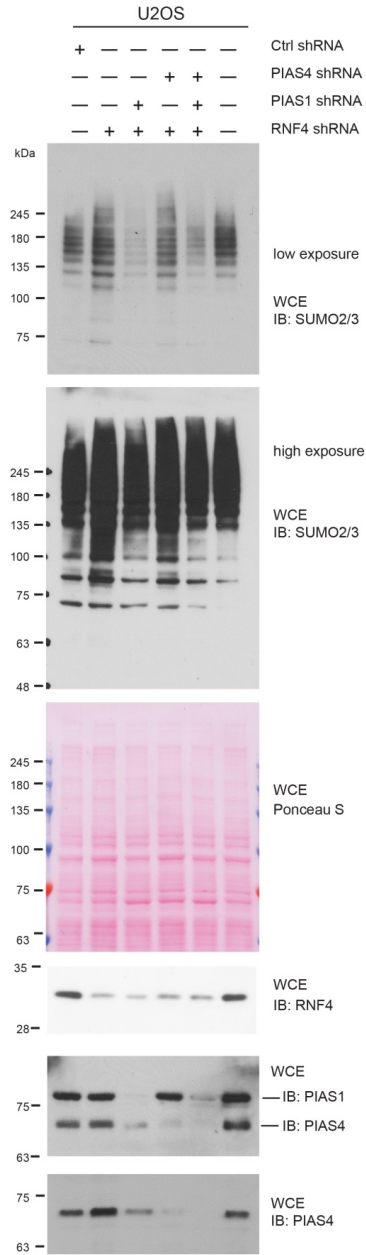
immunoblotting. U2OS cells stably expressing His10-SUMO2 were separately infected with lentiviruses expressing three different shRNAs directed against RNF4 or a control shRNA. Three days post infection, cells were harvested and His10-SUMO2 conjugates were purified from denaturing lysates. Validated RNF4 substrates include BARD1, SETDB and RAD18. Experiments were independently repeated. RNF4 knockdown efficiency and His10-SUMO2 conjugates were verified by immunoblotting. n.s. indicates non-specific bands.



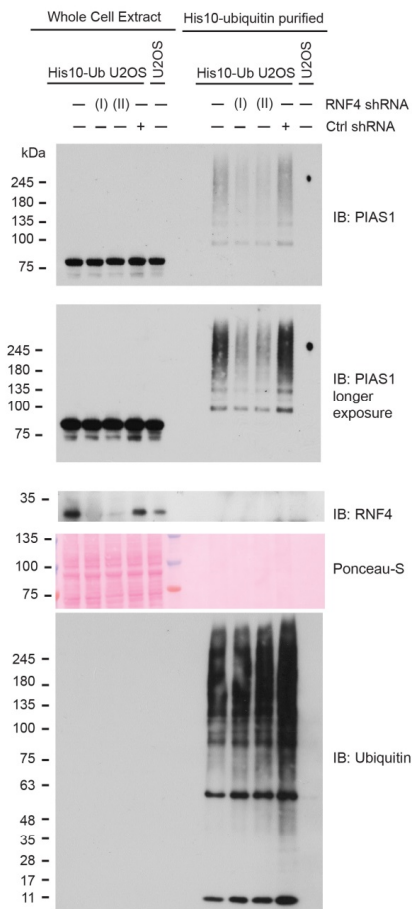
Supplementary Fig. 3

Supplementary figure 3. Gene ontology analysis of all RNF4-TULIP target proteins regarding biological process and molecular function. Full Gene ontology is shown in Supplementary Dataset 5.

a



b

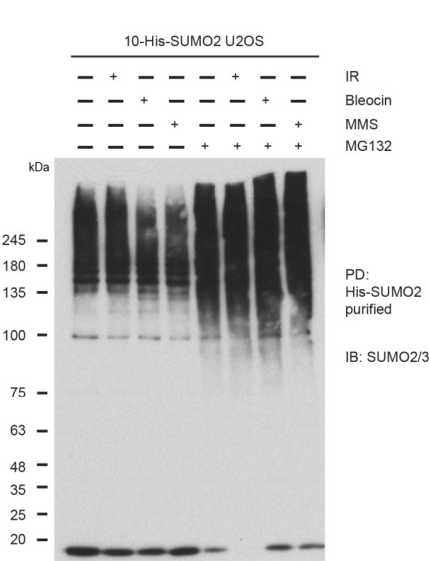


Supplementary Fig. 4

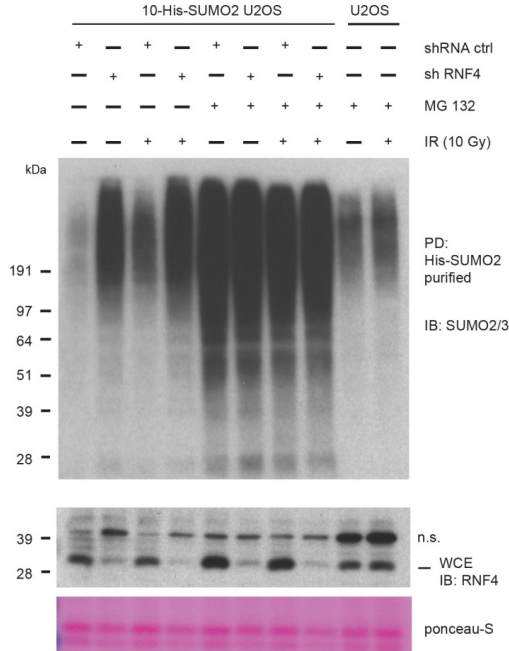
Supplementary figure 4. RNF4 targets the SUMO E3 ligase PIAS1. **(a)** The overall increase in protein SUMOylation upon RNF4 knockdown is counteracted by co-knock down of PIAS1. Independent confirmation of Fig. 4a. U2OS cells were (co)-infected with lentiviruses expressing shRNAs against RNF4, PIAS1 or PIAS4 or a control shRNA as indicated. Three days after infection, cells were lysed in a denaturing buffer and knockdown efficiencies and overall levels of SUMO2/3 were analysed by immunoblotting. **(b)** Independent confirmation of Fig. 4e. RNF4 regulates PIAS1 ubiquitylation. U2OS cells stably expressing His10-ubiquitin were infected with lentiviruses expressing shRNAs directed against RNF4 or a control

shRNA. Three days after infection, cells were lysed in a denaturing buffer and His10-ubiquitin conjugates were purified. The levels of ubiquitylated PIAS1 were verified by immunoblotting. Similarly, RNF4 knockdown efficiency was verified by immunoblotting.

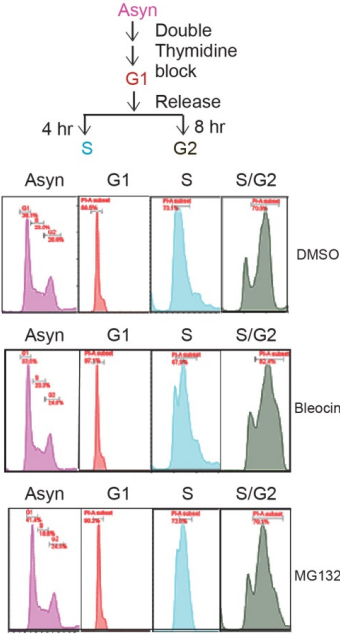
a



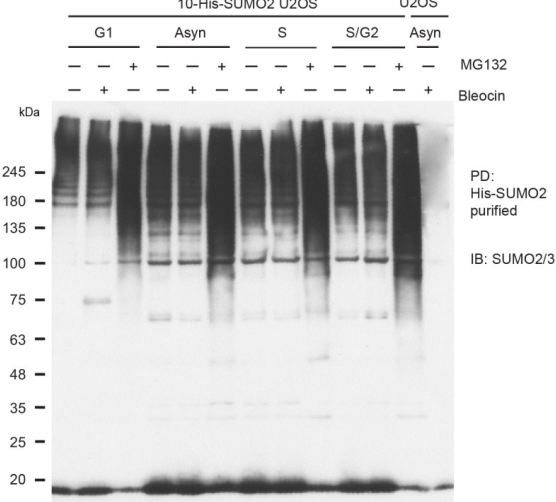
b



c

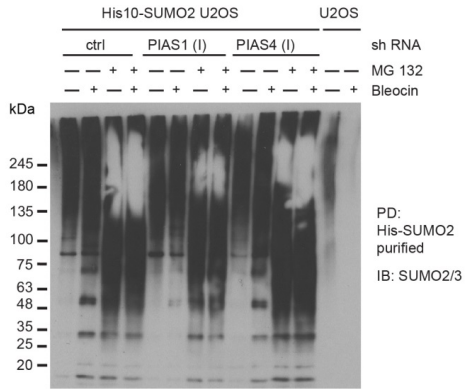
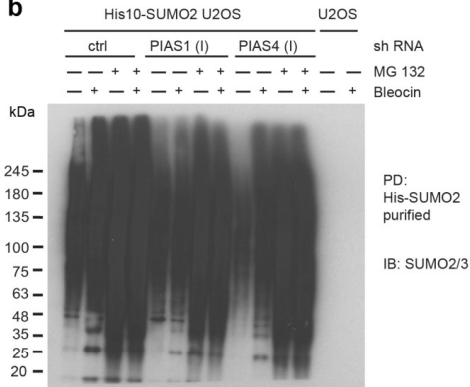
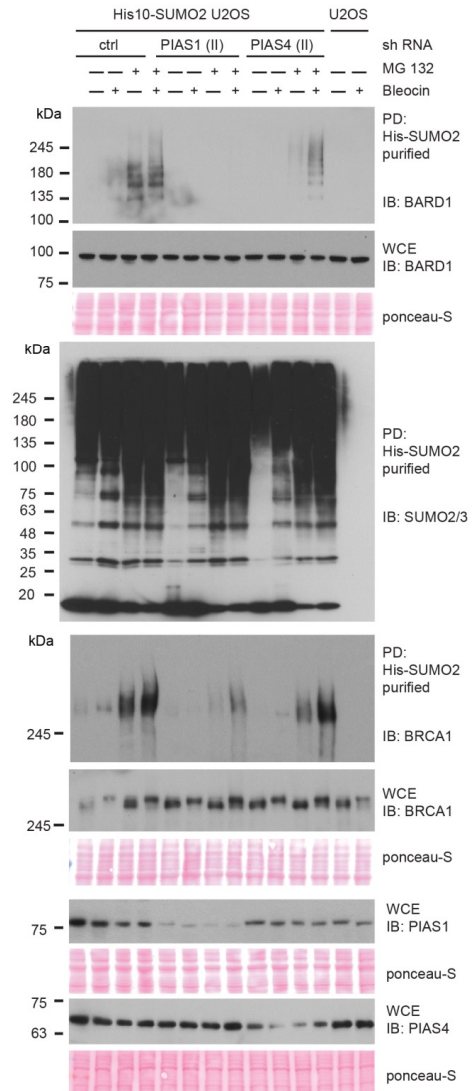


d



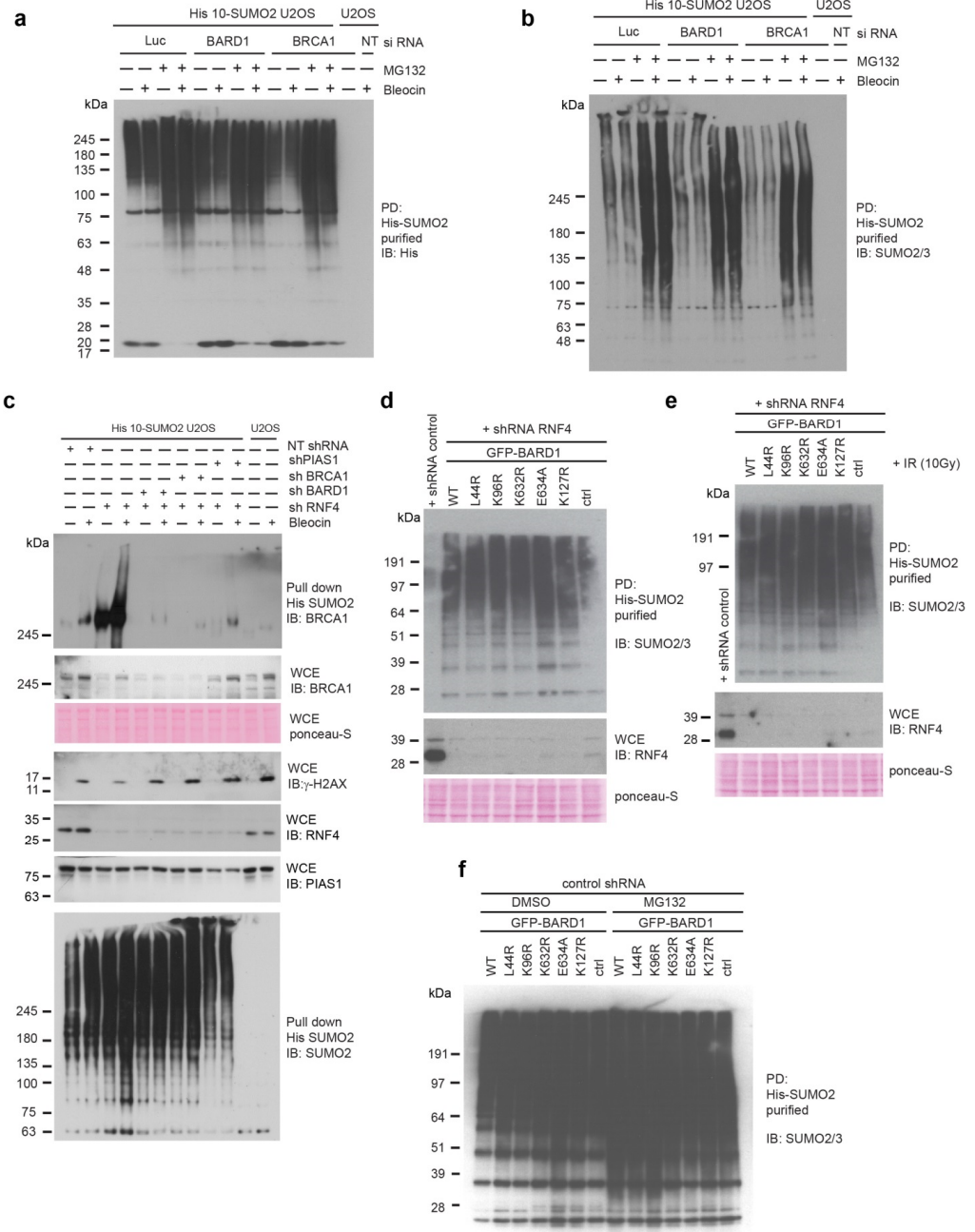
Supplementary Fig. 5

Supplementary figure 5. BARD1 is a RNF4 substrate regulated by DNA damage. **(a)** BARD1 is SUMOylated in response to DNA damage and SUMOylated BARD1 is degraded by the proteasome. Protein samples from the experiment described in Fig. 5a, were immunoblotted with SUMO2/3 antibody. **(b)** Protein samples from the experiment presented in Fig. 5b, were immunoblotted with SUMO2/3 antibody. **(c)** Experimental set up of the cell cycle arrest experiment. To arrest cells at the G1/S boundary, 30% confluent cells were treated with 2 mM thymidine for 19 hrs and then released for 9 hours, followed by a second thymidine (2 mM) block for 17 hrs. G1 arrested cells were washed and released for cell cycle progression by adding fresh cell culture medium. Cells were collected after 4 hours and 8 hours to obtain cells enriched in S-phase or G2/M respectively. FACS profiles of asynchronously growing, G1 arrested, S phase and G2/M cells were determined after propidium iodide (PI) staining. DMSO treated, Bleocin treated and MG132 cells at different stages of cell cycle **(b)**. **(d)** Protein samples from the experiments presented in Fig. 5c and Fig. 5d, were immunoblotted with antibody raised against SUMO2/3. Unprocessed full-size scans of blots are provided in Supplementary Fig. 9.

a**b****c**

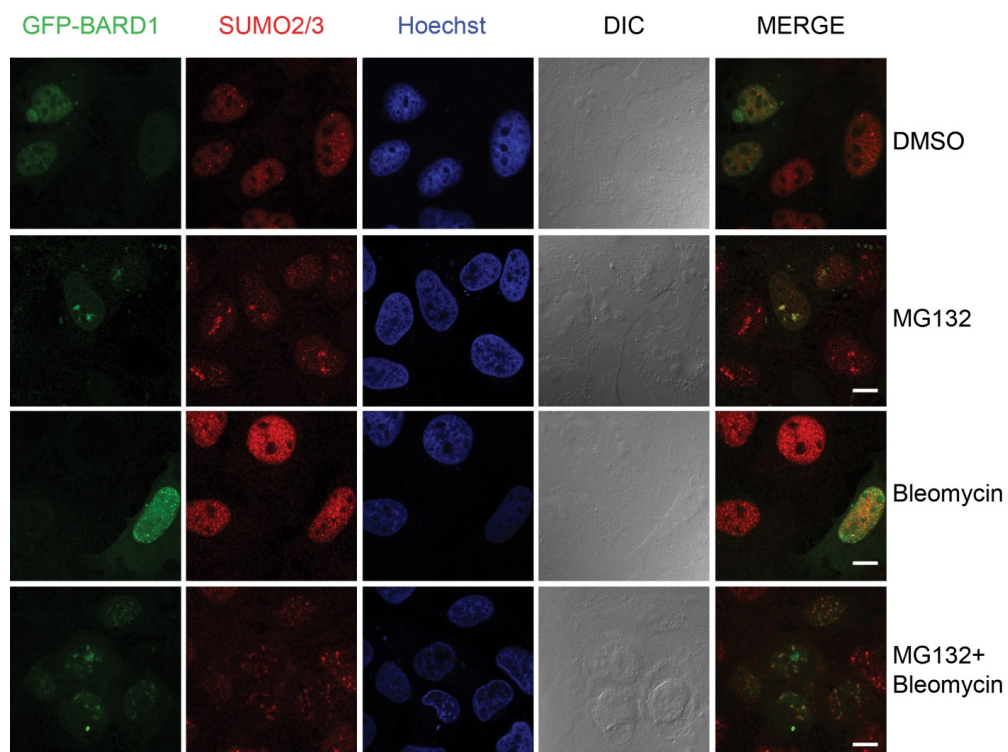
Supplementary Fig. 6

Supplementary figure 6. The SUMO E3 ligase PIAS1, is responsible for the SUMOylation of BARD1 and BRCA1. **(a-b)** Protein samples from experiment presented in Fig. 6a and 6b, were immunoblotted with antibody raised against SUMO2/3. **(c)** The experiment presented in Fig. 6 was repeated with independent sets of shRNAs (PIAS1/II and PIAS4/II) to deplete PIAS1 and PIAS4. Levels of SUMOylated BARD1 and total BARD1 were determined by immunoblotting, Levels of SUMOylated BRCA1 and total BRCA1 were determined by immunoblotting. Knockdown efficiencies of PIAS1 and PIAS4 were determined by immunoblotting. Additionally, the SUMO2 purification efficiency was determined by immunoblotting. Unprocessed full-size scans of blots are provided in Supplementary Fig. 9.



Supplementary Fig. 7

Supplementary figure 7. SUMOylation of BARD1 occurs in a BRCA1-dependent manner. **(a- b)**. Protein samples from experiments presented in Fig. 7a and 7b were immunoblotted with antibody raised against SUMO2/3. **(c)** Protein samples from the experiment presented in Fig. 7b were immunoblotted with antibody raised against SUMO2/3. To determine the RNF4 depletion level whole cell extracts were immunoblotted with RNF4 antibody. To determine de formation of DNA damage whole cell extracts were immunoblotted against γ -H2A.X **(d-f)** Protein samples from experiment presented in Fig. 7(c-e) respectively, were immunoblotted with SUMO2/3 antibody. Unprocessed full-size scans of blots are provided in Supplementary Fig. 9.

**Supplementary Fig. 8**

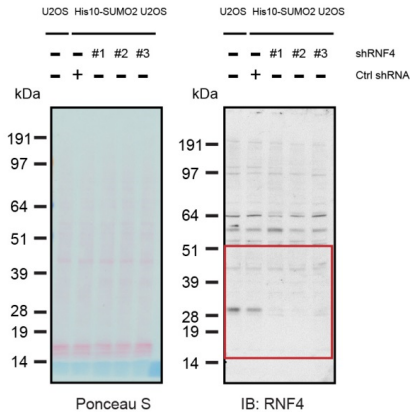
Supplementary figure 8. Co-localization of BARD1 and SUMO2/3 in response to DNA damage and proteasome inhibition.

Independent biological replicate of the experiment presented in Fig. 8.

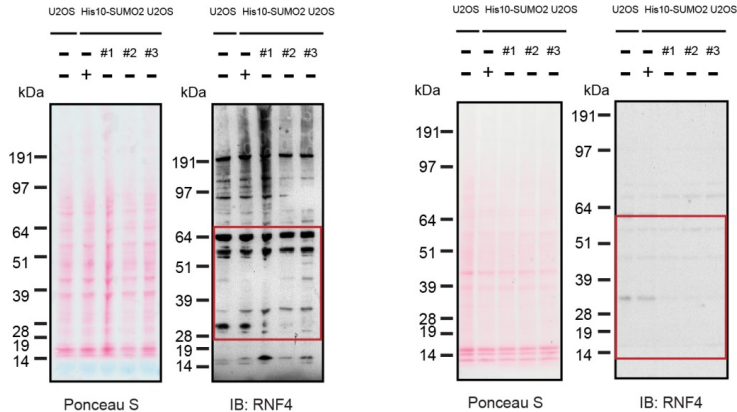
all full scale scans for immunoblots related to all figures

WCE

related to Fig. 1c



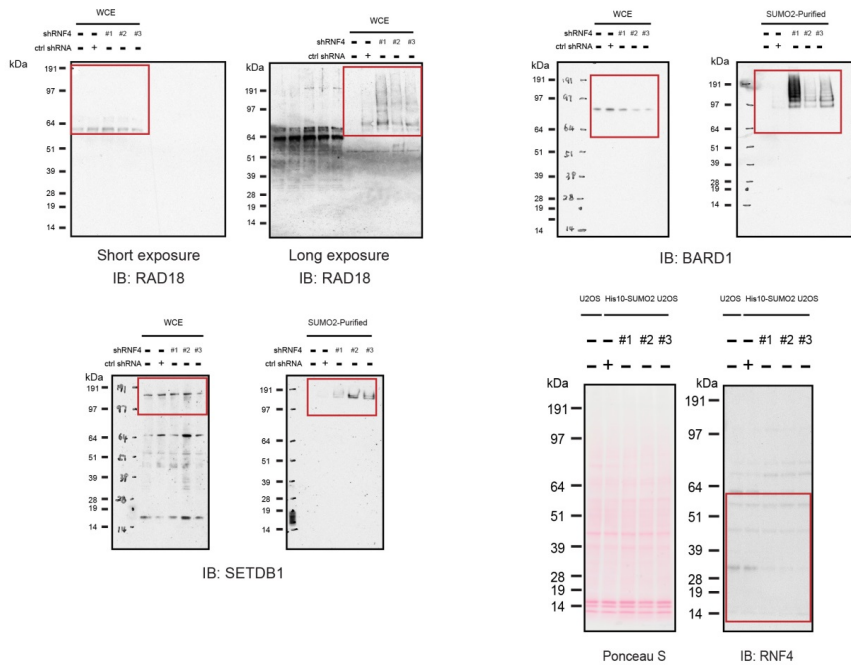
related to Supplementary Fig. 1a



Supplementary Fig. 9

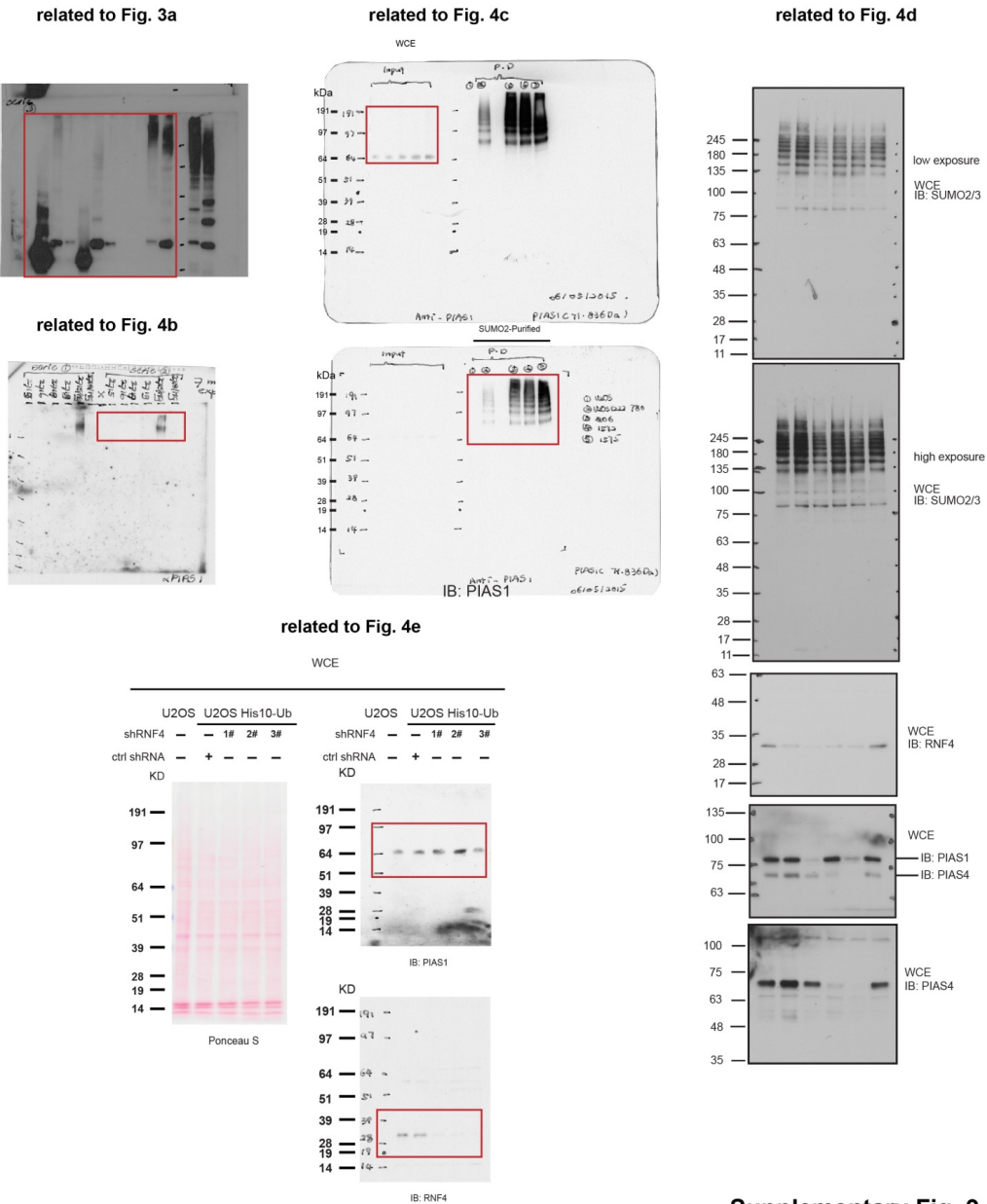
all full scale scans for immunoblots related to all figures

related to Supplementary Fig. 2b



Supplementary Fig. 9

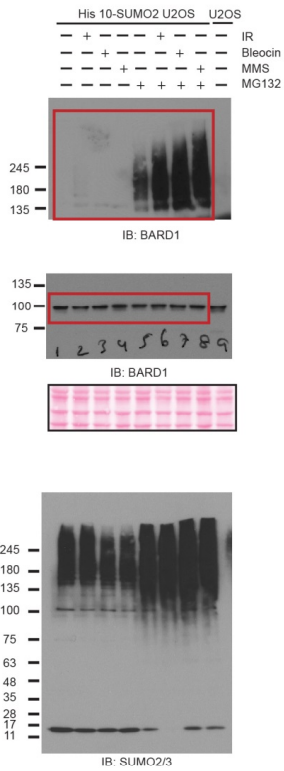
all full scale scans for immunoblots related to all figures



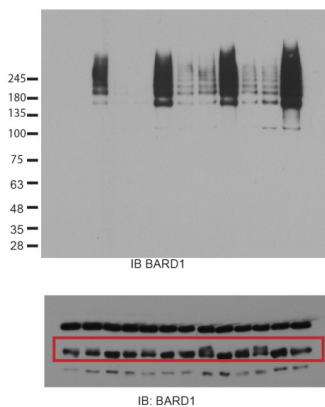
Supplementary Fig. 9

all full scale scans for immunoblots related to all figures

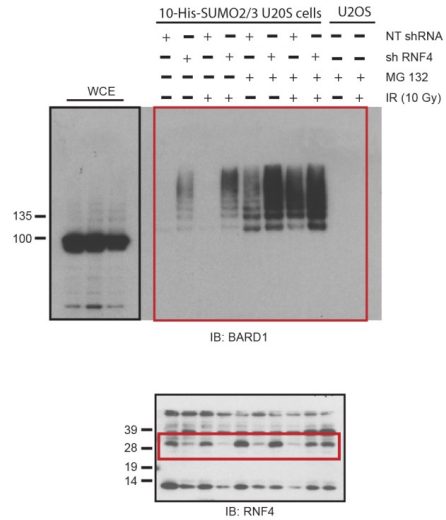
related to Fig. 5a / Supplementary Fig. 5a



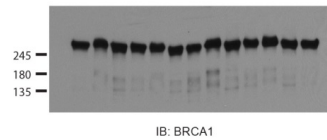
related to Fig. 5c



related to Fig. 5b / Supplementary Fig. 5b



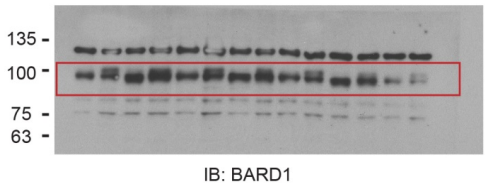
related to Fig. 5d



Supplementary Fig. 9

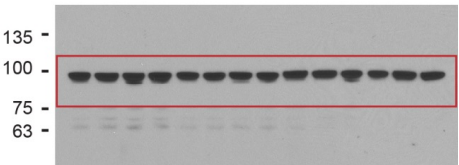
all full scale scans for immunoblots related to all figures

related to Fig. 6a



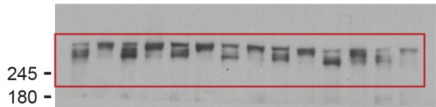
IB: BARD1

related to Supplementary Fig. 6c

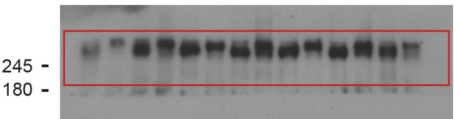


IB: BARD1

related to Fig. 6b

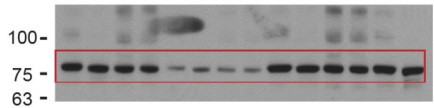


IB: BRCA1

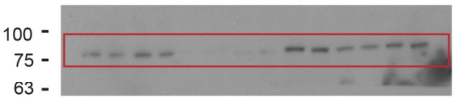


IB: BRCA1

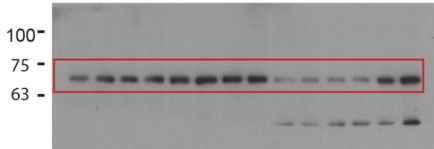
related to Fig. 6c



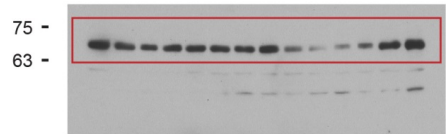
IB: PIAS1



IB: PIAS1



IB: PIAS4

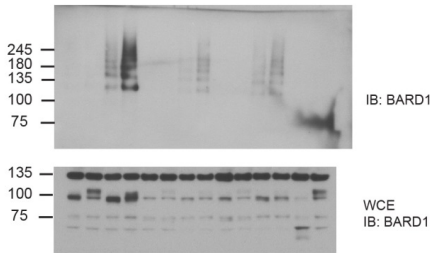


IB: PIAS4

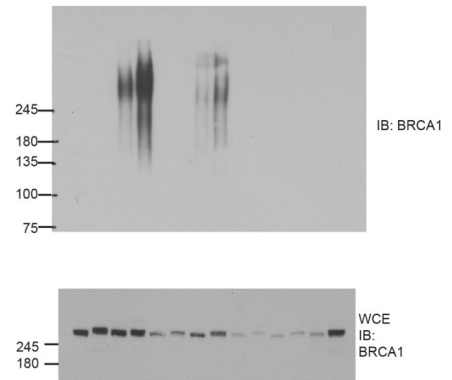
Supplementary Fig. 9

all full scale scans for immunoblots related to all figures

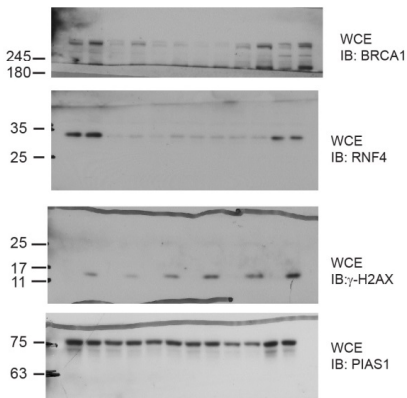
related to Fig. 7a



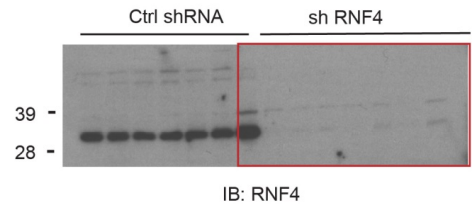
related to Fig. 7b.



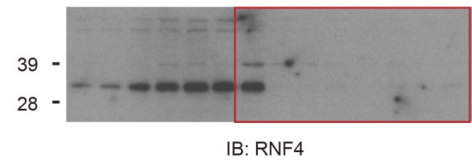
related to Supplementary Fig. 7c



related to Supplementary Fig. 7d



related to Supplementary Fig. 7e



Supplementary Fig. 9

Chapter 5

UFMylation: an emerging post-translational modification regulating protein synthesis

Katharina F. Witting^{1*}, Zhenyu Xiao^{1*}, Román González-Prieto¹, Huib Ovaa^{1, #}, Alfred C.O. Vertegaal^{1, #}

Affiliations: ¹Department of Cell and Chemical Biology, Division of Chemical Immunology, Leiden University Medical Center (LUMC), Einthovenweg 20, 2333 ZC, Leiden, The Netherlands

* These authors contributed equally

Address correspondence to: H.O. (h.ovaa@lumc.nl) and A.C.O.V. (e-mail: vertegaal@lumc.nl)

Summary

UFMylation is a reversible post-translational modification implicated in a variety of biological processes including ER homeostasis upon unfolded protein response (UPR), erythroid differentiation, and more recently coordinating interaction of mRNA and other proteins to the ribosome. We studied global UFMylation in a site-specific manner in both wildtype and UFSP2 knockout cells. We identified seven proteins (RPL26, RPL26L1, TUBA1B, MCM5, SLC26A7, SCYL2, WDR63) for which we could map the modified lysines. Interestingly, we found UFMylation of RPL26 (L24), a 60S ribosomal protein, that we demonstrate to specifically interact with the signal recognition particle receptor (SRPR). Together, these results shed a new light on the biological role of UFM1 and open new avenues of research to clarify the physiological role of the UFM1-system.

1 Introduction

Post-translational modifications (PTMs) tightly regulate protein functions either through the addition of chemical moieties like methylation, acetylation and phosphorylation or small protein modifiers that belong to the ubiquitin-like (Ubl) family to specific residues of their target proteins^{1,2}. Besides ubiquitin, other proteins such as Small Ubiquitin-like Modifiers (SUMOs), Atg8, Atg12, FUBI, FAT10, HUB1, ISG15, Nedd8, URM1 and UFM belong to the Ubl family^{3,4}. As a result of these PTMs, functional study on proteomics become extremely complex^{5,6}.

Phosphorylation and ubiquitination are the most common and well-studied post-translational modification mechanisms in eukaryotes⁷⁻⁹. Small ubiquitin-like modifiers (SUMOs), have also been extensively studied in the past two decades. Thousands of SUMO substrates as well as acceptor lysines were discovered by mass spectrometry (MS) analyses¹⁰⁻¹⁴, which provided a wealth of resources for further functional study of these modifications. However, only limited information is available about other Ubl conjugation systems such as UFM1.

UFM1, a recently identified ubiquitin-like protein, resembles ubiquitin, covalently modifies lysine residues of its substrates through a three components enzymatic cascade¹⁵⁻¹⁹. UFM1 is translated as an inactive precursor form (pro-UFM1) which has one or two additional amino acids beyond the conserved single glycine and are cleaved by UFSP2 to expose the UFM1- single active C-terminal glycine. Mature UFM1 is activated as an adenylate by UBA5 forming a high-energy thioester bond with the active-site cysteine of UBA5. This allows the transfer onto the catalytic site of the E2 enzyme-UFC1, where again a thioester bond is formed. With the action of the E3-like enzyme, UFL1, activated UFM1 is transferred onto the lysine residues of the protein substrates, forming an isopeptide bond. Similar to ubiquitin, UFMylation is reversed by cleavage of UFM1 specific protease, UFSP2. UFM1 has six lysines which are amenable to forming poly-UFMylation^{20,21}.

The main conserved function of the ribosome is translation of the genomic code into proteins^{22,23}. The eukaryotic ribosome consists of 80 ribosomal proteins and a ribosomal RNA shell^{24,25}. Secretory proteins are delivered through the secretory pathway. In this pathway proteins possessing signal sequences are recognized at the ribosome by the signal recognition particle (SRP) while they are still undergoing synthesis. With the help of SRP and its membrane bound receptor (SRPR or SR), co-translational secretory and membrane proteins are directly translocated into the translocation channel and finally into the ER²⁶.

UFMylation has been connected to biological processes including ER homeostasis, vesicle trafficking, autophagy and more recently it has been demonstrated to modify ribosomal proteins^{27, 29,30}. A failure to conjugate UFM1 to target proteins results in the promotion of tumor formation²⁰. Loss of function of UFMylation in mice leads to apoptosis in fetal liver cells and pancreatic beta cells²¹. Using an improved immunoprecipitation approach, Simsiek et al. reported UFMylation of a ribosomal protein (eL36) and its interaction with the UFM1 E3 ligase UFL1²⁷. More recently, Pirone et al reported their

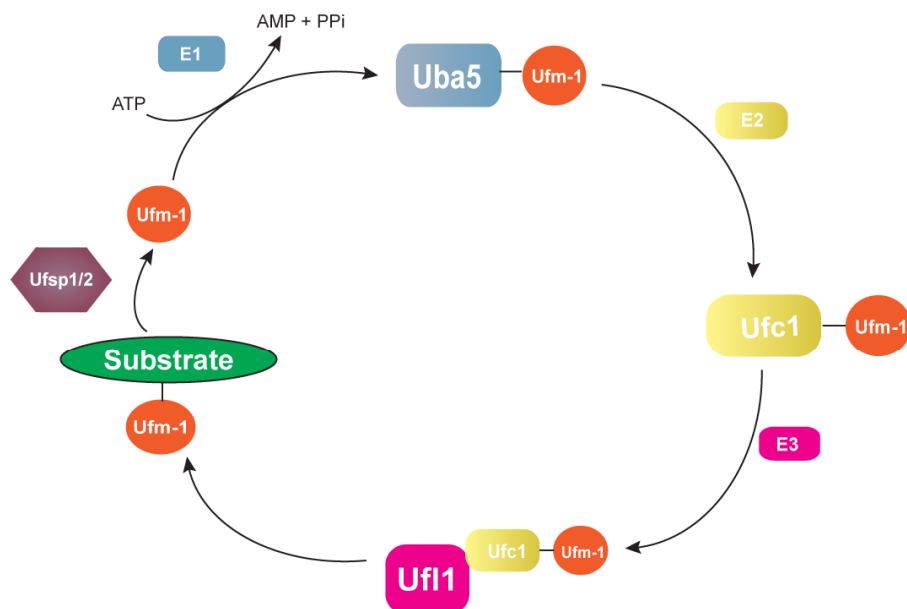
identification of 494 UFMylated proteins, of which 82 survived stringent filtering²⁸. However, none of them been independently validated and no biological function studies have been well conducted to elucidate the role of protein UFMylation. Despite these interesting findings, the biological consequences of UFMylation remain obscure.

Given the lack of understanding of biological processes involving UFMylation, as well as the transient nature of UFM1 modification, substrate identification has posed major technical challenges and urgently calls for the development of proteomics methods. To address this problem, we adapted the mass spec technology developed by Hendriks et al^{10,13}. by combining it with CRISPR/Cas9-mediated gene knockout of the only known UFM1 specific protease-UFSP2 that is active in mammalian cells, to enrich UFMylated substrates and to identify UFM1 specific acceptor lysines. Seven proteins with UFM1 acceptor lysines have been identified in our screen. One of the major UFM1 modified proteins was the ribosomal protein RPL26. We found that UFMylated RPL26 can efficiently interact with signal recognition particle receptor (SRPR). Our findings provide new insight in UFM1 signal transduction.

2 Results

UFM1 is a 9.1-kDa protein with a similar tertiary structure to ubiquitin and with a similar three component enzymatic cascade. Mature UFM1 can be attached to target proteins. Deconjugation of UFM1 is mediated by the two UFM1-specific proteases, UFSP1 and UFSP2 (**Figure 1 A and B**). Previously Ishimura et al. has shown that knockdown of UFSP2 leads to an increase in UFMylation, thereby enabling the enrichment of UFM1 modified proteins³¹. We employed CRISPR-Cas9 mediated genome editing to knock out UFSP2 in Hela cells (**Figure 2A**).

A



B

Alignment of Ufm1 and Ub



Figure 1. UFM-1 cascade and UFM-1. A) Ufm-1 conjugation machinery B) UFM-1 sequence alignment with Ubiquitin (Ub) shows that the C-terminal di-glycine motif characteristic of Ub is replaced by a valine- glycine motif in UFM-1.

2.1 Proteomics approach to study global UFMylation in UFSP2 knockout cells

Given that the knockout of UFSP2 in HeLa cells caused a significant increase in overall UFMylation and slightly decreased cell proliferation, we were interested in identifying UFM1 target proteins as well as acceptor lysines in UFSP2 knock out cells using a proteomics approach. Due to the nature of UFM1, one major challenge in addressing this is that analogous technologies used for identification of ubiquitinated substrates (e.g. di-Gly proteomics) are unavailable for UFM1, as antibodies targeting VG-sites are still not available. Furthermore, the abundance of UFMylated proteins is magnitudes lower than that of ubiquitination or SUMOylation, posing a major challenge for proteomics³².

To address this, we have adapted the biochemical purification of SUMO target proteins and SUMO sites developed by Hendriks et al¹⁰. We generated a HeLa cell line and a HeLa cell line with UFSP2 knockout stably expressing a His10-tagged lysine-deficient UFM1 (His10-UFM1-K0) to enable UFM1 acceptor lysine mapping (**Figure 2B**). To generate a stable cell line, we used puromycin as marker

for selection. Immunoblot analysis confirmed the slightly increased expression of His10-UFM1-K0 compared with endogenous levels of UFM1 in HeLa cells and the efficient purification of UFM1 conjugates by the His10-UFM1-K0 method (**Figure 2 C and D**).

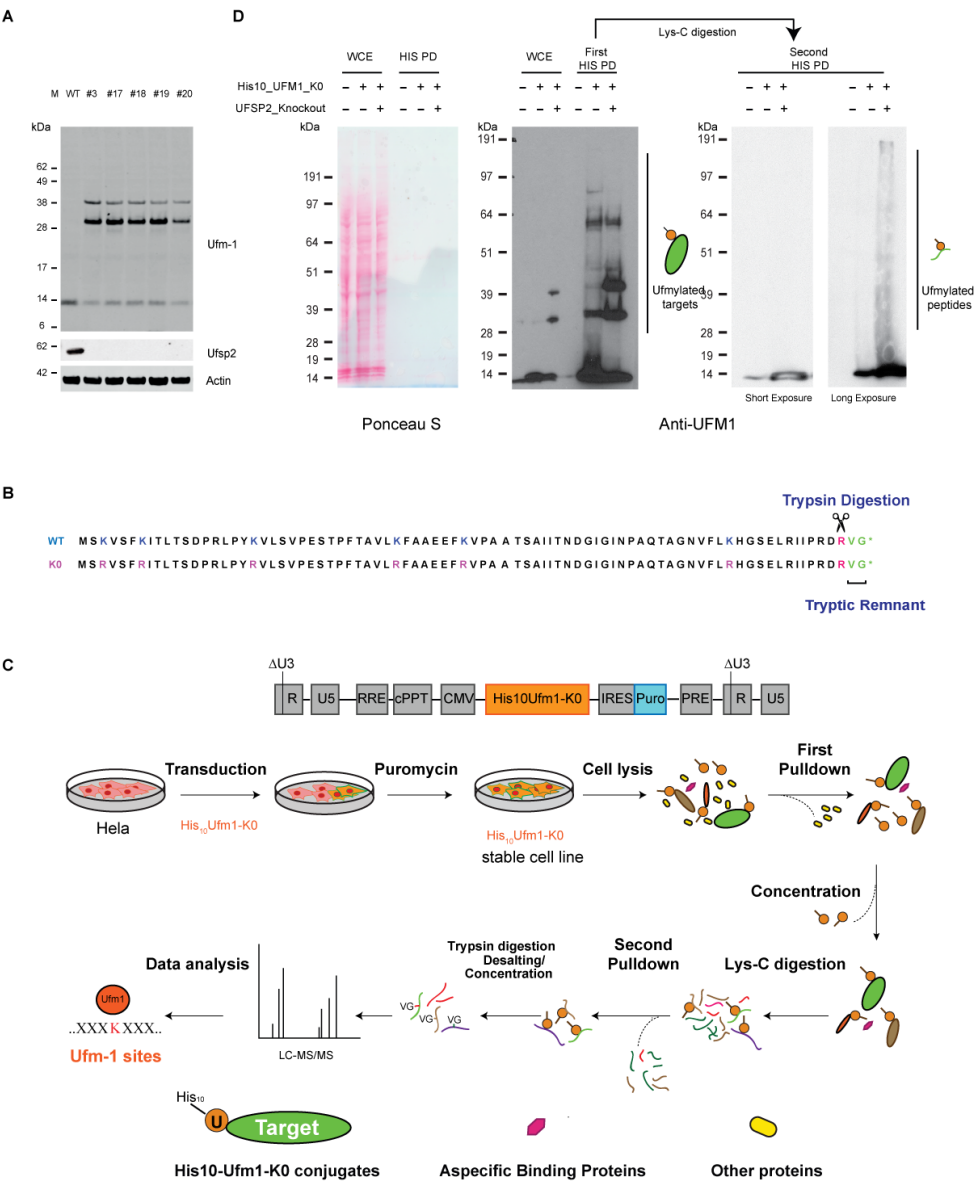


Figure 2. Generation and validation of HeLa cells stably expressing His10-UFM1-K0 with or without UFSP2. **A)** UFSP2 depletion increases UFMylation. UFSP2 was knocked out using CRISPR-CAS9 gene editing technology. The efficiency of UFSP2 knock out was assessed by immunoblotting using an antibody against UFSP2. Depletion of UFSP2 caused an increase of UFMylation as shown by immunoblotting using an antibody against UFM1. **B)** Sequence of His10-UFM1-K0 protein used to generate stable cell lines in either HeLa or HeLa UFSP2 knockout cells. **C)** Cartoon depicting the strategy to identify UFM1 targets as well as acceptor lysines. **D)** Purification of His10-UFM1-K0 conjugates via NTA purification was confirmed by immunoblotting. Whole cell extracts and UFM1 purified proteins of cells were run on 4-12% Bis-Tris polyacrylamide gels and levels of His10-UFM1 conjugates were compared by immunoblotting using anti-UFM1 antibody. Ponceau-S staining was performed to confirm the purity of the final fraction.

We used this cell line to enrich and purify UFMylation peptides from HeLa cells (His10-UFM1-K0) or HeLa cells (His10-UFM1-K0) without UFSP2. Wild type HeLa cells were used as the negative control. Cells were cultured using a lower puromycin concentration to ensure stable expression. To enrich the His10-UFM1 modified proteins, a two-step purification strategy was employed as described in Figure 2 C¹⁰. For proteomic analysis, three biological replicates with two technical repeats for all cell lines were analyzed in this study. In total, 13 sites were identified in seven proteins, which are RPL26, RPL26L1, TUBA1B, MCM5, SLC26A7, SCYL2, WDR63 (**Figure 3A**).

2.2 RPL26 is extensively UFMylated

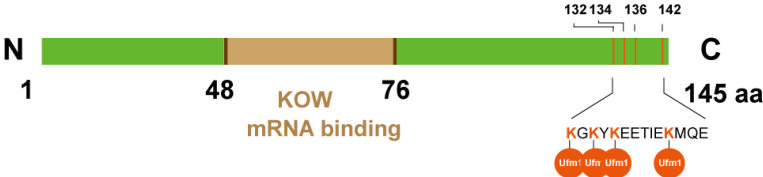
The ribosome production is a fundamental process in cell biology. RPL26 overexpression inhibits cell proliferation and induces cell cycle arrest through activating p53³³ and later more studies indicated that RPL26 controls p53 protein levels through binding to the 5' untranslated region of the p53 mRNA³⁴. However, little is known about its post translational modifications and its role in protein synthesis. We selected RPL26 for further validation to study its regulation by UFMylation because it is the top UFM1 target identified in our screen.

We found four UFM1 acceptor lysines in the C-terminus of RPL26 (K132, K134, K136 and K142) (**Figure 3 A and B**). Interestingly, alignment of orthologous genes showed that these four acceptor lysines are highly conserved across species probably due to their functional importance (**Figure 3C**). Immunoblotting analysis of endogenous RPL26 and Flag-RPL26 in His10-UFM1-K0 HeLa cells and UFSP2 knockout HeLa cells revealed that UFMylation of RPL26 strongly increased in the absence of UFSP2 (**Figure 3 D and E**). Increased RPL26-UFMylation thus can be explained by depletion of UFSP2.

A

Protein names	VG (UFM1) site within proteins	VG (UFM1) site positions	Score
RPL26 / RPL26L1	132	KILERKAKSRQVGKEKGYKEETIEKMQE	210.35/140.14
RPL26 / RPL26L1	134	LERKAKSRQVGKEKGYKEETIEKMQE	100.22/62.16
RPL26 / RPL26L1	136	RKAKSRQVGKEKGYKEETIEKMQE	69.03/60.16
RPL26	142	QVGKEKGYKEETIEKMQE	64.84
TUBA1B	370	INYPPTVPGGDLAKVQRAVCMSLNTTAA	92.93
MCM5	582	VRQLEAIVRIAEALSKMKLQPFATEADVEEA	61.16
SLC26A7	120	SIVLVLVKELNEQFKKIKVVLPLVDLVUI	60.91
SCYL2	73	KQEVAVFVDKLDIKYQKFEKDDQIDSLKR	59.91
SCYL2	76	VAVFVFDKLDIKYQKFEKDDQIDSLKRGVQ	59.91
WDR63	564	DHLCKTGQCKMLAGSKTEKAEEMMPYNNLES	45.28

B

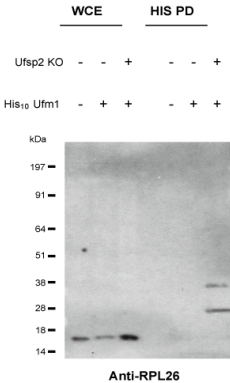


C

RPL26_YEAST11-127	MAQSLDVSSDRRKARKAYFTAPSSQRRVLLSAPLSKELRAQYGIKALPIRRDDEVLVVRGSKKGQEGKISSVY
Q9VU02_DROME1-149	MKQNPFFVSSSRKRNKRHFNAPSHIRRLMSAPLSKELRQKYNVRSMPIRRDDEVQVIRGHRKGNQVQVQAY
RPL26_MOUSE1-145	MKFNPFFVTSDRSKNRKRHFNAPSHIRRKIMSSPLSKELRQKYNVRSMPIRRDDEVQVVRGHHYKQQIGKVVQVY
RPL26_HUMAN1-145	MKFNPFFVTSDRSKNRKRHFNAPSHIRRKIMSSPLSKELRQKYNVRSMPIRRDDEVQVVRGHHYKQQIGKVVQVY
RPL26RAT1-145	MKFNPFFVTSDRSKNRKRHFNAPSHIRRKIMSSPLSKELRQKYNVRSMPIRRDDEVQVVRGHHYKQQIGKVVQVY
H2QC81_PANTR1-145	MKFNPFFVTSDRSKNRKRHFNAPSHIRRKIMSSPLSKELRQKYNVRSMPIRRDDEVQVVRGHHYKQQIGKVVQVY
RPL26_CAEEL1-145	MKVNPFFVSSDSGSKRKAHFNAPSHERRRIMSAPLTKELRTHGIRAIPIRTDDEVVVMRGRHKGNTGRVLRQYR
F7F4S1_MACMU1-145	MKFNPFFVTSDRSKNRKRHFNAPSHIRRKIMSSPLSKELRQKYNVRSMPIRRDDEVQVVRGHHYKQQIGKVVQVY
Q7SXA1_DANRE1-145	MKLNTFTVTSRRKRNKRHFNAPSHIRRKIMSSPLSKELRQKYNVRSMPIRRDDEVQVVRGHHYKQQIGKVVQVY
RPL26_SCHPO1-126	MKFSDVFTSSRRKRNKRKAHFGAPSSVVRVMSAPLSKELREQYKIRSLPVRDDQITIVIRGNSKGRGKITSVYR

RPL26_YEAST11-127	RLKFAVQVDKTYTEKVNAGASVPIINLHPSKLVITKLHLDKDRKALIQKGGKLE-----
Q9VU02_DROME1-149	RKKFVVVVEKIQRENAGTNVYVGIHPSKVLIVKLKLDKDRKAILERRGKGRALALSGKGGYETEETAQPMETA
RPL26_MOUSE1-145	RKKYVIYIERVQREKANGTTVHVGIRPSKVVITRLKLDKDRKKILERKAKSRQVGKEKGYKEETIEKMQE----
RPL26_HUMAN1-145	RKKYVIYIERVQREKANGTTVHVGIRPSKVVITRLKLDKDRKKILERKAKSRQVGKEKGYKEETIEKMQE----
RPL26RAT1-145	RKKYVIYIERVQREKANGTTVHVGIRPSKVVITRLKLDKDRKKILERKAKSRQVGKEKGYKEETIEKMQE----
H2QC81_PANTR1-145	RKKYVIYIERVQREKANGTTVHVGIRPSKVVITRLKLDKDRKKILERKAKSRQVGKEKGYKEETIEKMQE----
RPL26_CAEEL1-145	-KKFVIHIDKITREKANGSTVHIGIHPSKVAITKLKLDKDRRALVERKAAGRSRVGTILKGNHTDETVDN----
F7F4S1_MACMU1-145	RKKYVIYIERVQREKANGTTVHVGIRPSKVVITRLKLDKDRKKILERKAKSRQVGKEKGYKEETIEKMQE----
Q7SXA1_DANRE1-145	RKKYVIYIERVQREKANGTTVHVGIRPSKVVITRLKLDKDRKKILERKAKSRQDIEKGYKEETIEKMQE----
RPL26_SCHPO1-126	KKFLLLIERTVREKANGASAPVIGIDASKVVIITKLHLDKDRKDLIVRKGKGV-----

D



E

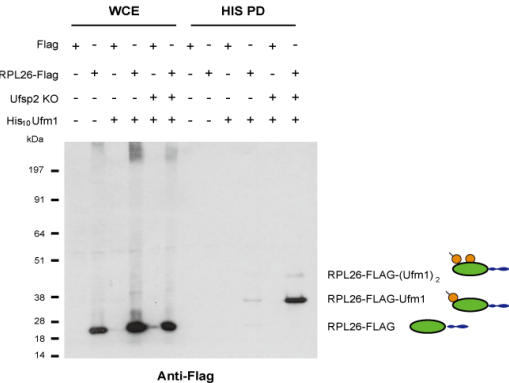


Figure 3. Label Free Quantification Results. **A)** VG-sites identified for UFM1 modified proteins by mass spectrometry. **B)** Domain structure of RPL26 highlighting the C-terminal UFM1 modification. **C)** Sequence Alignment of RPL26 reveals that the C-terminal lysine residues of this ribosomal protein are conserved across species except for yeast. **D)** HeLa cells with His10-UFM1-K0 and with or without UFSP2 were lysed and proteins were subjected to NTA purification. Whole cell extracts and UFM1 purified samples were analysed by immunoblotting using an antibody against RPL26. **E)** Flag-RPL26 was overexpressed in His10-UFM-K0 cell lines with or without UFSP2. Cells were lysed and proteins subjected to NTA purification. Whole cell extracts and UFM1 purified samples were analysed by immunoblotting using antibody against Flag.

2.3 UFMylated RPL26 can directly interact with SRPR

Next, we asked whether UFMylation of RPL26 affected ribosome associated protein interaction. To investigate such protein interactions, we performed immunoprecipitation of ribosomes with modified or unmodified RPL26 coupled with a label free quantitative proteomics approach to identify proteins preferentially binding to UFMylated RPL26. In this case, Flag-RPL26 was transiently overexpressed in both HeLa cells and HeLa cells UFSP2 knockout cells stably expressing His10-UFM1-K0 and immunoprecipitated using an anti-Flag antibody. Proteins were digested on beads with trypsin and four independent experiments were performed (**Figure 4A**). Proteins were identified by mass spectrometry and silver staining were performed to confirm the purity of the final fraction (**Figure 4B**). Volcano plots shown in Figure 4C indicate the proteins preferentially binding to UFMylated RPL26.

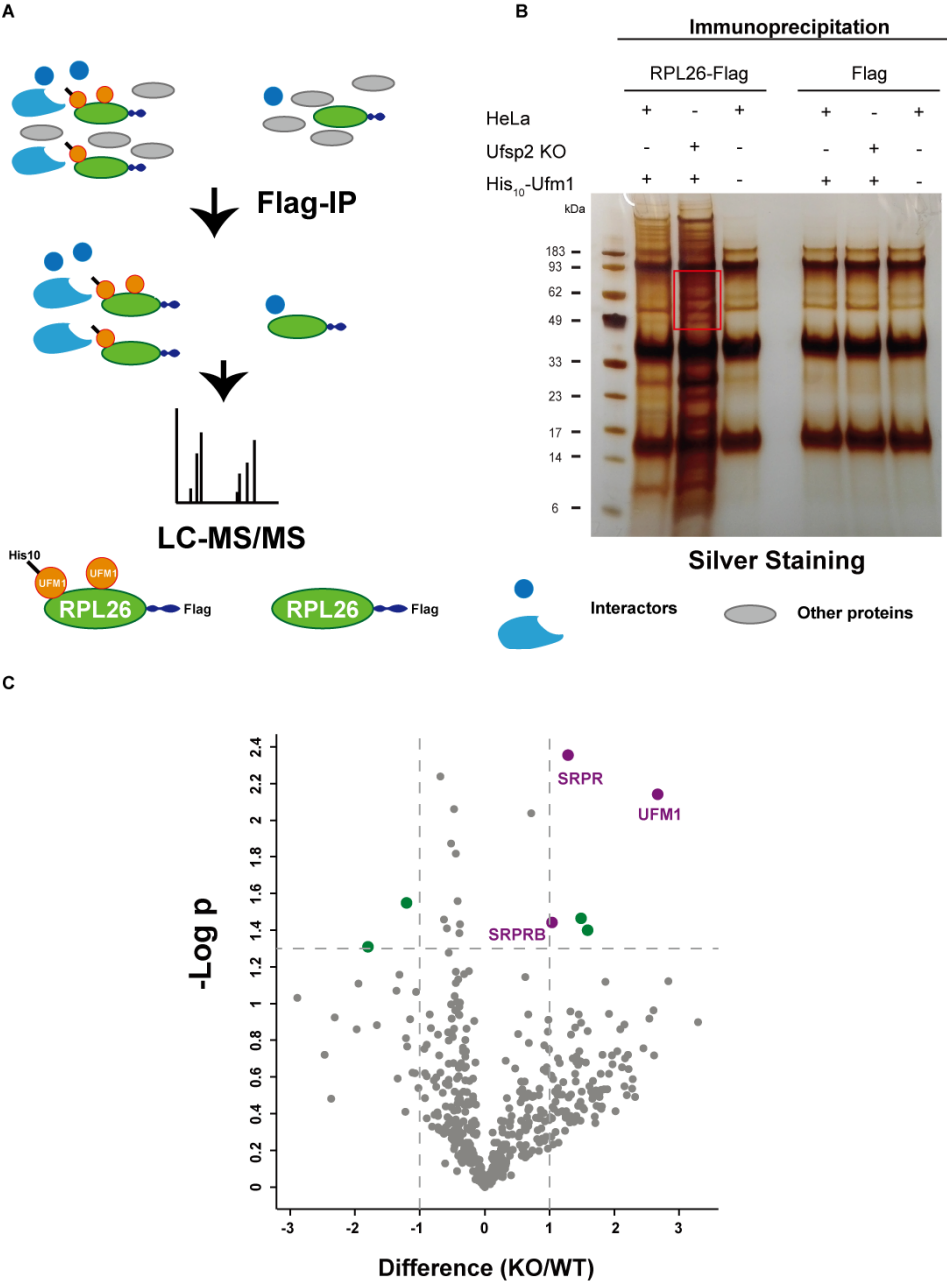


Figure 4. RPL26-UFMylation alters the interaction of ribosome associated proteins. **A)** Cartoon depicting the strategy to identify interactors of UFMylation RPL26. **B)** Silver staining of mass spec samples. Co-Immunoprecipitation of Flag-RPL26 from either the wild type or the UFSP2 knockout HeLa cells stably expressing His10-UFM-K0. The silver stained gels reveal that there are indeed differences between the modified and the unmodified RPL26 containing ribosomes. **C)** Volcano plots showing proteins preferentially binding to UFMylation RPL26.

Most notably, the signal recognition particle alpha (SRPR) interacts with UFM1 modified ribosomes in the UFSP2 KO His10-UFM1-K0 cells, but less when UFM1 is removed by the UFSP2 protease. This interaction could be confirmed by co-immunoprecipitation, which clearly showed that the UFM1 modified ribosome preferentially interacts with SRPR (**Figure 5 A and B**).

3 Discussion

It has been found that UFMylation is important to promote interactions between proteins¹⁵. Only little is known about the target proteins regulated by UFM1^{20,27,28,35}. To identify UFM1 target proteins as well as acceptor lysines, we have adopted a pioneering and well-developed method that has been proven to enable identification of SUMO acceptor lysines¹³. As a result, we successfully identified 8 novel UFM1 target proteins with 23 sites including RPL26 as a major UFM1 target in the absence of UFSP2. We confirmed RPL26 as a key UFM1 target and further confirmed that the UFMylation form of RPL26 can efficiently interact with SRPR. Functional analysis of RPL26 such as RNA profiling of UFSP2 knockout HeLa cells revealed a role for UFM1-modified RPL26 in promoting translation elongation (data not shown).

3.1 PTM of ribosomal proteins

The mammalian ribosomes are composed of two subunits: 40S (small) and 60S (large) subunits. The 40S subunit is composed of 33 proteins and 18S rRNA, whereas the 60S subunit contains 46 proteins and three rRNAs (5S, 5.8S and 28S). While the 40S subunit is involved in the decoding of mRNA, the 60S subunit is responsible for peptide bond formation and export of nascent peptides^{22-24,36-41}. The rim of the exit point is composed of RNA, and a ring of four ubiquitously conserved ribosomal proteins (L22, L23, L24 and L29). L24 is RPL26^{42,43}. (**Supplementary Table 1**)

Among all ribosomal proteins, RPS6 is the first reported and best characterized phosphorylated ribosomal protein which is involved in PI3K, mTOR and Ras signalling pathways^{30,44,45}. Later, mouse models linked phosphorylation of RPS6 to the initiation of pancreatic cancer, however the physiological roles are still elusive⁴⁶. Another PTM modifying ribosomal protein is ubiquitin, which forms K63-linked chains on the RPL28/uL15 protein in yeast⁴⁷. Proteomic analysis of cells with chemically induced unfolded protein stress (UPR), revealed that mono-ubiquitination of small ribosomal proteins Rps2/uS5 and Rps3/uS3 was altered dynamically, most likely affecting protein-protein interaction⁴⁸. Using an improved immunoprecipitation approach, Simsiek et al. first report UFMylation of a

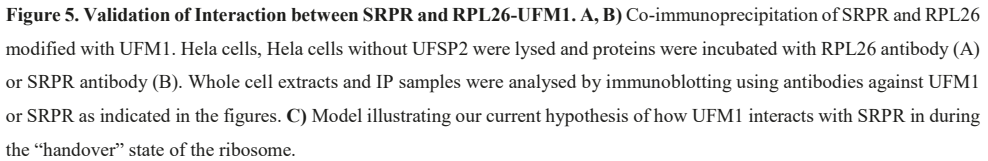
ribosomal protein (eL36) and its interaction with the UFM1 E3 ligase UFL1. However, they did not further investigate its biological impact on protein translation or protein-protein interaction²⁷.

3.2 UFMylation of RPL26 stimulates its interaction with SRPR.

Synthesis of new proteins is a very complex process. During translation of mRNA transcripts, various factors including enzymes and chaperones are involved in the processing, folding and targeting of nascent peptide chains to translocation pores at membranes⁴⁹. The ribosome plays a key role in these different tasks and serves as a platform for the regulated enzymes, targeting factors and chaperones that work on the newly translated polypeptides emerging from the exit tunnel. Secretory proteins are translated with an N-terminal signal sequence that can be recognized by the signal recognition particle (SRP). In order to target the ribosome-nascent chain complex to the endoplasmic reticulum membrane, SRP contacts the SRP receptor (SRPR) and leads the nascent chain transformation from the SRP into the lumen of the ER.

We immunoprecipitated ribosomes with modified or unmodified RPL26 coupled with a label free quantitative proteomics approach in order to identify associated factors. We identified the signal recognition particle receptor alpha (SRPR), to be significantly enriched in UFSP2 KO cells, suggesting that this protein might be preferentially regulated by UFMylated RPL26. As expected, immunoblot analysis revealed that SRPR coimmunoprecipitated more efficiently with UFMylated RPL26 than with the non-UFMylated RPL26. These results suggest that the ribosome contacts the SRPR directly in the presence of UFM1 modification of RPL26 and does not require the signal recognition particle (SRP) to mediate the interaction. Upon examination of literature⁵⁰, the scenario seems to be the case, during the “handover” state of the ribosome, that is when the ribosome has been properly positioned over the SEC translocon and the SRP is being displaced. Taking our results into consideration, we hypothesize that UFM1 seems to mediate direct contact of the tunnel exit of the ribosome with the SRPR during the last step of direct peptides to the Sec61 translocon (**Figure 5C**). Our data indicate that RPL26 is one of the UFM1 target proteins serving as a platform for ribosome associated protein interaction.

The observation that SRPR can preferentially interact with UFMylated RPL26 in our experiments further raises the question how UFMylation specifically regulates translation. Given the novelty of this discovery, more experiments addressing the functional consequences on translation need to be performed to fully understand the role of UFM1 in this context. To answer this question, more studies are being performed (i.e. ribosome profiling, data not shown).



4 Conclusion

In this study, we identified seven proteins (RPL26, RPL26L1, TUBA1B, MCM5, SLC26A7, SCYL2, WDR63) that are covalently modified by UFM1. We describe for the first time a novel mechanism by which UFMylation might regulate proteins during the last step of the translation (elongation), after (or shortly before) the SRP leaves the ribosome and the ribosome correctly binds to the Sec61 translocon. The identification of the ribosomal protein RPL26 being modified by UFM1 might help to understand how UFM1 regulates protein synthesis by coordinating its targets.

5 Experimental procedures

Antibodies

The primary antibodies used for this study are: anti-UFSP2 (Abcam ab185965), anti-Actin (Sigma Aldrich A5941, Clone AC-15, ascites fluid), anti-UFM1 (Abcam ab109305), anti-RPL26 (Thermo Fisher Scientific, PA5-17093), anti-Flag (Sigma Aldrich, M2), anti-SRPR (Abnova H00006734-B02P). Secondary antibodies used for this study are: anti-mouse-800 (1: 10000 dilution; LiCOR, 926-3210) and anti-rabbit-800 (1: 10000 dilution; LiCOR, 926-3211). Fluorescent second antibody were used for visualization of labeled proteins on LICOR Odyssey system v3.0.

For immunofluorescence, goat anti-rabbit and goat anti-mouse AlexaFluor-488/568 conjugates (Thermo Fisher Scientific) were used.

Cloning, cell culture and cell line generation

The His10-UFM1-K0 we described and used in this manuscript has the following amino acid sequence:

MHHHHHHHHHHHGGSMRSRVSRITLTSDPRLPYRVLSVPESTPFTAVLRFAAEEFRVPAA
TSAITNDGIGINPAQTAGNVFLRHGSELRIIPDRVG.

The corresponding nucleotide sequence was cloned in plasmid pLV-CMV-IRES-puro⁵¹. HeLa cell lines originated from ATCC, they were cultured under standard conditions (DMEM (Gibco) supplemented with 10% FCS (Sigma-Aldrich) at 37 °C with 5% CO₂). HeLa cells at 60% confluency were infected using a bicistronic lentivirus at an MOI of 3 encoding His10_UFM1_K0_IRES_puro. Following infection, cells were cultured under standard conditions as described before for two weeks under puromycin selection (2.5 µM). All cell lines used in this study were tested for mycoplasma contamination routinely with consistently negative outcome.

CRISPR-Cas mediated gene editing

UFSP2 guide RNA (gRNA) was designed using the CRISPR Design tool (<http://crispr.mit.edu/>), subcloned into a pX260 vector (Addgene).

UFSP2 guide RNA was subcloned into pX330-U6-Chimeric_BB-CBh-hSpCas9 (Addgene #42230, Cambridge, MA, USA), a human codon-optimized SpCas9 and chimeric guide RNA expression plasmid. CRISPR-mediated UFSP2 depletion in HeLa cells was performed by co-transfecting confluent HeLa with a vector harbouring the gRNA and the Cas9 and a construct conferring blasticidin resistance. After blasticidin selection and clonal expansion, UFSP2 depletion was verified by immunoblotting against using anti-UFSP2 antibody (1:1000 dilution, Abcam ab185965).

Electrophoresis and immunoblotting

Whole cell extracts or purified protein samples were separated on Novex Bolt 4-12% Bis-Tris Plus gradient gels (Life Technologies) using MOPS buffer or standard SDS-PAGE using a Tris-glycine

buffer and transferred onto Hybond-C nitrocellulose membranes (GE Healthcare Life Sciences) using a submarine system (Life Technologies). Membranes were stained with Ponceau S (Sigma) to visualize total protein amounts and subsequently blocked with 8% milk in PBS containing 0.05% Tween-20 before incubating with the primary antibodies as indicated.

Silver staining

For silver staining, gels were stained as described before⁵². Gels were incubated in Fixer (30% ethanol, 10% acetic acid) for 30 minutes and then washed two time in 20% ethanol and two times in water, 10 minutes each. Gels were sensitized in 0.8mM sodium thiosulfate for one minute and then rinsed in water twice for one minutes. Gels were then impregnated with 12 mM silver nitrate for 20 minutes and transferred to stop solution (40g of Tris and 20 ml of acetic acid per liter) for 30 minutes. Gels were then washed twice in water for 30 minutes.

Co-immunoprecipitation

For co-immunoprecipitation of RPL26-FLAG with endogenous interacting proteins, both HeLa cell lines stably expressing His10-UFM1 were transiently transfected with RPL26-FLAG or FLAG only as a background control. Cells were scraped on ice into ice cold co-immunoprecipitation buffer (50mM tris-HCl, pH 7.5, 1.5mM MgCl₂, 0.5% NP-40 and 5% Glycerol supplemented with 1mM DTT and a Complete Protease Inhibitor Tablet (Roche)). Lysates were clarified by centrifugation (4000rpm, 4°C, 20 min) and immunoprecipitated for 1h at 4°C with 3 µg Flag antibody followed by a 3h incubation with pre-equilibrated Protein-G Sepharose beads (GE Healthcare) at 4°C while rotating. After immunoprecipitation, samples were washed four times with ice-cold co-immunoprecipitation buffer and processed further for mass spectrometric analysis.

Purification of His10-UFM1 conjugates

Hela cells expressing His10-UFM1-K0 were washed, and scraped into ice-cold PBS. For total lysates, a small aliquot of cells was kept separately and lysed in 2% SDS, 1% N-P40, 50 mM TRIS pH 7.5, 150 mM NaCl. The remaining parts of the cell pellets were lysed in 6 M guanidine-HCl pH 8.0 (6 M guanidine-HCl, 0.1 M Na₂HPO₄/NaH₂PO₄, 10 mM TRIS, pH 8.0). The samples were snap frozen using liquid nitrogen, and stored at -80°C.

For His10-UFM1-K0 purification, the cell lysates were first thawed at room temperature and sonicated for 5 sec using a sonicator (Misonix Sonicator 3000) at 30 Watts to homogenize the lysate. Protein concentrations were determined using the bicinchoninic acid (BCA) protein assay reagent (Thermo Scientific) and lysate concentrations were equalized. Subsequently, imidazole as well as β-mercaptoethanol were added to a final concentration of 50 mM and 5 mM, respectively. His10-UFM1-K0 conjugates were enriched on nickel-nitrilotriacetic acid-agarose beads (Ni-NTA) (Qiagen), followed by washes with the following buffers A-D: Wash buffer A: 6 M guanidine-HCl, 0.1 M

Na₂HPO₄/NaH₂PO₄ pH 8.0, 0.01 M Tris-HCl pH 8.0, 10 mM imidazole pH 8.0, 5 mM β -mercaptoethanol, 0.1% Triton X-100. Wash buffer B: 8 M urea, 0.1 M Na₂HPO₄/NaH₂PO₄ pH 8.0, 0.01 M Tris-HCl pH 8.0, 10 mM imidazole pH 8.0, 5 mM β -mercaptoethanol, 0.1% Triton X-100. Wash buffer C: 8 M urea, 0.1 M Na₂HPO₄/NaH₂PO₄ pH 6.3, 0.01 M Tris-HCl pH 6.3, 10 mM imidazole pH 7.0, 5 mM β -mercaptoethanol, no Triton X-100. Wash buffer D: 8 M urea, 0.1 M Na₂HPO₄/NaH₂PO₄ pH 6.3, 0.01 M Tris-HCl, pH 6.3, 5 mM β -mercaptoethanol. For samples subjected to immunoblotting, wash buffers containing 0.2% Triton X-100, were used. Samples were eluted in 7 M urea, 0.1 M NaH₂PO₄/Na₂HPO₄, 0.01 M Tris/HCl, pH 7.0, 500 mM imidazole pH 7.0. For site-specific purification, we used the strategy developed previously^{10,13}.

Sample preparation and mass spectrometry

UFM1-enriched samples were supplemented with 1 M Tris-(2-carboxyethyl)-phosphine hydrochloride (TCEP) to a final concentration of 5 mM, and incubated for 20 minutes at room temperature. Iodoacetamide (IAA) was then added to the samples to a final concentration of 10 mM and samples were incubated in the dark for 15 minutes at room temperature. Lys-C and Trypsin digestions were performed according to the manufacturer's specifications. Lys-C was added in a 1:50 enzyme-to-protein ratio, samples were incubated at 37 °C for 4 hours, and subsequently 3 volumes of 100 mM Tris-HCl pH 8.5 were added to dilute urea to 2 M. Trypsin (V5111, Promega) was added in a 1:50 enzyme-to-protein ratio and samples were incubated overnight at 37 °C. For site-specific sample preparation, we used the strategy developed previously^{10,13}.

Then, digested samples were desalted, desalted samples were further concentrated on STAGE-tips as described previously^{10,13}. Samples were eluted with 0.1% formic acid in 80% acetonitrile. Elution step has been vacuum dried using a SpeedVac RC10.10 (Jouan, France) and before online nanoflow liquid chromatography-tandem mass spectrometry (nanoLC-MS/MS), final fractions were dissolved in 10 μ L 0.1% formic acid.

Experiments were performed on a Q-Exactive Orbitrap (Thermo Fisher Scientific, Germany) connected with an EASY-nLC 1000 system (Proxeon, Odense, Denmark) through a nano-electrospray ion source. Peptides were separated in a 15 cm analytical column with an inner-diameter of 75 μ m, in-house packed with 1.9AQ - μ m C18 beads (Reprospher-DE, Pur, Dr. Manish, Ammerbuch-Entringen, Germany).

The gradient length was 120 minutes from 2% to 95% acetonitrile in 0.1% formic acid at a flow rate of 200 nL/minute. The mass spectrometer was operated in a data-dependent acquisition mode with a top 10 method. Full-scan MS spectra were acquired at a target value of 3×10^6 and a resolution of 70,000, and the Higher-Collisional Dissociation (HCD) tandem mass spectra (MS/MS) were recorded at a target value of 1×10^5 and with a resolution of 17,500 with a normalized collision energy (NCE) of 25%. The maximum MS1 and MS2 injection times were 20 ms and 60 ms, respectively. The

precursor ion masses of scanned ions were dynamically excluded (DE) from MS/MS analysis for 60 sec. Ions with charge 1, and greater than 6 were excluded from triggering MS2 events.

For samples enriched for identification of UFM1 acceptor lysines, a 120-minute gradient was used for chromatography. Data dependent acquisition with a top 5 method was used. Maximum MS1 and MS2 injection times were 20 ms and 250 ms, respectively. Resolutions, normalized collision energy and automatic gain control target were set as mentioned previously. Dynamic exclusion was set to 20 sec.

Mass spec data analysis

Site level UFMylation data analysis

Site-specific purification was performed in three biological replicates, and all samples were measured in technical duplicates. All 18 RAW files were analyzed by MaxQuant (version 1.5.3.30). The first search was carried out with a mass accuracy of 20 ppm, while the main search used 64.5 ppm for precursor ions. Database searches were performed with Trypsin/P specificity, allowing three missed cleavages. Carbamidomethylation of cysteine residues was considered as a fixed modification. Mass tolerance of MS/MS spectra was set to 20 ppm to search against an in silico digested UniProt reference proteome for Homo sapiens (2017-30-01). Additionally, MS/MS data were searched against a list of 245 common mass spectrometry contaminants by Andromeda. Oxidation (M) and Acetyl (Protein N-term) were set as variable modifications. Additionally, VG modified lysine was introduced as a variable modification with a composition of $C_7H_{12}N_2O_2$ and a monoisotopic mass of 156.090 Da. Match between runs was used with 0.7 min match time window and 20 min alignment time window.

Interactors of Ufmylated RPL26

Protein lists generated by MaxQuant were further analyzed by Perseus (Version 1.5.5.3). Proteins identified only by site and as a common contaminant were filtered out, and then all the LFQ intensities were \log_2 transformed. Different experiments were annotated in to six group, Hela-Flag, Hela-Flag-RPL26, Hela-His10-Ufm1-K0-Flag, Hela-His10-Ufm1-K0-Flag-RPL26, Hela-His10-Ufm1-K0-Flag (without UFSP2), Hela-His10-Ufm1-K0-Flag-RPL26 (without UFSP2). Proteins identified in at least one treatment condition and found in at least three biological replicates were included for further analysis. For each experimental condition individually, missing values were imputed using Perseus software by normally distributed values with a 1.8 downshift (\log_2) and a randomized 0.3 width (\log_2). Final corrected P values were filtered to be less than 0.05. Then, average \log_2 ratios for Hela-His10-Ufm1-K0-Flag-RPL26 vs Hela-His10-Ufm1-K0-Flag were calculated and P values of each protein across all treatment conditions were calculated by T-tests. Proteins were selected as Ufmylated RPL26 interactors when their average \log_2 ratios are greater than 1 and corresponding p values were less than 0.05. Volcano plots to demonstrate identified proteins preferentially binding to UFMylated RPL26.

References

- 1 Deribe, Y. L., Pawson, T. & Dikic, I. Post-translational modifications in signal integration. *Nat Struct Mol Biol* **17**, 666-672, doi:10.1038/nsmb.1842 (2010).
- 2 Han, Z. J., Feng, Y. H., Gu, B. H., Li, Y. M. & Chen, H. The post-translational modification, SUMOylation, and cancer (Review). *Int J Oncol* **52**, 1081-1094, doi:10.3892/ijo.2018.4280 (2018).
- 3 Vertegaal, A. C. Uncovering ubiquitin and ubiquitin-like signaling networks. *Chem Rev* **111**, 7923-7940, doi:10.1021/cr200187e (2011).
- 4 Kerscher, O., Felberbaum, R. & Hochstrasser, M. Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu Rev Cell Dev Biol* **22**, 159-180, doi:10.1146/annurev.cellbio.22.010605.093503 (2006).
- 5 Choudhary, C. & Mann, M. Decoding signalling networks by mass spectrometry-based proteomics. *Nat Rev Mol Cell Biol* **11**, 427-439, doi:10.1038/nrm2900 (2010).
- 6 Schratzenholz, A., Klemm, M. & Cahill, M. Potential of comprehensive toxico-proteomics: quantitative and differential mining of functional proteomes from native samples. *Altern Lab Anim* **32 Suppl 1A**, 123-131 (2004).
- 7 Chen, Z., Zhou, Y., Zhang, Z. & Song, J. Towards more accurate prediction of ubiquitination sites: a comprehensive review of current methods, tools and features. *Brief Bioinform* **16**, 640-657, doi:10.1093/bib/bbu031 (2015).
- 8 Ghosh, S. & Dass, J. F. Study of pathway cross-talk interactions with NF-kappaB leading to its activation via ubiquitination or phosphorylation: A brief review. *Gene* **584**, 97-109, doi:10.1016/j.gene.2016.03.008 (2016).
- 9 Tsuchida, S., Satoh, M., Takiwaki, M. & Nomura, F. Ubiquitination in Periodontal Disease: A Review. *Int J Mol Sci* **18**, doi:10.3390/ijms18071476 (2017).
- 10 Hendriks, I. A. *et al.* Site-specific mapping of the human SUMO proteome reveals co-modification with phosphorylation. *Nat Struct Mol Biol* **24**, 325-336, doi:10.1038/nsmb.3366 (2017).
- 11 Hendriks, I. A. & Vertegaal, A. C. A high-yield double-purification proteomics strategy for the identification of SUMO sites. *Nat Protoc* **11**, 1630-1649, doi:10.1038/nprot.2016.082 (2016).
- 12 Xiao, Z. *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. *Mol Cell Proteomics* **14**, 1419-1434, doi:10.1074/mcp.O114.044792 (2015).
- 13 Hendriks, I. A. *et al.* Uncovering global SUMOylation signaling networks in a site-specific manner. *Nat Struct Mol Biol* **21**, 927-936, doi:10.1038/nsmb.2890 (2014).
- 14 Munk, S. *et al.* Proteomics Reveals Global Regulation of Protein SUMOylation by ATM and ATR Kinases during Replication Stress. *Cell Rep* **21**, 546-558, doi:10.1016/j.celrep.2017.09.059 (2017).
- 15 Komatsu, M. *et al.* A novel protein-conjugating system for Ufm1, a ubiquitin-fold modifier. *EMBO J* **23**, 1977-1986, doi:10.1038/sj.emboj.7600205 (2004).
- 16 Kang, S. H. *et al.* Two novel ubiquitin-fold modifier 1 (Ufm1)-specific proteases, UfSP1 and UfSP2. *J Biol Chem* **282**, 5256-5262, doi:10.1074/jbc.M610590200 (2007).
- 17 Tatsumi, K. *et al.* A novel type of E3 ligase for the Ufm1 conjugation system. *J Biol Chem* **285**, 5417-5427, doi:10.1074/jbc.M109.036814 (2010).

- 18 Hertel, P. *et al.* The ubiquitin-fold modifier 1 (Ufm1) cascade of *Caenorhabditis elegans*. *J Biol Chem* **288**, 10661-10671, doi:10.1074/jbc.M113.458000 (2013).
- 19 Daniel, J. & Liebau, E. The ufm1 cascade. *Cells* **3**, 627-638, doi:10.3390/cells3020627 (2014).
- 20 Yoo, H. M. *et al.* Modification of ASC1 by UFM1 is crucial for ERalpha transactivation and breast cancer development. *Mol Cell* **56**, 261-274, doi:10.1016/j.molcel.2014.08.007 (2014).
- 21 Yoo, H. M., Park, J. H., Jeon, Y. J. & Chung, C. H. Ubiquitin-fold modifier 1 acts as a positive regulator of breast cancer. *Front Endocrinol (Lausanne)* **6**, 36, doi:10.3389/fendo.2015.00036 (2015).
- 22 Garrett, R. A. & Wittmann, H. G. Structure and function of the ribosome. *Endeavour* **32**, 8-14 (1973).
- 23 Henderson, E. *et al.* A new ribosome structure. *Science* **225**, 510-512 (1984).
- 24 Khatter, H., Myasnikov, A. G., Natchiar, S. K. & Klaholz, B. P. Structure of the human 80S ribosome. *Nature* **520**, 640-645, doi:10.1038/nature14427 (2015).
- 25 Petrov, A. S. *et al.* RNA-magnesium-protein interactions in large ribosomal subunit. *J Phys Chem B* **116**, 8113-8120, doi:10.1021/jp304723w (2012).
- 26 Wild, K., Rosendal, K. R. & Sinning, I. A structural step into the SRP cycle. *Mol Microbiol* **53**, 357-363, doi:10.1111/j.1365-2958.2004.04139.x (2004).
- 27 Simsek, D. *et al.* The Mammalian Ribo-interactome Reveals Ribosome Functional Diversity and Heterogeneity. *Cell* **169**, 1051-1065 e1018, doi:10.1016/j.cell.2017.05.022 (2017).
- 28 Pirone, L. *et al.* A comprehensive platform for the analysis of ubiquitin-like protein modifications using in vivo biotinylation. *Sci Rep* **7**, 40756, doi: 10.1038/srep40756 (2017).
- 29 Zhang, Y., Zhang, M., Wu, J., Lei, G. & Li, H. Transcriptional regulation of the Ufm1 conjugation system in response to disturbance of the endoplasmic reticulum homeostasis and inhibition of vesicle trafficking. *PLoS One* **7**, e48587, doi:10.1371/journal.pone.0048587 (2012).
- 30 Joassard, O. R. *et al.* Regulation of Akt-mTOR, ubiquitin-proteasome and autophagy-lysosome pathways in response to formoterol administration in rat skeletal muscle. *Int J Biochem Cell Biol* **45**, 2444-2455, doi:10.1016/j.biocel.2013.07.019 (2013).
- 31 Ishimura, R. *et al.* A novel approach to assess the ubiquitin-fold modifier 1-system in cells. *FEBS Lett* **591**, 196-204, doi:10.1002/1873-3468.12518 (2017).
- 32 Merbl, Y., Refour, P., Patel, H., Springer, M. & Kirschner, M. W. Profiling of ubiquitin-like modifications reveals features of mitotic control. *Cell* **152**, 1160-1172, doi:10.1016/j.cell.2013.02.007 (2013).
- 33 Zhang, Y. *et al.* Negative regulation of HDM2 to attenuate p53 degradation by ribosomal protein L26. *Nucleic Acids Res* **38**, 6544-6554, doi:10.1093/nar/gkq536 (2010).
- 34 Chen, J., Guo, K. & Kastan, M. B. Interactions of nucleolin and ribosomal protein L26 (RPL26) in translational control of human p53 mRNA. *J Biol Chem* **287**, 16467-16476, doi:10.1074/jbc.M112.349274 (2012).
- 35 Cai, Y. *et al.* UFBP1, a Key Component of the Ufm1 Conjugation System, Is Essential for Ufm1ylation-Mediated Regulation of Erythroid Development. *PLoS Genet* **11**, e1005643, doi:10.1371/journal.pgen.1005643 (2015).
- 36 Singh, R. *et al.* Cryo-electron microscopic structure of SecA protein bound to the 70S ribosome. *J Biol Chem* **289**, 7190-7199, doi:10.1074/jbc.M113.506634 (2014).

- 37 Nierhaus, K. H. Nobel Prize for the elucidation of ribosome structure and insight into the translation mechanism. *Angew Chem Int Ed Engl* **48**, 9225-9228, doi:10.1002/anie.200905795 (2009).
- 38 Cisterna, B. & Biggiogera, M. Ribosome biogenesis: from structure to dynamics. *Int Rev Cell Mol Biol* **284**, 67-111, doi:10.1016/S1937-6448(10)84002-X (2010).
- 39 Dunkle, J. A. & Cate, J. H. Ribosome structure and dynamics during translocation and termination. *Annu Rev Biophys* **39**, 227-244, doi:10.1146/annurev.biophys.37.032807.125954 (2010).
- 40 Wilson, D. N. & Doudna Cate, J. H. The structure and function of the eukaryotic ribosome. *Cold Spring Harb Perspect Biol* **4**, doi:10.1101/cshperspect.a011536 (2012).
- 41 Morgan, D. G., Menetret, J. F., Neuhoef, A., Rapoport, T. A. & Akey, C. W. Structure of the mammalian ribosome-channel complex at 17A resolution. *J Mol Biol* **324**, 871-886 (2002).
- 42 Nakatogawa, H. & Ito, K. The ribosomal exit tunnel functions as a discriminating gate. *Cell* **108**, 629-636 (2002).
- 43 Menetret, J. F. *et al.* Ribosome binding of a single copy of the SecY complex: implications for protein translocation. *Mol Cell* **28**, 1083-1092, doi:10.1016/j.molcel.2007.10.034 (2007).
- 44 Roux, P. P. *et al.* RAS/ERK signaling promotes site-specific ribosomal protein S6 phosphorylation via RSK and stimulates cap-dependent translation. *J Biol Chem* **282**, 14056-14064, doi:10.1074/jbc.M700906200 (2007).
- 45 Grundy, M. *et al.* Early changes in rpS6 phosphorylation and BH3 profiling predict response to chemotherapy in AML cells. *PLoS One* **13**, e0196805, doi:10.1371/journal.pone.0196805 (2018).
- 46 Khalaileh, A. *et al.* Phosphorylation of ribosomal protein S6 attenuates DNA damage and tumor suppression during development of pancreatic cancer. *Cancer Res* **73**, 1811-1820, doi:10.1158/0008-5472.CAN-12-2014 (2013).
- 47 Spence, J. *et al.* Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain. *Cell* **102**, 67-76 (2000).
- 48 Higgins, R. *et al.* The Unfolded Protein Response Triggers Site-Specific Regulatory Ubiquitylation of 40S Ribosomal Proteins. *Mol Cell* **59**, 35-49, doi:10.1016/j.molcel.2015.04.026 (2015).
- 49 Kramer, G., Boehringer, D., Ban, N. & Bukau, B. The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nat Struct Mol Biol* **16**, 589-597, doi:10.1038/nsmb.1614 (2009).
- 50 Kobayashi, K. *et al.* Structure of a prehandover mammalian ribosomal SRP.SRP receptor targeting complex. *Science* **360**, 323-327, doi:10.1126/science.aar7924 (2018).
- 51 Vellinga, J. *et al.* A system for efficient generation of adenovirus protein IX-producing helper cell lines. *J Gene Med* **8**, 147-154, doi:10.1002/jgm.844 (2006).
- 52 Chevallet, M., Luche, S. & Rabilloud, T. Silver staining of proteins in polyacrylamide gels. *Nat Protoc* **1**, 1852-1858, doi:10.1038/nprot.2006.288 (2006).

Chapter 6

General Discussion

Post-translational modifications (PTMs) are essential to regulate the localization, conformation, interaction and stability of target proteins. PTMs do not play their roles solo but interact with each other¹.

Protein SUMOylation is essential for almost all organisms and regulates many nuclear process²⁻⁴. It is a highly dynamic process that enables fast and reversible cellular signal transduction. However due to the low modification stoichiometry, highly active proteases and a large remnant after trypsin digestion, functional proteomes of SUMO targets as well as SUMO acceptor lysines are difficult to characterize^{2,5-19}. The first three chapters of this thesis focus on deciphering the roles of SUMOs in response to replication stress, how the roles of SUMO and ubiquitin signalling are balanced by the STUB1 RNF4 and how SUMOylation crosstalks with phosphorylation during replication stress using global and unbiased proteomic strategies.

UFM1 is a recently identified ubiquitin-like protein that covalently modifies lysine residues of its substrates through a three component enzymatic cascade - UFMylation. UFMylation is conserved among nearly all of the eukaryotic organisms, except yeast, and associated with several cellular activities including the endoplasmic reticulum stress response and haematopoiesis. The last chapter of this thesis describes how ribosomal protein large subunit 26 (RPL26) is modified by UFM1, and we investigated its role in the last step of translation.

1 Dynamic SUMO targets and acceptor lysines in response to replication stress.

SUMOylation was first linked to DNA damage (DNA mismatch repair) by studies on the molecular mechanisms of SUMOylation induced protein transfer²⁰. Since then, studies were focused on the roles that SUMO play in response to DNA damage and SUMOylation turned out to be one of the most important processes to modify DNA repair proteins²¹. However, the relevant SUMO target proteins to maintain genome stability during replication have remained largely elusive due to difficulties of purifying and identifying relevant SUMO target proteins. Therefore, we wanted to identify SUMO target proteins involved in counteracting replication stress.

In chapter 2, we have performed a large-scale proteomic screen to identify proteins that are regulated by SUMO-2 upon treatment of cells with Hydroxyurea at two different time points (2 hours and 24 hours). In mammalian cells, agents like hydroxyurea can stall or collapse DNA replication forks and induce homologous recombination to counteract replication stress. It was reported that most DNA replication forks restart after 2 hours of Hydroxyurea (HU) while 24 hours of HU will cause the collapse of the replication fork²². By introducing these two-time points, we could study SUMO-2 conjugates from cells exposed to replication stress for a short period (2 hours) when repair proteins are involved to restart the replication fork or for a longer period (24 hours) when double strand breaks occur and homologous recombination is induced to repair it.

We used optimized purification procedures that were developed in our lab to enrich SUMO target proteins as well as SUMO acceptor lysines^{6,23}. The improved approach employed a tag of 10 histidines to the N-terminus of SUMO-2 or SUMO-2 that lacks all its lysines. The use of the His10 tag allows a purification procedure with a relative high concentration of Imidazole to largely get rid of the binding of non-specific proteins and allows the use of denaturing buffers to inactivate proteases. Lysine-deficient version of SUMO-2 protect SUMO-2 from Lys-C digestion and enabled a second round of purification of SUMOylated peptides after Lys-C digest, which reduced the copurification of non-SUMOylated peptides and decreased the complexity of the samples.

Using this method, 566 SUMOylated proteins and 1,043 SUMOylation sites were identified in this project. Among those identified SUMOylated proteins, when cells were treated with HU for two hours, 12 dynamic SUMO-2 targets were identified (for 10 proteins SUMOylation was increased and for 2 proteins SUMOylation was decreased), while for U2OS cells that were treated with HU for 24 hours, 48 dynamic SUMO-2 targets were identified (for 35 proteins SUMOylation was increased and for 13 proteins SUMOylation was decreased).

A plethora of known SUMOylated DNA damage response factors have been identified in our screen, such as MDC1²⁴, BLM²⁵, ATRIP²⁶ and BRCA1²⁷, which can serve as a proof of the validity of our approach. There are also many DNA damage response related proteins that had no previous known relationship with SUMOylation at the time when the paper was published such as RMI1, BARD1, EME1, CHAF1A. In 2016, there was a paper on RMI1 SUMOylation in yeast, mediated by Smc5/6²⁸. This paper showed that the STR complex (Sgs1, Top3, Rmi1) can interact with SUMO and Smc5/6 through Sgs1 and these interactions contribute to the SUMOylation of Sgs1, Top3 and Rmi1. Reduced STR SUMOylation leads to recombination intermediates accumulating and impairment in cell growth. RMI1 is the second most highly upregulated SUMO target protein in our screen in response to 24 hours of HU treatment. In human cells, TOP3 α , BLM and RMI1/2 forms the BTR complex that dissolves Double-Holliday junctions. Increased SUMOylation of RMI1 could represent a mechanism to increase interaction of RMI1 with BLM and TOP3 α to disentangle double Holliday junctions and to produce non-crossover products. Further studies about SUMOylation of RMI1 in homologous recombination will be of great interest.

We also identified some components of the DNA replisome in our screen such as Proliferating Cell-Nuclear Antigen (PCNA). After two hours of HU treatment, SUMOylation of PCNA was shown to be down-regulated. PCNA is a ring type homotrimer that embraces DNA. PCNA functions as a sliding clamp and is critical for faithful DNA replication and repair. The original modification of PCNA by SUMO was identified in budding yeast on two residues, lysine127 and lysine164. In the presence of DNA damage, the interaction between PCNA and UvrD-like helicase Srs2 is strongly enhanced at both residues to prevent access of the recombination machinery²⁹⁻³². In human cells, PCNA was also SUMOylated to inhibit double strand break initiated homologous recombination and also to prevent

double strand breaks formation at stalled replication forks^{33,34}. PCNA SUMOylation was down-regulated at the early time point of replication fork stalling.

Furthermore, SUMOylation of a group of centromeric proteins was identified in our screen after 24 hours of HU treatment, including MLF1IP (CENPU), CENPH and MIS18A, which play a role in the regulation of CENPA deposition³⁵. It was reported that the subcellular localization of SUMO-2/3 was at centromeres and kinetochores while the subcellular localization of SUMO-1 was at the mitotic spindle and spindle midzone³⁶. Overexpressing SUMO protease SENP2 inhibited the accumulation of SUMO conjugates and led to a prometaphase arrest due to a defect in targeting of CENPE to the kinetochores, which indicates a key role of SUMOylation in regulating the cell cycle³⁷. The functional significance of SUMOylation to centromeric proteins could further be explored.

Finally, it is of great interest to note that SUMOylation is involved in regulating other post translational modifications such as ubiquitination, phosphorylation and demethylation, demonstrating extensive crosstalk among different modifications. These proteins include lysine-specific demethylases 5D, 5C and 4A; ubiquitin E3 ligases RAD18 and BRCA1. Regarding crosstalk, mixed SUMO-ubiquitin chains have also been found in our study as well. There are three SUMO-2 acceptor lysines in ubiquitin, which are lysine11, lysine48 and lysine63. Moreover, there are 83 peptides that have been found to be modified by both SUMO-2 and phosphorylation indicating a crosstalk between SUMOylation and phosphorylation in the DNA damage response. Also, the optimized methodology could be broadly applied to study signal transduction by other ubiquitin-like modifiers.

2 Crosstalk between SUMO and other PTMs

2.1 The ubiquitin E3 ligase RNF4 regulates protein group SUMOylation by targeting the SUMO E2 and E3s.

We identified and characterized targets for the human STUbL RNF4 by using two complementary proteomics approaches. Firstly, levels of RNF4 targeted SUMO conjugates in the cell should be increased upon RNF4 knockdown. Secondly they were covalently trapped using RNF4-TULIP constructs. They were enriched in a SIM-dependent manner after treatment of cells with proteasome inhibitor MG132. Targets meeting both requirements were identified as SUMO E3 ligase PIAS1, PIAS2, PIAS3, NSMCE2, ZNF451, Ubiquitin E3 ligase RAD18, RNF216 and other proteins such as SLFN5, IMPDH2, TOP2A, SLX4, MORC3, SMC5, ERCC4. They were regulated by RNF4 with intact SIMs and they were enriched upon proteasome inhibitor MG132 treatment. Thus, we proposed a model of RNF4 to balance SUMO-ubiquitin signalling by targeting the SUMO conjugation machinery.

Many active SUMO E3 ligases could undergo auto-SUMOylation³⁸⁻⁴⁰. Our data indicate that in human cells, RNF4 could target auto-SUMOylated SUMO E3 ligases to limit SUMO signal transduction. In budding yeast, it was reported that the yeast STUbL Slx5/Slx8 mediates regulation of the SUMO E3 ligase Siz1 and affected SUMOylation and nuclear localization of Siz1 *in vivo* to prevent

the accumulation of SUMO conjugates in the nucleus⁴¹. In fission yeast, the SUMOylated SUMO E3 ligase Pli1 was also found to be targeted by the STUbL and its cofactor Ufd1⁴². It was previously found that there are many closely located SUMO acceptor lysines on SUMO E3 ligases⁶, which might be efficiently recognized by RNF4.

The SUMO E2 Ubc9 was also identified in our RNF4-TULIP screen, co-enriched with RNF4 wild type constructs from cells treated with proteasome inhibitor MG132. It has been reported that the mammalian E2-conjugating enzyme Ubc9 could undergo autoSUMOylation at Lys14. SUMOylated Ubc9 provides an additional interface for its SIM-containing targets and this regulates target discrimination⁴³. Ubc9 standard activity was severely reduced by SUMOylation but this stimulated SUMO chain formation via noncovalent SUMO interaction. Targeting SUMOylated Ubc9 by the STUbL RNF4 could therefore be a valid way to decrease SUMO polymer formation. In the RNF4 knockdown approach, Ubc9 was also identified in our screen but statistically not picked up because its value was below the cut off value.

Yeast homologs (Rfp1 and Rfp2) of RNF4 are essential for DNA repair⁴⁴. Mammalian RNF4 knockout mice die at the early embryonic stage with higher levels of methylation in genomic DNA⁴⁵. RNF4 has been implicated in arsenic-induced PML degradation⁴⁶. RNF4 also accumulated rapidly at sites of DNA damage and co-localized with γ H2AX. RNF4 deficient MEFs showed an increased size and number of phosphorylated H2AX foci.

PIAS1, 2, and 3 identified in our screen is reminiscent to two studies in 2009 that reported that PIAS1 and PIAS4 were responsible for the SUMO accrual at DNA damage foci^{27,47}. The BRCA1 ubiquitin ligase activity was reduced in PIAS1/4-depleted cells. Upon RNF4 knockdown, we identified BRCA1 as well as its partner BARD, but they did not show up in the TULIP screen. Therefore we hypothesized that BRCA1/BARD1 could be regulated by RNF4 in an indirect manner with the SUMO E3 ligase PIAS1 as its primary target. BARD1 is a binding partner of BRCA1⁴⁸. Mutations in BRCA1 increase the risk of female breast and ovarian cancer⁴⁹. BRCA1 is a human tumor suppressor protein which helps to repair damaged DNA and acts as a ubiquitin ligase⁵⁰. BRCA1-BARD1 plays an important role in protecting cells from DNA double strand breaks and maintains genome stability via homologous recombination (HR). Interestingly, depletion of RNF4 showed similar defects in homologous recombination, indicating that RNF4 may be functionally related to BRCA1-BARD1 in regulation of homologous recombination repair pathways. Moreover, in the TULIP RNF4 screen, PIAS4 was identified upon proteasome inhibitor-MG132 treatment in a SIM-dependent manner, but its value was below the statistical cut off.

When validating RNF4 targets, we did not observe similar changes in total protein levels. This could be explained by SUMOylation as a low stoichiometry PTM. Only a small portion of target proteins are SUMOylated. However, a small fraction of SUMOylation can have a profound effect. In the ubiquitin field, substoichiometric ubiquitination was also found in a study using an antibody

recognizing the diGly remnant of ubiquitin. It showed that the ubiquitinated portions of many proteins increased in the absence of changes in total protein levels⁵¹.

Although many approaches have been used for the impartial identification of target proteins for ubiquitin E3 ligases, including yeast two hybrid assays, co-immunoprecipitation approaches, Ubait approaches⁵², identifying E3 ligase substrates in cells is still very challenging. We have developed TULIP technology to meet this challenge. By using the TULIP technology, we can sufficiently distinguish non-covalent binding partners and covalently bound substrates and the TULIP technology could be widely used to identify substrates of other UBL E3s.

2.2 Proteomics Reveals Global Regulation of Protein SUMOylation by ATM and ATR Kinases during Replication Stress.

In this project, by employing unbiased proteomics approaches, almost 1400 regulated SUMOylated acceptor lysines and 3300 phosphorylation sites were found in proteins upon DNA damaged caused by MMC and HU. This work expands our understanding of the DNA damage response and repair system, especially related to DNA replication stresses. Interestingly, by integrating SUMOylation and phosphorylation datasets for further bioinformatics analysis, our work showed that proteins modified by both SUMO and phosphor, have many sites that are regulated in the response of DNA replication stress. This poses a challenge for the functional research on a given modification site.

It is already known that protein SUMOylation plays vital roles in the ATR pathway with the evidence of the SUMOylation of the regulatory partner of ATR, that is ATRIP at lysine 234 and lysine 289. Without SUMOylation, ATRIP failed to accumulate at the DNA damage sites and failed to support ATR activation²⁶. Many other ATRIP partners such as RPA70, TopBP1 and the MRE11-RAD50-NBS1 complex display a lower binding ability to ATRIP without SUMOylation. Later on, the SUMO E3 ligase PIAS3 was shown to be indispensable for ATR phosphorylation, activation and ATR downstream checkpoint signalling⁵³. Moreover, in budding yeast, RPA coated ssDNA enabled the damage site localization of the yeast SUMO ligase Siz2 and RPA is needed for Siz2 mediated Rad53/59 SUMOylation. In this study, our results showed that ATR and ATM kinases that are known to phosphorylate downstream proteins in responses to replication stress and replication fork breakage, through transmitting DNA damage signals with the ATR-Chk1 and ATM-Chk2 kinase cascade, respectively, also regulate SUMOylation signaling in the response of DNA replication stress derived DNA damage. More interestingly, we find that under conditions of induced replication fork breakage, inhibiting ATM kinase leads to decreased SUMOylation of many target proteins, suggesting that this kinase may play an important role in the maintaining of the physiological balance of SUMOylated proteins.

With the purpose of uncovering the crosstalk between protein phosphorylation and SUMOylation in DNA damage response signaling cascade caused by DNA replication stress, our proteomics analysis identified a large set of central replication stress and DDR responders like BRCA1, BARD1, Fanconi

Anemia proteins, DNA double strand response proteins including MDC1, NBN and CtIP and TOPBP1. These proteins were co-regulated by both phosphorylation and SUMOylation upon DNA damage. SUMOylation and phosphorylation of BRCA1 was proved to be important for its function, and SUMOylation of BRCA1 was reported to increase its ubiquitin ligase activity^{27,47}. Further research will be required to analyse the co-dependency of these modifications for the functions of BRCA1 and other co-modified proteins. Especially, in this study, we found that SUMOylation of a main ATR activator, TOPBP1, in response to MMC treatment causing replication stress. TOPBP1 was identified as the most highly phosphorylated and SUMOylated protein in our dataset in response to eight hours of MMC treatment, and using pharmacological inhibitors of ATR and ATM signalling cascade. Our study showed that SUMOylation of TOPBP1 was heavily regulated by ATR during DNA replication stress and ATM upon DNA replication fork collapse. Future studies should point out whether TOPBP1 SUMOylation is required for ATR activation.

The work in this chapter represents an analysis of global protein phosphorylation and SUMOylation in response to replication stress for the first time and presents the largest database in proteomics to date of SUMOylation regulated target proteins under these conditions. With our findings, we propose that SUMOylation increases on specific targets in response to distinct DNA lesions, as shown in our research by unchanged TOPBP1 SUMOylation upon IR-induced DSBs, whereas TOPBP1 is hyper-SUMOylated after replication stress induced DNA breakage. Our data suggest that this SUMOylation response is regulated by two key kinases ATR and ATM. It is of great interest to further investigate the co-regulation of these two modifications, since the induction of DSBs induced by DNA replication stress is increasingly used in killing cancer cells⁵⁴⁻⁵⁷. On the basis of the vital role of SUMOs in the maintenance of genomic stability, further research on how SUMOylation perturbation affects global signaling networks will be of great interests. Also the methodology used in this study is applicable to study co-regulation of proteins by other PTMs.

3 Ubiquitin like modifiers (UFM1)

Recently, post-translational protein modification (PTM) with Ufm-1 has gained increasing attention, but its biological function as well as its substrates have remained largely unknown. Although Ufmylation has been connected to biological processes including ER homeostasis, vesicle trafficking, blood progenitor development and differentiation, (G-coupled protein receptor) GPCR maturation, transcriptional control, mitosis and more recently autophagy^{58,59}, the underlying mechanisms and biological consequences remain obscure.

Previously, uS3, uS10, uL16, subunits of the ribosome have been identified by mass spectrometry to be novel UFM-1 targets⁶⁰ and more recently, 494 UFMylated proteins also were identified by Pirone et al, of which 82 passed the cut off⁶⁴. However, none of them have been independently validated and no biological functional studies have been conducted to elucidate the role of protein UFMylation⁶¹.

In this project, we employed a site-specific strategy, which was developed by our lab to enrich SUMO target proteins as well as SUMO acceptor lysines, to identify novel Ufm-1 acceptor lysines. We have identified RPL26 (lysine 132, 134, 136, 142), RPL26L1 (lysine 132, 134, 136), TUBA1B (lysine 370), TUBA4A (lysine 355), MCM5 (lysine582), SLC26A7 (lysine120), SCYL2 (lysine 73, 76) and WDR63 (lysine564) as novel Ufm-1 targets. We further confirmed RPL26 as a Ufm-1 modified protein by western blotting.

Ribosomal proteins were reported to be post-translationally modified in many studies. The phosphorylation of rpS6 has attracted considerable attention in numerous labs since its discovery in 1974, RpS6 phosphorylation has no obvious role in protein synthesis or other cellular functions in yeast, however in MEFs, protein synthesis was found to be downregulated by rpS6 phosphorylation^{62,63}.

The early events in the life of newly synthesized proteins in the cellular environment are remarkably complex⁶⁴. In order to address the biological function of RPL26 and identify its associated factors, we immunoprecipitated ribosomes with UFMylated or non-UFMylated RPL26 using a label free quantitative proteomics approach. We identified the signal recognition particle alpha (SRPR), to be significantly enriched in UFSP2 depleted cells, suggesting that this protein might be preferentially binding UFMylated RPL26. As expected, immunoblot analysis revealed that SRPR co-immunoprecipitated more efficiently with UFM-1 modified RPL26 than with the non-modified RPL26. These results suggest that in the presence of UFM-1 modification, RPL26 interacted with SRPR and seems to mediate direct contact of the exit tunnel of the ribosome with the SRPR during the last step of directing peptides to the Sec61 translocon. Our findings might help to understand how UFM1 regulates protein synthesis by coordinating its target. However, to fully understand the role of UFMylated RPL26 in regulating translation, more studies like ribosome profiling are needed to be performed.

Overall, the results of this thesis showed that SUMOs play a pivotal role in maintaining genome stability. SUMOs do not play their roles alone but in coordination with other PTMs such as ubiquitination and phosphorylation. By identifying substrates of UFM1 as well as their acceptor lysines, we also proved that the site-specific strategy developed in our lab can be applied in identifying substrates of other ubiquitin like proteins. This will help us to better understand the roles of ubiquitin like modifiers in eukaryotic life.

References

- 1 Uy, R. & Wold, F. Posttranslational covalent modification of proteins. *Science* **198**, 890-896 (1977).
- 2 Kuehn, M. R. Mouse Ubc9 knockout: many path(way)s to ruin. *Dev Cell* **9**, 727-728, doi:10.1016/j.devcel.2005.11.008 (2005).
- 3 Constanzo, J. D. *et al.* Pias1 is essential for erythroid and vascular development in the mouse embryo. *Dev Biol* **415**, 98-110, doi:10.1016/j.ydbio.2016.04.013 (2016).
- 4 Saracco, S. A., Miller, M. J., Kurepa, J. & Vierstra, R. D. Genetic analysis of SUMOylation in Arabidopsis: conjugation of SUMO1 and SUMO2 to nuclear proteins is essential. *Plant Physiol* **145**, 119-134, doi:10.1104/pp.107.102285 (2007).

- 5 Schimmel, J. *et al.* Uncovering SUMOylation dynamics during cell-cycle progression reveals FoxM1 as a key mitotic SUMO target protein. *Mol Cell* **53**, 1053-1066, doi:10.1016/j.molcel.2014.02.001 (2014).
- 6 Hendriks, I. A. *et al.* Uncovering global SUMOylation signaling networks in a site-specific manner. *Nat Struct Mol Biol* **21**, 927-936, doi:10.1038/nsmb.2890 (2014).
- 7 Hendriks, I. A., D'Souza, R. C., Chang, J. G., Mann, M. & Vertegaal, A. C. System-wide identification of wild-type SUMO-2 conjugation sites. *Nat Commun* **6**, 7289, doi:10.1038/ncomms8289 (2015).
- 8 Yang, W. & Paschen, W. SUMO proteomics to decipher the SUMO-modified proteome regulated by various diseases. *Proteomics* **15**, 1181-1191, doi:10.1002/pmic.201400298 (2015).
- 9 Filosa, G., Barabino, S. M. & Bachi, A. Proteomics strategies to identify SUMO targets and acceptor sites: a survey of RNA-binding proteins SUMOylation. *Neuromolecular Med* **15**, 661-676, doi:10.1007/s12017-013-8256-8 (2013).
- 10 Matafora, V., D'Amato, A., Mori, S., Blasi, F. & Bachi, A. Proteomics analysis of nucleolar SUMO-1 target proteins upon proteasome inhibition. *Mol Cell Proteomics* **8**, 2243-2255, doi:10.1074/mcp.M900079-MCP200 (2009).
- 11 Vertegaal, A. C. *et al.* A proteomic study of SUMO-2 target proteins. *J Biol Chem* **279**, 33791-33798, doi:10.1074/jbc.M404201200 (2004).
- 12 Tammsalu, T. *et al.* Proteome-wide identification of SUMO2 modification sites. *Sci Signal* **7**, rs2, doi:10.1126/scisignal.2005146 (2014).
- 13 Tammsalu, T. *et al.* Proteome-wide identification of SUMO modification sites by mass spectrometry. *Nat Protoc* **10**, 1374-1388, doi:10.1038/nprot.2015.095 (2015).
- 14 Bursomanno, S. *et al.* Proteome-wide analysis of SUMO2 targets in response to pathological DNA replication stress in human cells. *DNA Repair (Amst)* **25**, 84-96, doi:10.1016/j.dnarep.2014.10.011 (2015).
- 15 Galisson, F. *et al.* A novel proteomics approach to identify SUMOylated proteins and their modification sites in human cells. *Mol Cell Proteomics* **10**, M110 004796, doi:10.1074/mcp.M110.004796 (2011).
- 16 Matic, I. *et al.* In vivo identification of human small ubiquitin-like modifier polymerization sites by high accuracy mass spectrometry and an in vitro to in vivo strategy. *Mol Cell Proteomics* **7**, 132-144, doi:10.1074/mcp.M700173-MCP200 (2008).
- 17 Hendriks, I. A. & Vertegaal, A. C. A high-yield double-purification proteomics strategy for the identification of SUMO sites. *Nat Protoc* **11**, 1630-1649, doi:10.1038/nprot.2016.082 (2016).
- 18 Matic, I. & Hay, R. T. Detection and quantitation of SUMO chains by mass spectrometry. *Methods Mol Biol* **832**, 239-247, doi:10.1007/978-1-61779-474-2_17 (2012).
- 19 Hendriks, I. A. & Vertegaal, A. C. A comprehensive compilation of SUMO proteomics. *Nat Rev Mol Cell Biol* **17**, 581-595, doi:10.1038/nrm.2016.81 (2016).
- 20 Baba, D. *et al.* Crystal structure of thymine DNA glycosylase conjugated to SUMO-1. *Nature* **435**, 979-982, doi:10.1038/nature03634 (2005).
- 21 Jackson, S. P. & Durocher, D. Regulation of DNA damage responses by ubiquitin and SUMO. *Mol Cell* **49**, 795-807, doi:10.1016/j.molcel.2013.01.017 (2013).
- 22 Petermann, E., Orta, M. L., Issaeva, N., Schultz, N. & Helleday, T. Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair. *Mol Cell* **37**, 492-502, doi:10.1016/j.molcel.2010.01.021 (2010).
- 23 Matic, I. *et al.* Site-specific identification of SUMO-2 targets in cells reveals an inverted SUMOylation motif and a hydrophobic cluster SUMOylation motif. *Mol Cell* **39**, 641-652, doi:10.1016/j.molcel.2010.07.026 (2010).
- 24 Luo, K., Zhang, H., Wang, L., Yuan, J. & Lou, Z. Sumoylation of MDC1 is important for proper DNA damage response. *EMBO J* **31**, 3008-3019, doi:10.1038/emboj.2012.158 (2012).
- 25 Ouyang, K. J., Yagle, M. K., Matunis, M. J. & Ellis, N. A. BLM SUMOylation regulates ssDNA accumulation at stalled replication forks. *Front Genet* **4**, 167, doi:10.3389/fgene.2013.00167 (2013).
- 26 Wu, C. S. *et al.* SUMOylation of ATRIP potentiates DNA damage signaling by boosting multiple protein interactions in the ATR pathway. *Genes Dev* **28**, 1472-1484, doi:10.1101/gad.238535.114 (2014).

- 27 Morris, J. R. *et al.* The SUMO modification pathway is involved in the BRCA1 response to genotoxic stress. *Nature* **462**, 886-890, doi:10.1038/nature08593 (2009).
- 28 Bonner, J. N. *et al.* Smc5/6 Mediated Sumoylation of the Sgs1-Top3-Rmi1 Complex Promotes Removal of Recombination Intermediates. *Cell Rep* **16**, 368-378, doi:10.1016/j.celrep.2016.06.015 (2016).
- 29 Pfander, B., Moldovan, G. L., Sacher, M., Hoege, C. & Jentsch, S. SUMO-modified PCNA recruits Srs2 to prevent recombination during S phase. *Nature* **436**, 428-433, doi:10.1038/nature03665 (2005).
- 30 Burkovics, P. *et al.* Srs2 mediates PCNA-SUMO-dependent inhibition of DNA repair synthesis. *EMBO J* **32**, 742-755, doi:10.1038/emboj.2013.9 (2013).
- 31 Armstrong, A. A., Mohideen, F. & Lima, C. D. Recognition of SUMO-modified PCNA requires tandem receptor motifs in Srs2. *Nature* **483**, 59-63, doi:10.1038/nature10883 (2012).
- 32 Papouli, E. *et al.* Crosstalk between SUMO and ubiquitin on PCNA is mediated by recruitment of the helicase Srs2p. *Mol Cell* **19**, 123-133, doi:10.1016/j.molcel.2005.06.001 (2005).
- 33 Gali, H. *et al.* Role of SUMO modification of human PCNA at stalled replication fork. *Nucleic Acids Res* **40**, 6049-6059, doi:10.1093/nar/gks256 (2012).
- 34 Moldovan, G. L. *et al.* Inhibition of homologous recombination by the PCNA-interacting protein PARI. *Mol Cell* **45**, 75-86, doi:10.1016/j.molcel.2011.11.010 (2012).
- 35 Fujita, Y. *et al.* Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1. *Dev Cell* **12**, 17-30, doi:10.1016/j.devcel.2006.11.002 (2007).
- 36 Brown, P. W., Hwang, K., Schlegel, P. N. & Morris, P. L. Small ubiquitin-related modifier (SUMO)-1, SUMO-2/3 and SUMOylation are involved with centromeric heterochromatin of chromosomes 9 and 1 and proteins of the synaptonemal complex during meiosis in men. *Hum Reprod* **23**, 2850-2857, doi:10.1093/humrep/den300 (2008).
- 37 Zhang, X. D. *et al.* SUMO-2/3 modification and binding regulate the association of CENP-E with kinetochores and progression through mitosis. *Mol Cell* **29**, 729-741, doi:10.1016/j.molcel.2008.01.013 (2008).
- 38 Ritterhoff, T. *et al.* The RanBP2/RanGAP1*SUMO1/Ubc9 SUMO E3 ligase is a disassembly machine for Crm1-dependent nuclear export complexes. *Nat Commun* **7**, 11482, doi:10.1038/ncomms11482 (2016).
- 39 Goeke, C. B., Yu, H. & Kang, J. Systematic identification and analysis of mammalian small ubiquitin-like modifier substrates. *J Biol Chem* **280**, 5004-5012, doi:10.1074/jbc.M411718200 (2005).
- 40 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 U S A* **99**, 2872-2877, doi:10.1073/pnas.052559499 (2002).
- 41 Westerbeck, J. W. *et al.* A SUMO-targeted ubiquitin ligase is involved in the degradation of the nuclear pool of the SUMO E3 ligase Siz1. *Mol Biol Cell* **25**, 1-16, doi:10.1091/mbc.E13-05-0291 (2014).
- 42 Nie, M. & Boddy, M. N. Pli1(PIAS1) SUMO ligase protected by the nuclear pore-associated SUMO protease Ulp1SEN1/2. *J Biol Chem* **290**, 22678-22685, doi:10.1074/jbc.M115.673038 (2015).
- 43 Knipscheer, P. *et al.* Ubc9 sumoylation regulates SUMO target discrimination. *Mol Cell* **31**, 371-382, doi:10.1016/j.molcel.2008.05.022 (2008).
- 44 Kosoy, A., Calonge, T. M., Outwin, E. A. & O'Connell, M. J. Fission yeast Rnf4 homologs are required for DNA repair. *J Biol Chem* **282**, 20388-20394, doi:10.1074/jbc.M702652200 (2007).
- 45 Hu, X. V. *et al.* Identification of RING finger protein 4 (RNF4) as a modulator of DNA demethylation through a functional genomics screen. *Proc Natl Acad Sci U S A* **107**, 15087-15092, doi:10.1073/pnas.1009025107 (2010).
- 46 Lallemand-Breitenbach, V. *et al.* Arsenic degrades PML or PML-RARalpha through a SUMO-triggered RNF4/ubiquitin-mediated pathway. *Nat Cell Biol* **10**, 547-555, doi:10.1038/ncb1717 (2008).
- 47 Galanty, Y. *et al.* Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks. *Nature* **462**, 935-939, doi:10.1038/nature08657 (2009).

- 48 Wu, L. C. *et al.* Identification of a RING protein that can interact in vivo with the BRCA1 gene product. *Nat Genet* **14**, 430-440, doi:10.1038/ng1296-430 (1996).
- 49 Miki, Y. *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* **266**, 66-71 (1994).
- 50 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).
- 51 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).
- 52 O'Connor, H. F. *et al.* Ubiquitin-Activated Interaction Traps (UBAITS) identify E3 ligase binding partners. *EMBO Rep* **16**, 1699-1712, doi:10.15252/embr.201540620 (2015).
- 53 Wu, C. S. & Zou, L. The SUMO (Small Ubiquitin-like Modifier) Ligase PIAS3 Primes ATR for Checkpoint Activation. *J Biol Chem* **291**, 279-290, doi:10.1074/jbc.M115.691170 (2016).
- 54 Benada, J. & Macurek, L. Targeting the Checkpoint to Kill Cancer Cells. *Biomolecules* **5**, 1912-1937, doi:10.3390/biom5031912 (2015).
- 55 Rundle, S., Bradbury, A., Drew, Y. & Curtin, N. J. Targeting the ATR-CHK1 Axis in Cancer Therapy. *Cancers (Basel)* **9**, doi:10.3390/cancers9050041 (2017).
- 56 Manic, G., Obrist, F., Sistigu, A. & Vitale, I. Trial Watch: Targeting ATM-CHK2 and ATR-CHK1 pathways for anticancer therapy. *Mol Cell Oncol* **2**, e1012976, doi:10.1080/23723556.2015.1012976 (2015).
- 57 Zhang, Y. *et al.* Targeting radioresistant breast cancer cells by single agent CHK1 inhibitor via enhancing replication stress. *Oncotarget* **7**, 34688-34702, doi:10.18632/oncotarget.9156 (2016).
- 58 Joassard, O. R. *et al.* Regulation of Akt-mTOR, ubiquitin-proteasome and autophagy-lysosome pathways in response to formoterol administration in rat skeletal muscle. *Int J Biochem Cell Biol* **45**, 2444-2455, doi:10.1016/j.biocel.2013.07.019 (2013).
- 59 Zhang, Y., Zhang, M., Wu, J., Lei, G. & Li, H. Transcriptional regulation of the Ufm1 conjugation system in response to disturbance of the endoplasmic reticulum homeostasis and inhibition of vesicle trafficking. *PLoS One* **7**, e48587, doi:10.1371/journal.pone.0048587 (2012).
- 60 Simsek, D. *et al.* The Mammalian Ribo-interactome Reveals Ribosome Functional Diversity and Heterogeneity. *Cell* **169**, 1051-1065 e1018, doi:10.1016/j.cell.2017.05.022 (2017).
- 61 Pirone, L. *et al.* A comprehensive platform for the analysis of ubiquitin-like protein modifications using in vivo biotinylation. *Sci Rep* **7**, 40756, doi: 10.1038/srep40756 (2017).
- 62 Grundy, M. *et al.* Early changes in rpS6 phosphorylation and BH3 profiling predict response to chemotherapy in AML cells. *PLoS One* **13**, e0196805, doi:10.1371/journal.pone.0196805 (2018).
- 63 Roux, P. P. *et al.* RAS/ERK signaling promotes site-specific ribosomal protein S6 phosphorylation via RSK and stimulates cap-dependent translation. *J Biol Chem* **282**, 14056-14064, doi:10.1074/jbc.M700906200 (2007).
- 64 Kramer, G., Boehringer, D., Ban, N. & Bukau, B. The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nat Struct Mol Biol* **16**, 589-597, doi:10.1038/nsmb.1614 (2009).

Summary

Genomic instability serves as both a distinguishing character of tumor cells and a primary contributing factor to carcinogenesis. To ensure genome stability, eukaryotic cells harbor a global signal transduction cascade, termed the DNA damage response (DDR). Modification of chromatin and DDR machinery components by rapid and reversible posttranslational protein modifications (PTMs) including by Small Ubiquitin-like Modifiers (SUMOs) plays important roles in orchestrating cellular responses to DNA damage. Our lab has developed a generic lysine-deficient SUMO purification strategy combined with mass spectrometry for the identification of these SUMO target proteins and their acceptor lysines. My primary goal was to employ this quantitative proteomics approach to perform unbiased, proteome-wide and site-specific analyses of SUMOylation to identify individual target proteins that are regulated by SUMOs during DNA damage responses. I reviewed the SUMOylation process, roles of SUMO in DNA damage responses, cross talk among SUMO and ubiquitin, SUMO and phosphorylation, and technical developments in SUMO proteomics (**Chapter 1**).

SUMOs are known to play a key role in counteracting replication stress caused by genotoxic agents. In order to identify individual target proteins that are regulated by SUMOs in response to DNA replication stalling induced by DNA damage, we identified SUMO-2 targets as well as SUMO-2 modified sites from cells exposed to DNA damage due to DNA replication stress caused by Hydroxyurea (HU), using the lysine-deficient SUMO-2 purification approach coupled to mass spectrometry as developed in our lab. We found sets of SUMO-2 targets up-regulated or down-regulated as well as SUMO-2 modified acceptor lysines in response to short (two hours) or long (twenty-four hours) periods of HU treatment. Our results show that SUMO-2 target proteins are functionally connected as groups to ensure genome stability during replication stress. We also found that SUMOylation regulates sets of other factors related to DNA damage response (**Chapter 2**).

DNA replication is vital for eukaryotic organisms, and it is constantly threatened by endogenous and exogenous damages. To ensure efficient responses to genotoxic insults and maintain genomic integrity, different PTMs precisely coordinate with each other. Phosphorylation is widely known to be a key player in response to DNA damage caused by replication stress. In later years, SUMO has also been shown to play an essential role in the DNA damage response. In order to investigate cross-talk between phosphorylation and SUMOylation in response to DNA replication stress-induced DNA damage, we have employed complementary proteomics strategies and identified 3300 phosphorylation sites and 1400 SUMOylation acceptor lysines upon mitomycin C (MMC) and hydroxyurea (HU) treatment and found that a set of proteins are co-regulated by both SUMOylation and phosphorylation. We further proved that TOPBP1, a major partner and co-activator of ATR, is SUMOylated upon MMC

treatment-induced replication stress. This indicates that SUMOylated TOPBP1 together with ATR and ATRIP, may play a key role in the ATR dependent checkpoint signal transduction (**Chapter 3**).

The SUMO-Targeted Ubiquitin Ligase (STUBLs) RNF4 is known to ubiquitylate SUMOylated proteins upon DNA damage. In order to gain a global view of RNF4 regulated proteins and better understand the function of RNF4 in DNA damage repair, we have optimized methods to identify direct substrates and indirect interactors of ubiquitin E3 ligases. We named this method: TULIP (Targets for Ubiquitin Ligases Identified by Proteomics).

Our study showed that the SUMO E2 Ubc9 and five SUMO E3 ligases (ZNF451, NSMCE2, PIAS1, PIAS2 and PIAS3) are direct RNF4 substrates. This indicates a role for RNF4 in balancing SUMO signal transduction. Moreover, we identified the tumor suppressor ubiquitin E3 ligase BARD1, the partner of BRCA1, as an indirect RNF4 substrate. We propose that the ubiquitin E3 ligase RNF4 targets autoSUMOylated Ubc9 and SUMO E3 ligases to balance SUMO signal transduction. More importantly, the TULIP technology we developed is able to distinguish covalently bound substrates and non-covalent binding interactors and it is applicable for identifying targets for other E3 ligases (**Chapter 4**).

UFM1 is a recently identified ubiquitin-like protein that covalently modifies lysine residues of its substrates through a three component enzymatic cascade. We applied our lysine-deficient Ubiquitin-like (Ubl) purification strategy coupled to mass spectrometry as developed in our lab to study UFMylation in a site-specific manner in HeLa cells. We showed that RPL26, RPL26L1, TUBA1B, MCM5, SLC26A7, SCYL2, WDR63 are covalently modified by UFM1. We further confirmed RPL26 as a key UFM1 target and the UFMylated form of RPL26 can efficiently interact with SRPR. We propose that RPL26 is targeted by UFM1 and serves as a platform for ribosome associated protein interaction (**Chapter 5**).

In the last chapter, I discussed results obtained in this thesis. The SUMO-2 site-specific proteomics method is applicable to identify targets for other Ubls such as UFM1 as well as their acceptor lysines, which provides important resources to decipher the complex protein networks regulated by these PTMs. However, to fully understand the role of PTMs in the cells, more functional work is needed. It is also of great interest to note that SUMOylation is involved in coordinating with other post translational modifications such as ubiquitination and phosphorylation. The TULIP method developed in our lab is not only helpful for us to better understand the role of the STUBL RNF4 in regulating the SUMO conjugation machinery to maintain genomic stability but also provides a strong tool for identification of targets for E3 ligases in other regulatory pathways (**Chapter 6**).

Samenvatting

Genomische instabiliteit is een onderscheidende karaktereigenschap van tumorcellen en draagt in belangrijke mate bij aan oncogeniciteit. Eukaryoten zijn uitgerust met een cascade van eiwitten die beschadigd DNA repareren via de DNA schade respons. Deze eiwitten worden gereguleerd door dynamische en reversibele post-translationele modificaties (PTMs) zoals SUMOs. Onze onderzoeksgroep heeft een methode ontwikkeld om SUMO signalering te kunnen bestuderen door een nieuwe methode om SUMOs op te zuiveren te koppelen aan massa spectrometrie. Het doel van mijn onderzoek was om veranderingen in SUMO signalering door DNA schade in kaart te brengen.

In **hoofdstuk 1** heb ik de huidige kennis van SUMO signalering beschreven, de rol van SUMO in de DNA schade respons, crosstalk beschreven tussen SUMO en twee andere PTMs, ubiquitine en fosforylering, en ik heb een overzicht gemaakt van de methodologie om SUMO signalering te bestuderen.

Cellen kopiëren DNA tijdens celdeling; dit proces heet replicatie. SUMO co-reguleert een set eiwitten die verantwoordelijk zijn voor de reparatie van beschadigd DNA tijdens replicatie-stress. In **hoofdstuk 2** heb ik deze rol van SUMO tijdens replicatie-stress beschreven.

Tijdens replicatie-stress werkt SUMO samen met fosforylatie. Crosstalk tussen SUMOylering en fosforylatie heb ik beschreven in **hoofdstuk 3**. We hebben ontdekt dat het eiwit TOPBP1, een activator van het kinase ATR, een belangrijk doelwit is van SUMO signalering tijdens replicatie stress.

Crosstalk tussen SUMO en ubiquitine is beschreven in **hoofdstuk 4**. Het ubiquitine ligase RNF4 koppelt ubiquitine aan eiwitten die eerst gemodificeerd zijn door SUMO. RNF4 is belangrijk voor de DNA schade respons. In dit project hebben we de identiteit achterhaalt van de doeleiwitten van RNF4, waaronder de SUMO E2 Ubc9 en vijf SUMO E3 ligasen, ZNF451, NSMCE2, PIAS1, PIAS2 en PIAS3. Door het ubiquitineren van de actieve poule van deze enzymen remt RNF4 SUMO signalering efficiënt. Indirect wordt de tumor-suppressor BARD1, een partner van de tumor-suppressor BRCA1 door RNF4 gereguleerd.

De proteomics methodologie die we hebben ontwikkeld bleek ook geschikt voor het onderzoeken van een ander ubiquitine familielid, UFM1. Dit onderzoek is beschreven in **hoofdstuk 5**. Hierbij hebben we ontdekt dat het ribosomale eiwit RPL26 door UFM1 wordt gemodificeerd. Dit speelt een rol bij de koppeling van ribosomen aan het endoplasmatisch reticulum. Tenslotte heb ik in **hoofdstuk 6** deze nieuwe resultaten bediscussieerd.

List of Publications

1. Katharina F. Witting[#], ZY Xiao[#], Driss el Atmioui, Román González-Prieto, Fabricio Loayza-Puch, Reuven Agami, Alfred C.O. Vertegaal* and Huib Ovaa*. (2018) Ufmylation—an emerging post-translational modification regulating protein synthesis. *In submission*.
2. S Munk[#], JO Sigurðsson[#], ZY Xiao[#], TS Batth, G Franciosa, L Stechow, AJ Lopez-Contreras, ACO Vertegaal* and JV Olsen*. (2017) Proteomics reveals global regulation of protein SUMOylation by ATM and ATR kinases in replication stress. *Cell Reports*. 10;21(2):546-558.
3. R Kumar[#], R González-Prieto[#], ZY Xiao[#], M Verlaan-de Vries and ACO Vertegaal*. (2017) The STUbL RNF4 regulates protein group SUMOylation by targeting the SUMO conjugation machinery. *Nature Communications*. 27;8(1):1809.
4. ZY Xiao, JG Chang, IA Hendriks, JO Sigurðsson, JV Olsen* and ACO Vertegaal*. (2015) 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. *Mol Cell Proteomics*. 14(5):1419-34.
5. XJ Hou[#], W Zhang[#], ZY Xiao, HY Gan, Chunsheng Han*. (2012) Mining and characterization of ubiquitin E3 ligases expressed in the mouse testis. *BMC Genomics*. 13:495.

[#], These authors contributed equally to the work

^{*}, Corresponding author

Acknowledgements

With sincere gratitude, I look back on the opportunity to study under the supervision of Alfred. C. O. Vertegaal, all the memories of past five years emerged in my mind. I am extremely grateful to Alfred for his continuous guidance and support of my PhD thesis. Thank you for providing me the great opportunity to perform my PhD program under your supervision in the LUMC. Your scientific views, attitudes and enthusiasm have encouraged me all the time. I can still remember our first meet in Shanghai China.

I would like to thank Peter ten Dijke for his help in professional knowledge and support. Thank you for helping me overcome my hard times. Expressing my gratitude with only words is not enough.

I would also like to express my thanks to A. G. Jochemsen and David Baker for their kindly support of this thesis. I am very grateful for your kind help and suggestions regarding my thesis.

I would like to give my gratitude to the SUMO lab members, Bing Yang, Karo, Joost, Ivo, Joel, Matty, Román, Aleksandra, Frauke, Sabine, Edwin, Ekatherina, Sumit, Tamara and Jessie. I was so lucky to share so much good time with you.

I would also like to express my appreciation to my collaborators: Stephanie, Jón Otti and Jesper from University of Copenhagen, Novo Nordisk Foundation Center for Protein Research and Katharina and Huib from LUMC for interesting collaborations. Without you, the studies could not be performed and we could not have such nice publications. Thank you all.

Furthermore, I would like to express my thanks to my Chinese friends in the Netherlands. Yun Yang, Yudan Tan, Chun Wei, Mengmen Sun, Min He, Jian Yang, Yanming Guo, Jingxian Wang, Guangchao Chen, Guangshen Du, Yingguang Li and Yuchuan Qiao. We spent so many unforgettable moments together. It was my pleasure to share our experiences together. Thank you for your accompanying me in these years. Thank you for lighting my life abroad.

I would like to thank all my teachers and supervisors China. Chunsheng Han, Qi Zhou, Hongmei Wang. In spite of the distance, your warm encouragements are always with me.

I would like to thank my family members. Thank you so much for your kind concern. My parents, Youhong Xiao and Junying Pan, it was a pity that I could not be in your company for such a long time. Therefore, I could not fully feel your happiness and worries when you needed me. But I hope that you are proud of your daughter, and your unconditional love will inspire me forever. Yaojin Peng, my dearest husband, thanks for your all efforts to give us a happy and satisfactory life in these years. Our wonderful experiences and memories are my precious treasures, and will always give me power and courage for every challenge. Let us go on chasing our dream together.

Finally, to our loving daughters Heya Peng (Emilie) and Hena Peng (Hannah).

Curriculum Vitae

Zhenyu Xiao was born on December 24th, 1987 in Ningxiang, Hunan Province, P.R. China. In 2005, after finishing her high school education at Ningxiang No 1 High School, she started to study at University of Science and Technology, Beijing, majoring in biotechnology. She obtained her Bachelor's degree in 2009. Her GPA at University of Science and Technology, Beijing was very high and she was admitted as a master student of Developmental biology by Institute of Zoology, Chinese Academy of Sciences in 2009 without Graduate Entrance Examination. She received her Master's degree in 2013.

In 2013, she came to the Netherlands to start her PhD career under supervision of Prof. dr. A.C.O. Vertegaal, at the Molecular Cell Biology Department of Leiden University Medical Center. During her PhD training, the author had the chance to join the EU ITN Upstream, and presented her work in Europe, and established international collaborations with other institutes. The author will go back to China where she will continue her clinical and scientific research.

