Safeguarding genome integrity with small ubiquitin-like modifiers

Claessens, L.A.

Chapter 1

General introduction

Complexity of the human proteome

The human genome encodes around 20,000 genes. According to the central dogma of molecular biology, these genes are first transcribed into RNA and then translated into proteins. The entire set of proteins that is expressed in cells at a given time is termed 'the proteome'. Already on the RNA level, additional diversity is introduced through processes like alternative splicing and RNA editing. Then, once a protein is being translated, it can become co- or post-translationally modified to further increase the complexity and functionality of the proteome. These modifications are collectively termed posttranslational modifications (PTMs). PTMs can involve the covalent attachment of a chemical group (e.g., phosphorylation, acetylation and methylation) or of a small protein. Ubiquitin was the first discovered protein to covalently modify proteins and since then many more structurally related ubiquitin-like proteins have been identified [1]. PTMs are essential for the spatiotemporal regulation of proteins, meaning that proteins are only expressed and/or functional under given conditions, at given times, at given locations. Dysregulation of protein function has major implications for a plethora of cellular processes and is linked to many diseases, including cancer.

SUMO conjugation and deconjugation

Small ubiquitin like modifier (SUMO) proteins are members of the ubiquitin-like protein family and is a PTM that is critical in maintaining nuclear integrity [2]. SUMOs are ~11 kDa in size, slightly larger than ubiquitin. It is essential for many cellular processes, including gene expression, cell cycle progression and genome stability. SUMOs are covalently attached to lysine residues via an enzymatic cascade, involving a SUMO E1 activating enzyme (SAE1/UBA2), a SUMO E2 conjugating enzyme (UBC9) and SUMO E3 ligating enzymes. Mammalian cells encode four SUMO isoforms, SUMO1-4. SUMO2 and SUMO3 are nearly identical and often referred to as SUMO2/3, and share ~50% sequence identity with SUMO1. SUMOs are translated as inactive precursor proteins and are processed into mature SUMOs by SUMO proteases, exposing a C-terminal diGly motif for conjugation. The SUMO isoform SUMO4 does not seem to be processed by SUMO proteases and is thought to be non-conjugatable [3-5]. SUMO proteases also ensure timely removal of SUMOs from substrates [6]. Consequently, SUMOylation is a highly dynamic and reversible modification, allowing cells to adapt swiftly to changing conditions.

SUMOylation can result in the modification of a single lysine residue (monoSUMOylation) or multiple lysine residues (multi-monoSUMOylation). SUMOylation can also result in the formation of polymeric SUMO chains through internal lysine residues in SUMOs [7, 8].

SUMOylation can affect protein stability, activity, localization and noncovalent protein-protein interactions through SUMO-interaction motifs (SIMs) [9-11]. SUMOylation often targets multiple physically or functionally related proteins within a cellular pathway in a coordinated manner. Psakhye and Jentsch introduced this concept of 'protein group modification' a decade ago in the yeast DNA damage response and it is still the leading explanation in the field for the observed redundancy of single SUMOylation events [12, 13]. To date, thousands of different SUMO target proteins have been identified using mass spectrometry approaches [14]. However, for the majority of these, the physiological relevance and functional role of their SUMOylation still remains unclear.

Noncovalent SUMO interactions

The interaction between a SIM and SUMO is the most predominant and best characterized mode of noncovalent SUMO interaction. This is termed the class I SUMO interaction and involves interaction between a SIM and a conserved hydrophobic groove on SUMO, the SIM-binding groove. SIMs are short hydrophobic stretches in proteins consisting of isoleucine, leucine or valine residues, flanked by acidic residues [15]. To date, two other classes of SUMO interaction have been identified. The class II SUMO interaction involves a surface on SUMO opposite of the SIM-binding groove and class III SUMO interaction involves binding to a distinct region of SUMO1 via a ZZ zinc finger motif [16-21]. It is unclear whether there are yet unidentified modes of SUMO interaction, involving uncharacterized SUMO binding domains. In the last years, we and others have put increasing efforts into shedding more light on the SUMO interactome [22-27].

SUMO-targeted ubiquitin ligases

Extensive crosstalk exists between SUMOylation and ubiquitination. SUMO targeted ubiquitin ligases (STUbLs) are a class of ubiquitin E3 ligases that consist of one or multiple SIMs and recognize multi- or poly-SUMOylated substrates [28]. The paradigmatic signalling function of the STUbL pathway is to ubiquitinate SUMOylated proteins and target them for degradation by the proteasome, although non-proteolytic signalling functions also exist. In mammalian cells, two STUbLs have been characterized, RNF4 and RNF111/Arkadia [29, 30]. RNF4 is the most predominant and best characterized STUbL. RNF4 binding to polymeric SUMO2/3 chains on a substrate induces its dimerization and activation, which is followed by substrate ubiquitination as well as RNF4 autoubiquitination and degradation, the latter providing a negative feedback loop [31].

SUMO and the DNA damage response

Maintaining genomic integrity is critical for normal cellular function. Errors in DNA replication, repair of damaged DNA, chromosome segregation and cell cycle progression can lead to abnormal cellular behaviour and disease, including uncontrolled cell proliferation in cancer. Cells employ an integrated network of signalling pathways to sense and repair DNA damage induced by both endogenous and exogenous sources (e.g., ionizing radiation), collectively termed the DNA damage response. DNA double-strand breaks occur when both DNA strands are broken and are highly toxic lesions when unresolved. Cells have two major pathways for repairing double-strand breaks: non-homologous end joining (NHEJ) and homologous recombination (HR). HR is restricted to G2 and S phases of the cell cycle as it relies on a homologous DNA sequence for repair, leading to accurate repair. In NHEJ, the broken ends are modified and ligated together without the use of a homologous template, generating small deletions or inserts. Thus, in contrast to HR, NHEJ is error-prone, but can occur throughout the cell cycle.

Posttranslational modifications, including SUMOylation, are important regulators of the DNA damage response by facilitating recruitment, interaction and retention of the DNA repair machinery. Among the thousands of identified SUMO substrates by mass spectrometry, many are involved in DNA replication and repair [14]. Indeed, multiple DNA damage response proteins become SUMOylated or show increased SUMOylation levels in response to genotoxic stress [32]. Moreover, SUMO1, SUMO2/3 and components of the SUMO machinery (UBC9, PIAS1 and PIAS4) localize to sites of DNA damage to SUMOylate substrates involved in the repair response [33, 34]. The STUbL RNF4 is also recruited to sites of DNA damage and ubiquitinates SUMOylated substrates to target them for timely clearance and termination of repair signalling [35, 36]. Besides SUMO conjugation, SUMO deconjugation is also critical for the DNA damage response. Knockdown of all SENPs, except SENP3, impairs repair by NHEJ and HR in double-strand break reporter assays [37]. To date, there is some mechanistic insight on how deconjugation by SENP2, SENP3, SENP6 and SENP7 contributes to DNA repair [37-45]. Nonetheless, we are still limited in our understanding of the regulation of individual DNA damage response proteins by the SUMO system and how these integrate into the complex signalling networks that comprise the DNA damage response.

SUMO and nuclear condensates

Most cellular processes are compartmentalized by membranous or non-membranous organelles (i.e., molecular condensates) to enable the spatial organization of biochemical

pathways. Molecular condensates typically consist of intrinsically disordered proteins and/or RNA and are formed by a process called liquid-liquid phase separation (LLPS). Examples of molecular condensates in the nucleus include PML bodies, Cajal bodies and the nucleolus. Currently, there is an increasing body of evidence supporting the involvement of SUMO in LLPS and the dynamics and compositional regulation of molecular condensates (reviewed in [46]). Indeed, recent *in vitro* studies illustrate that SUMO polymers and SUMO-SIM interactions are the driving forces of PML body assembly [47]. According to the model introduced by Banani et al., the assembly of molecular condensates can be driven by multivalent SUMO-SIM interactions between proteins that have excess SUMO moieties and SIMs available for interaction [47]. Thus, changes in the SUMOylation levels of proteins through SUMO conjugation and deconjugation can influence the formation, dissolution and composition of molecular condensates. In recent years, LLPS and the formation of molecular condensates have gained a lot of attention and are increasingly recognized for their involvement in various cellular processes, including the DNA damage response, and it will be interesting to continue to explore the involvement of the SUMO system.

Scope of the thesis

In this thesis, we aim to gain a better understanding of how the SUMO system maintains nuclear integrity. To this end, we explore both the role of covalent protein SUMOylation and deSUMOylation (**chapter 2** and **chapter 3**), and the role of noncovalent SUMO interactions (**chapter 4** and **chapter 5**).

In **chapter 2,** we provide a comprehensive overview of the recent literature on SUMO proteases, specifically members of the SENP family. We discuss recent advances in our knowledge on the range of substrates that are regulated by these SUMO proteases, and the cellular and physiological processes in which they are involved. We also highlight their emergence as important players in human disease, specifically cancer, and discuss their potential as therapeutic targets. Currently, dysregulation of SENPs is also linked to cardiac, metabolic, muscle and neurological conditions. Lastly, we cover the advances that have been made in the discovery and development of SENP inhibitors, and which challenges still lie ahead.

One member of the SENP family of SUMO proteases is SENP6, which together with SENP7, is capable of removing SUMO2/3 chains from substrates. SENP6 was previously shown to be important for genome maintenance, but we are still limited in our understanding of the underlying mechanisms. In **chapter 3**, we provide novel insights into the role of SENP6 in the DNA damage response. Using mass spectrometry and biochemical validation,

we show that SENP6 deconjugates SUMO2/3 chains on multiple DNA damage response proteins under unstressed conditions and in response to DNA damage, pointing towards coordinated group modification. We explore how regulation of their SUMOylation levels by SENP6 contributes to their functioning during DNA damage signalling by investigating their subcellular localization, nuclear condensation state and kinetics at sites of DNA damage. Since the paradigmatic signalling function of polymeric SUMO2/3 chains is to target proteins for degradation by the proteasome through STUbLs, we also evaluate the interplay between SENP6 and RNF4 in the regulation of these proteins.

Besides SUMO conjugation and deconjugation, noncovalent SUMO interactions also play an important role in the DNA damage response. In **chapter 4**, we perform a noncovalent SUMO interaction mass spectrometry screen using different SUMO isoforms to gain more insight into the SUMO interactome. We screened for interactors of monomeric SUMO1, monomeric SUMO2 and a SUMO2 trimer. Among the interactors, we found many DNA repair proteins, of which XRCC4 was of specific interest as it was the only protein showing a preference for binding SUMO2 trimers over monomeric SUMO2. We identified a SIM in XRCC4 that is required for its binding to SUMO2 trimers. XRCC4 and DNA ligase IV form a complex that is recruited to sites of DNA damage during NHEJ. We explore the functional consequences of an XRCC4 SIM mutant by investigating its interaction with DNA ligase IV, as well as its recruitment to sites of DNA damage.

Interaction through a SIM is the best characterized and thought to be the most predominant mode of noncovalent SUMO interaction. However, we are still limited in our knowledge on other modes of interactions. In **chapter 5**, we perform a noncovalent SUMO interaction mass spectrometry screen with SIM-binding groove SUMO2/3 mutants to specifically capture proteins that interact with SUMO2/3 in a different manner than the typical SUMO-SIM interaction. We also extend our understanding of the SUMO interactome to include covalent SUMO modification in the context of the SIM-binding groove. We explore the SIM-dependency of the covalent SUMO modification of proteins through noncovalent interactions between SUMO and the SUMOylation machinery, including the SUMO E2 enzyme UBC9, SUMO E3 ligases and SUMO proteases of the SENP family. Altogether, **chapter 4** and **chapter 5** provide unique resources and offer valuable new insights on the intricacies of the SUMO interactome.

To conclude, the research in this thesis is summarized and discussed within the broader scope of the field in **chapter 6**. We position our research in the context of current literature and discuss future perspectives.

REFERENCES

1. Kerscher, O. et al. (2006) Modification of Proteins by Ubiquitin and Ubiquitin-Like Proteins. 22 (1), 159-180.
2. Vertegaal, A.C.O. (2022) Signalling mechanisms and cellular functions of SUMO. Nat. Rev. Mol. Cell Biol. 23 (11), 715-731.
3. Owerbach, D. et al. (2005) A proline-90 residue unique to SUMO-4 prevents maturation and sumoylation. Biochem. Biophys. Res. Commun. 337 (2), 517-520.
4. Hickey, C.M. et al. (2012) Function and regulation of SUMO proteases. Nat. Rev. Mol. Cell Biol. 13 (12), 755-66.
5. Mukhopadhyay, D. and Dasso, M. (2007) Modification in reverse: the SUMO proteases. Trends Biochem. Sci. 32 (6), 286-95.
6. Nayak, A. and Müller, S. (2014) SUMO-specific proteases/isopeptidases: SENPs and beyond. Genome Biol 15 (7), 422.
7. Jansen, N.S. and Vertegaal, A.C.O. (2021) A Chain of Events: Regulating Target Proteins by SUMO Polymers. Trends Biochem Sci 46 (2), 113-123.
8. Keiten-Schmitz, J. et al. (2019) SUMO Chains Rule on Chromatin Occupancy. Front Cell Dev Biol 7, 343.
9. Lascorz, J. et al. (2022) SUMO-SIM interactions: From structure to biological functions. Semin. Cell Dev. Biol. 132, 193-202.
10. Yau, T.Y. et al. (2021) SUMO Interacting Motifs: Structure and Function. Cells 10 (11), 2825.
11. Pichler, A. et al. (2017) SUMO conjugation - a mechanistic view. Biomolecular concepts 8 (1), 13-36.
12. Psakhye, I. and Jentsch, S. (2012) Protein Group Modification and Synergy in the SUMO Pathway as Exemplified in DNA Repair. Cell 151 (4), 807-820.
13. Jentsch, S. and Psakhye, I. (2013) Control of Nuclear Activities by Substrate-Selective and Protein-Group SUMOylation. Annu Rev Genet 47 (1), 167-186.
14. Hendriks, I.A. and Vertegaal, A.C. (2016) A comprehensive compilation of SUMO proteomics. Nat. Rev. Mol. Cell Biol. 17 (9), 581-95.
15. Song, J. et al. (2004) Identification of a SUMO-binding motif that recognizes SUMO-modified proteins. Proceedings of the National Academy of Sciences of the United States of America 101 (40), 14373-8.
16. Danielsen, J.R. et al. (2012) DNA damage-inducible SUMOylation of HERC2 promotes RNF8 binding via a novel SUMO-binding Zinc finger. J Cell Biol 197 (2), 179-87.
17. Diehl, C. et al. (2016) Structural Analysis of a Complex between Small Ubiquitin-like Modifier 1 (SUMO1) and the ZZ Domain of CREB-binding Protein (CBP/p300) Reveals a New Interaction Surface on SUMO. J Biol Chem 291 (24), 12658-12672.
18. Knipscheer, P. et al. (2007) Noncovalent interaction between Ubc9 and SUMO promotes SUMO chain formation. EMBO J 26 (11), 2797-807.
19. Capili, A.D. and Lima, C.D. (2007) Structure and analysis of a complex between SUMO and Ubc9 illustrates features of a conserved E2-Ubl interaction. J Mol Biol 369 (3), 608-18.
20. Pilla, E. et al. (2012) A novel SUMO1-specific interacting motif in dipeptidyl peptidase 9 (DPP9) that is important for enzymatic regulation. J Biol Chem 287 (53), 44320-9.
21. Sriramachandran, A.M. et al. (2019) Arkadia/RNF111 is a SUMO-targeted ubiquitin ligase with preference for substrates marked with SUMO1-capped SUMO2/3 chain. Nat Commun 10 (1), 3678.
22. Brüninghoff, K. et al. (2022) A Photo-Crosslinking Approach to Identify Class II SUMO-1 Binders. 10, 2022.900989.
23. Brüninghoff, K. et al. (2020) Identification of SUMO Binding Proteins Enriched after Covalent Photo-Cross-Linking. ACS Chem. Biol. 15 (9), 2406-2414.
24. Cabello-Lobato, M.J. et al. (2022) Microarray screening reveals two non-conventional SUMO-binding modules linked to DNA repair by non-homologous end-joining. Nucleic Acids Res. 50 (8), 4732-4754.

25. Gonzalez-Prieto, R. et al. (2021) Global non-covalent SUMO interaction networks reveal SUMO-dependent stabilization of the non-homologous end joining complex. Cell Rep 34 (4), 108691.
26. Cox, E. et al. (2017) Global Analysis of SUMO-Binding Proteins Identifies SUMOylation as a Key Regulator of the INO80 Chromatin Remodeling Complex. Mol Cell Proteomics 16 (5), 812-823.
27. Aguilar-Martinez, E. et al. (2015) Screen for multi-SUMO-binding proteins reveals a multi-SIM-binding mechanism for recruitment of the transcriptional regulator ZMYM2 to chromatin. Proc Natl Acad Sci U S A 112 (35), E4854-63.
28. Sriramachandran, A.M. and Dohmen, R.J. (2014) SUMO-targeted ubiquitin ligases. Biochim. Biophys. Acta 1843 (1), 75-85.
29. Poulsen, S.L. et al. (2013) RNF111/Arkadia is a SUMO-targeted ubiquitin ligase that facilitates the DNA damage response. J Cell Biol 201 (6), 797-807.
30. Tatham, M.H. et al. (2008) RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. Nat. Cell Biol. 10 (5), 538-46.
31. Rojas-Fernandez, A. et al. (2014) SUMO chain-induced dimerization activates RNF4. Mol. Cell 53 (6), 880-92.
32. Garvin, A.J. and Morris, J.R. (2017) SUMO, a small, but powerful, regulator of double-strand break repair. Philos. Trans. R. Soc. Lond., Ser. B: Biol. Sci. 372 (1731).
33. Galanty, Y. et al. (2009) Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks. Nature 462 (7275), 935-939.
34. Morris, J.R. et al. (2009) The SUMO modification pathway is involved in the BRCA1 response to genotoxic stress. Nature 462 (7275), 886-890.
35. Galanty, Y. et al. (2012) RNF4, a SUMO-targeted ubiquitin E3 ligase, promotes DNA double-strand break repair. Genes Dev 26 (11), 1179-95.
36. Chang, Y.C. et al. (2021) SUMO-Targeted Ubiquitin Ligases and Their Functions in Maintaining Genome Stability. Int J Mol Sci 22 (10), 5391.
37. Garvin, A.J. et al. (2013) The deSUMOylase SENP7 promotes chromatin relaxation for homologous recombination DNA repair. EMBO Rep. 14 (11), 975-83.
38. Garvin, A.J. et al. (2019) The deSUMOylase SENP2 coordinates homologous recombination and nonhomologous end joining by independent mechanisms. Genes Dev. 33 (5-6), 333-347.
39. Alghoul, E. et al. (2023) Compartmentalization of the SUMO/RNF4 pathway by SLX4 drives DNA repair. Mol. Cell 83 (10), 1640-1658.e9.
40. Claessens, L.A. et al. (2023) SENP6 regulates localization and nuclear condensation of DNA damage response proteins by group deSUMOylation. Nat. Commun. 14 (1), 5893.
41. Schick, M. et al. (2022) Genetic alterations of the SUMO isopeptidase SENP6 drive lymphomagenesis and genetic instability in diffuse large B-cell lymphoma. Nat. Commun. 13 (1), 281.
42. Wagner, K. et al. (2019) The SUMO Isopeptidase SENP6 Functions as a Rheostat of Chromatin Residency in Genome Maintenance and Chromosome Dynamics. Cell Rep. 29 (2), 480-494 e5.
43. Gibbs-Seymour, I. et al. (2015) Ubiquitin-SUMO circuitry controls activated fanconi anemia ID complex dosage in response to DNA damage. Mol. Cell 57 (1), 150-64.
44. Dou, H. et al. (2010) Regulation of DNA repair through deSUMOylation and SUMOylation of replication protein A complex. Mol. Cell 39 (3), 333-45.
45. Zhang, T. et al. (2022) Crosstalk between SUMOylation and ubiquitylation controls DNA end resection by maintaining MRE11 homeostasis on chromatin. Nat. Commun. 13 (1), 5133.
46. Keiten-Schmitz, J. et al. (2021) SUMO: Glue or Solvent for Phase-Separated Ribonucleoprotein Complexes and Molecular Condensates? Front Mol Biosci 8, 673038.
47. Banani, S.F. et al. (2016) Compositional Control of Phase-Separated Cellular Bodies. Cell 166 (3), 651-663.