Cover Page

The handle http://hdl.handle.net/1887/28466 holds various files of this Leiden University dissertation

Author: Hendriks, Ivo Alexander

Title: Global and site-specific characterization of the SUMO proteome by mass spectrometry

Issue Date: 2014-09-03

S

SUMMARY

System-wide characterization of proteomes by mass spectrometry has revolutionized the study of post-translational modifications. Small Ubiquitin-like Modifier (SUMO) is a post-translational modification with critical importance to all eukaryotic life. Proteomic study of SUMO at a system-wide level has lagged behind other modifications over the last decade, and my primary goal was to develop purification methods catered to mass spectrometric global study of SUMOylation. I reviewed SUMO with a focus on mass spectrometry, as compared to other post-translational modifications, and describe the various pitfalls that hamper SUMO proteomics (**Chapter 1**).

SUMOylation plays a key role in the deoxyribonucleic acid (DNA) damage response. In order to study the effect of DNA damage on SUMOylation at a system-wide level, we developed a method for efficient purification of epitope FLAG-tagged SUMO from cells. Utilizing this method in combination with mass spectrometry, we confirmed over 400 proteins to be SUMOylated, and furthermore found a subset of these proteins to be dynamically SUMOylated in response to the DNA damaging agent methyl methanesulfonate (MMS). Strikingly, we found a pair of histone demethylases, Lysine-Specific Demethylase 5B (JARID1B) and Lysine-Specific Demethylase 5C (JARID1C), to be inversely regulated in response to DNA Damage. We showed that SUMOylated JARID1B is ubiquitylated by the SUMO-targeted ubiquitin E3 ligase RING Finger Protein 4 (RNF4), and rapidly degraded in response to MMS, whereas JARID1C is relocated from the soluble nucleus to the chromatin, where it exists in its SUMOylated state (**Chapter 2**).

The ubiquitin E3 ligase RNF4 fulfills an important function in regulating Promyelocytic Leukemia Protein (PML) nuclear bodies, and has been increasingly implicated in the DNA damage response. We purified FLAG-tagged RNF4 and used mass spectrometry in order to identify potential RNF4 interaction partners, using a buffer condition where all SUMO is removed from proteins by endogenous proteases. Here, we discovered the Ubiquitin-Specific Peptidase 11 (USP11) to interact with RNF4. Interestingly, the core functions of the ligase and protease are intrinsically opposed, yet both have been reported to be important in DNA double-strand break repair. We demonstrated that depletion of USP11 leads to a decrease in SUMOylated PML, and a destabilization of PML nuclear bodies, whereas depletion of RNF4 yielded an increase. Furthermore, expression of exogenous USP11 was able to prevent the RNF4-mediated dissociation of PML bodies in response to the DNA damaging agent MMS (**Chapter 3**).

Summary

Mass spectrometry has been used to study SUMOylation at the protein-level, but failed to effectively study this modification at a site-specific level. Thus, SUMOylation separation-of-function mutants have traditionally been generated using trial-and-error mutagenesis; a laborious and inconvenient process. We developed methodology to study SUMOylation at the site-specific level in an unbiased fashion, using a lysine-deficient SUMO bearing an additional carboxyl-terminal mutation in order to avoid the bulky tryptic remnant. Using this approach, we identified 103 SUMOylated lysines in endogenous proteins. Sequence analysis provided novel understanding of the SUMO conjugation consensus motif, including an inverted SUMOylation motif and a hydrophobic cluster motif, as well as more insight into phosphorylation-dependent SUMOylation (**Chapter 4**).

SUMO is of paramount importance to cell cycle progression, and SUMO has been shown to modify many cell-cycle related factors. Antagonizing SUMOylation leads to chromosomal defects and aberrations in dividing cells. To further extend our knowledge on SUMOylation during cell cycle progression, we utilized the FLAG-SUMO purification methodology (Chapter 2), with the ability to identify SUMO sites (Chapter 4). Using mass spectrometry, we confirmed over 500 proteins to be SUMOylated, and over 100 proteins were found to be dynamically SUMOylated throughout the cell cycle. Furthermore, 203 SUMOylated lysines were pinpointed. The transcription factor Forkhead Box Protein M1 (FoxM1) was revealed to be upregulated in SUMOylation during mitosis, where SUMOylation serves to increase the activity of FoxM1 (**Chapter 5**).

In order to gain a global view on SUMOylation at the site-specific level, we greatly optimized the lysine-deficient SUMO purification methodology (Chapter 4). Optimizations included generation of a stable cell line, extension of the epitope tag, a double purification, and filtration steps where unconjugated SUMO is removed. We combined this greatly improved purification strategy with cutting-edge high-resolution mass spectrometry, and identified 3,246 SUMOylation sites, mapped to 1,364 SUMOylated proteins. We provided the first global analysis of the SUMO proteome, and greatly enhance insight into SUMOylation by extensive bioinformatical analysis of all SUMOylation sites and proteins. We revealed large functional clusters of interacting proteins to be co-SUMOylated, and demonstrated crosstalk and competition between SUMOylation and other major post-translational modifications; ubiquitylation, acetylation, methylation and phosphorylation (**Chapter 6**).

Identification of SUMOylated lysines has primarily been performed using mutant SUMO (Chapter 4, 5 and 6). Although this approach has proven rather effective and generally unbiased, it prohibits identification of lysines in endogenous proteins modified by endogenous SUMO. In order to take the first hurdle towards this goal, we developed the Protease-Reliant Identification of SUMO Modification (PRISM) methodology, which involves chemical blocking of all lysines in a sample, followed by removal of SUMO from lysines by specific proteases. The “freed” lysines could then be identified using high-resolution mass spectrometry. With this approach, while using His-tagged but otherwise completely wild-type SUMO, we identified 389 SUMOylation sites, mapped to just over 200 proteins. Further refinement of this method may lead to system-wide identification of completely endogenous SUMOylation sites, ultimately from patient samples or animal tissues (**Chapter 7**).

Finally, I discuss all data described in this thesis, and performed additional statistical comparison between all identified SUMOylated proteins and SUMO sites in this thesis, as well as externally identified SUMO proteomes and sites. Ultimately, we described 3,538 unique SUMOylation sites throughout this thesis, expanding nearly a 100-fold over the cumulative amount of SUMOylated lysines reported outside of this lab. Thus, we provide the SUMO community with an invaluable resource for the generation of SUMO-deficient proteins for further in-depth functional studies. With the development of PRISM, we also provide a promising option for extrapolating the site-specific study of SUMO to the truly endogenous level, where it can hopefully assist in elucidation of the role of endogenous SUMO in cancer and other diseases (**Chapter 8**).