SUMO: getting it on

J. Anckar† and L. Sistonen††

†Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, PO Box 123, FI-20521 Turku, Finland, and †Department of Biology, Åbo Akademi University, FI-20521 Turku, Finland

Abstract

Post-translational modification of cellular proteins by the SUMO (small ubiquitin-related modifier) is involved in numerous modes of regulation in widely different biological processes. In contrast with ubiquitination, SUMO conjugation is highly specific in terms of target lysine residues, but many aspects of substrate and lysine selection by the SUMO conjugating machinery are still poorly understood. SUMOylation events usually occur on the ψKXE SUMO consensus motifs, which mediate binding to Ubc9 (ubiquitin-conjugating enzyme 9), the SUMO E2 conjugating enzyme. Although most, if not all, SUMO conjugations are catalysed by Ubc9, far from all ψKXE tetrapeptides are modified, demonstrating a need for additional specificity determinants in SUMOylation. Recent results intimately link regulation of SUMOylation to other post-translational modifications, including phosphorylation and acetylation and reveal that certain lysine residues are marked for SUMOylation by negatively charged amino acid residues or phosphorylation events immediately downstream of the consensus site. In the present review, we explore the intriguing role of extended motifs in the regulation of SUMO conjugation.

Introduction

A central way of diversifying protein function is through the post-translational addition or removal of chemical groups. Among these modifiers, Ubls (ubiquitin-like proteins) such as SUMO (small ubiquitin-related modifier) are known to be key regulators of several different biological functions. Like ubiquitin, SUMO proteins are covalently and reversibly conjugated to specific lysine residues in the target proteins. However, SUMOylation does not target proteins for proteosomal degradation and can in fact stabilize the target protein by blocking ubiquitination of the same lysine residue, as has been shown for IkBα (inhibitory κBα) [1]. Lately, it has become clear that SUMOylation is involved in surprisingly diverse biological pathways, such as genome integrity, chromosome packing and dynamics, and various aspects of signal transduction [2,3]. In particular, SUMO has been shown to play crucial roles in meiosis and mitosis [4–8]. In yeast, disruption of the SUMO conjugation pathway prevents cell cycle progression beyond the G2/M-phase [9]. Similarly, mice that have a dysfunctional SUMO pathway exhibit defects in nuclear architecture as well as in mitotic chromosome organization and die at early stages of embryonic development [10].

The most prominent group of SUMO targets is proteins involved in transcriptional regulation. The common outcome of transcription factor SUMOylation is an inhibition of transcription capacity [11]. This is partially due to an inherent and quite potent ability of SUMO proteins to repress transcription, since an artificial recruitment of SUMO to promoters is enough to shut down transcription [12,13]. A characteristic of SUMO-mediated repression is that, although the fraction of SUMO-modified protein is small relative to the total pool of substrate, mutations that disrupt SUMOylation often have dramatic effects on transcriptional activity (for an in-depth discussion, see [3]). Although the precise mechanism of SUMO-mediated repression in many cases is unclear, accumulating evidence points towards a model in which SUMOylation allows the protein to interact with transcriptional co-repressors, for example chromatin-modifying enzymes [11,14–16]. Additionally, SUMO may exert its repressor function by inducing a conformational change of the substrate or by altering protein localization or DNA-binding activity [17–19].

SUMO conjugation is analogous to ubiquitination

The attachment of different Ubls to their target proteins proceeds in a stepwise manner through analogous, yet dedicated pathways (Figure 1) [20–22]. In the initial ATP-dependent step, the Ubl forms a thioester bond with an E1 activating enzyme. The activated Ubl moiety is subsequently passed to an E2 conjugating enzyme, which acts together with E3 ligases to attach the Ubl to the substrate through an isopeptide bond. The Ubl is finally recycled by isopeptidases, which cleave the Ubl from the substrate, maintaining a free pool of Ubls. In the ubiquitin system, substrate specificity is determined by the concerted actions of tens of E2 and hundreds of E3 enzymes, in which the E3 ligases play a dominant role in substrate specificiation [20]. The SUMO system utilizes only a single E2 enzyme, Ubc9 (ubiquitin-conjugating enzyme 9), and probably many fewer E3 ligases. This difference

Key words: acetylation, negatively charged amino acid-dependent SUMOylation motif (NDSM), phosphorylation-dependent SUMOylation motif (PDSM), small ubiquitin-related modifier (SUMO), transcription, ubiquitin-conjugating enzyme 9 (Ubc9).

Abbreviations used: Elk-1, ETS (E twenty-six)-like kinase 1; HIC1, hypermethylated in cancer 1; HSF, heat-shock factor; MEF2A, myocyte enhancer factor 2A; NDSM, negatively charged amino acid-dependent SUMOylation motif; PCNA, proliferating-cell nuclear antigen; PDSM, phosphorylation-dependent SUMOylation motif; SUMO, small ubiquitin-related modifier; Ubc9, ubiquitin-conjugating enzyme 9; Ubl, ubiquitin-like protein.

†To whom correspondence should be addressed (email lea.sistonen@btik.fi).
notwithstanding, SUMO conjugation is known to be highly lysine-specific, raising an obvious question of how this accuracy is achieved.

A distinctive feature of SUMO conjugation is that the target lysine residues are often found within a consensus tetrapeptide, ψKXE (ψ denoting a hydrophobic residue), which positions the catalytic site of Ubc9 on the target lysine residue [23,24]. Consequently, Ubc9 possesses not only the capacity to catalyse the transfer of the SUMO moiety to the substrate, but also the capacity to select the accurate lysine residue for modification. Indeed, the SUMO conjugation of a cellular protein can often be reconstituted in vitro by using only the E1 and E2 enzymes. Although ligases of the SUMO system are able to promote SUMOylation efficiency and are thought to be involved in most SUMOylation events in vivo, it is unknown whether they are actually involved in determining substrate specificity [2]. On the contrary, many SUMO ligases show promiscuous behaviour in cells and especially in vitro, indicating that Ubc9 is a central player in specifying SUMO substrate and target lysine [25].

For some SUMO substrates, the binding to Ubc9 is enhanced by additional interactions that occur outside the consensus sequence [18,24]. The consensus tetrapeptide is nevertheless necessary for efficient SUMOylation also under these circumstances, demonstrating the involvement of multiple co-operating interactions in regulating Ubc9–substrate binding. In this regard, the consensus sequence can be seen as a local mediator of substrate–Ubc9 interaction, which fine-tunes the SUMO conjugation event by facilitating the correct positioning of the target lysine residue to the active site of Ubc9.

Although apparently all SUMO conjugations in the cell are catalysed by Ubc9, many SUMO target sites do not conform to the ψKXE consensus tetrapeptide. A possible explanation of this conundrum is that the ψKXE motif serves as a consensus sequence only for a subclass of structurally defined target peptides. For example, the selection of SUMO target sites of the ubiquitin E2 enzyme E2-25K is dictated by the structural features of the surrounding region. In the context of the folded protein, E2-25K is SUMOylated on a non-consensus lysine residue within an α-helix, whereas neighbouring SUMO consensus lysine residues are not modified. In the context of unstructured peptides, however, only consensus sequence lysine residues are SUMOylated [26]. Similarly, PCNA (proliferating-cell nuclear antigen) is SUMOylated on two lysine residues: one found in a consensus sequence resides in a long loop structure, and another non-consensus lysine residue is found in a protruding tip of PCNA [27,28]. Nevertheless, as consensus tetrapeptides within a structurally flexible region are not always SUMOylated, additional specificity determinants in the SUMOylation of ψKXE tetrapeptides are required [18].
Figure 2 | Extended motifs for SUMO conjugation

(A) Sequence comparison between PDSM and NDSM. The PDSM is composed of a proline-directed phosphorylation site at position +3 downstream of the SUMO consensus site. In the NDSM, negatively charged residues are found predominantly at positions +3–6. (B) The motif was conserved between human and mouse species is shown (left column). Using the ScanProsite tool, the Swiss-Prot database was searched for proteins containing a SUMO consensus motif followed by a downstream SP site at positions +1–8. The number of proteins in which the motif was conserved between human and mouse species is shown (right column).

<table>
<thead>
<tr>
<th>SUMO consensus</th>
<th>motif</th>
<th>Proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΨKxExxSP</td>
<td>PDSM</td>
<td>12</td>
</tr>
<tr>
<td>ΨKxExxSP</td>
<td>NDSM</td>
<td>55</td>
</tr>
<tr>
<td>ΨKxExxExx</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>ΨKxExxExx</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>ΨKxExxExx</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>ΨKxExxExx</td>
<td></td>
<td>22</td>
</tr>
<tr>
<td>ΨKxExxExx</td>
<td></td>
<td>19</td>
</tr>
</tbody>
</table>

mutation of the negatively charged residues C-terminal to the consensus sequence inhibited the SUMOylation of Elk-1. An extensive analysis of SUMO substrates revealed an overrepresentation of negatively charged residues within a 10-amino-acid region downstream of the SUMO consensus site [40]. Such motifs, commonly termed NDSM, can be used to identify novel SUMO substrates (Figure 2A).

**PDSM and NDSM: variations on the same theme?**

Within PDSM, the phosphorylated serine residue can be replaced by a negatively charged amino acid, indicating that the SUMOylation-promoting action provided by the phosphate group is due to a gain of a negative charge at position +3. Therefore it is possible that PDSM and NDSM function through the same mechanism. Indeed, the loss of MEF2A SUMOylation by inhibition of PDSM phosphorylation could be rescued by inserting an acidic residue into position +6 [40]. Importantly, the negatively charged residues of the NDSM were found to interact directly with two positively charged residues of Ubc9, creating an additional binding surface between the substrate and the conjugating enzyme [40]. In the case of HSF1, a serine-to-aspartate mutant mimicking PDSM phosphorylation is a slightly better SUMO substrate than the wild-type HSF1 also in the absence of E3 ligases, which suggests that the mechanism of action of PDSM could be due to a more efficient binding of Ubc9 [30]. Interestingly, the positively charged residues of Ubc9 that bind to NDSM are not conserved in Saccharomyces cerevisiae [24], indicating that SUMO substrate recognition proceeds through another mechanism in yeast.

A distinction between the NDSM and PDSM motifs can be made based on the spacing between the SUMO target site and the negatively charged residues or the phosphorylated serine. In an alignment of known SUMO substrates, acidic residues are more or less equally overrepresented at positions +3–6 [40]. In the PDSM, however, the spacing between the SUMO consensus tetrapeptide and the SP site is particularly important, since the SP sites C-terminal to the SUMO consensus tetrapeptide seem not to be randomly distributed at positions +3–6. Among the proteins in the Swiss-Prot database, ΨKxExxSP is clearly more common than any other version of this motif (Figure 2B). Moreover, there is a dramatic enrichment of SUMOylated substrates and transcriptional regulators when the search parameters are restricted specifically to ΨKxExxSP [31]. This enrichment does not exclude the possibility that other negatively charged residues could also influence the SUMOylation efficiency, but it strongly suggests that a phosphorylation event at position +3 is most important for determining SUMO conjugation efficiency. Indeed, the phosphorylation of HSF1 on a serine residue at position +7 does not affect HSF1 SUMOylation [30,31]. Another, equally important distinction between NDSM and PDSM is that the negative charge gained by PDSM phosphorylation is of a transient nature; the negative charge persists only as long as the phosphate group is present. Moreover, as the negative charge of the phosphate group is added and removed by substrate-specific signalling pathways, the PDSM could present a refined module to couple extracellular signalling with substrate-specific SUMOylation. The placing of SUMO conjugation under the control of phosphorylation thus provides a dynamic and highly specific way of co-ordinating SUMO-mediated responses [41].

**Switches between SUMO and acetyl groups in transcriptional regulation**

The complexity of upstream signalling events regulating the PDSM-mediated SUMOylation was further increased by the finding that the same lysine residue within the PDSM of MEF2A can be targeted by both SUMO and acetyl groups [42,43]. An elegant study by Shalizi et al. [44] showed that calcineurin-mediated activation of MEF2A in depolarized neurons is due to a dephosphorylation and a subsequent loss of SUMO from the PDSM of MEF2A [44]. Interestingly, this decreased SUMOylation is coupled with an increased acetylation of the PDSM lysine residue, indicating that the PDSM regulates MEF2A activation through a switch between SUMO and acetyl groups [SAS (SUMOylation-acetylation switch)] [44]. An unresolved issue is whether dephosphorylation of the PDSM serine residue actively promotes acetylation or whether it renders more PDSM lysine residues accessible for acetylation by blocking their SUMOylation (Figure 3). For example, a phosphorylation-independent deacetylation/SUMOylation switch was recently reported in the tumour suppressor HIC1 (hypermethylated in cancer 1) [45]. In HIC1, a proline residue at position +1 downstream of a SUMO consensus sequence appears to be required for acetylation. Since many PDSMs contain a proline residue at position +1 [31], it is possible that similar switches between SUMO and acetyl groups could occur also on other PDSM-containing proteins.
The targeting of the same lysine residues by SUMO and acetyl groups is not restricted to PDSM- or ψKXE-containing proteins. The lysine residues of p300, oestrogen receptor and histones that are all targeted by both SUMO and acetyl groups are not found in either motif [46–49]. In fact, it is tempting to speculate that acetylation could function as a recurrent regulator of SUMOylation events. For example, histone deacetylases are known to be intimately involved in SUMO-mediated transcriptional repression, functioning not only as effectors of SUMO-mediated repression, but also as regulators of SUMOylation [11]. It is easy to envision a model where SUMO and acetyl groups would take turns at modifying lysine residues, functioning as SUMO-acetyl-regulated transcriptional switches. In this model, the repressive function of SUMO on transcription could be counteracted by the acetyl group, which by blocking SUMO attachment, facilitates transcriptional activation (Figure 3). Consequently, elucidation of the conjugating and deconjugating enzymes of phosphorylation, SUMOylation and acetylation pathways converging on certain PDSMs and PDSM-like motifs is likely to provide increased insights into how such SUMO-mediated transcriptional switches are regulated.

We apologize to our colleagues, whose work could only be cited indirectly owing to space limitations. Johanna Ahlsgård, Henri Blomster, John Eriksson, Aura Kaunisto and Anton Sandqvist are acknowledged for comments on this paper. Our work is financially indirectly owing to space limitations. Johanna Ahlskog, Henri Blomster, John Eriksson, Aura Kaunisto and Anton Sandqvist are acknowledged for comments on this paper. Our work is financially supported by the Academy of Finland, the Sigrid Juselius Foundation, the Finnish Cancer Organizations, the Finnish Life and Pension Insurance Companies and Åbo Akademi University. J.A. is supported by the TuBS (Turku Graduate School of Biomedical Sciences).

References

©The Authors Journal compilation ©2007 Biochemical Society

Received 31 May 2007
doi:10.1042/BST0351409