

YILIN YAN  
STEVEN J. ORCUTT  
JAMES E. STRICKLER

# The use of SUMO as a fusion system for protein expression and purification

## ABSTRACT

The preparation of large amounts of recombinant protein is a major bottleneck for many areas of scientific research, including structural proteomics and therapeutic drug discovery. Recently, the Small Ubiquitin-like Modifier (SUMO)-based fusion tag technology has gained increasing use for the expression of recombinant proteins. In *E. coli* recombinant proteins expressed as SUMO fusions have demonstrated enhanced stability and solubility, leading to greatly increased yields (5 to 20-fold) over constructs lacking this tag. Furthermore, efficient removal of the SUMO tag by SUMO protease allows the generation of a native N-terminus of the target protein. The use of SUMO-fusion technology in *E. coli*, termed SUMOpro, as well as the use of SUMOstar, an adaptation of SUMO for use in eukaryotic hosts, will be reviewed. This review will also provide a comparison between SUMOpro technology and other fusion tags commonly used to enhance protein production and facilitate purification.

## INTRODUCTION

The production of recombinant proteins has increased exponentially with the rapid growth of genomics, proteomics and bioinformatics (1). Efficient expression of recombinant proteins of interest (POIs) at high yield and purity is required for structural studies by X-ray crystallography and NMR spectroscopy (2). With the advent of biological therapeutics, the large scale production of recombinant protein drugs is gaining in importance, with particular interest being focused on the generation of shorter recombinant peptides (<50 amino acids) as potential tools for drug discovery or as therapeutic agents themselves. To this end, the optimization of conditions for the expression of POIs in a variety of different systems such as *Escherichia coli*, *Saccharomyces cerevisiae*, *Pichia pastoris*, baculoviral/insect cells, and mammalian may always be an ongoing process (3-7). Among these hosts, *E. coli* has been widely used in heterologous protein production due to its low cost, ease of manipulation, and the existence of well established isotope labelling protocols for NMR (8). However, *E. coli* has been found to express membrane-bound proteins poorly and often the majority of

the recombinant protein collects as insoluble inclusion bodies. This problem is particularly acute for short peptides, which are often either insoluble or subject to proteolytic degradation by the expression host. Extensive efforts have been aimed at the generation of a number of fusion tag technologies, with those that are commercially available listed in Table 1 (9-11). Of these, the SUMO-fusion system has been shown to be superior in the enhancement of both expression and solubility of difficult-to-express POIs, while protecting against proteolytic degradation, facilitating protein purification, and most importantly allowing for the generation of a native N-terminus following tag removal (12).

## SUMO AND SUMO PATHWAYS

SUMO is a highly conserved protein in eukaryotes belonging to a group of proteins called ubiquitin-like proteins (Ubls), due to their shared structural homology with ubiquitin. SUMO is absent from prokaryotes (13-15). *Saccharomyces cerevisiae* has only one SUMO gene (*SMT3*), whereas the mammalian SUMO family contains three SUMO homologous (SUMO-1, SUMO-2, and SUMO-3). Although there is only 18 percent sequence identity between ubiquitin and SUMO, the global folding and overall structures are similar, with the exception of a flexible N-terminal extension in SUMO (16). The mechanisms underlying conjugation (sumoylation) and deconjugation (desumoylation) of SUMO to target proteins

are similar to those described for the ubiquitin pathway (17). Similar to ubiquitinylation, sumoylation of target proteins occurs through conjugation of the C-terminal glycine of SUMO to the ε-amino group of specific lysine residues. Free SUMO is initially activated by SUMO activating enzyme E1, which is composed of Aso-1 and Uba2 subunits (18, 19). This conjugation is followed by covalent transfer of SUMO to its target protein via the SUMO specific E2-type conjugating enzyme (Ubc9) and

E3 ligase (20). The resultant isopeptide bond, forming the junction between SUMO and the target protein, is then readily cleaved by SUMO proteases, liberating the target protein from SUMO. The SUMO pathway is also similar to the ubiquitin pathway in that sumoylation/ desumoylation is a very dynamic process, with the status of target proteins as SUMO conjugates determining their cellular fate.

Tag	Size (kDa)	Purification method
His <sub>6</sub> -SUMO	12.2	Ni or Co-IMAC
His <sub>6</sub>	0.84	Ni or Co-IMAC
FLAG	1	Anti-Flag antibodies
His <sub>6</sub> -Ub	8.6	Ni or Co-IMAC
TRX	14.3	ThioBond resin
GST	27.3	Glutathione-Sepharose
MBP	45	Amylose resin
His <sub>6</sub> -NusA	54.8	Ni or Co-IMAC

Table 1. Commercial available tags for protein expression and purification.

Sumoylation has been shown to play an important role in targeting proteins to the nuclear subspace, formation of subnuclear structures, transcriptional regulation, control of protein stability, DNA repair and signal transduction pathways (21-25).

Desumoylation is achieved by the SUMO proteases Ulp1 and Ulp2 in yeast, and SENP1 and SENP2 in humans (25-27). These proteases recognize SUMO through tertiary interactions, prior to active site docking and cleavage at the C-terminus (28, 29). Recently, complexes of Ulp1-SMT3 and SENP2-preSUMO-2 were resolved by x-ray crystallography, illustrating the mechanism of SUMO deconjugation (30, 31). Common features within the active site include a 90 degree kink proximal to the scissile bond, which would force the C-terminal amino acid residues on the lysine side chain (on the target protein) toward the protease surface, apparently optimizing steric interactions and allowing for isopeptide hydrolysis. This mechanism of directed cleavage through tertiary recognition is presumably how both deubiquitinylases (DUBs) and desumoylases maintain such a narrow substrate specificity, despite the presence of a highly conserved Gly-Gly cleavage site among Ubls.

#### THE USE OF SUMO<sub>pro</sub> ENHANCES PROTEIN EXPRESSION AND SOLUBILITY

The most commonly used method to increase the soluble yield of recombinant protein in *E. coli* is expression of the POI as a fusion partner with a solubility-enhancing tag (32). Although the mechanism of how these tags maintain protein solubility remains largely unknown, they have been shown to be very effective for a large number of recombinant proteins. The most widely used fusion tags of this kind are maltose-binding protein (MBP), glutathione-S-transferase (GST), N-utilization substance A (NusA), and SUMO (33-36). MBP and GST tags also conveniently function as affinity tags. Recently, several recombinant proteins were expressed as N-terminal fusion constructs containing either a hexahistidine tag (HIS6) alone, or in combination with a SUMO tag (33, 37). Matrix metalloprotease (MMP13), green fluorescent protein (GFP), and SARS-CoV 3CL protease were all successfully expressed and purified using the SUMO-fusion system. In particular, the expression of MMP13 as a SUMO fusion was found to demonstrate dramatically increased solubility, compared to the untagged species. This strongly suggests that SUMO may be acting in a chaperon-like manner, resulting in the observed enhancement of protein stability and solubility. In order to compare the ability of SUMO-fusion technology to enhance protein expression and solubility with more traditional fusion systems, three POIs were chosen: enhanced green fluorescent protein (eGFP), MMP13, and myostatin (36). These POIs were expressed as C-terminal fusions to either MBP, GST, thioredoxin (TRX), NusA, ubiquitin, or SUMO. Of all the affinity tags tested, SUMO displayed the strongest ability to enhance recombinant protein expression and solubility. The SUMO<sub>pro</sub> system has also been compared to the GST fusion system for the production of recombinant human aspartyl-tRNA synthetase (hDRS) in *E. coli* (38). Although fusion of both GST and SUMO to hDRS greatly enhanced active yields of the final product, relative to expression of an untagged species, fusion of SUMO to hDRS resulted in nearly 10-fold higher product yields over GST as a fusion partner.

Protease	Recognition site	Residual N-terminus	Enzyme:substrate	Cost (dollar/mg) <sup>1</sup>
SUMO protease 1	x-G-G x and the tertiary structure	None	0.1% (wt/wt)	3.15 <sup>2</sup>
TEV	E-x-x-Y-x-Q (G/S)	G/S	1-3%(wt/wt)	45.00
Thrombin	L-V-P-R G	G	1-10% (wt/wt)	4.00
Factor Xa (fXa)	I-E-G-R x	None	1-10% (wt/wt)	4.00
Enterokinase	D-D-D-D-K x	None	0.1% (wt/wt)	45.00
Immobilized subtilisin BPN	C-terminus of propeptide x	None	1:1 (mol/mol)	26.00

<sup>1</sup>approximate pricing based on manufacture's recommended protocol for processing.

<sup>2</sup>based on LifeSensors' current pricing and protocol for processing.

Table 2. Comparison of various proteolytic strategies for affinity tag removal.

#### PROTEOLYTIC PROCESSING WITH SUMO PROTEASE YIELDS SPECIES WITH NATIVE N-TERMINUS

Fusion tag removal from recombinant POIs is typically desired due to concerns of potential interference by these tags with POI structure and/or activity (32, 39). For most fusion tag systems, proteolytic removal of the affinity tag requires engineering a cleavage site between the tag and the POI (Table 2).

A key advantage of SUMO fusion over competing technologies is the ease, efficiency, and fidelity with which the SUMO fusion tag is removed through the use of SUMO protease 1 (LifeSensors, Inc.). SUMO protease 1 has been demonstrated to efficiently cleave recombinant SUMO fusion proteins ranging in molecular weight from roughly 6 to 110 kDa, and to date approximately 100 SUMO fusions have been cleaved without erroneous digestion (12, 33, 37). SUMO protease 1 activity has been exhaustively tested to determine conditions that maintain both efficiency and fidelity of cleavage for SUMO fused POIs. (33). SUMO protease has been shown to tolerate a wide range of reaction conditions, including pH from 6 to 9, high ionic strength (e.g. 0.5M NaCl, 300mM imidazole) and even the presence of low concentrations of chaotropes and detergents (e.g. 2M Urea, 0.1 M Guanidine-HCl, 1 percent Triton).

An added advantage of SUMO fusion technology is the potential savings in both reagent use and time during tag removal and purification of processed POI. The use of factor Xa and thrombin for fusion protein digestion typically requires enzyme-substrate ratios of 1:20 (wt/wt) to achieve efficient digestion, and is often associated with non-specific cleavage of the recombinant POI (40). Non-specific cleavage in the POI can be avoided through optimization of cleavage conditions and/or a mutagenesis strategy to eliminate potential cleavage sites within the POI. However, either approach adds additional time and cost in generation of recombinant protein. Although tobacco etch virus protease (TEV) displays high fidelity relative to thrombin and fXa (41), the additional cost of this reagent is significant. SUMO protease 1, however, has demonstrated both specificity and efficiency, with efficient removal of the SUMO tag only requiring a 1:1000 (wt/wt) ratio of enzyme to substrate. Therefore SUMO fusion technology provides for the robust generation of processed POIs, without the added expense of large amounts of enzyme, or the time and effort required to avoid non-specific cleavage.

Perhaps, though, the most significant advantage that SUMO fusion technology holds over other fusion technologies is the generation of a native N-terminus for the POI. This is particularly critical for producing proteins that are dependent upon a specific N-terminus for activity. For most commercial proteases listed in Table 2, removal of the affinity tag results in an undesired N-terminal extension of processed POIs. This is due to the minimal requirements (at the site of cleavage)

necessary for efficient enzyme recognition and proteolysis. For example, efficient cleavage by thrombin requires the amino acid sequence Leu-Val-Pro-Arg-Gly, with cleavage after the arginine residue leaving an extra glycine (at the so-called P1' position (42)) at the N-terminus of the POI. A detailed analysis of the specificity of TEV protease processing revealed that a wide range of residues beyond the canonical Gly/Ser are tolerated at the P1' position, with varying degrees of efficiency (41). However, for some residues (Leu, Lys, and Gln in particular) enzyme concentrations of up to 1:2 (wt/wt) were required, with efficiency approaching only 60 percent. In comparison, efficient recognition and cleavage by SUMO protease tolerates nearly any amino acid immediately following the scissile bond, the exception being proline (33). Therefore, by not having the strict requirements that are inherent with thrombin, or the additional cost associated with TEV, almost any desired N-terminus can be generated in the POI by SUMO protease with minimal investments in time and reagents.

This key advantage of SUMO protease processing has recently been demonstrated to have been fundamental to the generation of recombinant gene products for use in both basic research and therapeutic drug discovery. For example, Lu, et al. employed the use of SUMO fusion technology in the production of biologically active chemokines (43). Biological activity of chemokines is critically dependent upon their native N-termini, and the use of SUMO technology enabled the generation of 15 active chemokines from *E. coli*, with yields of up to 1.5mg/ml. Similarly, the use of SUMOpro has also allowed for low cost production of functional peptide ligands for the G-protein coupled receptors (GPCRs) PTHR1 and GLP-1R (44). Similar to chemokines, high affinity interactions of these peptides with their cognate receptors is dependent upon the presence of a native N-terminus, with the enhanced yields achieved through SUMO fusion (20-40mg per litre culture) greatly facilitating downstream modification (e.g. isotope labelling) of the recombinant peptides for both GPCR characterization and antagonist/agonist assay development. The use of SUMOpro also greatly enhanced expression and purification of recombinant brazzein, which in natural form is a sweet protein derived from an African fruit (45). Recombinant production of brazzein promises to allow for further research in the field of taste receptors, as well as a potential artificial commercial sweetener. The production of brazzein required overcoming two main obstacles, a high di-sulfide content and the requirement for a native N-terminus. The expression of brazzein with the SUMOpro system greatly enhanced the ultimate yield of the soluble, folded, form of the recombinant protein (>5mg/per gram wet cells). Digestion of the fusion protein was >90 percent with SUMO protease 1 (1hr, 30°C), and functional analysis revealed it to be indistinguishable from native brazzein. SUMO protease 1 is clearly established in its efficiency, fidelity, and its ability to generate a recombinant POI with its native N-terminus.

#### PURIFICATION OF RECOMBINANT POIS WITH SUMOPRO

The purification of high quantity and quality POIs can be easily achieved with the SUMOpro system, following the scheme developed in our laboratory (Figure 1). Once expressed, small scale (0.5-4 mg) protein purification, including proteolytic

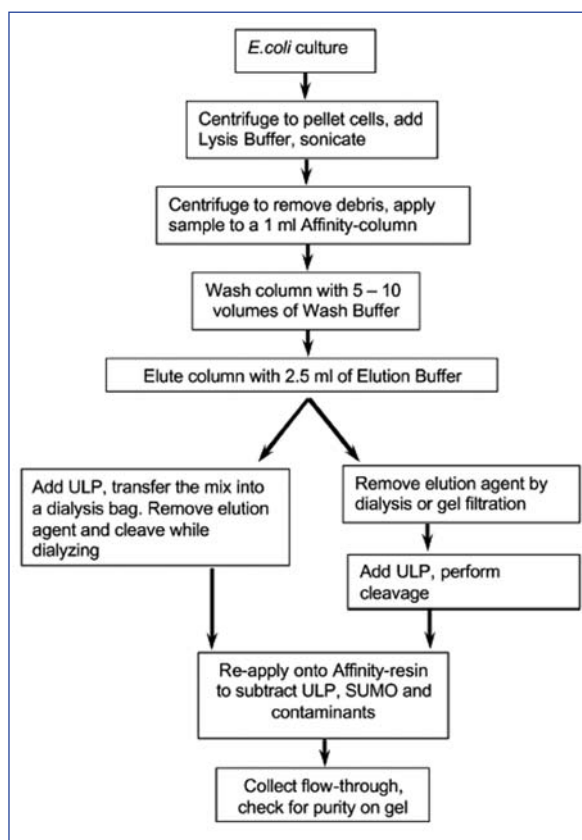


Figure 1. Recombinant protein purification flow chart using SUMO-fusion technology (32).

removal of the SUMO tag, can be completed within 4 hours. Larger batches of recombinant protein can be processed, usually with overnight cleavage and dialysis, requiring essentially the same amount of labour and time (33). In addition, the recombinant form of SUMO protease includes the N-terminal His<sub>6</sub> tag, facilitating its removal by IMAC chromatography, after tag removal. Often the initial IMAC affinity chromatography step results in >80 percent purity. Final purity, including the subtractive purification step for removal of SUMO protease 1 and the free SUMO tag, often results in >90 percent purity (38, 43, 45).

#### SUMOSTAR ALLOWS FOR ENHANCED PROTEIN EXPRESSION IN EUKARYOTIC SYSTEMS

Due to the lack of endogenous SUMO proteases in prokaryotes, the SUMOpro expression system is virtually free of *in vivo* proteolytic processing of the SUMO fusions generated in *E. coli*. However, the nature of SUMO tagged POIs is such that they would likely be cleaved by endogenous SUMO proteases, hampering the tag's ability to facilitate enhanced protein production (12) in eukaryotes. In order to achieve the production of otherwise difficult to express POIs (that may also require post-translational modification) in eukaryotic systems, a novel SUMO fusion technology termed "SUMOstar" was recently developed in our laboratory. The rationale for the

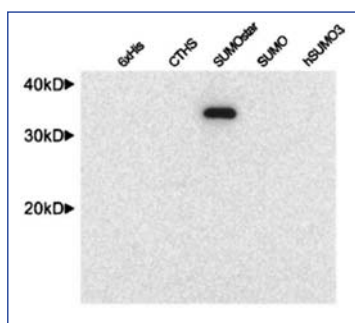


Figure 2. sPLA<sub>2</sub>-X expression level examined by western blot. Sample at 48 hours after post-transfection in HEK-293T cell medium were loaded on SDS-PAGE and transferred to nitrocellulose membrane for western blot detection. The blots were probed with monoclonal anti-6xHis antibody and five N-terminal tags, including 6xHis, 6xHis-CTHS, 6xHis-SUMOstar, 6xHis-SUMO, and 6xHis-hSUMO3 were examined (32). Abbreviations: CTHS, C-terminal half; hSUMO3, human SUMO 3.

development of SUMOstar was in accordance to the following rules: 1) there would be minimal changes to the protein sequence, thereby minimizing changes to the global structure, 2) the fusion should display little to no processing by native SUMO proteases (both *in vivo* and *in vitro*), 3) the new tag should significantly enhance the expression of recombinant POIs.

Detailed structural analysis resulted in the discovery of electrostatic interactions between two positively charged residues (Arg64 and Arg71) on the surface of SUMO and a corresponding negatively charged residue on the surface of the protease, Ulp1 (31). A SUMO variant containing mutations at both positions (R64T and R71E) was generated and shown to be resistant to cleavage by Ulp1 (46). In order to test the ability of SUMOstar to enhance protein expression, GFP, mouse UBP43, Ubiquitin specific protease (USP) 4, USP15, and trypsin were chosen as model proteins to test in a baculovirus/insect cell system (47).

The SUMOstar tag was found to prevent cleavage of the fusion proteins *in vivo* and greatly enhance the expression of all model proteins. In addition, the expression of mouse phospholipase A<sub>2</sub> (sPLA<sub>2</sub>-X) was evaluated as a fusion to various SUMO tags (Smt3, hSUMO3, SUMOstar) in HEK-293T cells (46). SUMOstar displayed a significant increase of sPLA<sub>2</sub>-X expression compared with both Smt3 and hSUMO3 (Figure 2).

The expression of a phospholipase, such as sPLA<sub>2</sub>-X would be expected to be toxic to cells, however when fused to SUMO the toxicity should be reduced as the native N-terminus (normally required for the phospholipase activity) is unavailable for co-ordination with the active site. When fused with native SUMO tags, sPLA<sub>2</sub>-X showed moderate toxicity, whereas expression of sPLA<sub>2</sub>-X as a SUMOstar fusion appeared to protect cells from death.

This is presumed to be due to the inability of endogenous SUMO proteases to cleave the fusion protein, and subsequently generate free sPLA<sub>2</sub>-X.

Therefore, SUMOstar appears to be resistant to digestion by endogenous SUMO proteases, allowing for the use of this fusion technology for enhanced production of recombinant proteins in eukaryotic systems.

## SUMMARY

Over the past few decades, the development and use of numerous fusion tag technologies, has become common place in the production of recombinant proteins. Although these technologies have demonstrated some level of success, often the addition of a fusion tag still does not allow for sufficient amounts of recombinant material to be produced. In addition, these technologies are subject to major limitations with respect to efficient fusion tag removal, non-specific cleavage, and the ability to generate the POI with a native N-terminus. Not only has the use of SUMO as fusion partner been demonstrated to increase protein solubility and prevent proteolytic degradation, but the SUMO system allows for the efficient removal of the tag, leaving a recombinant POI with its native N-terminus. This is accomplished through the use of SUMO protease 1 for tag removal, which specifically recognizes both the C-terminal sequence "Gly-Gly" and tertiary structure of SUMO. Our SUMO constructs have the added advantage of containing an N-terminal His<sub>6</sub> tag, allowing for straightforward affinity purification of the fusion construct, as well as final purification of the POI from the SUMO tag and SUMO protease after digestion. We recently extended our studies to generate SUMOstar, a new version of SUMO, which is resistant to endogenous SUMO protease but retains the ability to enhance protein expression in both mammalian and insect cell systems. In summary, SUMO and

SUMOstar fusion systems provide robust recombinant protein expression and purification in both prokaryotic and eukaryotic hosts. SUMOpro and SUMOstar fusion technologies, available from LifeSensors, Inc., represent invaluable tools for superior recombinant protein expression, and have the potential to help facilitate the furthering of scientific research and therapeutic drug discovery.

## REFERENCES AND NOTES

1. J. Arnau et al., "Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins", *Protein Expr Purif.* **48**(1), pp. 1-13 (2006).
2. A.M. Edwards et al., "Protein production: feeding the crystallographers and NMR spectroscopists", *Nat Struct Biol.* **7** Suppl, pp. 970-972 (2000).
3. F. Baneyx, M. Mujacic, "Recombinant protein folding and misfolding in *Escherichia coli*", *Nat Biotechnol.* **22**(11), pp. 1399-1408 (2004).
4. L. Borsig, E.G. Berger et al., "Expression and purification of His-tagged beta-1,4-galactosyltransferase in yeast and in COS cells", *Biochem Biophys Res Commun.* **240**(3), pp. 586-589 (1997).
5. R. Bhikhabhai et al., "Production of milligram quantities of affinity tagged-proteins using automated multistep chromatographic purification", *J Chromatogr A* **1080**(1), pp. 83-92 (2005).
6. R. Daly, M.T. Hearn, "Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production", *J Mol Recognit* **18**(2), pp. 119-138 (2005).
7. B. Prinz et al., "Establishing a versatile fermentation and purification procedure for human proteins expressed in the yeasts *Saccharomyces cerevisiae* and *Pichia pastoris* for structural genomics", *J Struct Funct Genomics* **5**(1-2), pp. 29-44 (2004).
8. P. Braun, J. LaBaer, "High throughput protein production for functional proteomics", *Trends Biotechnol.* **21**(9), pp. 383-388 (2003).
9. B. Miroux, J.E. Walker, "Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels", *J Mol Biol.* **260**(3), pp. 289-298 (1996).
10. U. Brinkmann, R.E. Mattes et al., "High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the dnaY gene product", *Gene* **85**(1), pp. 109-114 (1989).
11. W.A. Prinz et al., "The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm", *J Biol Chem.* **272**(25), pp. 15661-16667 (1997).
12. T.R. Butt et al., "SUMO fusion technology for difficult-to-express proteins", *Protein Expr Purif.* **43**(1), pp. 1-9 (2005).
13. M. Hochstrasser, "Evolution and function of ubiquitin-like protein-conjugation systems", *Nat Cell Biol.* **2**(8), pp. E153-157 (2000).
14. S. Jentsch, G. Pyrowolakis, "Ubiquitin and its kin: how close are the family ties?", *Trends Cell Biol.* **10**(8), pp. 335-342 (2000).
15. S. Muller et al., "SUMO, ubiquitin's mysterious cousin", *Nat Rev Mol Cell Biol.* **2**(3), pp. 202-210 (2001).
16. P. Boyer et al., "Structure determination of the small ubiquitin-related modifier SUMO-1", *J Mol Biol.* **280**(2), pp. 275-286 (1998).
17. K.I. Kim, S.H. Baek et al., "Chung, Versatile protein tag, SUMO: its enzymology and biological function", *J Cell Physiol.* **191**(3), pp. 257-268 (2002).
18. E.S. Johnson et al., "The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer", *Embo J.* **16**(18), pp. 5509-5519 (1997).
19. J.M. Desterro, M.S. Rodriguez et al., "SUMO-1 modification of IκappaBα inhibits NF-κappaB activation", *Mol Cell.* **2**(2), pp. 233-239 (1998).
20. E.S. Johnson, G. Blobel, "Ubc9p is the conjugating enzyme for the ubiquitin-like protein Smt3p", *J Biol Chem.* **272**(43), pp. 26799-26802 (1997).
21. Y. Shiio, R.N. Eisenman, "Histone sumoylation is associated with transcriptional repression", *Proc Natl Acad Sci U S A* **100**(23), pp. 13225-13230 (2003).
22. M.J. Matunis, "On the road to repair: PCNA encounters SUMO and ubiquitin modifications", *Mol Cell* **10**(3), pp. 441-442 (2002).
23. L.D. Wood et al., "Small ubiquitin-like modifier conjugation regulates nuclear export of TEL, a putative tumor suppressor", *Proc Natl Acad Sci U S A* **100**(6), pp. 3257-3262 (2003).
24. T.T. Huang et al., "Sequential modification of NEMO/IKKγ by SUMO-1 and ubiquitin mediates NF-κappaB activation by genotoxic stress", *Cell* **115**(5), pp. 565-576 (2003).
25. G. Zheng, Y.C. Yang, "ZNF76, a novel transcriptional repressor

- targeting TATA-binding protein, is modulated by sumoylation", *J Biol Chem*. **279**(41), pp. 42410-24241 (2004).
26. T. Kadoya et al., "Desumoylation activity of Axam, a novel Axin-binding protein, is involved in downregulation of beta-catenin", *Mol Cell Biol*. **22**(11), pp. 3803-3819 (2002).
  27. E.T. Yeh, L. Gong et al., "Ubiquitin-like proteins: new wines in new bottles", *Gene* **248**(1-2), pp. 1-14 (2000).
  28. S.J. Li, M. Hochstrasser, "A new protease required for cell-cycle progression in yeast", *Nature* **398**(6724), pp. 246-251 (1999).
  29. S.J. Li, M. Hochstrasser, "The yeast ULP2 (SMT4) gene encodes a novel protease specific for the ubiquitin-like Smt3 protein", *Mol Cell Biol*. **20**(7), pp. 2367-2377 (2000).
  30. D. Reverter, C.D. Lima, "Structural basis for SENP2 protease interactions with SUMO precursors and conjugated substrates", *Nat Struct Mol Biol*. **13**(12), pp. 1060-1068 (2006).
  31. Mossessova, E. and C.D. Lima, Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol Cell* 2000. **5**(5): p. 865-76.
  32. Waugh, D.S., Making the most of affinity tags. *Trends Biotechnol*, 2005. **23**(6): p. 316-20.
  33. Malakhov, M.P., et al., SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *J Struct Funct Genomics*, 2004. **5**(1-2): p. 75-86.
  34. Kapust, R.B. and D.S. Waugh, Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci*, 1999. **8**(8): p. 1668-74.
  35. Davis, G.D., et al., New fusion protein systems designed to give soluble expression in *Escherichia coli*. *Biotechnol Bioeng*, 1999. **65**(4): p. 382-8.
  36. Marblestone, J.G., et al., Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO. *Protein Sci*, 2006. **15**(1): p. 182-9.
  37. Zuo, X., et al., Expression and purification of SARS coronavirus proteins using SUMO-fusions. *Protein Expr Purif*, 2005. **42**(1): p. 100-10.
  38. Guzzo, C.M. and Yang, D.C.H., Systematic analysis of fusion and affinity tags using human aspartyl-tRNA synthetase expressed in *E. coli*, *Protein Expr Purif*, 2007. **54**(1): p. 166-175.
  39. Balbas, P., Understanding the art of producing protein and nonprotein molecules in *Escherichia coli*. *Mol Biotechnol*, 2001. **19**(3): p. 251-67.
  40. Jenny, R.J., et. al., A critical review of the methods for cleavage of fusion proteins with thrombin and factor Xa, *Protein Expr Purif*, 2003. **31**(1): p. 1-11.
  41. Kapust, R.B. et. al., The P1' specificity of tobacco etch virus protease, *Biochem Biophys Res Commun*, 2002. **294**(5): p. 949-955.
  42. Schechter, I. and Berger, A., On the size of the active site in proteases. I. Papain, *Biochem Biophys Res Commun*, 1967. **27**(2): p. 157-162.
  43. Lu, Q., et. al., Optimized procedures for producing biologically active chemokines, *Protein Expr Purif*, 2009. **65**(2): p. 251-260.
  44. Bosse-Doenecke, E., et. al., High yield production of recombinant native and modified peptides exemplified by ligands for G-protein coupled receptors, *Protein Expr Purif*, 2008. **58**(1): p. 114-121.
  45. Assadi-Porter, F., et.al., Efficient and rapid protein expression and purification of small high disulfide containing sweet protein brazzein in *E. coli*, *Protein Expr Purif*, 2008. **58**(2): p. 263-268.
  46. Peroutka, R.J., et al., Enhanced protein expression in mammalian cells using engineered SUMO fusions: secreted phospholipase A2. *Protein Sci*, 2008. **17**(9): p. 1586-95.
  47. Liu, L., et al., Enhanced protein expression in the baculovirus/insect cell system using engineered SUMO fusions. *Protein Expr Purif*, 2008.

Readers interested in a complete list of references are kindly invited to write to the author at: [stickler@lifesensors.com](mailto:stickler@lifesensors.com)

YILIN YAN, STEVEN J. ORCUTT, JAMES E. STRICKLER\*

\*Corresponding author  
Lifesensors Inc.  
271 Great Valley Parkway  
Malvern, PA 19355, USA