Title

Rapid addition of unlabeled silent solubility tags to proteins using a new substrate-fused sortase reagent

Journal

Journal of Biomolecular NMR, 64(3)

ISSN

0925-2738

Authors

Amer, Brendan R
Macdonald, Ramsay
Jacobitz, Alex W
et al.

Publication Date

2016-03-01

DOI

10.1007/s10858-016-0019-z

Peer reviewed

HHS Public Access

Author manuscript

J Biomol NMR. Author manuscript; available in PMC 2016 November 15.

Published in final edited form as:

J Biomol NMR. 2016 March ; 64(3): 197–205. doi:10.1007/s10858-016-0019-z.

Rapid Addition of Unlabeled Silent Solubility Tags to Proteins Using a New Substrate-Fused Sortase Reagent

Brendan R. Amer[1,2], **Ramsay Macdonald**[1,2], **Alex W. Jacobitz**[1,2], **Brandon Liauw**[1,2], and **Robert T. Clubb**[1,2,3,4]

[1]Department of Chemistry and Biochemistry, University of California, Los Angeles, 602 Boyer Hall, Los Angeles, CA 90095, USA

[2]UCLA-DOE Institute of Genomics and Proteomics, University of California, Los Angeles, 611 Charles Young Drive East, Los Angeles, CA 90095, USA

[3]Molecular Biology Institute, University of California, Los Angeles, 611 Charles Young Drive East, Los Angeles, CA 90095, USA

Abstract

Many proteins can't be studied using solution NMR methods because they have limited solubility. To overcome this problem, recalcitrant proteins can be fused to a more soluble protein that functions as a solubility tag. However, signals arising from the solubility tag hinder data analysis because they increase spectral complexity. We report a new method to rapidly and efficiently add a non-isotopically labeled Small Ubiquitin-like Modifier protein (SUMO) solubility tag to an isotopically labeled protein. The method makes use of a newly developed SUMO-Sortase tagging reagent in which SUMO and the Sortase A (SrtA) enzyme are present within the same polypeptide. The SUMO-Sortase reagent rapidly attaches SUMO to any protein that contains the sequence LPXTG at its C-terminus. It modifies proteins at least 15-times faster than previously described approaches, and does not require active dialysis or centrifugation during the reaction to increase product yields. In addition, silently tagged proteins are readily purified using the well-established SUMO expression and purification system. The utility of the SUMO-Sortase tagging reagent is demonstrated using PhoP and green fluorescent proteins, which are ~90% modified with SUMO at room temperature within four hours. SrtA is widely used as a tool to construct bioconjugates. Significant rate enhancements in these procedures may also be achieved by fusing the sortase enzyme to its nucleophile substrate.

Keywords

Silent solubility tag; sortase; protein ligation; SUMO

NMR spectroscopy is a powerful tool in which to study the structure and dynamics of proteins in solution (Cavanagh et al. 2010). However, in order to perform detailed studies, proteins must be soluble and stable in NMR compatible buffers at concentrations of

[4]To whom correspondence should be addressed: Dept. of Chemistry and Biochemistry, University of California, Los Angeles, 602 Boyer Hall, Los Angeles, CA 90095. Tel.: 310-206-2334; Fax: 310-206-4749; rclubb@mbi.ucla.edu. .

~250-500 μM or higher. Moreover, even higher protein concentrations are preferred as the signal-to-noise scales linearly with concentration. A significant problem limiting the study of many biologically interesting proteins is their limited solubility. Although a range of buffer conditions can be explored to improve protein solubility, many times these efforts prove unsuccessful. Fusing recalcitrant proteins to a more soluble domain (e.g. thioredoxin, maltose binding protein (MBP), protein G B1 domain (GB1), and glutathione-S-transferase domains (GST)) can dramatically improve overall solubility and are frequently used to facilitate expression and purification (di Guan et al. 1988; Smith and Johnson 1988; Huth et al. 1997; LaVallie et al. 2000). Wagner and colleagues pioneered the use of protein "solubility tags" for NMR studies by demonstrating that the solubility and overall spectral quality of a poorly behaving protein can be improved by fusing it to GB1 (Zhou et al. 2001; Zhou and Wagner 2009). However, the NMR spectra of the GB1 fusion contained additional, unwanted signals from the solubility tag (GB1) that increased spectral complexity. To avoid this problem, newer "silent" solubility tagging methods have been developed in which the isotopically enriched protein of interest is first purified, and then ligated *in vitro* to an unlabeled, more soluble protein. Because the added solubility domain is not isotopically labeled, it is NMR "silent" and does not increase the spectral complexity (Züger and Iwai 2005; Kobashigawa et al. 2009). Similar, *in vitro* ligation approaches are employed to segmentally isotope label multi-domain containing proteins, thereby reducing their spectral complexity (Yamazaki et al. 1998; Muona et al. 2010; Refaei et al. 2011; Freiburger et al. 2015).

Two approaches have been developed to attach silent solubility tags to proteins, intein trans-splicing and sortase catalyzed transpeptidation. In the intein trans-splicing method, split inteins self-associate to catalyze the splicing event (Yamazaki et al. 1998; Xu et al. 1999; Züger and Iwai 2005; Muona et al. 2010). Silent tagging is performed *in vitro*, and requires both the target protein and solubility tag first be expressed and purified as a split-intein fusion. These fusion proteins are then joined via trans-splicing in which the split-intein components are eliminated. Although high yields are obtainable, reaction times of 1-2 days, and special protein reagents (the split-intein fusions) are required. A second approach uses the *S. aureus* sortase (SrtA) enzyme that joins via a peptide bond the protein of interest to the silent solubility tag. SrtA is a cysteine transpeptidase that catalyzes peptide bond formation between the threonine residue within the sequence LPXTG (where X is any amino acid), and the amino group of a peptide that contains a penta-glycine (Gly_5) sequence at its N-terminus (Mazmanian 1999; Perry et al. 2002; Spirig et al. 2011; Schneewind and Missiakas 2014). In elegant work by Kobashigawa and colleagues, SrtA was used to append a silent tag to the C-terminus of an isotopically labeled protein (Kobashigawa et al. 2009). The reaction required three purified protein components: the unlabeled solubility tag containing an N-terminal Gly_5 sequence (Gly_5-GB1 in their study), the isotopically enriched target protein containing a C-terminal LPXTG sequence, and the SrtA transpeptidase. Using this approach, ~90% of the isotope labeled target protein was silently tagged with GB1 by performing the reaction for 3 days at room temperature. To improve yields the reaction was performed during dialysis to facilitate cleavage product removal. More recently, Sattler and colleagues demonstrated the utility of this approach for segmental labeling, and achieved

faster modification rates by removing the product during the reaction by centrifugal concentration (Freiburger et al. 2015).

Our objective was to create a more efficient and rapid solubility tagging method that satisfied the following criteria. First, proteins should be rapidly modified with the tag within a day, preferably in a process that can be conducted at lower temperatures to avoid protein aggregation and/or degradation of the isotopically labeled protein. Second, >90% of the labeled protein should be converted into the silently tagged protein product. This is critical as the target protein may be difficult to obtain and/or expensive to produce. Third, the tagging reaction should be a simple process that requires a minimum number of reagents, and should require only minor modification of the protein of interest. Finally, the procedure should be readily integrated into an established protein expression and purification scheme, enabling silently tagged proteins to be purified using conventional approaches.

We developed a SUMO-tagging system that silently tags poorly soluble proteins with an unlabeled Small Ubiquitin-like Modifier (SUMO) protein. SUMO was chosen because of its low molecular weight (12 kDa), and because it has previously been shown to improve protein solubility (Malakhov et al. 2004; Marblestone et al. 2006; Panavas et al. 2009; Peroutka Iii et al. 2011). Moreover, using SUMO as a solubility tag enables the new system to be integrated into the widely used, and commercially available, SUMO affinity tag purification (LifeSensors (Malakhov et al. 2004, Catalog no. 1001K) or ThermoFisher (Catalog no. K300-01)). In the SUMO affinity tag purification procedure, proteins are produced as fusion with an N-terminal 6xHis-SUMO affinity tag (Fig. 1a). The 6xHis-SUMO-Protein fusion is purified using standard immobilized metal affinity chromatography (IMAC) resin (e.g. Ni^{2+} or Co^{2+} columns). The 6xHis-SUMO tag is then removed by adding the highly specific 6xHis-Ulp1 protease, and the protein of interest is then purified by reapplying the mixture to the IMAC column. The protein of interest appears in the flow-through, while 6xHis-Ulp1 protease and 6xHis-SUMO are retained on the column. The SUMO-tagging system described here is readily integrated into this purification scheme because it makes use of 6xHis-SUMO-Protein fusion, as well as the robust 6xHis-Ulp1 protease. In our procedure, the isotopically enriched protein is expressed as a 6xHis-SUMO-Protein fusion, an unlabeled SUMO tag is then added in a single step to the protein, and the desired Protein-SUMO fusion is generated using the conventional SUMO affinity tag purification scheme (Fig. 1b). This approach is advantageous, as it simplifies procedures required to obtain the final product and because at all points in the process the protein remains fused to a solubility tag so as to minimize losses caused by aggregation.

The SUMO-tagging system is based on the SrtA-mediated approach developed by Kobashigawa and colleagues, but simplifies the process by creating a single SUMO-SrtA reagent that silently tags any protein containing the sequence LPXTG at its C-terminus (Kobashigawa et al. 2009). Moreover, because SrtA and its SUMO substrate are located within the same polypeptide, the rate of modification (“tagging”) is significantly increased (*vide infra*). The tagging reagent consists of a SUMO-SrtA fusion protein harboring five glycine residues at its N-terminus, and a histidine tag at its C-terminus (Gly_5-SUMO-SrtA-6xHis). Distinct from previous reported approaches, the ligation reaction requires only two components, the new tagging reagent and a purified protein that contains a C-terminal

LPXTG sequence. To facilitate integration into SUMO affinity tag purification system, the protein to be tagged is produced as a 6xHis-SUMO-Protein-LPETG fusion (Fig. 1b). When the tagging reagent and fusion protein are incubated with one another, the reagent joins itself via a peptide bond to the C-terminal end of 6xHis-SUMO-Protein-LPETG to create a 6xHis-SUMO-Protein-LPET-Gly_5-SUMO-SrtA-6xHis ligation product. The ligation mixture is then incubated with 6xHis-Ulp1 protease, releasing the N-terminal SUMO tag (6xHis-SUMO), the SrtA enzyme (SrtA-6xHis), and the desired tagged protein product (Protein-LPET-Gly_5-SUMO). Because only the product lacks a histidine tag, it is purified by immobilized metal affinity chromatography (IMAC).

The utility of the SUMO-tagging system was initially demonstrated using the PhoP protein from *M. tuberculosis.* PhoP is a two-component response regulator that binds DNA through its C-terminal domain (PhoPC, residues 142–247of PhoP) (Pathak et al. 2010; Macdonald et al. 2015). PhoPC was chosen because the isolated domain exhibits limited solubility in NMR compatible buffers, and it requires high salt concentrations; PhoPC is initially soluble in buffer containing 50 mM sodium phosphate pH 6.5 and 300 mM NaCl, but the protein begins to aggregate and precipitate to below ~500 μM after ~24-48 hours. To tag PhoPC, a 6xHis-SUMO-PhoPC-LPETG fusion protein was expressed from a commercially available pSUMO plasmid and purified by IMAC (Fig 1a). The protein was then incubated at 25°C for varying amounts time with unlabeled tagging reagent (Gly_5-SUMO-SrtA-6xHis) at a 1:5 molar ratio of protein to reagent. As shown in Fig. 2a, approximately 90% of PhoPC is converted into the ligated product within 4 hours. We also verified the robustness of the modification procedure using Green-fluorescent protein (GFP) containing the sequence LPETG at its C-terminus (GFP-LPETG). As shown in Fig. 2b, when it is incubated with Gly_5-SUMO-SrtA-6xHis using similar reaction conditions, ~90% of GFP-LPETG is ligated to the tagging reagent within five hours. For some proteins, it may be desirable to add silent solubility tags using lower reaction temperatures to avoid protein aggregation and/or proteolysis. Fig. 2c shows that this can readily be accomplished using the reagent, as it only requires 18 hours to convert ~90% of GFP-LPETG into the tagged product when the reaction is performed at 4°C (1:5 ratio molar ratio of protein to reagent).

An attractive feature of the SUMO-tagging system is that the silently tagged protein product can readily be purified for NMR studies using the SUMO affinity tag protein purification system (Fig. 1). This was demonstrated by silently tagging U-[^{15}N]-PhoPC with unlabeled SUMO protein. The U-[^{15}N]-PhoPC protein was produced using the pSUMO vector as a His-SUMO-PhoPC-LPETG fusion and purified using IMAC (Fig. 1a). The protein was then modified using the tagging reagent and processed by adding the Ulp1 protease, which selectively cleaves the peptide bond between residues in the Gly-Ser (or Gly-Gly) dipeptide sequence located at the C-terminus of SUMO. Ulp1 cleaves the ligation product at two sites, one cleavage event removes the labeled N-terminal SUMO domain, while the second cleavage removes SrtA. This releases the desired silently tagged U-[^{15}N]-PhoPC-[^{14}N]-SUMO protein, and the histidine tagged U-[^{15}N]-SUMO and SrtA proteins. Because U-[^{15}N]-PhoPC-[^{14}N]-SUMO is not histidine tagged, it is readily separated from the other reactants by IMAC; U-[^{15}N]-PhoPC-[^{14}N]-SUMO does not bind the IMAC column in the absence of imidazole, while the histidine tagged U-[^{15}N]-SUMO and SrtA reaction components bind the column. Fig. 3 shows that silent tagging eliminates spectral complexity

caused by the solubility tag. The ^{1}H-^{15}N HSQC spectrum of the U-[^{15}N]-SUMO-PhoPC-LPETG protein obtained prior to silent tagging contains signals arising from both PhoPC and the N-terminal SUMO domain, whereas only signals from PhoPC are observed in the silently tagged U-[^{15}N]-PhoPC-LPET-[^{14}N]-Gly_5-SUMO protein (signals enclosed by a rectangle in fig. 3 demonstrate the removal of SUMO resonances after the silent tagging reaction). Importantly, addition of the C-terminal silent SUMO tag to PhoPC enables protein concentrations in excess of 1 mM to be obtained, while halving the amount of NaCl that is required.

The SUMO-tagging reagent reported here silently tags proteins at least 15-times faster than a previously described method that used the isolated SrtA enzyme; using our reagent ~90% reaction yields are obtained in 4 hours, whereas reaction times of three days are required when the isolated enzyme is employed (Kobashigawa et al. 2009). The accelerated rate of modification presumably occurs because the effective concentration of the Gly_5 nucleophile near the enzyme active site in SrtA enzyme is increased because it is located within the same polypeptide (Krishnamurthy et al. 2007). This is beneficial, as it overcomes SrtA's intrinsically weak affinity for the Gly_5 nucleophile (K_m = 140 μM) (Chen et al. 2011). To directly explore the utility of fusing the enzyme to its nucleophile substrate, we performed a conventional three component SrtA-catalyzed reaction in which the reactants and SrtA were located on separate polypeptides. The isolated enzyme (SrtA-6xHis), and Gly_5-SUMO and GFP-LPETG substrates were mixed at a molar ratio of 5:5:1 at 25°C (Fig. 2d). In this three component reaction, only ~50% of the GFP-LPETG substrate is converted into product after 6 hours, a modification rate that is significantly slower than the rate obtained using the SUMO-tagging reagent that contains the nucleophile substrate fused to SrtA (compare figs 1b and 2d). In the SUMO-SrtA reagent, the component proteins are connected by a ~12 amino acid linker that is presumably structurally disordered (Ilangovan et al. 2001; Naik et al. 2006). Interestingly, this connector is critical for activity, as a deletion mutant of the reagent that removes the connector has greatly diminished activity (data not shown). This is presumably because in the reagent the connector segment is conformationally flexible, enabling the Gly_5 nucleophile located at the N-terminal end of the SUMO domain in the SUMO-SrtA reagent to enter the active site of SrtA. Based on the established reaction mechanism of SrtA, the reagent can be expected to first cleave the LPXTG component of the substrate to form a thio-acyl bond to the target protein, which is then resolved by enhanced intramolecular nucleophilic attack by the Gly_5 component of the reagent. In principle, even faster reaction rates may be obtainable by optimizing the connector segment to further increase the effective concentration of the Gly_5 component of the reagent near the active site. Moreover, further optimization may be beneficial, since even though it is possible to produce 40 milligrams of pure Gly_5-SUMO-SrtA-6xHis reagent per liter of culture, the connector segment within the reagent is susceptible to proteolytic degradation necessitating a three-step purification procedure. However, it should be stressed, that once the reagent is purified it is stable for ~2-3 weeks when stored at −20°C (or −80°C for longer).

In this paper we described a new reagent that efficiently adds unlabeled SUMO solubility enhancement tags to proteins that are prone to aggregate at concentrations required for NMR studies. There are several advantages to our approach as compared to previously reported methods. First, proteins are tagged more rapidly, with ~90% of the target protein modified

within ~4 hours, versus 2-3 days when other methods are used. Additionally, we have demonstrated that efficient tagging can be achieved at lower temperatures (4°C). Second, it requires minimal modification of the target protein as only the LPXTG sequence needs to be added, unlike other methods that require the production of split intein protein fusions. Third, the new method is designed to be used with the well-established SUMO expression and purification system (fig. 1). This facilitates purification of the tagged product and enables the potential utility of solubility tagging to be easily assessed without having to modify the protein of interest. This is because the NMR spectra of the 6xHis-SUMO-Protein intermediate that is produced during the standard SUMO expression and purification procedure are expected to be similar to the spectra of the final silently tagged protein, albeit the spectra of the intermediate will contain additional signals from SUMO. Thus, if the spectra of 6xHis-SUMO-Protein intermediate are of good quality it is worthwhile to perform the silent tagging reaction. Finally, as compared to previously described methods that use SrtA, the new method requires fewer protein components in the reaction (two instead of three), and rapid modification is obtained without the need for dialysis or centrifugation during the ligation reaction (Kobashigawa et al. 2009; Refaei et al. 2011; Freiburger et al. 2015). In addition to silent tagging, SrtA has been successfully used to label cells, add proteins to solid supports, and to generate antibody conjugates, nucleic acid-protein fusions, and PEGylation/Lipidated proteins (Parthasarathy et al.; Chan et al. 2007; Pritz et al. 2007; Antos et al. 2008; Sakamoto et al. 2010; Levary et al. 2011). In principle, the approach described here could be used to significantly increase the rate and extent of modifications in these procedures by fusing the nucleophile substrate to the sortase enzyme.

Methods

Preparation of the SUMO-SrtA tagging reagent and substrates

The SUMO-tagging reagent (Gly_5-SUMO-SrtA-6xHis) used in this study contained SUMO connected to the N-terminus of residues 59-206 of SrtA. Appended to the N- and C-terminus of this polypeptide was the sequence GGGGG (Gly5) and HHHHHH (6xHis). The plasmid overexpressing the SUMO-SrtA tagging reagent was generated by traditional restriction digest cloning. This was accomplished by ligating the SrtA gene insert into the pSUMO vector (LifeSensors) using *BamH1* and *XhoI* restriction enzymes and using the Quikchange mutagenesis reaction to add N-terminal Gly_5 sequence (Agilent). The vector encoding the reagent produces the following protein:
MGGGGGSSEHEHEHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVS
DGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPEDLDMEDN
DIIEAHREQIGGSQAKPQIPKDKSKVAGYIEIPDADIKEPVYPGPATPEQLNRGVSFAEE
NESLDDQNISIAGHTFIDRPNYQFTNLKAAKKGSMVYFKVGNETRKYKMTSIRDVK
PTDVGVLDEQKGKDKQLTLITCDDYNEKTGVWEKRKIFVATEVKLEHHHHHH. The sequence of the polypeptide that connects the SUMO and SrtA is underlined. The GFP and PhoPC substrates that were modified by the SUMO-SrtA tagging reagent were expressed from the pSUMO plasmid and contained the amino acid sequence LPETGGEST at their C-termini introduced using the Quikchange mutagenesis method. Here we include the sequence of the SUMO-PhoPC-LPETG construct utilized in these experiments:
MGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSS

EIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPEDLDMEDNDIIEA
HREQIGGSKGNKEPRNVRLTFADIELDEETHEVWKAGQPVSLSPTEFTLLRYFVINAG
TVLSKPKILDHVWRYDFGGDVNVVESYVSYLRRKIDTGEKRLLHTLRGVGYVLREP
RLPETGGEEST.

All proteins were expressed in appropriately transformed *Escherichia coli* (*E. coli*) BL21 (DE3) cells. The unlabeled U-[^{14}N]-Gly$_5$-SUMO-SrtA tagging reagent cultures were grown at 37°C to an OD600 of 0.6 before induction with isopropyl β-D-thiogalactopyranoside (IPTG) (Goldbio) to a final concentration of 1 mM in LB media. Induction proceeded at 18°C overnight before harvesting the cells by centrifugation at 7000g for 10 min at 4°C. The cell pellet was then resuspended in lysis buffer consisting of: 50 mM Tris-HCl, 300 mM NaCl (pH 8.0) containing protease inhibitor cocktail (Calbiochem) and phenylmethanesulfonyl fluoride (PMSF) (Sigma). Cells were lysed by sonication and centrifuged at 15,000g for 50 min at 4°C. The supernatant was then incubated with 6.0 mL of pre-equilibrated Co^{2+} resin (Thermo) for 10 min at 4°C on a rotisserie before being transferred to a gravity column. The resin was then washed with 100 mL lysis buffer, and the protein was finally eluted by adding lysis buffer containing 150 mM imidazole. Fractions containing the Gly$_5$-SUMO-SrtA Δ59 tagging reagent were then buffer exchanged via concentrator (Amicon) into 50 mM MES pH6.0 buffer for HiTrap SP HP ion exchange chromatography (GE Healthcare). The Gly$_5$-SUMO-SrtA reagent was eluted from the 5 mL SP HP column by a linear gradient over 40mins into 50 mM MES pH6.0, 1 M NaCl. A final gel-filtration (Superdex 75pg, GE Healthcare) step is carried out to remove any further contaminants and buffer exchange into the ligation reaction buffer, 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$. The SUMO-GFP-LPETG and SUMO-PhoPC-LPETG proteins were expressed and purified in a similar fashion. However they did not require further downstream ion exchange or gel-filtration chromatography, a one-step Co^{2+} affinity purification yielded highly pure protein. The proteins were buffer exchanged into ligation reaction buffer, 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$ by dialysis.

Ligation Reactions

All ligation reactions were performed in 1.7 mL Eppendorf tubes rotating at either room temperature (25°C) or in the cold at 4°C. Reactions were performed in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$ buffer and contained 5:1 molar ratio of the SUMO-SrtA tagging reagent to labelled target protein. Typically, proteins concentrations of ~125 μM and ~25 μM were used for SUMO-SrtA and its substrate protein, respectively. The progress of the reactions was followed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and the band intensity quantified by densitometry using the program ImageJ (NIH). Post-ligation purification was carried out within 24 hours of the reaction. Briefly, the ligated product of the reaction, U-[^{15}N] SUMO-Protein-LPET-[^{14}N]-Gly$_5$-SUMO-SrtA, was dialyzed into 50 mM Tris-HCl pH 8.0, 150 mM NaCl buffer and Ulp1 protease was added at a 25:1 molar ratio. After incubation overnight at 4°C, the final ligated product was obtained by application to a Co^{2+} affinity column. It is important to note that in the absence of a nucleophile substrate the sortase enzyme functions as a protease and can cleave the final protein product (U-[^{15}N] SUMO-Protein-LPET-[^{14}N]-Gly$_5$-SUMO-SrtA) unless the enzyme is removed. For our reactions we found that the product was stable for 24-36 hours in the

presence of SUMO-tagging reagent (Gly_5-SUMO-SrtA-6xHis) and that proteolysis of the substrate was not a problem if the ligated product was immediately processed by the addition of Ulp1 protease and IMAC chromatography. However, it is important to note that the ligated product can readily be stabilized by adding a sortase inhibitor (p-hydroxymecuribenzoic acid, Sigma) after the ligation reaction is complete, as this small molecule inhibits the sortase enzyme but does not affect the activity of Ulp1 protease.

NMR data acquisition and processing

^{1}H-^{15}N-HSQC NMR experiments were performed at 298 K on a Bruker Avance HD 600-MHz spectrometer and Bruker Avance II 500-MHz spectrometer equipped with triple resonance cryogenic probes. NMR spectra were processed using NMRPipe (Delaglio et al. 1995).

Acknowledgements

We would like to thank Dibyendu Sakar (Institute of Microbial Technology, India) for the plasmid encoding PhoP and members of the Clubb group for useful discussion. B.R.A was supported by a Whitcome Pre-doctoral training grant from the UCLA Molecular Biology Institute, and R.M. was supported by a Cellular and Molecular Biology Training Grant (Ruth L. Kirschstein National Research Service Award GM007185). This work was supported by the National Institutes of Health grant AI52217 to RTC. As well as support through the US Department of Energy Office of Science, Office of Biological and Environmental Research program under Award Number DE-FC02-02ER63421.

References

Antos JM, Miller GM, Grotenbreg GM, Ploegh HL. Lipid modification of proteins through sortase-catalyzed transpeptidation. J Am Chem Soc. 2008; 130:16338–43. DOI: 10.1021/ja806779e [PubMed: 18989959]

Cavanagh, J.; Fairbrother, WJ.; Palmer, AG., et al. Protein NMR Spectroscopy: Principles and Practice. Academic Press; 2010.

Chan L, Cross HF, She JK, et al. Covalent attachment of proteins to solid supports and surfaces via Sortase-mediated ligation. PLoS One. 2007; 2:e1164.doi: 10.1371/journal.pone.0001164 [PubMed: 18000537]

Chen I, Dorr BM, Liu DR. A general strategy for the evolution of bond-forming enzymes using yeast display. Proc Natl Acad Sci U S A. 2011; 108:11399–404. DOI: 10.1073/pnas.1101046108 [PubMed: 21697512]

Delaglio F, Grzesiek S, Vuister GW, et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR. 1995; 6:277–93. [PubMed: 8520220]

di Guan C, Li P, Riggs PD, Inouye H. Vectors that facilitate the expression and purification of foreign peptides in Escherichia coli by fusion to maltose-binding protein. Gene. 1988; 67:21–30. [PubMed: 2843437]

Freiburger L, Sonntag M, Hennig J, et al. Efficient segmental isotope labeling of multi-domain proteins using Sortase A. J Biomol NMR. 2015; 63:1–8. DOI: 10.1007/s10858-015-9981-0 [PubMed: 26319988]

Huth JR, Bewley CA, Jackson BM, et al. Design of an expression system for detecting folded protein domains and mapping macromolecular interactions by NMR. Protein Sci. 1997; 6:2359–64. DOI: 10.1002/pro.5560061109 [PubMed: 9385638]

Ilangovan U, Ton-That H, Iwahara J, et al. Structure of sortase, the transpeptidase that anchors proteins to the cell wall of Staphylococcus aureus. Proc Natl Acad Sci U S A. 2001; 98:6056–61. DOI: 10.1073/pnas.101064198 [PubMed: 11371637]

Kobashigawa Y, Kumeta H, Ogura K, Inagaki F. Attachment of an NMR-invisible solubility enhancement tag using a sortase-mediated protein ligation method. J Biomol NMR. 2009; 43:145–50. DOI: 10.1007/s10858-008-9296-5 [PubMed: 19140010]

Krishnamurthy VM, Semetey V, Bracher PJ, et al. Dependence of effective molarity on linker length for an intramolecular protein-ligand system. J Am Chem Soc. 2007; 129:1312–20. DOI: 10.1021/ja066780e [PubMed: 17263415]

LaVallie ER, Lu Z, Diblasio-Smith EA, et al. Thioredoxin as a fusion partner for production of soluble recombinant proteins in Escherichia coli. Methods Enzymol. 2000; 326:322–40. [PubMed: 11036651]

Levary DA, Parthasarathy R, Boder ET, Ackerman ME. Protein-protein fusion catalyzed by sortase A. PLoS One. 2011; 6:e18342.doi: 10.1371/journal.pone.0018342 [PubMed: 21494692]

Macdonald R, Sarkar D, Amer BR, Clubb RT. Solution structure of the PhoP DNA-binding domain from Mycobacterium tuberculosis. J Biomol NMR. 2015; 63:111–7. DOI: 10.1007/s10858-015-9965-0 [PubMed: 26209027]

Malakhov MP, Mattern MR, Malakhova OA, et al. SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. J Struct Funct Genomics. 2004; 5:75–86. DOI: 10.1023/B:JSFG.0000029237.70316.52 [PubMed: 15263846]

Marblestone JG, Edavettal SC, Lim Y, et al. Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO. Protein Sci. 2006; 15:182–9. DOI: 10.1110/ps.051812706 [PubMed: 16322573]

Mazmanian SK. Staphylococcus aureus Sortase, an Enzyme that Anchors Surface Proteins to the Cell Wall. Science. 1999; 285:760–763. (80-). DOI: 10.1126/science.285.5428.760 [PubMed: 10427003]

Muona M, Aranko AS, Raulinaitis V, Iwaï H. Segmental isotopic labeling of multi-domain and fusion proteins by protein trans-splicing in vivo and in vitro. Nat Protoc. 2010; 5:574–87. DOI: 10.1038/nprot.2009.240 [PubMed: 20203672]

Naik MT, Suree N, Ilangovan U, et al. Staphylococcus aureus Sortase A transpeptidase. Calcium promotes sorting signal binding by altering the mobility and structure of an active site loop. J Biol Chem. 2006; 281:1817–26. DOI: 10.1074/jbc.M506123200 [PubMed: 16269411]

Panavas T, Sanders C, Butt TR. SUMO fusion technology for enhanced protein production in prokaryotic and eukaryotic expression systems. Methods Mol Biol. 2009; 497:303–17. DOI: 10.1007/978-1-59745-566-4_20 [PubMed: 19107426]

Parthasarathy R, Subramanian S, Boder ET. Sortase A as a novel molecular “stapler” for sequence-specific protein conjugation. Bioconjug Chem. 18:469–76. DOI: 10.1021/bc060339w

Pathak A, Goyal R, Sinha A, Sarkar D. Domain structure of virulence-associated response regulator PhoP of Mycobacterium tuberculosis: role of the linker region in regulator-promoter interaction(s). J Biol Chem. 2010; 285:34309–18. DOI: 10.1074/jbc.M110.135822 [PubMed: 20814030]

Peroutka RJ Iii, Orcutt SJ, Strickler JE, Butt TR. SUMO fusion technology for enhanced protein expression and purification in prokaryotes and eukaryotes. Methods Mol Biol. 2011; 705:15–30. DOI: 10.1007/978-1-61737-967-3_2 [PubMed: 21125378]

Perry AM, Ton-That H, Mazmanian SK, Schneewind O. Anchoring of surface proteins to the cell wall of Staphylococcus aureus. III. Lipid II is an in vivo peptidoglycan substrate for sortase-catalyzed surface protein anchoring. J Biol Chem. 2002; 277:16241–8. DOI: 10.1074/jbc.M109194200 [PubMed: 11856734]

Pritz S, Wolf Y, Kraetke O, et al. Synthesis of biologically active peptide nucleic acid-peptide conjugates by sortase-mediated ligation. J Org Chem. 2007; 72:3909–12. DOI: 10.1021/jo062331l [PubMed: 17432905]

Refaei MA, Combs A, Kojetin DJ, et al. Observing selected domains in multi-domain proteins via sortase-mediated ligation and NMR spectroscopy. J Biomol NMR. 2011; 49:3–7. DOI: 10.1007/s10858-010-9464-2 [PubMed: 21188472]

Sakamoto T, Sawamoto S, Tanaka T, et al. Enzyme-mediated site-specific antibody-protein modification using a ZZ domain as a linker. Bioconjug Chem. 2010; 21:2227–33. DOI: 10.1021/bc100206z [PubMed: 21069999]

Schneewind O, Missiakas D. Sec-secretion and sortase-mediated anchoring of proteins in Gram-positive bacteria. Biochim Biophys Acta. 2014; 1843:1687–97. DOI: 10.1016/j.bbamcr.2013.11.009 [PubMed: 24269844]

Smith DB, Johnson KS. Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. Gene. 1988; 67:31–40. [PubMed: 3047011]

Spirig T, Weiner EM, Clubb RT. Sortase enzymes in Gram-positive bacteria. Mol Microbiol. 2011; 82:1044–59. DOI: 10.1111/j.1365-2958.2011.07887.x [PubMed: 22026821]

Xu R, Ayers B, Cowburn D, Muir TW. Chemical ligation of folded recombinant proteins: Segmental isotopic labeling of domains for NMR studies. Proc Natl Acad Sci. 1999; 96:388–393. DOI: 10.1073/pnas.96.2.388 [PubMed: 9892643]

Yamazaki T, Otomo T, Oda N, et al. Segmental Isotope Labeling for Protein NMR Using Peptide Splicing. J Am Chem Soc. 1998; 120:5591–5592. DOI: 10.1021/ja980776o

Zhou P, Lugovskoy AA, Wagner G. A solubility-enhancement tag (SET) for NMR studies of poorly behaving proteins. J Biomol NMR. 2001; 20:11–4. [PubMed: 11430750]

Zhou P, Wagner G. Overcoming the solubility limit with solubility-enhancement tags: successful applications in biomolecular NMR studies. J Biomol NMR. 2009; 46:23–31. DOI: 10.1007/s10858-009-9371-6

Züger S, Iwai H. Intein-based biosynthetic incorporation of unlabeled protein tags into isotopically labeled proteins for NMR studies. Nat Biotechnol. 2005; 23:736–40. DOI: 10.1038/nbt1097 [PubMed: 15908942]

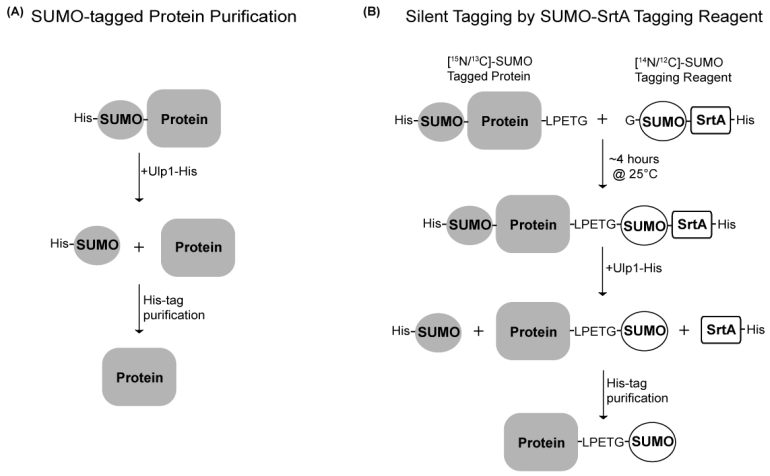

Figure 1.

Schematic of a typical 6xHis-SUMO-Protein purification and of the Silent Tagging reaction by SUMO-SrtA

(A) A schematic of a typical SUMO fusion purification **(B)** A schematic of the Silent Tagging reaction by SUMO-SrtA. The target protein is expressed in isotopically enriched media producing a labeled fusion protein. The target protein is expressed as a SUMO fusion protein with a C-terminal sortase, LPETG, recognition motif, aiding in overall expression and solubility. The unlabeled one-reagent SUMO-SrtA fusion tagging reagent will rapidly recognize the LPETG motif and cleave between the carbonyl of the threonine and glycine residues present in the motif. This reaction is resolved when the sortase Gly_5 secondary substrate nucleophile, fused to the N-terminus of the unlabeled SUMO, enters the active site and the enzyme performs a transpeptidation reaction yielding the amino terminus of the penta-glycine motif appended to the LPET motif present in the target protein. This ligation mixture is then incubated with Ulp1, the SUMO protease, releasing both the labelled SUMO solubility tag, and the sortase enzyme from the ligation product. This allows for a rapid purification of the ligation product by cobalt (or nickel) affinity purification.

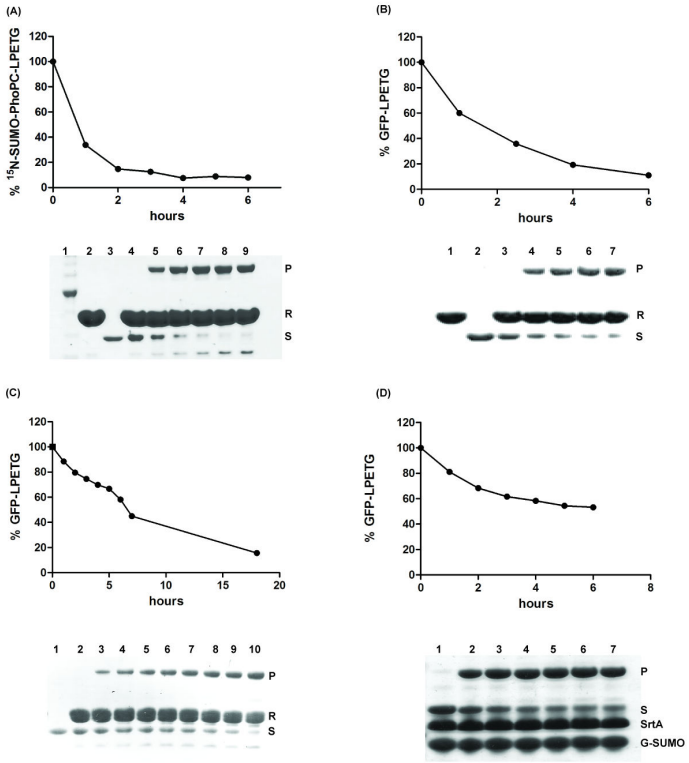

Figure 2.

A. ImageJ analysis (top) of the SDS-PAGE results (bottom) of conversion of ^{15}N-SUMO-PhoPC-LPETG into ^{15}N-SUMO-PhoPC-LPETG-^{14}N-SUMO. This reaction was carried out as a 5:1 mixture of the SUMO-SrtA single tagging reagent to ^{15}N-SUMO-PhoPC-LPETG resulting in ~90% conversion to product in 4 hours at room temperature (25°C). P, ligation product, R, SUMO-SrtA reagent, and S, SUMO-PhoPC-LPETG substrates. Lane 1, molecular weight ladder, lane 2, SUMO-SrtA tagging reagent, lane 3, ^{15}N-SUMO-PhoPC-LPETG control, lane 4, 5:1 reagent: ^{15}N-SUMO-PhoPC-LPETG ligation reaction 0 hour, lane 5, 1 hour, lane 6, 2 hours, lane 7, 3 hours, lane 8, 4 hours, lane 9, 6 hours. **B.** ImageJ analysis (top) of the representative SDS-PAGE results (bottom) indicate that the 5:1 incubation of the SUMO-SrtA ligation reagent with GFP-LPETG results in the conversion of approximately 90% of the initial substrate into ligated product in approximately 5 hours at room temperature (25°C). Lane 1, SUMO-SrtA tagging reagent, lane 2, GFP-LPETG, lane 3, 0 hour reaction, lane 4, 1 hour, lane 5, 2.5 hour, lane 6, 4 hours, lane 7, 6 hours. **C.** ImageJ analysis (top) of the SDS-PAGE results (bottom) indicate that the 5:1 incubation of the SUMO-SrtA ligation reagent with GFP-LPETG results in the conversion of approximately 90% of the initial substrate into ligated product in approximately 18 hours at low temp (4°C). Reactions done in the cold prevent degradation of the ligated product, tagging reagent, and target protein. Lane 1, GFP-LPETG control, lane 2, 5:1 reagent:GFP-LPETG ligation reaction 0 hour, lane 3, ligation reaction 1 hour, lane 4, 2 hours, lane 5, 3 hours, lane 6, 4 hours, lane 7, 5 hours, lane 8, 6 hours, lane 9, 7 hours, lane 10, 18 hours. **D.** 3 component reaction does not have high yield or efficiency. ImageJ analysis (top) of the SDS-PAGE results (bottom) of conversion of GFP-LPETG into ligation product. This reaction was carried out as a 5:5:1 mixture of the Gly_5-SUMO and SrtA-6xHis to GFP-

LPETG resulting in ~55% conversion to product in 6 hours at room temperature (25°C). Lane 1, reaction 0 hour, lane 2, reaction 1 hour, lane 3, reaction 2 hours, lane 4, 3 hour, lane 5, 4 hours, lane 6, 5 hours, lane 7, 6 hours.

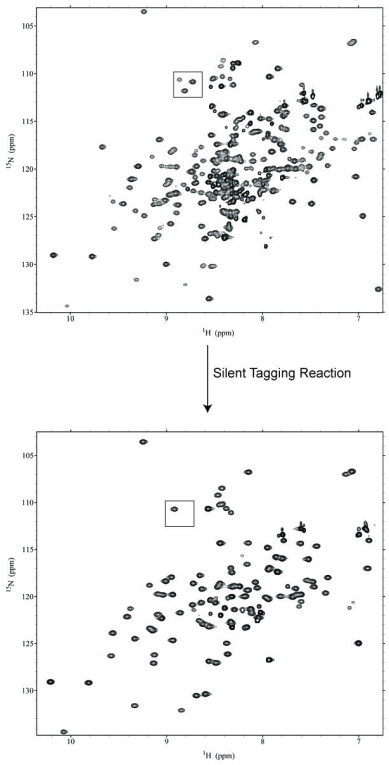

Figure 3.

HSQCs of SUMO tagged ^{15}N-PhoPC. ^{15}N-SUMO-PhoPC-LPETG prior to the SUMO-SrtA silent ligation reaction (top), and after SUMO-SrtA silent tagging reaction (bottom). The boxed regions show regions where signals from the 15N-SUMO tag are removed by replacing it with the silent SUMO tag.