



Published in final edited form as:

Curr Protoc Protein Sci. 2009 April ; CHAPTER 15: Unit–15.3. doi:10.1002/0471140864.ps1503s56.

Site-Specific Protein Labeling via Sortase-Mediated Transpeptidation

John M. Antos¹, Jessica Ingram², Tao Fang³, Novalia Pishesha^{3,4}, Matthias C. Truttmann³, and Hidde L. Ploegh³

¹Department of Chemistry, Western Washington University, Bellingham, WA 98225

²Department of Cancer Immunology and Virology, Dana Farber Cancer Institute, Boston, MA 02215

³Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA 02115

⁴Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02142

Abstract

Strategies for site-specific protein modification are highly desirable for the construction of conjugates containing non-genetically encoded functional groups. Ideally, these strategies should proceed under mild conditions, and be compatible with a wide range of protein targets and non-natural moieties. The transpeptidation reaction catalyzed by bacterial sortases is a prominent strategy for protein derivatization that possesses these features. Naturally occurring or engineered variants of sortase A from *Staphylococcus aureus* catalyze a ligation reaction between a five amino acid substrate motif (LPXTG) and oligoglycine nucleophiles. By pairing proteins and synthetic peptides that possess these ligation handles, it is possible to install modifications onto the protein N- or C-terminus in site-specific fashion. As described in this unit, the successful implementation of sortase-mediated labeling involves straightforward solid-phase synthesis and molecular biology techniques, and this method is compatible with proteins in solution or on the surface of live cells.

Keywords

sortase; transpeptidation; site-specific labeling; chemoenzymatic labeling

INTRODUCTION

Protein conjugates possessing non-genetically encoded functional groups are fundamentally important tools for the biochemical sciences. Through the careful addition of non-natural moieties, one is able to access unique protein derivatives that enable activities ranging from the development of novel therapeutics to the study of complex biochemical processes. With these applications in mind, this unit describes the use of sortase-catalyzed transpeptidation for the site-specific modification of protein termini. This chemoenzymatic method, which involves a ligation between a five amino acid recognition motif (LPXTG) and an oligoglycine nucleophile, is suitable for modifying proteins in solution, in complex lysates, or even on the surface of living cells (Mao et al., 2004; Popp et al., 2007; Parthasarathy et

al., 2007; Tanaka et al., 2008; Ritzefeld, 2014; Schmohl and Schwarzer, 2014). This experimentally straightforward approach is also notable for its mild reaction conditions, as well as its capacity to tolerate a wide range of protein targets and modifications. To enable implementation of sortase-based labeling, this unit includes multiple protocols covering the design and preparation of all necessary reagents, and a protocol for the sortase-catalyzed ligation itself. In addition, a supplemental protocol for using sortase to modify protein targets on the surface of live cells is included.

STRATEGIC PLANNING

Selection of Sortase Enzyme

The majority of sortase-based protein labeling is reliant on sortase A from *Staphylococcus aureus* (SrtA_{staph}) or related mutants (Ritzefeld, 2014; Schmohl and Schwarzer, 2014; Antos et al., 2016; Popp and Ploegh, 2011). As such, this unit will focus on this subset of enzymes. In all cases, these sortases recognize substrates displaying a five amino acid substrate motif (LPXTG), cleaving between the threonine and glycine residues to generate a thioester-linked acyl enzyme intermediate (Figure 1). The acyl enzyme is then attacked by an oligoglycine nucleophile (typically three to five glycines), resulting in site-specific ligation of the substrate and nucleophile via a native amide bond. If the nucleophile is omitted, the acyl enzyme intermediate will hydrolyze under the aqueous reaction conditions. Currently, two engineered variants of SrtA_{staph} containing either five (SrtA_{staph} pentamutant) or seven (SrtA_{staph} heptamutant) point mutations have risen to prominence as the preferred enzymes for this process by offering key advantages over the wild-type enzyme. These include enhanced reaction rates (penta- and heptamutant) and lack of the requirement for a Ca²⁺ cofactor (heptamutant) in the reaction mixture (Chen et al., 2011; Hirakawa et al., 2015; Witte et al., 2015; Wuethrich et al., 2014). Given these features, the authors recommend that initial labeling studies employ either of these engineered SrtA_{staph} derivatives, and complementary labeling protocols for the use of either the penta- or heptamutant are provided below. Protocols for sortase expression are also included here, and expression plasmids for both enzymes are conveniently available through the online gene repository Addgene, or via request from the authors.

Design of Proteins Compatible with Sortase-Mediated Transpeptidation

Proteins compatible with sortase-mediated labeling are accessible using standard molecular biology and cloning techniques. The specific design features required depend on which terminus of the protein is being targeted for modification. For C-terminal labeling, the protein target will serve as the initial sortase substrate, and so the LPXTG motif is simply inserted onto the C-terminal end of the protein (Figure 2A). Most amino acids are tolerated in the X position of the substrate motif, though glutamic acid (E) is frequently used in this position due to its prevalence in *in vivo* sortase A substrates (Boekhorst et al., 2005; Litou et al., 2008). The accessibility of the LPXTG motif is also critical for successful transpeptidation, and so it is common to add a short, flexible linker composed of Gly₄Ser repeats between the body of the protein and the LPXTG site (Wagner et al., 2014; Chen et al., 2015; Heck et al., 2014). Whether this linker is required, and the exact of length of the linker sequence, is protein dependent and must be determined empirically. Other design

features that are important include not placing the LPXTG motif at the very C-terminus of the protein target. Best results are obtained when at least one additional residue is in amide linkage with the final glycine. In practice, this is typically achieved by adding an affinity purification handle (e.g. His₆) following the LPXTG site. The loss of this tag following transpeptidation provides not only a convenient means of monitoring the efficiency of labeling by immunoblotting and/or a Coomassie stained gel, but can also be leveraged for protein conjugate purification (Levary et al., 2011; Antos et al., 2008; Swee et al., 2013; Massa et al., 2016).

With respect to N-terminal labeling, the target protein will serve as the oligoglycine nucleophile, and therefore at least one N-terminal glycine residue is inserted at the protein N-terminus (Figure 2B) (Antos et al., 2009a; Yamamoto and Nagamune, 2009; Hess et al., 2012; Williamson et al., 2014; Schoonen et al., 2015). As in the case of C-terminal labeling, accessibility of the reactive terminus is critical, and so multiple glycines (3–5 residues) are often utilized to ensure optimal results. For substrates produced in *E. coli*, glycines should be placed immediately following the initiator methionine, which should be removed by methionylaminopeptidase to liberate a free aminoglycine terminus (Hirel et al., 1989). In the event that methionine removal is incomplete, protease cleavage sites that generate a free glycine N-terminus can provide an alternate strategy. Appropriate proteases for this purpose include thrombin, TEV, Factor Xa, or the SUMO protease Ulp1 (Waugh, 2011; Malakhov et al., 2004).

One final note about protein target design concerns derivatives with both a C-terminal LPXTG motif and N-terminal glycines. If the end goal is selective modification of a specific protein terminus, then such dual-functionalized derivatives should be avoided as they may be susceptible to intramolecular transpeptidation, giving rise to circular protein products (Antos et al., 2009b; Jia et al., 2014; van 't Hof et al., 2015). In certain cases, proteins containing a C-terminal LPXTG site and an N-terminal glycine may be amenable to site-specific modification at both sites. Examples of controlled modification at both protein termini using a sortase-based approach have been reported, and will not be discussed further here (Antos et al., 2009a; Wuethrich et al., 2014).

Design of Peptide Probes Compatible with Sortase-Mediated Transpeptidation

As in the case of the design principles for compatible protein targets, the structural requirements for the associated peptide probes depends on whether these materials are intended for C-terminal or N-terminal modification. In the case of C-terminal ligations, LPXTG-containing proteins are paired with a synthetic oligoglycine nucleophile consisting of at least one N-terminal glycine covalently tethered to the modification of choice (Figure 2A). While a single glycine can be sufficient for successful transpeptidation (Row, et al. 2015; Baer et al., 2014; Alt et al., 2015), most nucleophiles contain three to five glycine residues (for examples see Policarpo et al., 2014; Fang et al., 2016; Hui et al., 2015; Rashidian et al., 2015). This more closely mimics the pentaglycine nucleophile utilized *in vivo* by *S. aureus* sortase A (Ton-That et al., 1997), and presumably prevents the peptide payload from obstructing the reaction of the peptide amine terminus with the acyl enzyme intermediate. It should also be noted that alternate primary amine (Parthasarathy et al., 2007;

Baer et al., 2014; Glasgow et al., 2016) or hydrazine nucleophiles (Li et al., 2014) have been successfully used in transpeptidation reactions, however these non-glycine nucleophiles are generally less successful than glycine derivatives.

For N-terminal labeling, protein targets with one or more N-terminal glycines are ligated to peptides containing the LPXTG motif (Figure 2B). The modification of choice is positioned on the N-terminus of the LPXTG peptide probe. Like the LPXTG-containing proteins described above, LPXTG peptides often contain glutamic acid in the X position, and possess an additional residue following the LPXTG sequence (e.g. LPXTGG) or terminate in a primary amide (e.g. LPXTG-NH₂) to promote optimal reactivity (for examples see Hess et al., 2013; Yamamoto and Nagamune, 2009; Swee et al., 2015; Schoonen et al., 2015).

An attractive feature of sortase-based methods is that peptide probes for both N-terminal and C-terminal labeling are conveniently available via standard solid-phase peptide synthesis. The exceptional versatility of the technique arises from the ability to decorate these probes with a variety of non-genetically encoded functional groups. To this end, numerous commercial sources provide access to non-natural amino acid building blocks and amine- or thiol-reactive reagents that can be readily incorporated into a solid-phase synthesis scheme. As examples, synthetic protocols for both N-terminal and C-terminal labeling probes are presented in this unit. While we recommend manual peptide synthesis as a cost-effective means for obtaining the required peptide materials, suitably modified peptide materials may also be obtained via commercial peptide synthesis or university core facilities.

BASIC PROTOCOL

Site-Specific Protein Labeling via Sortase-Mediated Transpeptidation (N- or C-terminal Labeling)—This protocol describes the labeling of proteins that have been purified prior to the transpeptidation reaction. Generally, this involves target proteins that can be produced and isolated from *E. coli* or other standard expression hosts. Progress of the reaction is best monitored in accordance with the peptide probe to be installed; immunoblotting can be used for probes with a biotin moiety, while in-gel fluorescence scanning is convenient for probes containing a fluorophore. SDS-PAGE with Coomassie staining may also be used as transpeptidation products often exhibit mobilities that are distinct from the unlabeled input protein. Mass spectrometry provides an additional alternative for reaction monitoring. This protocol also takes advantage of engineered sortase variants that show increased activity relative to wild type SrtA_{staph}. This includes the SrtA_{staph} pentamutant, which exhibits increased activity but still requires the Ca²⁺ cofactor (Chen et al., 2011), and the SrtA_{staph} heptamutant, which also has increased activity but does not require the presence of Ca²⁺ for transpeptidation (Hirakawa et al., 2015; Witte et al., 2015; Wuethrich et al., 2014).

Materials

General

- 1) Purified sortase A pentamutant or heptamutant stock solution (see Support Protocol

10x sortase reaction buffer: pentamutant (500 mM Tris-HCl, pH 7.5, 1.5 M NaCl, 100 mM CaCl₂), heptamutant (500 mM Tris-HCl, pH 7.5, 1.5 M NaCl or 10x phosphate buffered saline (PBS))

1.5 mL microcentrifuge tubes

4 °C or 25 °C incubator

Zeba desalting column (ThermoFisher)

Centrifugal concentrator with MWCO below the molecular weight of target protein (Millipore)

Additional reagents and equipment for SDS-PAGE, immunoblot and Coomassie stain.

C-terminal Labeling

Purified LPXTG-containing target protein (cannot be in phosphate containing buffer when using pentamutant variant of sortase)

Oligoglycine peptide probe stock solution (5–10 mM in DMSO or H₂O)

Ni-NTA column (*optional*)

Imidazole (*optional*)

N-terminal Labeling

Purified target protein containing 1–5 N-terminal glycine residues (cannot be in phosphate containing buffer when using pentamutant variant of sortase)

LPXTGG peptide probe stock solution (5–10 mM in DMSO or H₂O)

1. Mix the following in a 1.5-mL microcentrifuge tube such that the final concentrations are as follows:

C-terminal Labeling

50 μM purified target protein (possessing C-terminal LPXTG motif)

2.5 μM sortase A (pentamutant or heptamutant)

250 μM oligoglycine peptide probe

1x sortase reaction buffer (1:10 dilution of 10x sortase reaction buffer in the final reaction solution)

N-terminal Labeling

50 μM purified target protein (possessing 1–5 N-terminal glycine residues)

2.5 μM sortase A (pentamutant or heptamutant)

250 μM LPXTGG peptide probe

1x sortase reaction buffer (1:10 dilution of 10x sortase reaction buffer in the final reaction solution)

The choice of reaction buffer depends on whether pentamutant or heptamutant sortase is used. See Materials above for buffer components and concentrations.

2. Incubate the reaction for 2 hrs at 4 °C or 30 min at 25 °C. Analyze by SDS-PAGE (Coomassie staining or in-gel fluorescence), immunoblotting, or mass spectrometry.

The reaction conditions described in steps 1–2 provide a suitable starting point for optimizing the labeling protocol. Slightly different reaction rates are observed for each purified target protein, and consequently the reaction parameters may require optimization to achieve high levels of protein labeling. Often, the concentrations of sortase and peptide probe may be reduced without sacrificing labeling efficiency. Reactions can be carried out at 4 °C or 25 °C, depending on target protein stability.

3. Remove excess peptide probe using commercial desalting columns (Zeba) or spin concentrators. If required, additional purification of the labeled protein may be achieved via size-exclusion or ion exchange chromatography.
4. (Optional step for C-terminal labeling) If a His₆ tag has been included C-terminal to the LPXTG motif in the target protein, then unlabeled target protein and His₆-tagged sortase may be conveniently removed by passing the crude reaction mixture over a Ni-NTA column. The column flow-through will contain the C-terminally labeled protein. To minimize non-specific binding, the reaction mixture can be diluted in 1x sortase reaction buffer supplemented with imidazole (10–30 mM final imidazole concentration) prior to exposure to the Ni-NTA resin.

ALTERNATE PROTOCOL

LABELING CELL-SURFACE PROTEINS IN LIVING CELLS VIA SORTASE-MEDIATED TRANSEPTIDATION

—In this protocol, we describe strategies for labeling proteins displayed on the surface of live cells using the Ca²⁺-independent sortase A heptamutant. This procedure is suitable for targets with extracellularly exposed N-termini (type I membrane proteins) or C-termini (type II membrane proteins). To allow labeling of type I membrane proteins, the target should be engineered to display 1–5 glycines at its N-terminus. For type II membrane proteins, the sortase recognition motif (LPXTG) is fused at the C-terminus. Once expressed on the target cell, these proteins can be labeled by exposing them to sortase and the appropriate peptide probe in tissue culture medium. After washing to remove excess peptide probe and sortase, labeled cells can be directly imaged, characterized by flow cytometry, or subjected to other analyses.

Materials

Target cells

Plasmid encoding target protein (type II membrane protein with C-terminal LPXTG motif or type I membrane protein with 1–5 N-terminal glycines)

Transfection reagent (Lipofectamine, Invitrogen; Trans IT, Mirus; FuGENE 6, Promega)

Plastic tissue culture dishes

Culture medium such as DMEM (presence of 10% serum does not inhibit the sortase reaction)

10x sortase reaction buffer: 500 mM Tris-HCl, pH 7.5, 1.5 M NaCl

1 mM purified sortase A heptamutant stock solution

10 mM peptide probe stock in DMSO or H₂O (oligoglycine probe for labeling type II membrane proteins, LPXTGG probe for labeling type I membrane proteins)

1.5 mL microcentrifuge tubes

Ice-cold phosphate-buffered saline (PBS)

Additional reagents and equipment for microscopy, flow cytometry, and immunoblot

1. Transfect target cells with a plasmid encoding the engineered target protein according to transfection reagent manufacturer's directions. Cells may be cultured and transfected on plastic dishes, subsequently detached and collected in 1.5 mL microcentrifuge tubes for labeling.
2. Collect $\sim 1 \times 10^6$ cells and resuspend in 70 μ L of DMEM. Place cell suspension on ice.
3. Label cells by preparing the following reaction mixtures:

N-terminal labeling of type I membrane proteins

Combine 10 μ L of 1 mM heptamutant sortase, 5 μ L of 10 mM LPXTGG peptide probe, 5 μ L of 10x sortase reaction buffer, and 10 μ L of sterile H₂O. Mix well and let the mixture incubate on ice for 10 minutes prior to adding this mixture to the cells suspended in DMEM.

C-terminal labeling of type II membrane proteins

Combine 10 μ L of 1 mM heptamutant sortase, 5 μ L of 10 mM oligoglycine probe, 5 μ L of 10x sortase reaction buffer, and 10 μ L of sterile H₂O. Mix well and directly add to the cells suspended in DMEM.

4. Incubate the cell suspension and the appropriate sortase reaction mixture on ice for 45 min with occasional gentle mixing.

As in the case of labeling purified proteins in solution (see Basic Protocol), the above conditions provide an appropriate starting point for further optimization. Sortase and probe concentrations, as well as reaction times, should be varied to achieve the desired level of labeling. In particular, peptide probe concentrations may be lowered without significant loss of labeling efficiency, though this must be determined empirically. Sortase-mediated cell labeling may also be carried out at higher temperatures (room temperature or 37 °C), though ice-cold temperatures are often preferable if endocytosis or other membrane trafficking

events affect the target protein of interest. Finally, to label a higher or lower number of cells, simply adjust the volume of reactions proportionally.

5. Wash cells at least three times with 1 mL portions of ice-cold PBS.

This step removes both sortase and excess probe. Extensive washing is necessary to remove unbound probe and decrease the background signal for microscopy applications.

6. Add either phenol-red free tissue culture medium or PBS to labeled cells, and replat to observe by microscopy. Labeled cells can also be resuspended, treated with relevant antibody or affinity reagents, and analyzed by flow cytometry. Alternatively, cell lysates may be characterized using SDS-PAGE and immunoblot.

SUPPORT PROTOCOL 1

Expression and Purification of SrtA_{staph} Pentamutant and Heptamutant—

Several mutant versions of SrtA_{staph} are now available, including those with increased catalytic activity and lack of dependence on a Ca²⁺ cofactor (Chen et al., 2011; Gianella et al., 2016; Hirakawa et al., 2015; Witte et al., 2015; Chen et al., 2016). In nearly all cases, the N-terminal transmembrane domain present in the wild-type enzyme has been truncated to improve solubility. Early in the development of sortase-mediated protein labeling, transpeptidations were often performed with a version of wild-type SrtA_{staph} lacking the first 59 amino acids (Ilango et al., 2001). The SrtA_{staph} pentamutant and heptamutant utilized in the Basic and Alternate Protocols of this unit are based on this original 59 amino acid deletion construct. Both mutant versions can be obtained in large quantities (20–40 mg/L) following expression and purification from *E. coli*. These enzymes also possess C-terminal His₆ tags to aid in purification.

Materials

Sortase pentamutant or heptamutants expression plasmid: pET30b (available from Addgene or via request from the authors)

E. coli BL-21(DE3)

Terrific broth (TB) medium with and without appropriate antibiotic:

Kanamycin (1000x stock = 30 mg/mL in H₂O)

1 M isopropyl β-D-thiogalactopyranoside (IPTG) in H₂O

Culture plates and tubes

4 L culture flask

37 °C incubator with shaking

Spectrophotometer and appropriate cuvettes

Lysis buffer: 50 mM Tris-HCl, pH 8, 150 mM NaCl, 20 mM imidazole, 10% glycerol (v/v)

10 mg/mL DNaseI in H₂O

French press (pre-chilled) or probe sonicator (Branson Sonifier 450)

Ni-NTA agarose slurry (Qiagen)

1.5×12 cm disposable polypropylene column (Biorad)

Elution buffer: 50 mM Tris HCl, pH 8, 150 mM NaCl, 350 mM imidazole, 10% glycerol (v/v)

Centrifugal concentrator with low <10 kD MWCO (optional, Millipore)

1. Transform sortase expression plasmid into *E. coli* BL-21 (DE3) and plate on selective medium.
2. Pick a single colony and inoculate 25 mL of TB containing the appropriate antibiotic. Grow overnight as a starter culture.
3. Add 25 mL of the overnight culture into 1 L of TB with antibiotics in a 4 L culture flask. Grow with agitation to ensure proper aeration. Monitor the OD₆₀₀ until 0.6. Induce by adding IPTG to a final concentration of 1 mM and shake overnight at 30 °C.
4. Harvest the bacterial pellet by centrifuging for 20 min at 6000 × g, 4 °C. Decant TB and resuspend pellet in 35 mL of ice-cold lysis buffer supplemented with 20 µg/mL DNase I.

Do not add protease inhibitors to lysis buffer as these may interfere with sortase activity.
5. Lyse bacteria by passing through a pre-chilled French press cell two times at 1250 psi. Alternatively, lyse by sonication (four cycles of 1 min with duty cycle set to 50%, and output control set to 4). Chill samples in between sonication cycles to avoid excessive warming.
6. Clarify the lysate by centrifuging for 30 min at 12,000 × g, 4 °C.
7. Pack 2–5 mL of Ni-NTA agarose slurry into a polypropylene column and wash with 10 slurry volumes of lysis buffer. Apply the clarified supernatant to the column and allow it to flow through. Wash column with 50 slurry volumes of lysis buffer.
8. Elute sortase with 2.5 slurry volumes of elution buffer.
9. (Optional) Dialyze two times against 4 L of lysis buffer without imidazole to remove excess imidazole.
10. (Optional) Concentrate the protein further in a centrifugal concentrator with a low-molecular-weight cutoff (<10 kDa).

The sortase constructs described are soluble and have been concentrated to 1 mM without signs of aggregation. Stock solutions can be stored for up to 1 month at 4 °C, or several months at –80 °C without loss of activity.

SUPPORT PROTOCOL 2

Synthesis of Oligoglycine Probes for C-terminal Labeling and LPXTG Probes for N-terminal Labeling

—Sortase-based methods for protein labeling impose very few limitations on modifications that can be tethered to the oligoglycine or LPXTG peptide probes. In a case-by-case manner, readers may find more efficient synthetic routes to access desired probes based on the available orthogonal protecting groups or the special structural requirement of payloads to be conjugated. Nevertheless, in this section, we will introduce three representative examples describing the synthesis of peptide substrates compatible with N- or C-terminal transpeptidation reactions using solid-phase synthesis and established peptide conjugation chemistries. Procedure **A** describes the solid-phase synthesis of GGGK(TAMRA), a triglycine-derived organic fluorophore for C-terminal protein labeling (see Figure 2C). Procedure **B** also outlines the synthesis of a C-terminal labeling probe, in this case via solution-phase labeling of a GGGC peptide with widely available maleimide-derived payloads through thiol-Michael addition. Finally, Procedure **C** describes an N-terminal labeling probe synthesized via solution-phase conjugation of an LPETGG peptide to amine-reactive NHS esters (see Figure 2C for a related example). All procedures require minimum instrumentation and should be feasible to laboratories not specialized in peptide or synthetic chemistry.

Materials

Solid-phase peptide synthesis vessel with associated caps and filter frits

Wrist-Action shaker or similar equipment for mixing solid-phase reaction vessels

Rink amide resin, 100–200 mesh, 1% (w/w) divinylbenzene (DVB)

Fmoc protected amino acids: Fmoc-Gly₃-OH; Fmoc-Cys(Trt)-OH; Fmoc-Lys(Mtt)-OH; Fmoc-Leu-OH; Fmoc-Pro-OH; Fmoc-Glu(tBu)-OH; Fmoc-Thr(tBu)-OH; Fmoc-Gly-OH

HBTU (2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate)

PyBOP (benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate)

DMF (*N,N*-dimethylformamide)

DIPEA (diisopropylethylamine)

TFA (trifluoroacetic acid)

TIPS (triisopropylsilane)

EDT (1,2-ethanedithiol)

DCM (dichloromethane)

TAMRA (5(6)-carboxytetramethylrhodamine)

Maleimide- or NHS-derived modification

Phosphate buffered saline (without Ca²⁺ or Mg²⁺)

HPLC system with C18 reversed-phase column

Lyophilizer

Instrument and consumables for liquid chromatography/mass spectrometry (LC-MS)

Procedure A: Solid-phase synthesis of GGGK(TAMRA)

1. Add 0.2 mmol of Rink amide resin to a peptide synthesis vessel and solvate the resin by adding DCM (10 mL). Gently shake the reaction vessel in a wrist-action shaker at RT for 15 min and remove the DCM by vacuum filtration.
2. Add 20% (v/v) piperidine solution in DMF (5 mL) to the resin and shake for 30 min at RT. Remove the solution by vacuum filtration and wash the resin with DMF (10 mL) three times with filtration and occasional shaking in between.
3. Dissolve Fmoc-Lys(Mtt)-OH (1 mmol), HBTU (1 mmol) and DIPEA (2 mmol) in DMF (5–10 mL) and add the mixture to the resin. Shake the suspension for 2 h at RT.
4. Remove the reaction solution by vacuum filtration, wash resin three times with DMF (5 mL).
5. Repeat step 2 to remove the Fmoc group.
6. Repeat step 3, substituting Fmoc-Gly₃-OH in place of Fmoc-Lys(Mtt)-OH.
7. Remove the reaction solution by vacuum filtration, wash the resin three times with DMF (5 mL).
8. Add a solution of 1% TFA and 2% TIPS in DCM (v/v) (5 mL) to selectively remove the 4-methyl trityl (Mtt) protecting group from the previously coupled Lys residue. Incubate the resin with this cleavage solution for 30 min at RT. Remove the cleavage solution via vacuum filtration. Add an additional 5 mL of MTT cleavage solution and incubate for 30 minutes to ensure removal of the MTT group.

The colorimetric Kaiser test, a very sensitive method to detect the presence of primary amines, can be used to confirm the successful deprotection of MTT. Similarly, this test can be used after each amino acid coupling step to ensure reaction completion.

9. Wash the resin with three 5 mL portions of DCM followed by three 5 mL portions of DMF.
10. Add a solution of 5(6)-TAMRA (0.2 mmol), PyBOP (0.2 mmol) and DIPEA (0.4 mmol) in DMF (7 mL) to the resin and shake overnight at RT. Avoid exposure to light and keep reaction vessel covered when possible.

If the dye is readily available, the amount of TAMRA may be increased to improve coupling efficiency. Whether this is done or not, the non-coupled peptide will ultimately be removed during HPLC purification.
11. Remove the reaction solution by vacuum filtration, then wash the resin three times with DMF (5 mL).

12. Remove the terminal Fmoc group by repeating step 2. Include two additional washes with DCM (5 mL) prior to cleavage of the peptide from the resin.
13. Cleave the peptide from the resin by treatment with a solution consisting of 95% (v/v) TFA, 2.5% (v/v) H₂O and 2.5% (v/v) TIPS (5 mL) for 3 h at RT.
14. Elute the cleavage solution into 35 mL of ice-cold diethyl ether, and rinse the resin with an additional 3 mL of the cleavage solution into the ether.
15. Spin down the precipitate at 1500 rcf or higher for 10 min. Remove the supernatant, and evaporate the excess diethyl ether under a constant air flow for ~20 min in a fume hood.
16. Redissolve the pellet in water. A small aliquot can be sent for mass spectrometry analysis. Purify the remaining material by reversed-phase HPLC on a C18 column.
17. Lyophilize the pure peptide fractions and redissolve as needed for sortase-catalyzed protein labeling (see Basic Protocol or Alternate Protocol)

Probes can be stored as solutions in DMSO at -20 °C for several months without loss of activity.

Procedure B: Synthesis of GGGC probes via maleimide coupling

1. Manually synthesize the GGGC peptide using the solid-phase technique outlined in Procedure A.

Proper protected Fmoc amino acid building blocks are listed in the Materials for Supporting Protocol 2. Alternatively, this short peptide is readily available from commercial sources or core facilities dedicated to peptide synthesis. (Caution: avoid extended storage of GGGC peptide in solution as intermolecular disulfide bonds will be formed, which in turn will compromise cysteine side chain reactivity. For the following steps, only dissolve the amount needed or remove solvent immediately after use if recovery of excess GGGC peptide is desired.)

2. Dissolve the desired amount of GGGC in PBS (without Ca²⁺ and Mg²⁺) to a final concentration of 10 mg/mL. In a separate tube, dissolve the maleimide-functionalized payload in a minimum amount of aqueous or water-miscible organic solvent (DMF, acetonitrile, PBS, H₂O, etc.). Combine the solutions to give a final molar ratio of 2 equivalents of peptide to 1 equivalent of maleimide reagent. Gently shake at RT.

An excess of peptide is typically used to fully consume costly maleimide-functionalized reagents. Maleimide-derived reagents containing modification such as fluorophores, biotin, or biorthogonal reaction handles are available from various commercial sources (Conju-probe, ThermoFisher, Lumiprobe, etc.).

3. Monitor the maleimide coupling reaction by mass spectrometry. The reaction is usually complete within 4 h at RT, but may also be left to incubate overnight.

4. After the coupling is complete, the reaction may be directly injected into a reversed-phase HPLC system (C18 column) for purification of the GGGC-maleimide peptide.
5. Lyophilize the pure peptide fractions and redissolve as needed for sortase-catalyzed protein labeling (see Basic Protocol or Alternate Protocol)

Probes can be stored as solutions in DMSO at -20°C for several months without loss of activity.

Procedure C: Synthesis of LPETGG probes via NHS-ester coupling

Manually synthesize the LPETGG peptide using the solid-phase technique outlined in Procedure A.

Proper protected Fmoc amino acid building blocks are listed in the Materials for Supporting Protocol 2. Alternatively, this short peptide is readily available from commercial sources or core facilities dedicated to peptide synthesis. While other amino acids may be substituted in place of Glu (E), lysine should be avoided to prevent competing reactivity with the peptide N-terminus in the following NHS-ester coupling steps.

Dissolve the desired amount of LPETGG peptide in DMSO to a final concentration of 5 mg/mL. Add the dissolved peptide directly to the NHS ester-activated payload to give a final molar ratio of 2 equivalents of peptide to 1 equivalent of NHS reagent. If the payload does not fully dissolve in organic solvent, a small amount of aqueous solvent at neutral pH can be added without significantly compromising yield. Gently shake at RT.

Monitor the NHS-ester coupling reaction by mass spectrometry. The reaction is usually complete within 4 h at RT, but may also be left to incubate overnight.

After the coupling is complete, dilute the reaction with water to ~50% DMSO (v/v). The reaction may then be directly injected into a reversed-phase HPLC system (C18 column) for purification of the modified LPETGG peptide.

Lyophilize the pure peptide fractions and redissolve as needed for sortase-catalyzed protein labeling (see Basic Protocol or Alternate Protocol)

Probes can be stored as solutions in DMSO at -20°C for several months without loss of activity.

COMMENTARY

Background Information

The construction of site-specifically modified proteins remains a significant challenge, and numerous strategies for this purpose are now available. Existing methods span a range from chemical approaches, such as the use of lysine- or cysteine-reactive small molecule reagents, to genetic techniques involving the installation of reactive natural amino acids (e.g. cysteine), unnatural amino acids, or entire protein domains (e.g. GFP, SNAP-tag) (for

relevant reviews see Spicer and Davis, 2014; Kim et al., 2013; Hinner and Johnsson, 2010). Within this spectrum of options, sortase-catalyzed transpeptidation represents a powerful hybrid strategy that combines the most advantageous features of chemical and genetic strategies, namely flexibility in the types of modifications installed and the site-specificity afforded by a genetically encoded short peptide tag. Sortase is not alone in this regard, and an expanding repertoire of conceptually similar chemoenzymatic labeling strategies have emerged that are built on the ability of natural or engineered enzymes to recognize short peptide tags, and to facilitate the subsequent installation of non-genetically encoded modifications. Representative examples of this group include butelase (Nguyen et al., 2014), transglutaminase (Lin and Ting, 2006), biotin ligase (Chen et al., 2005), lipoic acid ligase (Fernández-Suárez et al., 2007), trypsiligase (Liebscher et al., 2014), formyl-glycine generating enzyme (Carrico et al., 2007), phosphopantetheinyl transferase (Yin et al., 2005), farnesyltransferase (Duckworth et al., 2007), and tubulin tyrosine ligase (Schumacher et al., 2015). The key attributes of many of these enzymes have been reviewed (Rabuka, 2010; Rashidian et al., 2013). Among these strategies, sortase stands out due to its simplicity. All that is required for successful implementation are relatively simple techniques in synthetic peptide chemistry and protein cloning and expression. With these tools in place, sortase-catalyzed transpeptidation can be leveraged to decorate proteins with an extensive range of useful modifications.

Given the flexibility of sortase-mediated labeling, this method has proven effective for building a diverse range of polypeptide conjugates. Notable examples include the ligation of proteins or peptides to fluorophores (Popp et al., 2007; Yamamoto and Nagamune, 2009), carbohydrates (Wu et al., 2010; Samantaray et al., 2008), polymers (Qi et al., 2013; Parthasarathy et al., 2007), solid supports (Chan et al., 2007; Sinisi et al., 2012; Le et al., 2015), therapeutics (Beerli et al., 2015; Fang et al., 2016), lipids (Antos et al., 2008; Wu et al., 2010), nucleic acids (Koussa et al., 2014; Pritz et al., 2007), metal chelators (Westerlund et al., 2015; Paterson et al., 2014), viral particles (Schoonen et al., 2015; Hess et al., 2013), live cells (Yamamoto and Nagamune, 2009; Shi et al., 2014; Popp et al., 2007), or even intramolecular ligations for generating cyclic proteins and peptides (Antos et al., 2009b; Jia et al., 2014; van 't Hof et al., 2015). A full discussion of all relevant applications is beyond the scope of this unit, and we refer to the reader to reviews on this topic (Popp and Ploegh, 2011; Ritzefeld, 2014; Schmohl and Schwarzer, 2014; Haridas et al., 2014). Furthermore, while this unit does not outline specific procedures for generating each of the conjugates listed here, it should be stressed that most are attainable using simple variations of the Basic and Alternate Protocol outline above.

The proliferation of sortase-mediated applications has also been accompanied by refinements in the method itself. Early critiques of using sortase for protein modification included poor reaction rates, reaction reversibility, the requirement for Ca^{2+} in the reaction, as well as the fact that transpeptidation was restricted to LPXTG substrates. Notably, progress has been made in circumventing these limitations. The protocols in this unit, for example, rely on engineered variants of $\text{SrtA}_{\text{staph}}$ with enhanced catalytic activity. This includes the $\text{SrtA}_{\text{staph}}$ pentamutant, which was identified in a directed evolution screen and exhibits an approximately 120-fold increase in enzyme activity versus wild-type $\text{SrtA}_{\text{staph}}$ (Chen et al., 2011). The heptamutant builds on this derivative by including two additional

point mutations that alleviate Ca^{2+} dependency (Hirakawa et al., 2015; Wuethrich et al., 2014; Witte et al., 2015). In terms of reaction reversibility, a handful of strategies have been introduced that drive reaction equilibrium through selective removal or deactivation of transpeptidation products. These strategies may be as simple as running the reaction under dialysis conditions to separate out low molecular weight by-products (Pritz et al., 2007; Kobashigawa et al., 2009; Refaei et al., 2011), but also include the use of β -hairpins (Yamamura et al., 2011), depsipeptides (Williamson et al., 2012; Liu et al., 2014), or masked metal binding peptides (Row, et al. 2015) that prevent certain transpeptidation products from re-entering the catalytic cycle. Finally, advances have been made in expanding the substrate scope of sortase-catalyzed transpeptidations beyond the LPXTG motif. For example, sortase A from *Streptococcus pyogenes* is known to accept both LPXTG and LPXTA substrates (Antos et al., 2009a; Wuethrich et al., 2014; Hess et al., 2013). In addition, sortase derivatives from phage display or yeast display libraries have been identified that process XPKTG (numerous residues tolerated in X position), LAETG, or LPEXG (X = A, C, S) substrates (Piotukh et al., 2011; Dorr et al., 2014). This increased substrate scope has already enabled unique applications such as the semi-synthesis of histone H3 (Piotukh et al., 2011), orthogonal labeling of different proteins in the same viral capsid (Hess et al., 2013), and selective labeling of fetuin A in human plasma (Dorr et al., 2014).

Critical Parameters

For successful transpeptidation, the reactive protein termini (C-terminal LPXTG motif or N-terminal glycine) must be accessible to the sortase enzyme. Where possible, examination of the target protein's crystal or NMR structure is helpful in assessing solvent accessibility. In cases where steric interference is problematic, flexible linkers may be inserted between the body of protein and the sortase-reactive terminus (see Strategic Planning).

Optimal labeling is also achieved when the protein target serves as the limiting reagent and the peptide ligation partner is used in molar excess. This reaction stoichiometry helps ensure full conversion of the protein target to its modified form, which otherwise might be limited by the reversibility of the transpeptidation reaction. Excess peptide also minimizes the effect of competing acyl enzyme intermediate hydrolysis, which leads to unreactive protein or peptide side products.

Finally, the activity of the sortase A pentamutant requires the inclusion of Ca^{2+} in the reaction buffer. As outlined above, a 10 mM final concentration of calcium chloride is sufficient for optimal sortase A pentamutant activity. If heptamutant sortase A is used, then the calcium cofactor may be omitted.

Troubleshooting

For the Protocols in this unit, the most likely issue encountered will be low protein labeling efficiency. As discussed in the Basic Protocol, the reaction conditions provided serve as a starting point from which the user can optimize variables to achieve the desired labeling outcome. Therefore, in most cases higher reaction conversion may be achieved by simply extending reaction times, increasing sortase enzyme loading, or increasing the ratio of peptide probe to protein target. Running reactions at 37 °C may also prove beneficial,

though this is typically not required for pentamutant or heptamutant sortase A. Other recommended steps include checking the reaction pH to ensure that it is near neutral pH for optimal sortase activity, and independently characterizing the reaction components (sortase enzyme, protein target, peptide probe) to confirm that they are fully intact and free from degradation.

In the event that simple optimization is not sufficient, other possible causes of poor labeling include inaccessible protein termini (see Critical Parameters), or competing acyl enzyme hydrolysis. If poor accessibility of protein termini is suspected, extended flexible linker sequences should be inserted between the body of the protein and either the C-terminal LPXTG motif or the N-terminal oligoglycine (see Strategic Planning for details). For extensive hydrolysis, which may be evident when characterizing reactions by SDS-PAGE or mass spectrometry, the recommended course of action is to increase the ratio of peptide probe to protein target. It is also important to consider the reversible nature of the transpeptidation reaction. Successful transpeptidation leads to products with intact LPXTG motifs, which can in turn be repeatedly cleaved by sortase. On long enough time scales these products will be irreversibly hydrolyzed, and therefore it is important to carefully monitor reaction progress and to halt the reaction when the desired product has reached peak levels.

Finally, precipitation may be observed in some transpeptidation reactions. One cause may be the combination of calcium with phosphate buffers. Ca^{2+} is required to promote the activity pentamutant sortase, thus phosphate should be avoided in favor of Tris or HEPES. Alternatively, the sortase heptamutant may be used as it does not require the Ca^{2+} cofactor. Precipitation may also indicate issues with the solubility of target or labeled proteins. If this is encountered, overall protein concentrations can be reduced in the reaction mixture, and the transpeptidation reaction can be performed at lower temperature (4 °C) to limit protein precipitation.

Anticipated Results

The Basic Protocol described in this unit can typically be optimized to achieve excellent conversion (>90%) of the input protein to the desired transpeptidation product. Depending on the modification that is installed, formation of the site-specifically modified protein may be detected by mass spectrometry, SDS-PAGE with Coomassie staining or in-gel fluorescence, or immunoblotting. In the case of SDS-PAGE characterization, transpeptidation products often show distinct electrophoretic mobility as compared to the input protein, which facilitates detection of product formation. Mass spectrometry also provides an excellent technique for confirming successful labeling and monitoring reaction progress.

While robust transpeptidation is anticipated for most proteins, certain reaction by-products may be observed. For C-terminal protein labeling, in which the protein target contains the LPXTG motif, SDS-PAGE may reveal small quantities of hydrolyzed material at a lower apparent molecular weight, or residual acyl enzyme intermediate at a molecular weight appropriate for this covalent complex. Substrate hydrolysis may also be detected by mass spectrometry. For N-terminal labeling, given that the LPXTG peptide probe is used in excess, it is common to observe the formation of the acyl enzyme intermediate in significant

quantities via in-gel fluorescence or immunoblotting if the LPXTG peptide contains a fluorophore or biotin.

For proteins displayed on live cells, the Alternate Protocol provided here typically results in detectable protein labeling after a 45 minute incubation. Installation of the desired probe, which is often fluorescent or biotinylated, can then be confirmed by microscopy or flow cytometry. Alternatively, cells may be lysed and transpeptidation products may be separated by SDS-PAGE followed by in-gel fluorescence or immunoblotting.

Time Considerations

The expression and purification of pentamutant and heptamutant SrtA_{staph} is typically completed in 1–2 days. The synthesis and purification of peptide probes for C-terminal and N-terminal labeling may vary depending on peptide sequence, but generally this can be accomplished in 3–5 days. With all required materials in hand, the setup and analysis of sortase-catalyzed ligation reactions can be performed within 1 day. Transpeptidation reactions employing the pentamutant or heptamutant sortase are rapid, and labeling of purified protein targets in solution may be detected within minutes. However, full conversion to the labeled protein conjugate may require a few hours of incubation. In the case of cell surface labeling, 45 minutes of incubation is usually sufficient to achieve robust labeling. Importantly, for transpeptidation reactions both in solution and involving live cells, the specific reaction conditions will vary depending on the protein target, and should therefore be optimized for the particular application.

Acknowledgments

The authors want to thank the Antos and Ploegh labs for helpful input and suggestions. This work was supported by an NIH Award to H.L.P. (2R01AI087879-06)

LITERATURE CITED

- Alt K, Paterson BM, Westein E, Rudd SE, Poniger SS, Jagdale S, Ardipradja K, Connell TU, Krippner GY, Nair AKN, Wang X, Tochon-Danguy HJ, Donnelly PS, Peter K, Hagemeyer CE. A Versatile Approach for the Site-Specific Modification of Recombinant Antibodies Using a Combination of Enzyme-Mediated Bioconjugation and Click Chemistry. *Angew Chem Int Ed Engl.* 2015; 54:7515–7519. [PubMed: 25962581]
- Antos JM, Chew GL, Guimaraes CP, Yoder NC, Grotenbreg GM, Popp MWL, Ploegh HL. Site-specific N- and C-terminal labeling of a single polypeptide using sortases of different specificity. *J Am Chem Soc.* 2009a; 131:10800–10801. [PubMed: 19610623]
- Antos JM, Miller GM, Grotenbreg GM, Ploegh HL. Lipid modification of proteins through sortase-catalyzed transpeptidation. *J Am Chem Soc.* 2008; 130:16338–16343. [PubMed: 18989959]
- Antos JM, Popp MWL, Ernst R, Chew GL, Spooner E, Ploegh HL. A straight path to circular proteins. *J Biol Chem.* 2009b; 284:16028–16036. [PubMed: 19359246]
- Antos JM, Truttmann MC, Ploegh HL. Recent advances in sortase-catalyzed ligation methodology. *Curr Opin Struct Biol.* 2016; 38:111–118. [PubMed: 27318815]
- Baer S, Nigro J, Madej MP, Nisbet RM, Suryadinata R, Coia G, Hong LPT, Adams TE, Williams CC, Nuttall SD. Comparison of alternative nucleophiles for Sortase A-mediated bioconjugation and application in neuronal cell labelling. *Org Biomol Chem.* 2014; 12:2675–2685. [PubMed: 24643508]

- Beerli RR, Hell T, Merkel AS, Grawunder U. Sortase Enzyme-Mediated Generation of Site-Specifically Conjugated Antibody Drug Conjugates with High In Vitro and In Vivo Potency. *PLoS ONE*. 2015; 10:e0131177. [PubMed: 26132162]
- Boekhorst J, de Been MW, Kleerebezem M, Siezen RJ. Genome-wide detection and analysis of cell wall-bound proteins with LPxTG-like sorting motifs. *J Bacteriol*. 2005; 187:4928–4934. [PubMed: 15995208]
- Carrico IS, Carlson BL, Bertozzi CR. Introducing genetically encoded aldehydes into proteins. *Nat Chem Biol*. 2007; 3:321–322. [PubMed: 17450134]
- Chan L, Cross HF, She JK, Cavalli G, Martins HFP, Neylon C. Covalent attachment of proteins to solid supports and surfaces via Sortase-mediated ligation. *PLoS ONE*. 2007; 2:e1164. [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 USA*. 2011; 108:11399–11404. [PubMed: 21697512]
- Chen I, Howarth M, Lin W, Ting AY. Site-specific labeling of cell surface proteins with biophysical probes using biotin ligase. *Nat Methods*. 2005; 2:99–104. [PubMed: 15782206]
- Chen L, Cohen J, Song X, Zhao A, Ye Z, Feulner CJ, Doonan P, Somers W, Lin L, Chen PR. Improved variants of SrtA for site-specific conjugation on antibodies and proteins with high efficiency. *Sci Rep*. 2016; 6:31899. [PubMed: 27534437]
- Chen Q, Sun Q, Molino NM, Wang SW, Boder ET, Chen W. Sortase A-mediated multi-functionalization of protein nanoparticles. *Chem Commun*. 2015; 51:12107–12110.
- Row DR, Roark TJ, Philip MC, Perkins LL, Antos JM. Enhancing the efficiency of sortase-mediated ligations through nickel-peptide complex formation. *Chem Commun*. 2015; 51:12548–12551.
- Dorr BM, Ham HO, An C, Chaikof EL, Liu DR. Reprogramming the specificity of sortase enzymes. *Proc Natl Acad Sci USA*. 2014; 111:13343–13348. [PubMed: 25187567]
- Duckworth BP, Zhang Z, Hosokawa A, Distefano MD. Selective labeling of proteins by using protein farnesyltransferase. *ChemBioChem*. 2007; 8:98–105. [PubMed: 17133644]
- Fang T, Duarte JN, Ling J, Li Z, Guzman JS, Ploegh HL. Structurally Defined α MHC-II Nanobody-Drug Conjugates: A Therapeutic and Imaging System for B-Cell Lymphoma. *Angew Chem Int Ed Engl*. 2016; 55:2416–2420. [PubMed: 26840214]
- Fernández-Suárez M, Baruah H, Martínez-Hernández L, Xie KT, Baskin JM, Bertozzi CR, Ting AY. Redirecting lipoic acid ligase for cell surface protein labeling with small-molecule probes. *Nat Biotechnol*. 2007; 25:1483–1487. [PubMed: 18059260]
- Gianella P, Snapp EL, Levy M. An in vitro compartmentalization-based method for the selection of bond-forming enzymes from large libraries. *Biotechnol Bioeng*. 2016; 113:1647–1657. [PubMed: 26806853]
- Glasgow JE, Salit ML, Cochran JR. In Vivo Site-Specific Protein Tagging with Diverse Amines Using an Engineered Sortase Variant. *J Am Chem Soc*. 2016; 138:7496–7499. [PubMed: 27280683]
- Haridas V, Sadanandan S, Dheepthi NU. Sortase-based bio-organic strategies for macromolecular synthesis. *ChemBioChem*. 2014; 15:1857–1867. [PubMed: 25111709]
- Heck T, Pham PH, Hammes F, Thöny-Meyer L, Richter M. Continuous Monitoring of Enzymatic Reactions on Surfaces by Real-Time Flow Cytometry: Sortase A Catalyzed Protein Immobilization as a Case Study. *Bioconjugate Chem*. 2014; 25:1492–1500.
- Hess GT, Cragnolini JJ, Popp MW, Allen MA, Dougan SK, Spooner E, Ploegh HL, Belcher AM, Guimaraes CP. M13 Bacteriophage Display Framework That Allows Sortase-Mediated Modification of Surface-Accessible Phage Proteins. *Bioconjugate Chem*. 2012; 23:1478–1487.
- Hess GT, Guimaraes CP, Spooner E, Ploegh HL, Belcher AM. Orthogonal labeling of M13 minor capsid proteins with DNA to self-assemble end-to-end multiphage structures. *ACS Synth Biol*. 2013; 2:490–496. [PubMed: 23713956]
- Hinner MJ, Johnsson K. How to obtain labeled proteins and what to do with them. *Curr Opin Biotech*. 2010; 21:766–776. [PubMed: 21030243]
- Hirakawa H, Ishikawa S, Nagamune T. Ca²⁺-independent sortase-A exhibits high selective protein ligation activity in the cytoplasm of *Escherichia coli*. *Biotechnol J*. 2015; 10:1487–1492. [PubMed: 25864513]

- Hirel PH, Schmitter MJ, Dessen P, Fayat G, Blanquet S. Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. *Proc Natl Acad Sci USA*. 1989; 86:8247–8251. [PubMed: 2682640]
- Hui JZ, Tamsen S, Song Y, Tsourkas A. LASIC: Light Activated Site-Specific Conjugation of Native IgGs. *Bioconjugate Chem*. 2015; 26:1456–1460.
- Ilangovan U, Ton-That H, Iwahara J, Schneewind O, Clubb R. Structure of sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*. *Proc Natl Acad Sci U S A*. 2001; 98:6056–6061. [PubMed: 11371637]
- Jia X, Kwon S, Wang CIA, Huang YH, Chan LY, Tan CC, Rosengren KJ, Mulvenna JP, Schroeder CI, Craik DJ. Semiozymatic cyclization of disulfide-rich peptides using Sortase A. *J Biol Chem*. 2014; 289:6627–6638. [PubMed: 24425873]
- Kim CH, Axup JY, Schultz PG. Protein conjugation with genetically encoded unnatural amino acids. *Curr Opin Chem Biol*. 2013; 17:412–419. [PubMed: 23664497]
- 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; 43:145–150. [PubMed: 19140010]
- Koussa MA, Sotomayor M, Wong WP. Protocol for sortase-mediated construction of DNA-protein hybrids and functional nanostructures. *Methods*. 2014; 67:134–141. [PubMed: 24568941]
- Le RK, Raeeszadeh-Sarmazdeh M, Boder ET, Frymier PD. Sortase-mediated ligation of PsaE-modified photosystem I from *Synechocystis* sp. PCC 6803 to a conductive surface for enhanced photocurrent production on a gold electrode. *Langmuir*. 2015; 31:1180–1188. [PubMed: 25535846]
- Levary DA, Parthasarathy R, Boder ET, Ackerman ME. Protein-Protein Fusion Catalyzed by Sortase A. *PLoS ONE*. 2011; 6:e18342–1718. [PubMed: 21494692]
- Li YM, Li YT, Pan M, Kong XQ, Huang YC, Hong ZY, Liu L. Irreversible Site-Specific Hydrazinolysis of Proteins by Use of Sortase. *Angew Chem Int Ed Engl*. 2014; 53:2198–2202. [PubMed: 24470054]
- Liebscher S, Kornberger P, Fink G, Trost-Gross EM, Höss E, Skerra A, Bordusa F. Derivatization of antibody Fab fragments: a designer enzyme for native protein modification. *ChemBioChem*. 2014; 15:1096–1100. [PubMed: 24782039]
- Lin CW, Ting AY. Transglutaminase-catalyzed site-specific conjugation of small-molecule probes to proteins in vitro and on the surface of living cells. *J Am Chem Soc*. 2006; 128:4542–4543. [PubMed: 16594669]
- Litou ZI, Bagos PG, Tsirigos KD, Liakopoulos TD, Hamodrakas SJ. Prediction of cell wall sorting signals in gram-positive bacteria with a hidden markov model: application to complete genomes. *J Bioinform Comput Biol*. 2008; 6:387–401. [PubMed: 18464329]
- Liu F, Luo EY, Flora DB, Mezo AR. Irreversible sortase A-mediated ligation driven by diketopiperazine formation. *J Org Chem*. 2014; 79:487–492. [PubMed: 24378034]
- Malakhov MP, Mattern MR, Malakhova OA, Drinker M, Weeks SD, Butt TR. SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *J Struct Funct Genomics*. 2004; 5:75–86. [PubMed: 15263846]
- Mao H, Hart SA, Schink A, Pollok BA. Sortase-mediated protein ligation: a new method for protein engineering. *J Am Chem Soc*. 2004; 126:2670–2671. [PubMed: 14995162]
- Massa S, Vikani N, Betti C, Ballet S, Vanderhaegen S, Steyaert J, Descamps B, Vanhove C, Bunschoten A, van Leeuwen FWB, Hernot S, Caveliers V, Lahoutte T, Muyltermans S, Xavier C, Devoogdt N. Sortase A-mediated site-specific labeling of camelid single-domain antibody-fragments: a versatile strategy for multiple molecular imaging modalities. *Contrast Media Mol Imaging*. 2016; 11:328–339. [PubMed: 27147480]
- Nguyen GKT, Wang S, Qiu Y, Hemu X, Lian Y, Tam JP. Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nat Chem Biol*. 2014; 10:732–738. [PubMed: 25038786]
- Parthasarathy R, Subramanian S, Boder ET. Sortase A as a novel molecular “stapler” for sequence-specific protein conjugation. *Bioconjugate Chem*. 2007; 18:469–476.
- Paterson BM, Alt K, Jeffery CM, Price RI, Jagdale S, Rigby S, Williams CC, Peter K, Hagemeyer CE, Donnelly PS. Enzyme-mediated site-specific bioconjugation of metal complexes to proteins:

- sortase-mediated coupling of copper-64 to a single-chain antibody. *Angew Chem Int Ed Engl.* 2014; 53:6115–6119. [PubMed: 24777818]
- Piotukh K, Geltinger B, Heinrich N, Gerth F, Beyermann M, Freund C, Schwarzer D. Directed Evolution of Sortase A Mutants with Altered Substrate Selectivity Profiles. *J Am Chem Soc.* 2011; 133:17536–17539. [PubMed: 21978125]
- Policarpo RL, Kang H, Liao X, Rabideau AE, Simon MD, Pentelute BL. Flow-based enzymatic ligation by sortase A. *Angew Chem Int Ed Engl.* 2014; 53:9203–9208. [PubMed: 24989829]
- Popp MW, Antos JM, Grotenbreg GM, Spooner E, Ploegh HL. Sortagging: a versatile method for protein labeling. *Nat Chem Biol.* 2007; 3:707–708. [PubMed: 17891153]
- Popp MWL, Ploegh HL. Making and Breaking Peptide Bonds: Protein Engineering Using Sortase. *Angew Chem Int Ed Engl.* 2011; 50:5024–5032. [PubMed: 21538739]
- Pritz S, Wolf Y, Kraetke O, Klose J, Bienert M, Beyermann M. Synthesis of biologically active peptide nucleic acid-peptide conjugates by sortase-mediated ligation. *J Org Chem.* 2007; 72:3909–3912. [PubMed: 17432905]
- Qi Y, Amiram M, Gao W, McCafferty DG, Chilkoti A. Sortase-catalyzed initiator attachment enables high yield growth of a stealth polymer from the C terminus of a protein. *Macro Rapid Commun.* 2013; 34:1256–1260.
- Rabuka D. Chemoenzymatic methods for site-specific protein modification. *Curr Opin Chem Biol.* 2010; 14:790–796. [PubMed: 21030291]
- Rashidian M, Dozier JK, Distefano MD. Enzymatic Labeling of Proteins: Techniques and Approaches. *Bioconjugate Chem.* 2013; 24:1277–1294.
- Rashidian M, Keliher E, Dougan M, Juras PK, Cavallari M, Wojtkiewicz GR, Jacobsen J, Edens JG, Tas JMG, Victora G, Weissleder R, Ploegh H. The use of (18)F-2-fluorodeoxyglucose (FDG) to label antibody fragments for immuno-PET of pancreatic cancer. *ACS Cent Sci.* 2015; 1:142–147. [PubMed: 26955657]
- Refaei MA, Combs AI, Kojetin DJ, Cavanagh J, Caperelli C, Rance M, Sapitro J, Tsang P. Observing selected domains in multi-domain proteins via sortase-mediated ligation and NMR spectroscopy. *J Biomol NMR.* 2011; 49:3–7. [PubMed: 21188472]
- Ritzefeld M. Sortagging: a robust and efficient chemoenzymatic ligation strategy. *Chem Eur J.* 2014; 20:8516–8529. [PubMed: 24954404]
- Samantaray S, Marathe U, Dasgupta S, Nandicoori VK, Roy RP. Peptide-sugar ligation catalyzed by transpeptidase sortase: a facile approach to neoglycoconjugate synthesis. *J Am Chem Soc.* 2008; 130:2132–2133. [PubMed: 18229923]
- Schmohl L, Schwarzer D. Sortase-mediated ligations for the site-specific modification of proteins. *Curr Opin Chem Biol.* 2014; 22C:122–128.
- Schoonen L, Pille J, Borrmann A, Nolte RJM, van Hest JCM. Sortase A-Mediated N-Terminal Modification of Cowpea Chlorotic Mottle Virus for Highly Efficient Cargo Loading. *Bioconjugate Chem.* 2015; 26:2429–2434.
- Schumacher D, Helma J, Mann FA, Pichler G, Natale F, Krause E, Cardoso MC, Hackenberger CPR, Leonhardt H. Versatile and Efficient Site-Specific Protein Functionalization by Tubulin Tyrosine Ligase. *Angew Chem Int Ed Engl.* 2015; 54:13787–13791. [PubMed: 26404067]
- Shi J, Kundrat L, Pishesha N, Bilate A, Theile C, Maruyama T, Dougan SK, Ploegh HL, Lodish HF. Engineered red blood cells as carriers for systemic delivery of a wide array of functional probes. *Proc Natl Acad Sci USA.* 2014; 111:10131–10136. [PubMed: 24982154]
- Sinisi A, Popp MWL, Antos JM, Pansegrau W, Savino S, Nissum M, Rappuoli R, Ploegh HL, Buti L. Development of an Influenza virus Protein Array Using Sortagging Technology. *Bioconjugate Chem.* 2012; 23:1119–1126.
- Spicer CD, Davis BG. Selective chemical protein modification. *Nat Commun.* 2014; 5:4740. [PubMed: 25190082]
- Swee LK, Guimaraes CP, Sehrawat S, Spooner E, Barrasa MI, Ploegh HL. Sortase-mediated modification of α DEC205 affords optimization of antigen presentation and immunization against a set of viral epitopes. *Proc Natl Acad Sci USA.* 2013; 110:1428–1433. [PubMed: 23297227]

- Swee LK, Lourido S, Bell GW, Ingram JR, Ploegh HL. One-step enzymatic modification of the cell surface redirects cellular cytotoxicity and parasite tropism. *ACS Chem Biol*. 2015; 10:460–465. [PubMed: 25360987]
- Tanaka T, Yamamoto T, Tsukiji S, Nagamune T. Site-specific protein modification on living cells catalyzed by Sortase. *ChemBioChem*. 2008; 9:802–807. [PubMed: 18297670]
- Ton-That H, Faull KF, Schneewind O. Anchor structure of staphylococcal surface proteins. A branched peptide that links the carboxyl terminus of proteins to the cell wall. *J Biol Chem*. 1997; 272:22285–22292. [PubMed: 9268378]
- van 't Hof W, Ma ásková SH, Veerman ECI, Bolscher JGM. Sortase-mediated backbone cyclization of proteins and peptides. *Biol Chem*. 2015; 396:283–293. [PubMed: 25581753]
- Wagner K, Kwakkenbos MJ, Claassen YB, Maijor K, Böhne M, van der Sluijs KF, Witte MD, van Zoelen DJ, Cornelissen LA, Beaumont T, Bakker AQ, Ploegh HL, Spits H. Bispecific antibody generated with sortase and click chemistry has broad antiinfluenza virus activity. *Proc Natl Acad Sci USA*. 2014; 111:16820–16825. [PubMed: 25385586]
- Wagh DS. An overview of enzymatic reagents for the removal of affinity tags. *Protein Expr Purif*. 2011; 80:283–293. [PubMed: 21871965]
- Westerlund K, Honarvar H, Tolmachev V, Eriksson Karlström A. Design, Preparation, and Characterization of PNA-Based Hybridization Probes for Affibody-Molecule-Mediated Pretargeting. *Bioconjugate Chem*. 2015; 26:1724–1736.
- Williamson DJ, Fascione MA, Webb ME, Turnbull WB. Efficient N-terminal labeling of proteins by use of sortase. *Angew Chem Int Ed Engl*. 2012; 51:9377–9380. [PubMed: 22890696]
- Williamson DJ, Webb ME, Turnbull WB. Depsipeptide substrates for sortase-mediated N-terminal protein ligation. *Nat Protoc*. 2014; 9:253–262. [PubMed: 24407354]
- Witte MD, Wu T, Guimaraes CP, Theile CS, Blom AEM, Ingram JR, Li Z, Kundrat L, Goldberg SD, Ploegh HL. Site-specific protein modification using immobilized sortase in batch and continuous-flow systems. *Nat Protoc*. 2015; 10:508–516. [PubMed: 25719269]
- Wu Z, Guo X, Wang Q, Swarts BM, Guo Z. Sortase A-Catalyzed Transpeptidation of Glycosylphosphatidylinositol Derivatives for Chemoenzymatic Synthesis of GPI-Anchored Proteins. *J Am Chem Soc*. 2010; 132:1567–1571. [PubMed: 20078124]
- Wuethrich I, Peeters JGC, Blom AEM, Theile CS, Li Z, Spooner E, Ploegh HL, Guimaraes CP. Site-specific chemoenzymatic labeling of aerolysin enables the identification of new aerolysin receptors. *PLoS ONE*. 2014; 9:e109883. [PubMed: 25275512]
- Yamamoto T, Nagamune T. Expansion of the sortase-mediated labeling method for site-specific N-terminal labeling of cell surface proteins on living cells. *Chem Commun*. 2009:1022–1024.
- Yamamura Y, Hirakawa H, Yamaguchi S, Nagamune T. Enhancement of sortase A-mediated protein ligation by inducing a β -hairpin structure around the ligation site. *Chem Commun*. 2011; 47:4742–4744.
- Yin J, Straight PD, McLoughlin SM, Zhou Z, Lin AJ, Golan DE, Kelleher NL, Kolter R, Walsh CT. Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proc Natl Acad Sci USA*. 2005; 102:15815–15820. [PubMed: 16236721]

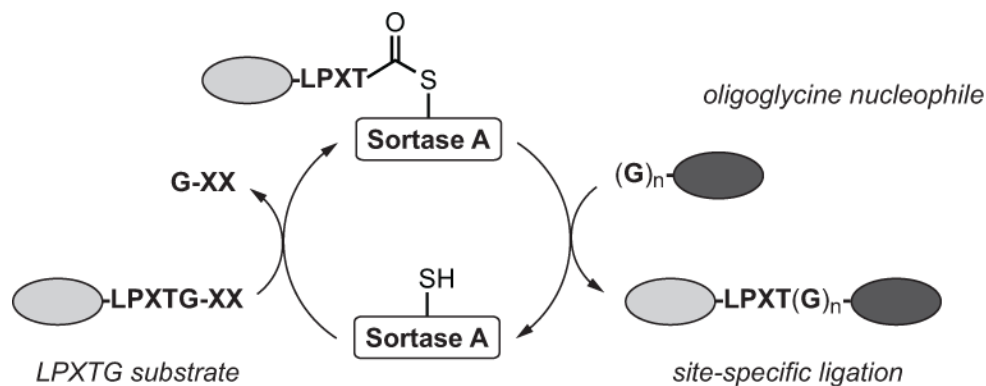


Figure 1. Sortase-catalyzed transpeptidation. Wild-type and engineered variants of sortase A from *Staphylococcus aureus* recognize substrates containing an LPXTG motif. The active site cysteine of sortase A cleaves between the threonine and glycine residues to generate a thioester-linked acyl enzyme intermediate. This intermediate is then attacked by an oligoglycine nucleophile, which releases the sortase enzyme and generates a site-specific ligation product linked via a native amide bond. This reaction can be harnessed to generate proteins site-specifically labeled at the N- or C-terminus.

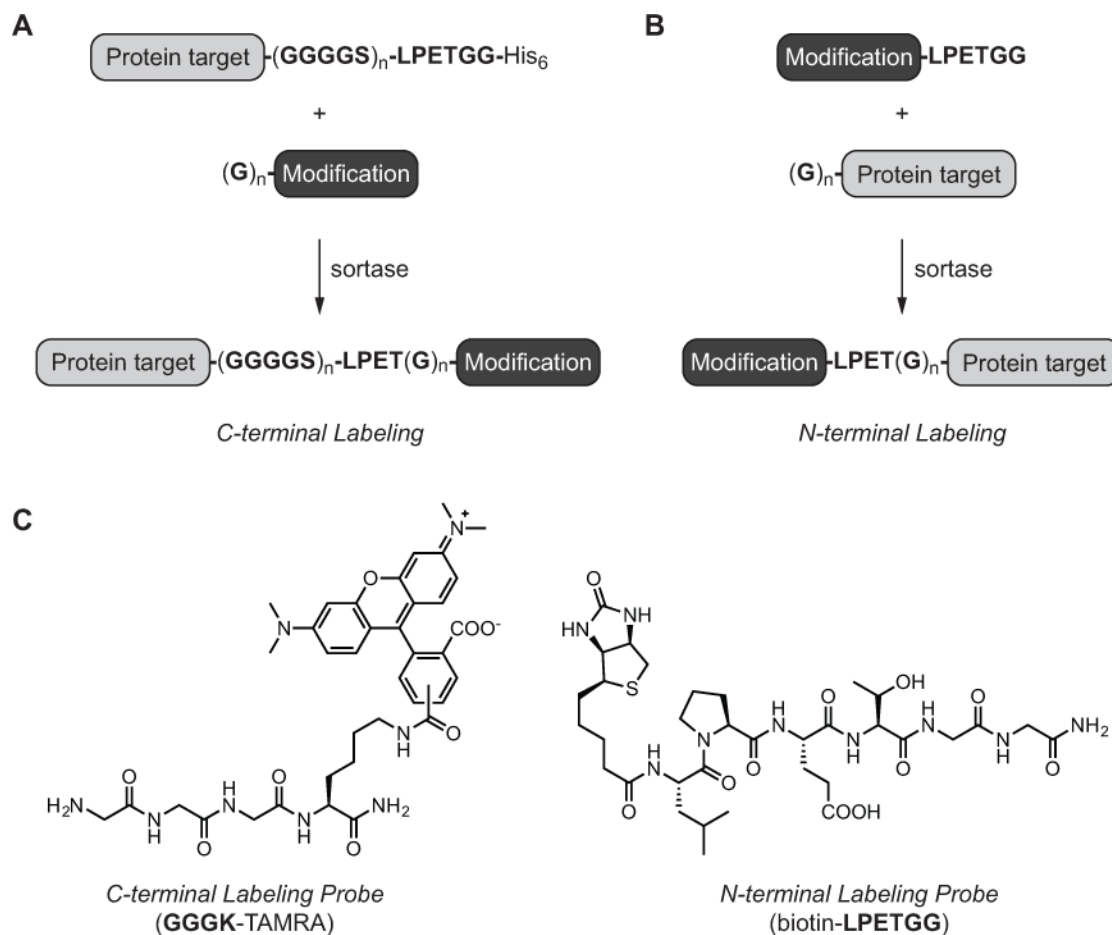


Figure 2. Site-specific labeling of target proteins using sortase-catalyzed transpeptidation. **(A)** Reaction scheme for installing modifications at the protein C-terminus. Target proteins contain the requisite sortase substrate motif (e.g. LPETGG) separated from the body of the protein with an optional GGGGS linker. The sortase cleavage site is followed by an optional tag, such as His₆, to assist with reaction monitoring and labeled protein purification. For C-terminal labeling, the target protein is paired with an oligoglycine nucleophile tethered to the desired modification. **(B)** Reaction scheme for installing modifications at the protein N-terminus. Target proteins serve as the reaction nucleophile, and contain one or more N-terminal glycines. Protein targets are paired with synthetic peptides containing the sortase recognition site (e.g. LPETGG) and the modification of interest. **(C)** Representative examples of peptide probes compatible with sortase-mediated C-terminal (left) or N-terminal (right) labeling.