

Site-Specific N- and C-Terminal Labeling of a Single Polypeptide Using Sortases of Different Specificity

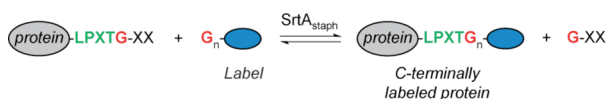
John M. Antos, Guo-Liang Chew, Carla P. Guimaraes, Nicholas C. Yoder, Gijsbert M. Grotenbreg, Maximilian Wei-Lin Popp, and Hidde L. Ploegh*

Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142

Received April 3, 2009; E-mail: ploegh@wi.mit.edu

Methods for site-specific modification of proteins remain in high demand. The transpeptidation reaction catalyzed by sortase A from *Staphylococcus aureus* (SrtA_{staph}) allows site-specific derivatization of proteins with virtually any type of functional material.¹ Target proteins are engineered to contain the SrtA_{staph} recognition site (LPXTG) near their C terminus, thus allowing a transacylation reaction in which the residues C-terminal to threonine are exchanged for a synthetic oligoglycine peptide (Scheme 1). While the range of applications for this technology has expanded considerably,¹ the ligation chemistry itself has seen relatively few modifications or improvements. Since nearly all Gram-positive bacteria possess sortases, many with reactivity distinct from SrtA_{staph}, there is an exciting opportunity to develop complementary strategies for protein engineering using other members of this enzyme family.^{2,3} Here we present a strategy for placing discrete labels at both termini in the same polypeptide through the use of multiple sortases. We first describe the ability of SrtA_{staph} to append labels at the protein N terminus and then demonstrate how this can be used in conjunction with the activity of sortase A from *Streptococcus pyogenes* (SrtA_{strep}) to yield dual-labeled proteins.

Scheme 1. C-terminal Labeling Using SrtA_{staph}



With regard to N-terminal labeling mediated by SrtA_{staph}, we reasoned that labeled synthetic peptides containing the LPXTG recognition motif, or structural analogues thereof, could generate the requisite acyl–enzyme intermediate necessary for transpeptidation. In combination with protein nucleophiles containing one or more N-terminal glycines, this should result in transfer of the label to the protein N terminus (Figure 1a). We synthesized FITC (**1**) and biotin (**2**) derivatives of an LPRT peptide in which the glycine of the normal LPXTG motif was replaced by a methyl ester (Figure 1b and Figure S1 in the Supporting Information). The use of an ester derivative rather than the entire LPXTG motif was motivated by our concern that the glycine residue released after LPXTG cleavage might compete with the protein nucleophile, potentially complicating the desired N-terminal-labeling reaction. In contrast, transacylation with **1** and **2** would generate MeOH, a poor nucleophile for transacylation compared with glycine. It should be noted that concurrently with the work described here, it was demonstrated that labeled LPETGG peptides are viable tools for N-terminal labeling using SrtA_{staph}.⁴

With **1** and **2** in hand, we expressed a series of model proteins containing N-terminal glycine residues. We prepared variants of the cholera toxin B subunit (CtxB) with one, three, or five N-terminal glycines. In order to verify the selectivity for glycine,

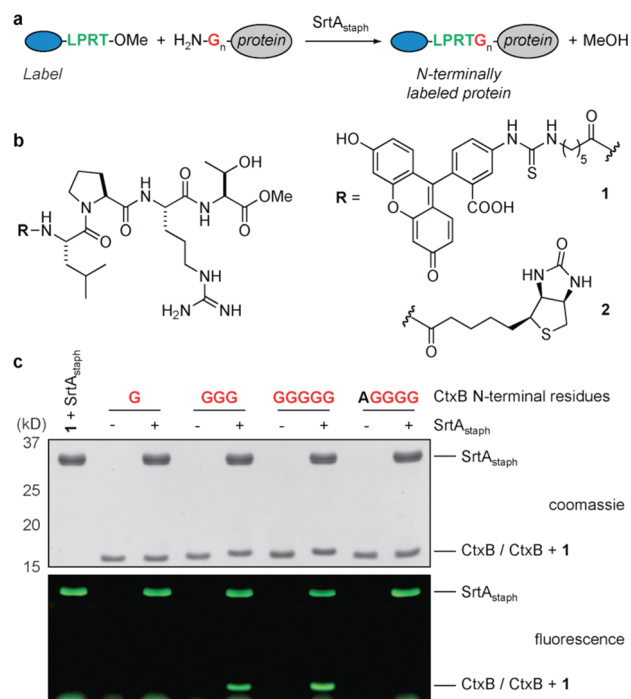


Figure 1. N-terminal labeling using SrtA_{staph}. (a) SrtA_{staph} catalyzes a transacylation reaction using labeled LPRT methyl esters as substrates. The labeled LPRT fragment is transferred to proteins containing N-terminal glycines in a site-specific fashion. (b) FITC (**1**) and biotin (**2**) LPRT methyl esters for N-terminal transacylation. (c) CtxB derivatives (50 μ M) were treated with 500 μ M **1** and 50 μ M SrtA_{staph} for 2 h at 37 $^{\circ}$ C in 50 mM Tris (pH 7.5), 150 mM NaCl, and 10 mM CaCl₂. Reactions were analyzed by SDS-PAGE with visualization by coomassie staining and fluorescent gel scanning.

we prepared a control construct containing an N-terminal alanine residue. In the presence of SrtA_{staph}, we observed robust labeling of G₃-CtxB and G₅-CtxB using 500 μ M **1** for 2 h at 37 $^{\circ}$ C, with no apparent labeling of the alanine-containing control (Figure 1c). Electrospray ionization mass spectrometry (ESI-MS) revealed quantitative labeling of G₃-CtxB and G₅-CtxB, with no modification observed for G₁-CtxB and AG₄-CtxB (Figure S2). Similar experiments with biotinylated derivative **2** and G₅-CtxB yielded comparable results, as verified by ESI-MS and streptavidin immunoblot (Figure S3). In all cases, residual labeling of SrtA_{staph} itself, attributable to the formation of a covalent acyl–enzyme intermediate, was detected. N-terminal transpeptidation was also successful for two additional protein substrates, eGFP with five N-terminal glycines and UCHL3 containing a single N-terminal glycine (Figure S4).

With the ability of SrtA_{staph} to append labels at either terminus, we pursued the possibility of installing two modifications within

the same protein. Attempts to execute this type of transformation using SrtA_{staph} alone were unsuccessful, as intramolecular transpeptidation between N-terminal glycines and the C-terminal LPXTG motif was unavoidable in most cases. Therefore, we considered the possibility of using a second, distinct sortase, an idea that has been suggested but never reduced to practice.^{1,5} We initially sought to use sortase B (SrtB) from either *Staph. aureus* or *Bacillus anthracis* as enzymes with recognition sequences (NPQTN and NPKTG, respectively) orthogonal to that of SrtA_{staph}.^{6,7} Both SrtB enzymes were easily produced in *Escherichia coli* and purified to homogeneity. We reproduced the reported in vitro enzyme activity using a FRET-based assay to measure cleavage of short peptide substrates.^{6,7} However, to date we have failed to obtain transpeptidation with either SrtB on protein substrates modified with the appropriate recognition sequences on a time scale or with yields that compare favorably with SrtA_{staph} (data not shown).

We ultimately arrived at a successful orthogonal strategy using SrtA_{strep}, which recognizes the same LPXTG sequence used by SrtA_{staph} but can accept alanine-based nucleophiles.⁸ This leads to the formation of an LPXTA sequence at the site of ligation, a motif refractory to cleavage by SrtA_{staph}.⁹ This allows SrtA_{staph} to act on the N terminus without affecting the C-terminal modification installed with SrtA_{strep}.

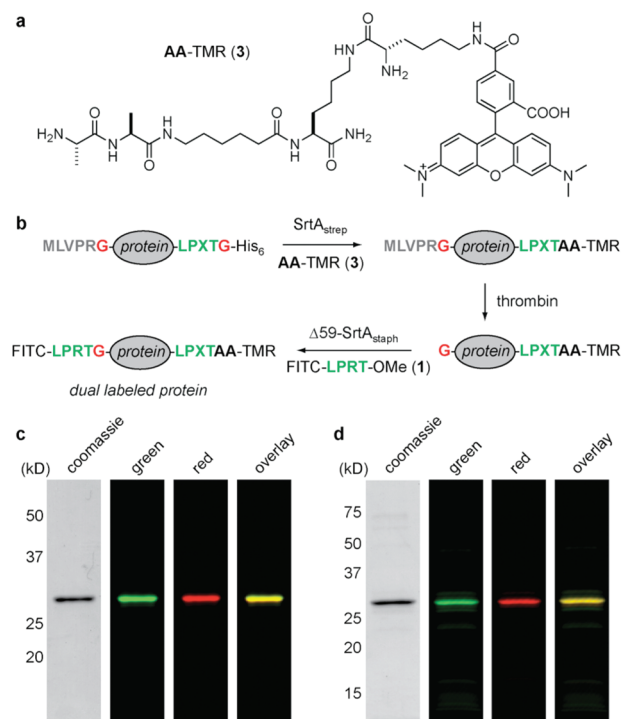


Figure 2. Site-specific N- and C-terminal labeling using multiple sortases. (a) Tetramethylrhodamine-labeled dialanine nucleophile (**3**) for SrtA_{strep}-mediated transpeptidation. (b) Strategy for the installation of discrete labels at both termini of the same protein using Δ 59-SrtA_{staph} and SrtA_{strep}. (c, d) SDS-PAGE characterization with fluorescent gel scanning of dual-labeled (c) eGFP and (d) UCHL3.

Our final strategy for dual-terminus labeling is outlined in Figure 2b. We first synthesized a tetramethylrhodamine-labeled peptide (**3**) containing two N-terminal alanine residues to serve as the nucleophile for SrtA_{strep}-mediated protein ligation (Figure 2a and Figure S5). We prepared two model substrates (eGFP and UCHL3) containing masked N-terminal glycines that are exposed only upon thrombin cleavage. Masking was required because SrtA_{strep} was

observed to ligate both glycine and alanine nucleophiles (data not shown). Substrates also contained an LPXTG motif at the C terminus to allow a first round of labeling with SrtA_{strep}. For both eGFP and UCHL3, C-terminal labeling using **3** and SrtA_{strep} resulted in >90% conversion to the desired adduct, as revealed by ESI-MS (Figure S6). SrtA_{strep} was quenched by the addition of MTSET followed by removal of His₆-tagged SrtA_{strep} using Ni-NTA. Residual **3** was then removed using a disposable desalting column. Thrombin cleavage proceeded in quantitative fashion using commercial thrombin agarose resin (Figure S6). The exposed N-terminal glycines were then labeled by treatment with 500 μ M **1** and 50 μ M Δ 59-SrtA_{staph}¹⁰ for \sim 1 h at 37 $^{\circ}$ C. ESI-MS of crude reaction mixtures showed the dual-labeled material as the major component, with only minor amounts of byproduct (Figure S6). A final separation by anion-exchange chromatography yielded dual-labeled eGFP and UCHL3 with excellent purity, as determined by both SDS-PAGE and ESI-MS (Figure 2c,d and Figure S6). In the case of UCHL3, we observed some additional low-intensity bands in the fluorescent gel scan (Figure 2d). However, quantitative densitometric analysis of coomassie-stained gels indicated purity in excess of 95% for both dual-labeled eGFP and UCHL3.

In summary, we have developed a strategy for placing different chemical labels at the two ends of the same polypeptide using two sortase enzymes with unique reactivities. We anticipate that this method will be applicable to the preparation of protein conjugates for refolding studies or for the construction of protein sensors, where measuring conformational changes by FRET is a common mode of detection. In more general terms, this work begins to explore the range of protein modifications that can be accessed using alternative sortases. The number of sortases that have been produced in recombinant form with retention of activity is continually increasing, and we are exploring the use of these unique enzymes as tools for protein engineering.

Acknowledgment. This work was supported by grants from the National Institutes of Health (R01-AI057182, R01-AI033456, R21-EB008875). J.M.A. and N.C.Y. acknowledge Clay Postdoctoral Fellowships. C.P.G. thanks the Fundacao para a Ciencia e Tecnologia (Portugal). G.-L.C. was supported by MIT's Undergraduate Research Opportunities Program (UROP). The authors thank Mark J. Banfield and colleagues for providing the expression plasmid for SrtA_{strep} and Olaf Schneewind for providing the SrtB constructs.

Supporting Information Available: Full experimental details and characterization data for peptide derivatives and protein conjugates. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Tsukiji, S.; Nagamune, T. *ChemBioChem* **2009**, *10*, 787–798.
- (2) Marraffini, L. A.; Dedent, A. C.; Schneewind, O. *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 192–221.
- (3) Pallen, M. J.; Lam, A. C.; Antonio, M.; Dunbar, K. *Trends Microbiol.* **2001**, *9*, 97–102.
- (4) Yamamoto, T.; Nagamune, T. *Chem. Commun.* **2009**, 1022–1024.
- (5) Popp, M. W.; Antos, J. M.; Grotenbreg, G. M.; Spooner, E.; Ploegh, H. L. *Nat. Chem. Biol.* **2007**, *3*, 707–708.
- (6) Maresso, A. W.; Chapa, T. J.; Schneewind, O. *J. Bacteriol.* **2006**, *188*, 8145–8152.
- (7) Mazmanian, S. K.; Ton-That, H.; Su, K.; Schneewind, O. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2293–2298.
- (8) Race, P. R.; Bentley, M. L.; Melvin, J. A.; Crow, A.; Hughes, R. K.; Smith, W. D.; Sessions, R. B.; Kehoe, M. A.; McCafferty, D. G.; Banfield, M. J. *J. Biol. Chem.* **2009**, *284*, 6924–6933.
- (9) Kruger, R. G.; Otvos, B.; Frankel, B. A.; Bentley, M.; Dostal, P.; McCafferty, D. G. *Biochemistry* **2004**, *43*, 1541–1551.
- (10) Δ 59-SrtA_{staph} is a truncated form of SrtA_{staph} that has identical reactivity.

JA902681K