Protein Thioester Synthesis Enabled by Sortase

Citation: Ling, Jingjing J., Rocco L. Policarpo, Amy E. Rabideau, Xiaoli Liao, and Bradley L. Pentelute. "Protein Thioester Synthesis Enabled by Sortase." Journal of the American Chemical Society 134, no. 26 (July 4, 2012): 10749-10752.

As Published: http://dx.doi.org/10.1021/ja302354v

Publisher: American Chemical Society

Persistent URL: http://hdl.handle.net/1721.1/82572

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Published in final edited form as:
J Am Chem Soc. 2012 July 4; 134(26): 10749–10752. doi:10.1021/ja302354v.

Protein Thioester Synthesis Enabled by Sortase

Jingjing Ling, **Rocco L. Policarpo**, **Amy E. Rabideau**, **Xiaoli Liao**, and **Bradley L. Pentelute***
Department of Chemistry, Massachusetts Institute of Technology, 16-573a, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Abstract

Proteins containing a C-terminal thioester are important intermediates in semi-synthesis. Currently there is one main method for the synthesis of protein thioesters that relies upon the use of engineered inteins. Here we report a simple strategy, utilizing Sortase A, for routine preparation of recombinant proteins containing a C-terminal $^{\alpha}$thioester. We used our method to prepare two different anthrax toxin cargo proteins: one containing an $^{\alpha}$thioester and another containing a D-polypeptide segment situated between two protein domains. We show that both variants can translocate through protective antigen pore. This new method to synthesize a protein thioester allows for interfacing of sortase-mediated ligation and native chemical ligation.

Chemical tailoring of proteins is a powerful approach to investigate structure function relationships and the role of post-translational modifications.[1–6] Protein semisynthesis [4,7] and total synthesis[8] are commonly used to introduce novel functionalities into proteins. Both approaches rely on native chemical ligation (NCL)–a chemoselective amide forming reaction between an $^{\alpha}$thioester and an N-terminal cysteine moiety.[9] When protein $^{\alpha}$thioesters are generated from engineered inteins and then modified with NCL the process is referred to as expressed protein ligation.[10–12] Inteins are protein self-splicing elements that can be engineered to generate protein $^{\alpha}$thioesters after self-cleavage in the presence of small molecule thiol.[10–12] Despite a number of new methods for the chemical synthesis of peptide $^{\alpha}$thioesters, intein mediated synthesis of protein $^{\alpha}$thioesters is the only direct route to generate this important functionality needed for semi-synthesis.

Here we report a sortase-mediated approach for the facile synthesis of protein $^{\alpha}$thioesters (Scheme 1). We show that the calcium dependent enzyme sortase A (SrtA) from *Staphylococcus aureus* can be used to attach synthetic oligoglycine $^{\alpha}$thioesters to a number of different proteins with good yield and efficiency. In addition, this approach allowed us to prepare two different model cargo proteins and probe their translocation through anthrax toxin protective antigen. We found that the anthrax toxin pore can translocate cargo proteins into the cell that either contained an $^{\alpha}$thioester or a D-polypeptide segment linking two protein domains.

Sortases are a class of thiol-containing transpeptidases that anchor proteins to the bacterial cell wall.[13] SrtA recognizes a C-terminal LPXTG sequence and cleaves the threonine-glycine bond to form a thioacyl-linked intermediate. [14] This intermediate is primed to react with the N-terminal amino group of an oligoglycine motif. Recent effort has shown that sortases can be co-opted for the site-specific modification of proteins at the N or C-terminus. [15–25] This sortagging, transpeptidation reaction has been used extensively to attach

*Corresponding Author: blp@mit.edu.

Supporting Information Available: Details of the peptide synthesis and purification, model SrtA and SrtA* reactions, protein construct preparation, and anthrax toxin cells assay. This material is available free of charge via the internet at http://pubs.acs.org.

virtually any water-soluble molecule to a protein of interest. Sortagging reactions are executed in calcium containing aqueous buffer (pH 7–8.5) at nanomolar to micromolar concentration of SrtA.[24] To carry out the sortagging reactions in water the N-terminal membrane spanning region of SrtA is removed.

We prepared variants of an oligoglycine $^{\alpha}$thioester and carried out model studies with peptide KLPETGG. During the initial stages of our studies Chen et al. reported an evolved SrtA (SrtA*) that had improved enzymatic kinetics. [26] Once we confirmed SrtA* and SrtA gave similar product yield in model peptide studies, we chose to work exclusively with SrtA* because it substantially shortened reaction time. Reactions with SrtA* were complete in 30 minutes as opposed to 2 hours with SrtA (Figure S1–S2). The analytical RP-HPLC traces for a model reaction between KLPETGG and G_5F-COSR in the presence of SrtA* are shown in Figure 1. We used RP-HPLC to purify the sortagged reaction product (6.9 mg, 54% yield) and then performed an NCL reaction under standard conditions. After purification, the NCL product was isolated (83 % yield, 2 mg) and characterized by high resolution LCMS (Figure 1b–d). We conducted additional sortagging studies to investigate two aspects of the oligoglycine $^{\alpha}$thioester: varying the C-terminal amino acid (Xaa) or varying the number of glycine residues. It has been demonstrated that the reactivity of a peptide $^{\alpha}$thioester is dependent upon the identity of the C-terminal amino acid; Gly is more reactive than Leu, for instance.[27,28] Product yields determined from analytical RP-HPLC ranged from 54–68 % with SrtA (Table S1) and 74–84 % with SrtA* (Table 1) suggesting the C-terminal amino acid Xaa does not significantly affect the reaction yield. We next investigated the relationship between reaction yield and the number of N-terminal glycine residues. The best yields were obtained when three or more glycines were used (Table 2).

With the ability to sortag oligopeptide $^{\alpha}$thioesters to peptides bearing a C-terminal acceptor sequence, we carried out investigations with three different model proteins. We prepared variants of eGFP, lethal factor N-terminal domain (LF_N) from anthrax toxin,[29] and lethal factor N-terminal domain fused to diphtheria toxin A-chain (LF_N-DTA).[29,30] DTA is the catalytic domain from diphtheria toxin and catalyzes the ADP-ribosylation of elongation factor-2 within the cytosol thereby halting protein synthesis and causing cell death.[31,32] LF_N and LF_N-DTA are moieties used to probe the molecular basis of anthrax toxin protein translocation.[33] We prefer to express proteins as SUMO-protein fusions because expression yields are enhanced and the native N-terminus is generated after removal of SUMO.

We first investigated, by use of high resolution LCMS, whether SrtA* could tag thioesters onto our three different model proteins and found that protein thioester products are readily formed in 30 minutes (Figure S4). We then proceeded to prepare protein thioesters on milligram scale using the approach shown in Figure 2a. In particular, a one-pot method was employed whereby we first removed the N-terminal SUMO tag with SUMO protease and subsequently added SrtA*, Ni-NTA agarose beads, and oligoglycine $^{\alpha}$thioester peptide. After completion of the SrtA*-mediated ligation (SML) reaction, pure protein thioester was isolated by simple filtration and concentration because all unreacted material remained bound to the Ni-NTA agarose beads. We obtained good yields of pure protein $^{\alpha}$thioester and observed minimal amounts of SrtA-mediated hydrolysis when analyzed by high resolution LCMS (Figure 2b–d). Our isolated yields for the three different model protein thioesters ranged from 40–80% (Table S3), which is consistent with prior reports for SML. When we monitored the sortagging reaction by SDS-PAGE and LCMS, we found some product still bound to the Ni-NTA agarose beads for LF_N and LF_N-DTA but not eGFP (Figure S5) suggesting the reaction yields are in part determined by the intrinsic properties of the protein. For eGFP-COSR we carried out NCL with the model peptide CFRALKAA under standard conditions (pH 7, TCEP, MPAA catalyst) and obtained 0.8 mg product (98% yield) (Figure S6).[34]

One of the proteins we have chosen to work with in these studies is LF_N-DTA. This chimera serves as an excellent model cargo protein to probe translocation through the anthrax toxin protective antigen (PA) pore. Having facile chemical access to this molecule allows for the incorporation of non-natural moieties to further elucidate the mechanisms in which this protein enters the cell through the PA pore. To begin these investigations, we aimed to study the translocation of LF_N-DTA-COSR and LF_N-D-linker-DTA, where "D-linker" refers to a small D-peptide tether between LF_N and DTA. The preparation of these synthetic constructs is simplified by our new approach to generate protein thioesters. We prepared LF_N-D-linker-DTA using the approach shown in Figure 3a. The synthetic approach involves first sortagging an oligoglycine thioester containing D-amino acids to LF_N followed by NCL to ligate on the C-terminal DTA domain. We also prepared LF_N-L-linker-DTA using the same approach to serve as a control in our translocation assays (Figure S7). By comparing the translocation efficiency of LF_N-D-linker-DTA against the L-amino acid variant, we were able to evaluate the stereochemical requirement for successful translocation through the pore.

With our modified LF_N-DTA variants in hand, we tested if they could translocate through PA.[29] In this protein translocation assay, anthrax toxin PA and LF_N-DTA were added to CHO-K1 cells and the amount of LF_N-DTA delivered to the cytosol was inferred by measuring protein synthesis via ^{3}H-Leu incorporation into the cellular proteome. [30,33] Once LF_N-DTA accesses the cytosol it inhibits protein synthesis. The protein synthesis levels for the variants are shown in Figure 4. We found all variants to translocate at levels similar to wild-type LF_N-DTA.

The translocation of LF_N-DTA-COSR suggests that the amide bond can be replaced with a thioester without affecting passage through the pore. We plan to explore the possibility of using the thioester functionality to capture translocated proteins and their possible binding partners in the cytosol.

Our findings indicate that a model cargo protein containing non-natural modifications in the linker region is capable of translocation through the PA pore. Successful translocation of LF_N-D-linker-DTA demonstrates that the stereochemical constraints on PA-mediated translocation are minimal, provided that requirements for prepore binding and translocation initiation are met. This finding is in-line with a prior report investigating replacement of the N-terminus of LF_N with a D-peptide segment.[35]

This method to generate an $^{\alpha}$thioester provides flexibility for covalent modification of proteins by interfacing key ligation approaches. The termini or linker regions of proteins can be site-specifically modified in a modular manner by the use of sortase-mediated ligation and native chemical ligation. Linker regions between two protein domains can be easily modified with natural or unnatural moieties by synthesizing various peptide thioesters. We demonstrated the utility of this approach by preparing a variant of LF_N-DTA in containing a D-peptide segment in the linker region. To the best of our knowledge, this is the first time that a D-peptide fragment has been installed between two recombinantly expressed protein domains.

Our method of protein thioester generation could be used to overcome solubility limitations of sortagging reactions. We have found that sortagging efficiency directly parallels oligoglycine nucleophile concentration. Hydrolysis of LPXTG is a competing side-reaction for sortagging endeavors.[36,37] We found that optimal concentrations of oligogylcine needed to be 300 μM and above; at lower concentrations, we observed significant LPSTGG hydrolysis (Figure S8). However, some oligoglycine peptides and proteins are insoluble at these concentrations. Our approach could be used to overcome this solubility limitation

because short oligoglycine peptide thioesters tend to be highly soluble in aqueous solution. Moreover, if a given peptide thioester is still found to be insoluble it can be further modified to increase solubility by installing an Arg tag on the thioester leaving group.[38] After sortagging with a peptide $^{\alpha}$thioester, native chemical ligation can be carried out in solvents and buffers that denature sortase but solubilize the coupling partners. Common solubilizing agents that can be used in native chemical ligation reactions include denaturants (urea or guanidinium), detergents, and organic solvents.

However, it should be noted that the utility of sortagging is limited to ligations in which introduction of the $LPXTG_n$ moiety does not significantly alter protein structure. In cases when the $LPXTG_n$ segment may alter the properties of the protein, other ligation methods should be considered. Despite this limitation, sortagging has been used in numerous instances to modify proteins for biological study.[15–25] In our case, the $LPSTG_5$ linker did not affect the translocation of our LF_N-DTA constructs. One way to overcome this limitation would be to evolve sortase to recognize different and possibly shorter sequences. Recently, Piotukh et al. evolved SrtA to recognize FPXTG or APXTG motif, suggesting this may be possible. [39]

In summary, we have developed a SrtA-mediated ligation approach for the synthesis of recombinant protein thioesters. Protein thioesters are generated in 30 minutes with good yields, and pure products are isolated without elaborate purification steps.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported the MIT startup funds for B.L.P., Agency for Science, Technology and Research National Science Scholarship (PhD) for J.L., a National Science Foundation Graduate Research Fellowship for A.E.R. We also thank R. John Collier for providing some of the laboratory equipment used to carry out the experiments. Some of the recombinant proteins employed in the study were prepared in the Biomolecule Production Core for the New England Regional Center of Excellence, supported by NIH grant number AI057159. We thank R. Ross and Ben T. Seiler for these services. We thank Mark Simon and Xiuyuan Li for for valuable comments and discussions. We thank Hidde L. Ploegh and Carla Guimaraes for the SrtA plasmid.

References

1. Murakami M, Okamoto R, Izumi M, Kajihara Y. Angew Chem Int Ed Engl. 2012; 51:3494.
2. Nagorny P, Sane N, Fasching B, Aussedat B, Danishefsky SJ. Angew Chem Int Ed Engl. 2012; 51:975. [PubMed: 22162182]
3. Bavikar SN, Spasser L, Haj-Yahya M, Karthikeyan SV, Moyal T, Kumar KS, Brik A. Angew Chem Int Ed Engl. 2012; 51:758. [PubMed: 22131237]
4. Vila-Perello M, Muir TW. Cell. 2010; 143:191. [PubMed: 20946979]
5. Scheuermann JC, de Ayala Alonso AG, Oktaba K, Ly-Hartig N, McGinty RK, Fraterman S, Wilm M, Muir TW, Muller J. Nature. 2010; 465:243. [PubMed: 20436459]
6. Chatterjee C, McGinty RK, Fierz B, Muir TW. Nat Chem Biol. 2010; 6:267. [PubMed: 20208522]
7. Muir TW. Annu Rev Biochem. 2003; 72:249. [PubMed: 12626339]
8. Kent SB. Chem Soc Rev. 2009; 38:338. [PubMed: 19169452]
9. Dawson PE, Muir TW, Clark-Lewis I, Kent SB. Science. 1994; 266:776. [PubMed: 7973629]
10. Severinov K, Muir TW. J Biol Chem. 1998; 273:16205. [PubMed: 9632677]
11. Muir TW, Sondhi D, Cole PA. Proc Natl Acad Sci USA. 1998; 95:6705. [PubMed: 9618476]
12. Evans TC Jr, Benner J, Xu MQ. Protein Sci. 1998; 7:2256. [PubMed: 9827992]
13. Mazmanian SK, Liu G, Ton-That H, Schneewind O. Science. 1999; 285:760. [PubMed: 10427003]

14. Navarre WW, Schneewind O. Mol Microbiol. 1994; 14:115. [PubMed: 7830549]
15. Wu Z, Guo X, Guo Z. Chem Commun (Camb). 2011; 47:9218. [PubMed: 21738926]
16. Popp MW, Ploegh HL. Angew Chem Int Ed Engl. 2011; 50:5024. [PubMed: 21538739]
17. Wu Z, Guo X, Wang Q, Swarts BM, Guo Z. J Am Chem Soc. 2010; 132:1567. [PubMed: 20078124]
18. Proft T. Biotechnol Lett. 2010; 32:1. [PubMed: 19728105]
19. Clancy KW, Melvin JA, McCafferty DG. Biopolymers. 2010; 94:385. [PubMed: 20593474]
20. Yamamoto T, Nagamune T. Chem Commun (Camb). 2009; 1022
21. Tsukiji S, Nagamune T. Chembiochem. 2009; 10:787. [PubMed: 19199328]
22. Guo X, Wang Q, Swarts BM, Guo Z. J Am Chem Soc. 2009; 131:9878. [PubMed: 19583255]
23. Samantaray S, Marathe U, Dasgupta S, Nandicoori VK, Roy RP. J Am Chem Soc. 2008; 130:2132. [PubMed: 18229923]
24. Popp MW, Antos JM, Grotenbreg GM, Spooner E, Ploegh HL. Nat Chem Biol. 2007; 3:707. [PubMed: 17891153]
25. Mao H, Hart SA, Schink A, Pollok BA. J Am Chem Soc. 2004; 126:2670. [PubMed: 14995162]
26. Chen I, Dorr BM, Liu DR. Proc Natl Acad Sci USA. 2011; 108:11399. [PubMed: 21697512]
27. Bang D, Pentelute BL, Kent SB. Angew Chem Int Ed Engl. 2006; 45:3985. [PubMed: 16639756]
28. Hackeng TM, Griffin JH, Dawson PE. Proc Natl Acad Sci USA. 1999; 96:10068. [PubMed: 10468563]
29. Young JA, Collier RJ. Annu Rev Biochem. 2007; 76:243. [PubMed: 17335404]
30. Milne JC, Blanke SR, Hanna PC, Collier RJ. Mol Microbiol. 1995; 15:661. [PubMed: 7783638]
31. Collier RJ, Cole HA. Science. 1969; 164:1179. [PubMed: 4305968]
32. Collier RJ. J Mol Biol. 1967; 25:83. [PubMed: 4291872]
33. Pentelute BL, Sharma O, Collier RJ. Angew Chem Int Ed Engl. 2011; 50:2294. [PubMed: 21351339]
34. Johnson ECB, Kent SBH. J Am Chem Soc. 2006; 128:6640. [PubMed: 16704265]
35. Pentelute BL, Sharma O, Collier RJ. Angew Chem Int Ed Engl. 2011; 50:2294. [PubMed: 21351339]
36. Ton-That H, Mazmanian SK, Faull KF, Schneewind O. J Biol Chem. 2000; 275:9876. [PubMed: 10734144]
37. Mohlmann S, Mahlert C, Greven S, Scholz P, Harrenga A. Chembiochem. 2011; 12:1774. [PubMed: 21656631]
38. Johnson EC, Kent SB. Tetrahedron Lett. 2007; 48:1795. [PubMed: 19177172]
39. Piotukh K, Geltinger B, Heinrich N, Gerth F, Beyermann M, Freund C, Schwarzer D. J Am Chem Soc. 2011; 133:17536. [PubMed: 21978125]

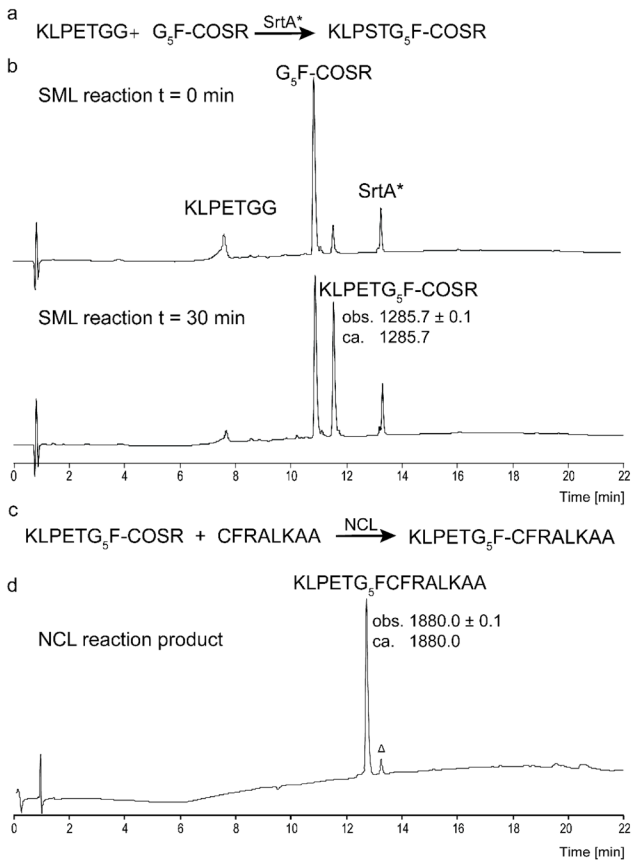


Figure 1.

LC data for peptide $^{\alpha}$thioester synthesis using SrtA* followed by NCL. **(a)** SrtA*-mediated ligation (SML) of model peptide. **(b)** Analytical RP-HPLC of crude SML reaction for t = 0 min and t = 30 min. Reaction conditions: 1 mM G_5F-COSR, 500 μM KLPETGG, 5 μM SrtA*, SrtA* buffer, pH 7.5. **(c)** Native chemical ligation (NCL) with sortagged thioester reaction product. (**d)** Analytical RP-HPLC of purified NCL reaction product. Reaction conditions are reported in the supporting information. All analytical RP-HPLC signals are measured at 214 nm. Δ: impurity peak.

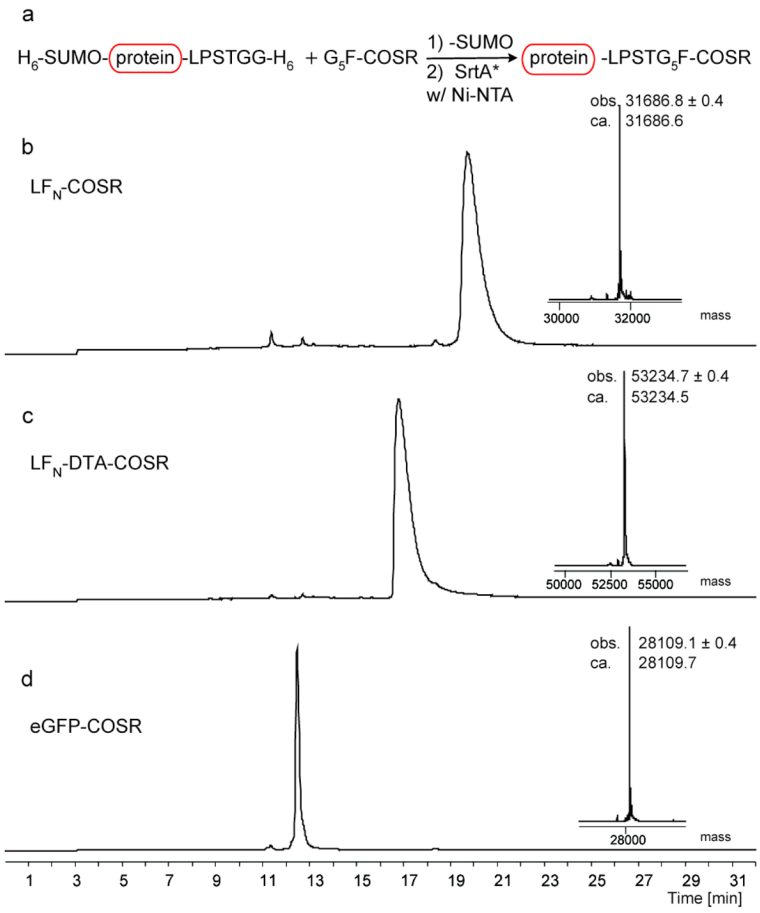


Figure 2.

LCMS data for protein $^{\alpha}$thioester synthesis using SrtA*. **(a)** SML of model proteins. **(b)–(d)** LCMS and deconvoluted MS (inset) of the main component for LF_N, LF_N-DTA and eGFP thioesters. Reaction conditions: **(1)** SUMO cleavage: 1 μg SUMO protease per 1 mg protein for 90 min at room temperature. **(2)** SML: 500 μM G_5F-COSR, 100 μM protein-LPSTGG, 5 μM SrtA*, 75 μl Ni-NTA agrose slurry per mg protein, SrtA* buffer, pH 7.0, 30 min. LCMS traces are shown as total ion current.

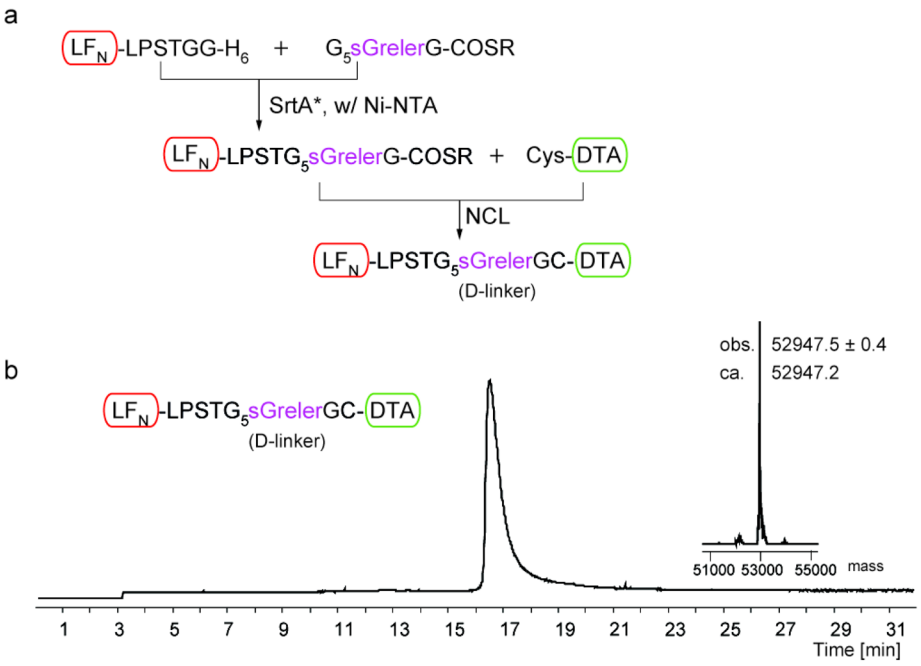


Figure 3.

Installing a D-peptide segment between LF_N and DTA. **(a)** Synthetic strategy used to modify the linker region (lower case = D-amino acid). **(b)** LCMS and deconvoluted MS (inset) for LF_N-D-linker-DTA. The reaction conditions are reported in the supporting information. LCMS traces are shown as total ion current.

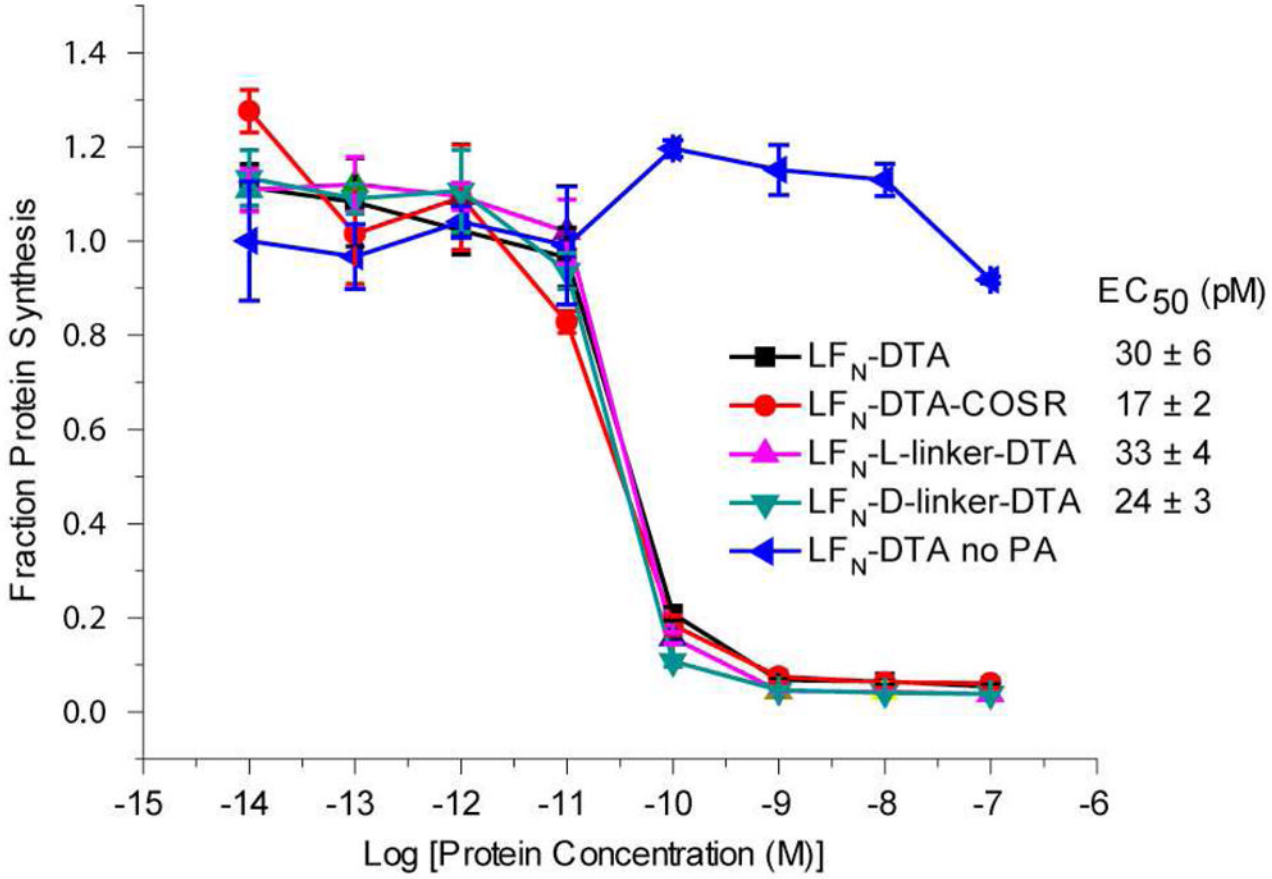
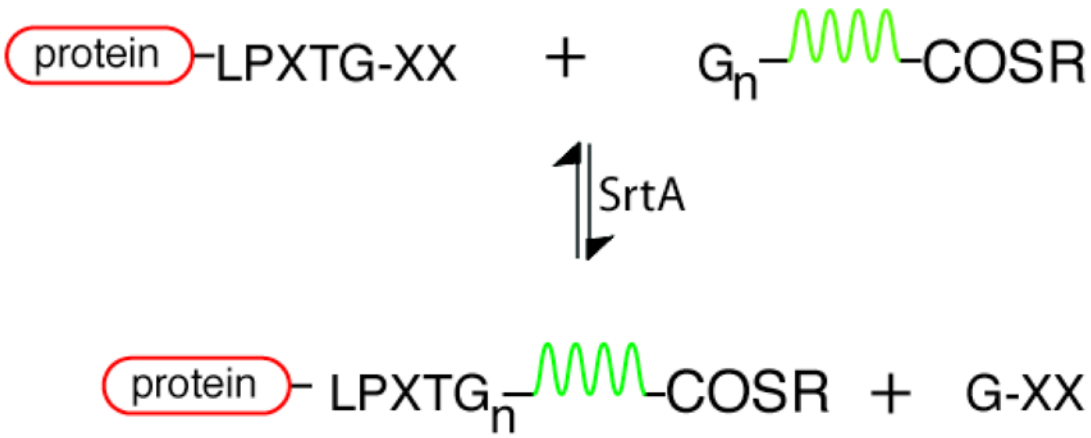


Figure 4.

Translocation of LF_N-DTA variants into CHO-K1 cells. Cells were incubated with LF_N-DTA, LF_N-DTA-COSR, LF_N-L-linker-DTA, or LF_N-D-linker-DTA at different concentrations in the presence or absence of 10 nM PA for 30 minutes. The media was then replaced with leucine-free medium supplemented with 1 μCi/mL ^{3}H-Leu, and incubated for 1 hour. After incubation, the cells were washed three times with cold PBS, scintillation fluid was added, and incorporated ^{3}H-Leu was determined by scintillation counting. Each data point represents the average of three trials at the specified concentration.



Scheme 1.
Protein $^{\alpha}$thioester synthesis using SrtA

Table 1

SrtA*-mediated ligation reaction yields with different G_5-Xaa-COSR to model peptide

Xaa	Gly	Phe	Ser	Leu
Yield (%)	76(2)	84(3)	74(3)	75(2)

Reagents and conditions: 500 μM KLPETGG was reacted with 1 mM G_5-Xaa-COSR for 30 min in the presence of 5 μM SrtA* and SrtA* buffer (10 mM $CaCl_2$, 50 mM Tris, 150 mM NaCl) pH 7.5. The thioester R group was -CH_2-CH_2-L-$RCONH_2$. Standard deviations are shown in parentheses.

Table 2

SrtA*-mediated ligation reaction yields with increasing number of glycines in G_n-COSR to model peptide.

Number of Gly	1	2	3	4	5	6
Yield(%)	62(1)❖	46(8)	68(6)	69(3)	70(1)	59(8)

Reagents and conditions: 500 μM KLPETGG was reacted with 1 mM G_n-COSR for 30 min in the presence of 5 μM SrtA* and SrtA* buffer, pH 7.5.

❖Both KLPETGG-COSR and KLPETG-COSR were formed when G-COSR was used as the nucleophile. The yield represents the sum of both reactions.