Enhancing the Efficiency of Sortase-Mediated Ligations through Nickel-Peptide Complex Formation

R. David Row, Travis J. Roark, Marina C. Philip, Lorena L. Perkins, and John M. Antos

Western Washington University, Department of Chemistry, Bellingham, WA

Supporting Information

Contents:

I. Materials and instrumentation

Chemicals and Reagents. Unless otherwise noted, all chemicals were obtained from commercial sources and used without further purification. Fmoc-Gly-Wang resin (100-200 mesh, 0.79 mmole/g) was purchased from Novabiochem. Fmoc Rink Amide MBHA resin (0.45 mmole/g) was purchased from Anaspec. Fmoc-L-Leu-OH, Fmoc-L-Pro-OH, Fmoc-L-Glu(OtBu)-OH, Fmoc-L-Thr(OtBu)-OH, Fmoc-Gly-OH, Fmoc-L-His(Trt)-OH, Fmoc-L-Lys(Boc)-OH, and *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HBTU) were purchased from Chem-Impex International. Piperidine and $NiSO_4 \bullet 6H_2O$ were obtained from Alfa Aesar. *N,N'*-diisopropylethylamine (DIPEA), peptide synthesis grade trifluoroacetic acid (TFA), dimethyl sulfoxide (DMSO), and methylene chloride (CH_2Cl_2) were obtained from Fisher Scientific. *N*-methyl-2-pyrrolidone (NMP) was purchased from either Fisher Scientific or Advanced Chemtech. A colorimetric ninhydrin test kit and Fmoc-L-Lys(DNP)-OH were purchased from Anaspec. Dansyl chloride (Dns-Cl) and triisopropylsilane (TIPS) were obtained from Acros. Diethyl ether was purchased from J.T. Baker. ChromAR® grade acetonitrile (MeCN) was obtained from Macron Fine Chemicals. Bovine insulin solution (I0516) and diglycine (GG) peptide were purchased from Sigma-Aldrich. Water used in biological procedures or as a reaction solvent was purified using a Milli-Q Advantage A10 system (Millipore).

RP-HPLC. HPLC purifications and analyses were achieved using a Dionex Ultimate 3000 HPLC system. Analytical scale separations were performed with a Phenomenex Kinetex 2.6 µm, 100 Å C18 column (3.0 x 100 mm). Analytical scale analyses were achieved using either **Method A** or **Method B**. **Method A**: MeCN (0.1% TFA) / H_2O (0.1% TFA) mobile phase. Flow rate = 0.7 mL/min. Gradient = 10% MeCN (0.0-0.5 min), 10% MeCN → 90% MeCN (0.5-6.0 min), hold 90% MeCN (6.0-7.0 min), 90% MeCN → 10% MeCN (7.0-7.1 min), re-equilibrate at 10% MeCN (7.1-10.0 min). **Method B**: MeCN (0.1% TFA) / H_2O (0.1% TFA) mobile phase. Flow rate = 0.7 mL/min. Gradient = 20% MeCN (0.0-1.0 min), 20% MeCN → 40% MeCN (1.0-9.0 min), 40% MeCN → 90% MeCN (9.0-10.0 min), hold 90% MeCN (10.0-12.0 min), 90% MeCN → 20% MeCN (12.0-12.1 min), re-equilibrate at 20% MeCN (12.1-15.0 min). Semi-preparative purifications were performed with a Phenomenex Luna 5 µm, 100 Å C18 column (10 x 250 mm). Semi-preparative purifications were achieved using either **Method C** or **Method D**. **Method C**: MeCN (0.1% TFA) / H_2O (0.1% TFA) mobile phase. Flow rate = 0.5 → 4.0 mL/min (0.0-2.0 min), hold 4.0 mL/min (2.0-17.01 min), 4.0 → 0.5 mL/min (17.01-19.0 min). Gradient = 20% MeCN (0.0-2.0 min), 20% MeCN → 90% MeCN (2.0-15.0 min), hold 90% MeCN (15.0-17.0 min), 90% MeCN → 10% MeCN (17.0-17.01 min), re-equilibrate at 10% MeCN (17.01-19.0 min). **Method D**: MeCN (0.1% TFA) / H_2O (0.1% TFA) mobile phase. Flow rate = 0.5 → 3.0 mL/min (0.0-2.0 min), hold 3.0 mL/min (2.0-17.01 min), 3.0 → 0.5 mL/min (17.01-19.0 min). Gradient = 20% MeCN (0.0-2.0 min), 20% MeCN → 90% MeCN (2.0-15.0 min), hold 90% MeCN (15.0-17.0 min), 90% MeCN → 10% MeCN (17.0-17.01 min), re-equilibrate at 10% MeCN (17.01-19.0 min).

Mass Spectrometry. Liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS) analysis was carried out with an API 2000 Triple Quadrupole mass spectrometer (Applied Biosystems) and an Agilent 1100 HPLC system equipped with a Phenomenex Aeris Widepore 3.6 µm 200 Å XB-C8 column (4.6 x 150 mm). All samples were analyzed using the following method: H_2O (0.1% formic acid) / organic (95% MeCN, 5% isopropanol, 0.1% formic acid) mobile phase. Flow rate = 1.25 mL/min. Gradient = 10% organic (0.0-1.0 min), 10% organic → 90% organic (1.0-10.0 min). Deconvolution of protein charge ladders was achieved using a maximum entropy algorithm provided by Analyst 1.4.2 software.

UV-vis Spectroscopy. UV-vis spectroscopy was performed on a Nanodrop™ ND-1000 spectrophotometer (Thermo Scientific, USA).

II. Synthesis and characterization of peptide model substrates

General Methods. All peptides were synthesized using manual Fmoc solid phase peptide synthesis (SPPS). Syntheses were performed in Extract Clean™ SPE 15 mL reservoirs fitted with appropriate frits and inlet/outlet caps. All manipulations (washing, coupling, deprotection, etc.) were conducted at room temperature and included gentle agitation on a bench-top rocking platform unless otherwise indicated.

*Synthesis of Ac-K(DNP)LPETGGHG (**1**) and Ac-K(DNP)LPETGGGG.* Solid phase synthesis reservoirs were loaded with 126 mg (0.10 mmol) each of Fmoc-Gly-Wang resin (100-200 mesh, 0.79 mmol/g). The resins were initially washed/swollen with ~10 mL of NMP (3x, 10 min per wash). Fmoc removal was achieved by treatment with 10 mL of 80:20 NMP/piperidine at room temperature (2x, 20 min per treatment). Resins were then washed with ~10 mL of NMP (3x, 3-5 min per wash). Suitably protected amino acid building blocks were then coupled as follows: Fmoc-protected amino acid (0.30 mmol, 3.0 equivalents relative to resin loading), HBTU (114 mg, 0.30 mmol), and DIPEA (155 µL, 0.89 mmol) were dissolved in 3 mL of NMP. This solution was mixed thoroughly, and then added to the deprotected resins. Couplings were incubated for 1-24 hours at room temperature. Resins were then washed with ~10 mL of NMP (3x, 3-5 min per wash). The success of each coupling reaction was assessed using a colorimetric ninhydrin test. In the event that the coupling was incomplete, the amino acid coupling step was repeated. Fmoc deprotection was then carried out by washing with 10 mL of 80:20 NMP/piperidine (2x, 20 min per treatment), followed by washing with ~10 mL of NMP (3x, 3-5 min per wash each). Repeated cycles of amino acid coupling and Fmoc deprotection were then performed to assemble the desired peptide sequences. **Note:** A commercially available lysine building block (Fmoc-L-Lys(DNP)-OH, Anaspec) was used to install the 2,4-dinitrophenyl (DNP) chromophore. This residue was coupled following the standard protocol described above. Following removal of the final Fmoc protecting group, the peptide N-termini were capped using acetic anhydride. Capping was achieved by treating the resins with a solution of acetic anhydride (472 µL, 5.0 mmol) and DIPEA (860 µL, 4.9 mmol) in 3.0 mL of NMP for 3 h at room

temperature. Resins were then washed with ~10 mL of NMP (3x, 3-5 min per wash), followed by ~10 mL of CH_2Cl_2 (3x, 3-5 min per wash). Peptides were cleaved from the resin by treatment with 5 mL of 95:2.5:2.5 TFA/TIPS/H_2O (2x, ~30 min per treatment). The combined cleavage solutions were then concentrated by rotary evaporation. The remaining residues were then dissolved in 1 mL of NMP and precipitated by dropwise addition to 40 mL of rapidly stirring, dry-ice chilled diethyl ether. The precipitated peptides were then collected by centrifugation, and the crude products were dried overnight under vacuum. Peptides **1** and Ac-K(DNP)LPETGGGG were then purified by RP-HPLC (Method C). Pure fractions were pooled and lyophilized. The purity and identity of **1** and Ac-K(DNP)LPETGGGG were confirmed by RP-HPLC and LC-ESI-MS (**Figure S1**). Stock solutions were prepared in H_2O and the concentrations were estimated by absorbance at 365 nm using an extinction coefficient of 17,400 $M^{-1}cm^{-1}$.

Synthesis of Ac-K(DNP)LPETGG. Ac-K(DNP)LPETGG was synthesized from Ac-K(DNP)LPETGGGG using a $SrtA_{staph}$ catalyzed transacylation. Crude Ac-K(DNP)LPETGGGG (precipitated from diethyl ether, not RP-HPLC purified) was dissolved in H_2O to generate a 4.8 mM stock solution. 104 µL of this stock solution was then combined with 10 equivalents of diglycine (GG) (47.6 µL of a 105 mM stock solution in H_2O), $SrtA_{staph}$ (11.3 µL of a 444 µM stock solution in 90% 50 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol), and 10x sortase reaction buffer (50 µL of 500 mM Tris pH 7.5, 1500 mM NaCl, 100 mM $CaCl_2$). H_2O was then added to bring the final volume to 500 µL. The reaction was incubated at 37 ^{o}C for 2 h. Ac-K(DNP)LPETGG was purified directly from this reaction mixture by RP-HPLC (Method C). Pure fractions were pooled and lyophilized. The purity and identity of Ac-K(DNP)LPETGG was confirmed by RP-HPLC and LC-ESI-MS (**Figure S1c**). A stock solution was prepared in H_2O, and the concentration was estimated by absorbance at 365 nm using an extinction coefficient of 17,400 $M^{-1}cm^{-1}$.

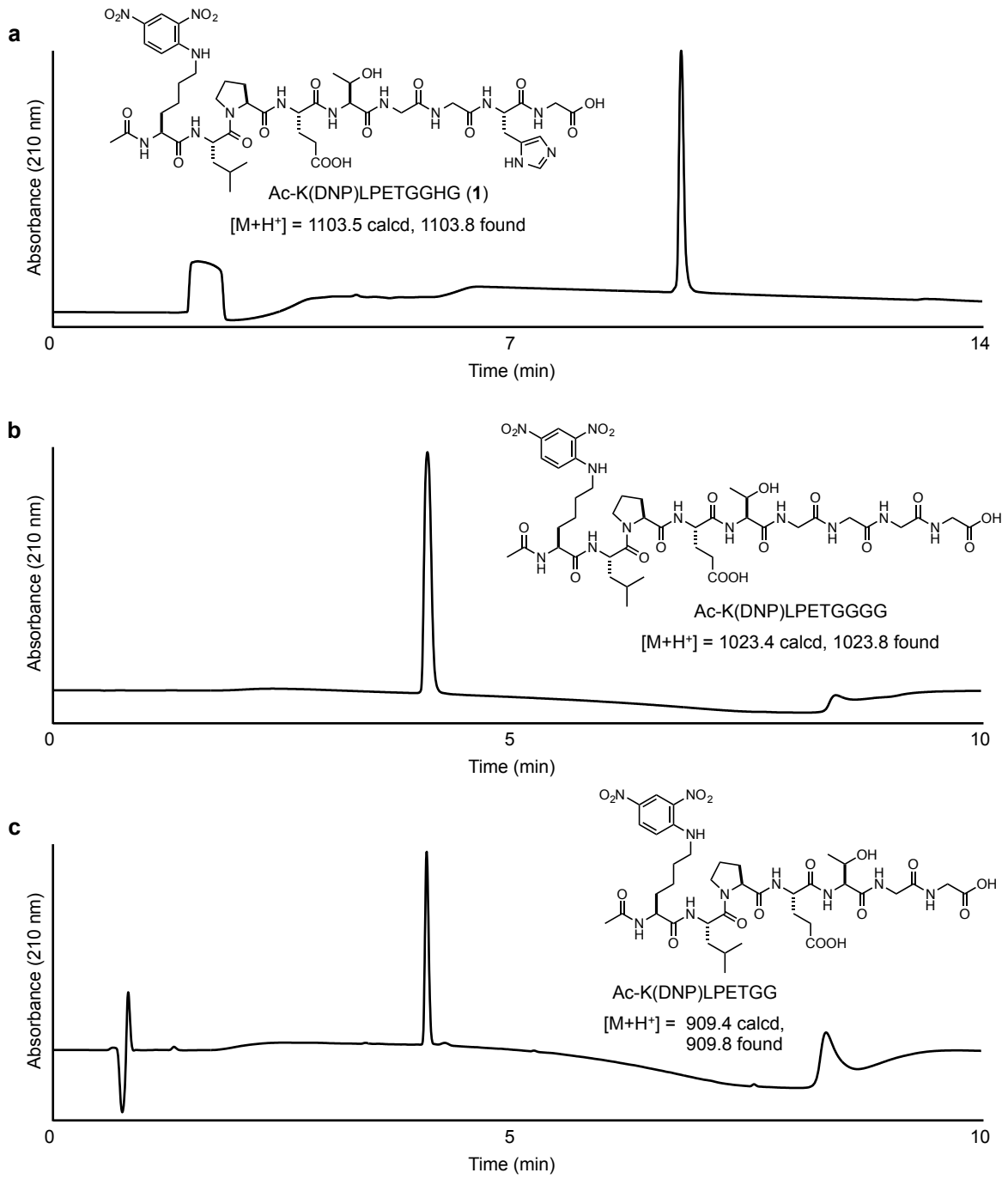

Figure S1. (**a**) RP-HPLC chromatogram (Method C, 210 nm) for purified **1** and LC-ESI-MS characterization. (**b**) RP-HPLC chromatogram (Method A, 210 nm) for purified Ac-K(DNP)LPETGGGG and LC-ESI-MS characterization. (**c**) RP-HPLC chromatogram (Method A, 210 nm) for purified Ac-K(DNP)LPETGG and LC-ESI-MS characterization.

*Synthesis of GK(Dns) (**2**), GGK(Dns), and GGGK(Dns).* Dansylated glycine nucleophiles were prepared according to the scheme shown in **Figure S2**. Solid phase synthesis reservoirs were loaded with 440 mg (0.20 mmol) each of Fmoc-protected Rink amide MBHA resin (0.45 mmol/g). Resins were initially washed/swollen with ~10 mL of NMP (3x, 10 min per wash). Fmoc removal was achieved by treatment with 10 mL of 80:20 NMP/piperidine at room temperature (2x, 10 min per treatment). Resins were then washed with ~10 mL of NMP (3x, 5-10 min per wash). Suitably protected amino acid building blocks were then coupled as follows: Fmoc-protected amino acid (0.60 mmol, 3.0 equivalents relative to resin loading), HBTU (228 mg, 0.60 mmol), and DIPEA (320 µL, 1.8 mmol) were dissolved in 10 mL of NMP. This solution was mixed thoroughly, and then added to the deprotected resins. Couplings were incubated for 1.5 hours at room temperature. Resins were then washed with ~10 mL of NMP (3x, 5-10 min per wash). The success of each coupling reaction was assessed using a colorimetric ninhydrin test. In the event that the coupling was incomplete, the amino acid coupling step was repeated. Fmoc deprotection was then carried out by washing with 10 mL of 80:20 NMP/piperidine (2x, 10 min per treatment), followed by washing with ~10 mL of NMP (3x, 5-10 min per wash). Repeated cycles of amino acid coupling and Fmoc deprotection were then performed to assemble resin bound intermediates **S1**-**S3**.

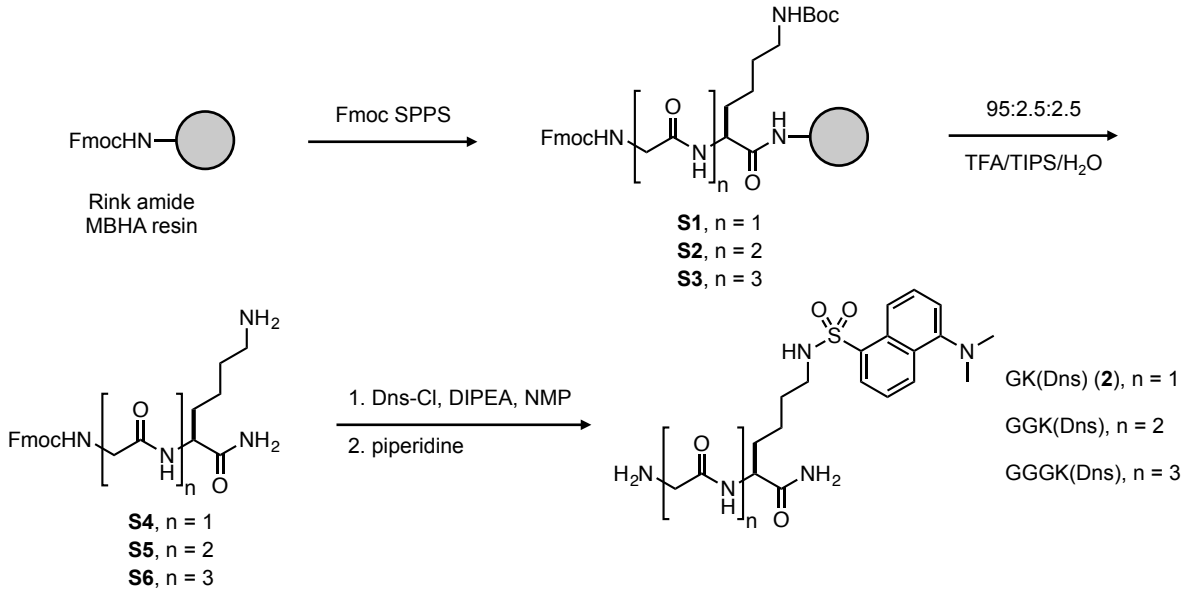

Figure S2. Synthesis of dansylated glycine nucleophiles GK(Dns) (**2**), GGK(Dns), and GGGK(Dns).

Peptides were cleaved from the solid support by treatment with 5 mL of 95:2.5:2.5 TFA/TIPS/H_2O (2x, ~30 min per treatment). The combined cleavage solutions were then concentrated by rotary evaporation. The remaining residues were resuspended in 10 mL of CH_2Cl_2 and again evaporated. This process was repeated one additional time to remove residual TFA. Crude peptides were then resuspended in 5 mL of water and lyophilized to yield peptide intermediates **S4**-**S6**. The identities of **S4**-**S6** were confirmed by LC-ESI-MS (**S4**: $[M+H^+]$ = 425.2 calcd, 425.3 found, **S5**: $[M+H^+]$ = 482.2 calcd, 482.4 found, **S6**: $[M+H^+]$ = 539.3 calcd, 539.5 found). These materials were used without further purification.

Crude **S4**-**S6** were dansylated using the following general protocol: 20-25 mg of crude **S4**, **S5**, or **S6** was combined with 2 molar equivalents of dansyl chloride (Dns-Cl), and 5 molar equivalents of DIPEA in NMP. The final concentration of **S4**-**S6** in each dansylation reaction was 100 mM, and final reaction volumes ranged from 415 µL to 470 µL. Reactions were incubated at room temperature for 1 h. Piperidine was then added to a final concentration of 20% (v/v) and the reactions were incubated for an additional 10 min. The reactions were then added dropwise to 40 mL of rapidly stirring, dry-ice chilled diethyl ether in order to precipitate crude GK(Dns) (**2**), GGK(Dns) and GGGK(Dns). Precipitated peptides were collected by centrifugation and purified by RP-HPLC (Method C). Pure fractions were pooled and lyophilized. The purity and identity of GK(Dns) (**2**), GGK(Dns), and GGGK(Dns) were confirmed by RP-HPLC and LC-ESI-MS (**Figure S3**). Stock solutions of all nucleophiles were prepared in 1:1 DMSO / H_2O and the concentrations were estimated by absorbance. Specifically, an absorbance maximum was observed in the range 330-336 nm for all nucleophiles. This absorbance value was used with an extinction coefficient of 4300 $M^{-1}cm^{-1}$ to estimate concentration.[1,2]

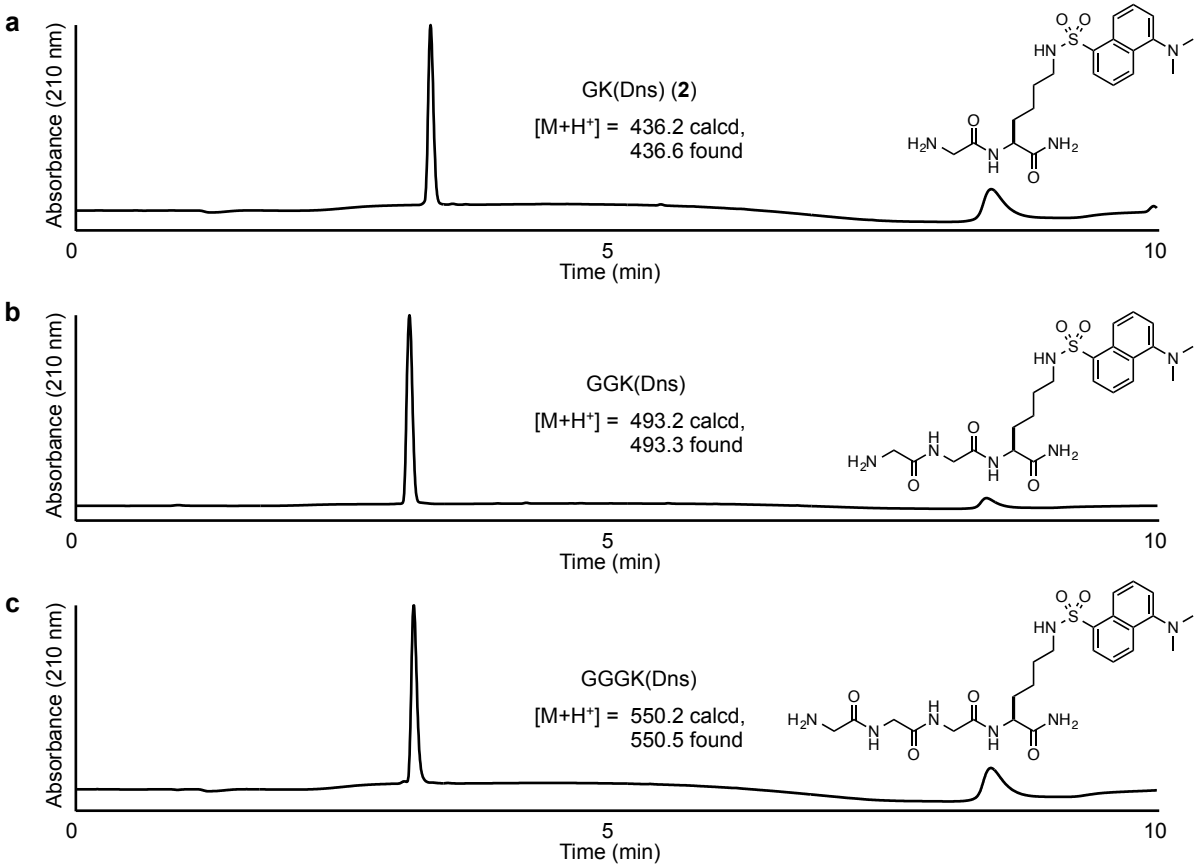

Figure S3. (**a**) RP-HPLC chromatogram (Method A, 210 nm) for purified GK(Dns) (**2**) and LC-ESI-MS characterization. (**b**) RP-HPLC chromatogram (Method A, 210 nm) for purified GGK(Dns) and LC-ESI-MS characterization. (**c**) RP-HPLC chromatogram (Method A, 210 nm) for purified GGGK(Dns) and LC-ESI-MS characterization.

*Synthesis of Bz-LPETGGHG (**7**).* A solid phase synthesis reservoir was loaded with 127 mg (0.10 mmol) of Fmoc-Gly-Wang resin (100-200 mesh, 0.79 mmol/g). The resin was initially washed/swollen with ~10 mL of NMP (3x, 10 min per wash). Fmoc removal was achieved by treatment with 10 mL of 80:20 NMP/piperidine at room temperature (2x, 20 min per treatment). The resin was then washed with ~10 mL of NMP (3x, 3-5 min per wash). Suitably protected amino acid building blocks were then coupled as follows: Fmoc-protected amino acid (0.30 mmol, 3.0 equivalents relative to resin loading), HBTU (114 mg, 0.30 mmol), and DIPEA (86 µL, 0.49 mmol) were dissolved in 3 mL of NMP. This solution was mixed thoroughly, and then added to the deprotected resin. Couplings were incubated for 1 hour at room temperature. Resins were then washed with ~10 mL of NMP (3x, 3-5 min per wash). The success of each coupling reaction was assessed using a colorimetric ninhydrin test. In the event that the coupling was incomplete, the amino acid coupling step was repeated. Fmoc deprotection was then carried out by washing with 10 mL of 80:20 NMP/piperidine (2x, 20 min per treatment), followed by washing with ~10 mL of NMP (3x, 3-5 min per wash each). Repeated cycles of amino acid coupling and Fmoc deprotection were then performed to assemble the desired peptide sequence. Following removal of the final Fmoc protecting group, the benzoyl (Bz) moiety was added by treating the resin with a solution of benzoic acid (37 mg, 0.30 mmol), HBTU (114 mg, 0.30 mmol), and DIPEA (86 µL, 0.49 mmol) in 3 mL of NMP for 1 h at room temperature. The resin was then washed with ~10 mL of NMP (3x, 3-5 min per wash), followed by ~10 mL of CH_2Cl_2 (3x, 3-5 min per wash). Peptide **7** was cleaved from the solid support by treatment with 5 mL of 95:2.5:2.5 TFA/TIPS/H_2O (2x, ~30 min per treatment). The combined cleavage solutions were then concentrated by rotary evaporation. The remaining residue was then dissolved in 500 µL of NMP and precipitated by dropwise addition to 40 mL of rapidly stirring, dry-ice chilled diethyl ether. Crude, precipitated **7** was then collected by centrifugation and dried overnight under vacuum. Peptide **7** was finally purified by RP-HPLC (Method D). Pure fractions were pooled and lyophilized. The purity and identity of **7** was confirmed by RP-HPLC and LC-ESI-MS (**Figure S4**). Pure **7** was weighed into a microcentrifuge tube and dissolved in nanopure H_2O to yield a 10 mM stock solution.

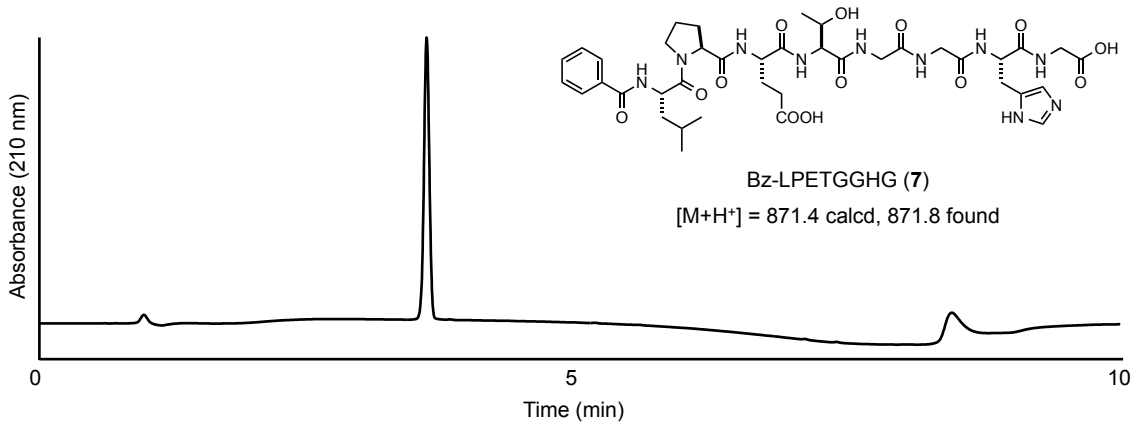

Figure S4. RP-HPLC chromatogram (Method A, 210 nm) for purified Bz-LPETGGHG (**7**) and LC-ESI-MS characterization.

III. Protein cloning and expression

Sequences of proteins used in this study:

$SrtA_{staph}$ (*initiator methionine absent in purified protein*)
GSSHHHHHHSSGLVPRGSHMQAKPQIPKDKSKVAGYIEIPDADIKEPVYPGPATPEQLNRGVSFAEENES
LDDQNISIAGHTFIDRPNYQFTNLKAAKKGSMVYFKVGNETRKYKMTSIRDVKPTDVGVLDEQKGKDKQL
TLITCDDYNEKTGVWEKRKIFVATEVK

MBP-LPETGGHG
MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGG
YAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKA
KGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYS
IAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLEN
YLLTDEGLEAVNKDKPLGAVALKSYEEELVKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAAS
GRQTVDEALKDAQTNSSSNNNNNNNNNNLGIEGRISHMSMGLPETGGHG

Bovine insulin
A Chain: GIVEQCCASVCSLYQLENYCN
B Chain: FVNQHLCGSHLVEALYLVCGERGFFYTPKA
(A chain contains one intrachain disulfide, A and B chains joined by two interchain disulfides)

Expression and purification of $SrtA_{staph}$. A pET-28a(+) expression vector for sortase A from *Staphylococcus aureus* (residues 60-206) containing a N-terminal hexahistidine tag was generously provided by Prof. Hidde Ploegh (Whitehead Institute for Biomedical Research). The vector was first transformed (heat shock) into *Escherichia coli* BL21(DE3) cells and the cells were plated on agar plates supplemented with kanamycin (50 µg/mL). After overnight incubation at 37 ^{o}C, a single colony was selected and used to inoculate a 100 mL starter culture containing sterile LB media and kanamycin (50 µg/mL). The starter culture was grown at 37 ^{o}C overnight. Approximately 50 mL of this starter culture was used to inoculate 1 L of sterile LB containing kanamycin (50 µg/mL), and the cells were grown at 37 ^{o}C to an optical density of ~0.8 at 600 nm. Protein expression was then induced with IPTG (1 mM final concentration) for 3 hours at 37 ^{o}C. Cells were isolated by centrifugation, and pellets were stored overnight at -80 ^{o}C. The pellets were then thawed, and resuspended in 35 mL of 20 mM Tris pH 8.0, 150 mM NaCl, and 0.5 mM EDTA. Lysozyme was added to a final concentration of 1 mg/mL, and this mixture was incubated at room temperature on a rocking platform for 1 hour. The suspension was then sonicated and centrifuged. The clarified lysate was then applied to a nickel nitrilotriacetic acid (Ni-NTA) column containing 5 mL of Ni-NTA agarose slurry (Thermo Scientific) that had been pre-equilibrated with 10 mL of 20 mM Tris pH 8.0, 150 mM NaCl, and 20 mM imidazole. After binding $SrtA_{staph}$ to the column, it was washed with 50 mL of 20 mM Tris pH 8.0, 150 mM NaCl, and 20 mM imidazole. $SrtA_{staph}$ was eluted with 10 mL of 20 mM Tris pH 8.0, 150 mM NaCl, and 300 mM imidazole. Purified $SrtA_{staph}$ was dialyzed against 50 mM Tris pH 8.0, 150 mM NaCl at 4 ^{o}C to remove imidazole. Glycerol was then added to the dialyzed solution (final glycerol concentration 10% (v/v)), and aliquots were stored at -80 ^{o}C. $SrtA_{staph}$ concentration was estimated by absorbance at 280 nm using a calculated extinction coefficient of 14,440 $M^{-1}cm^{-1}$ (Expasy ProtParam tool). The purity of $SrtA_{staph}$ was verified by SDS-PAGE analysis (**Figure S5**).

Construction of plasmid for MBP-LPETGGHG. The expression vector for MBP-LPETGGHG was constructed from the commercially available pMAL-c5x plasmid (New England Biolabs), which contains the coding sequence for MBP followed by a polyasparagine linker and a multiple cloning site. The plasmid was linearized through double digestion at the multiple cloning site using EcoRI (Fisher Optizyme™) and NcoI (Fisher Optizyme™) following the manufacturer's instructions. The linearized plasmid was then purified by agarose gel electrophoresis and isolated using a GeneJET Gel Extraction Kit (Thermo Scientific). A cassette encoding the LPETGGHG insert flanked by 5′-overhangs compatible with the EcoRI and NcoI sticky ends of the linearized pMal-c5x plasmid was then prepared by annealing the following 5′-phosphorylated DNA oligonucleotides (obtained from IDT):

Forward 5′-/5Phos/CATGGGTCTGCCGGAAACCGGTGGCCACGGTTAAG-3′

Reverse 5′-/5Phos/AATTCTTAACCGTGGCCACCGGTTTCCGGCAGACC-3′

Each oligo was resuspended in 10 mM Tris pH 7.5, 50 mM NaCl, 1 mM EDTA at a final concentration of 100 μM. Annealing was achieved by combining 50 μL of each oligo solution in a microcentrifuge tube followed by heating at 94 ^{o}C for 4 minutes. The mixture was then allowed to slowly cool to room temperature. This annealed oligo solution (1.0 μL) was then combined with linearized pMAL-c5x (16 μL of a 10 ng/μL solution), T4 DNA ligase (2.0 μL, Fisher Optizyme™), and T4 DNA ligase buffer (2.0 μL). The ligation reaction was incubated for 45 minutes at room temperature. The ligation reaction was then transformed (heat shock) into *Escherichia coli* BL21(DE3) cells and the cells were plated on agar plates supplemented with ampicillin (100 μg/mL). After overnight incubation at 37 ^{o}C, individual colonies were selected and used to inoculate 3 mL cultures containing sterile LB media and ampicillin (100 μg/mL). Cultures were grown at 37 ^{o}C overnight, and plasmids were then isolated using a GeneJET Plasmid Miniprep Kit (Thermo Scientific). Plasmids encoding the desired MBP-LPETGGHG construct were confirmed by sequencing.

Expression and purification of MBP-LPETGGHG. The MBP-LPETGGHG expression vector was first transformed (heat shock) into *Escherichia coli* BL21(DE3) cells and the cells were plated on agar plates supplemented with ampicillin (100 μg/mL). After overnight incubation at 37 ^{o}C, a single colony was selected and used to inoculate a 50 mL starter culture containing sterile LB media and ampicillin (100 μg/mL). The starter culture was grown at 37 ^{o}C overnight. A portion of this starter culture (~25 mL) was used to inoculate 1 L of sterile LB containing ampicillin (100 μg/mL), and the cells were grown at 37 ^{o}C to an optical density of ~0.8 at 600 nm. Protein expression was then induced with IPTG (1 mM final concentration) for 4 hours at 37 ^{o}C. Cells were then harvested by centrifugation. The pellets were resuspended in 40 mL of 20 mM Tris pH 8.0, 150 mM NaCl, and 0.5 mM EDTA. Lysozyme was added to a final concentration of 1 mg/mL, and the mixture was incubated at room temperature on a rocking platform for 30 minutes. The suspension was then sonicated and centrifuged. The clarified lysate was then applied to a column containing 4 mL of amylose resin (New England Biolabs) that had been pre-equilibrated with 12 mL of 20 mM Tris pH 7.5, 150 mM NaCl, and 1 mM EDTA. After binding MBP-

LPETGGHG to the column, it was washed with 20 mL of 20 mM Tris pH 7.5, 150 mM NaCl, and 1 mM EDTA. MBP-LPETGGHG was eluted with 5 mL of 20 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, and 10 mM maltose. Purified MBP-LPETGGHG was then dialyzed against 20 mM Tris pH 7.5, 150 mM NaCl at 4 ^{o}C. Protein concentration was estimated by absorbance at 280 nm using a calculated extinction coefficient of 66,350 $M^{-1}cm^{-1}$ (Expasy ProtParam tool). Aliquots were stored at -80 ^{o}C. The purity of MBP-LPETGGHG was verified by SDS-PAGE analysis (**Figure S5**).

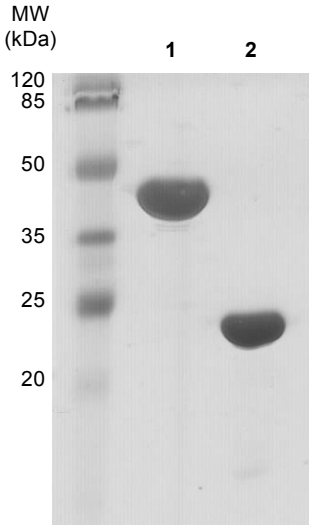

Figure S5. SDS-PAGE characterization of pure MBP-LPETGGHG (**lane 1**) and pure $SrtA_{staph}$ (**lane 2**).

IV. Sortase-mediated ligation studies

General ligation protocol for peptide substrates. Model peptide ligation reactions were generally conducted on 50 µL or 100 µL scale. The stock solutions shown below (**Table S1**) were combined in appropriate ratios, and then diluted with water to give the reactant concentrations described in this manuscript. All reactions were incubated at 37 ^{o}C unless otherwise indicated. All reactions contained 10% (v/v) 10x SrtA reaction buffer (500 mM Tris pH 7.5, 1500 mM NaCl, 100 mM $CaCl_2$). Additional notes: the $SrtA_{staph}$ stock solution contained 10% (v/v) glycerol, and therefore all reactions contained trace amounts of glycerol ≤ 1% (v/v). Additionally, the stock solutions for all dansylated glycine nucleophiles were prepared in 1:1 DMSO / H_2O, and therefore all reactions employing these nucleophiles contained DMSO at a final concentration of 5-17% (v/v). $SrtA_{staph}$ has been shown to tolerate DMSO concentrations in this range.[3] Reactions were analyzed by RP-HPLC and LC-ESI-MS. For RP-HPLC, unless otherwise indicated reactions were directly injected into the HPLC system and separated using Method A (see above). For LC-ESI-MS, reactions were diluted 5-10 fold with water (w/ 0.1% formic acid) and analyzed using the method described above. Fractions collected during RP-HPLC analysis were also analyzed by LC-ESI-MS.

Table S1. Stock solutions used in sortase-mediated peptide ligation reactions.

Reagent	Concentration	Solvent / Buffer
Ac-K(DNP)LPETGGHG (**1**)	840 μM	H_2O
Ac-K(DNP)LPETGGGG	711 μM	H_2O
Ac-K(DNP)LPETGG	488 μM	H_2O
GK(Dns) (**2**)	[a]963 μM or 887 μM	1:1 DMSO / H_2O
GGK(Dns)	928 μM	1:1 DMSO / H_2O
GGGK(Dns)	977 μM	1:1 DMSO / H_2O
$SrtA_{staph}$	444 μM	90% 50 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol (v/v)
10x SrtA reaction buffer	-	500 mM Tris pH 7.5, 1500 mM NaCl, 100 mM $CaCl_2$
$NiSO_4$•$6H_2O$	10 mM	H_2O
$CoCl_2$•$6H_2O$	10 mM	H_2O
$CuCl_2$•$2H_2O$	10 mM	H_2O

[a]Two independently synthesized batches of GK(Dns) (**2**) were used in these studies.

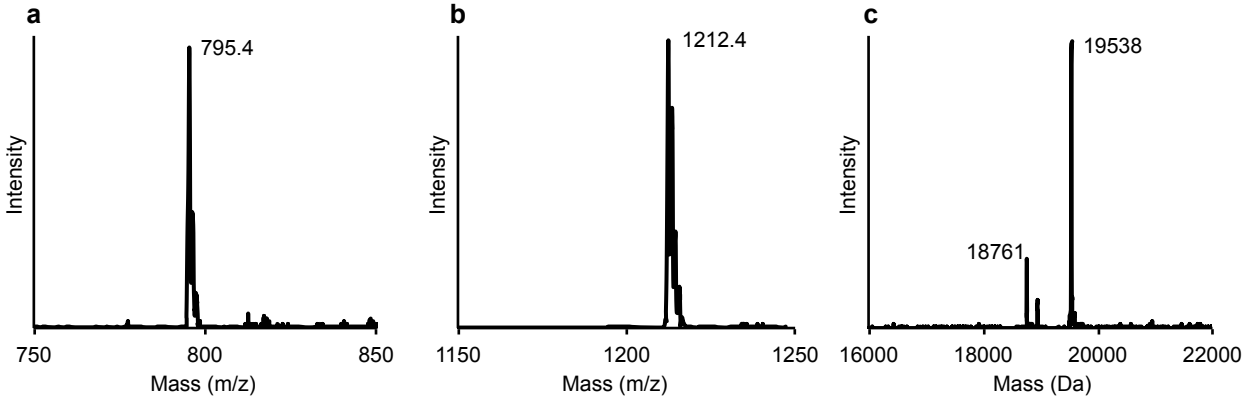

Figure S6. LC-ESI-MS characterization of reaction products from model peptide ligation. (**a**) Hydrolysis product (**4**) ([M+H$^+$] = 795.3 calcd, 795.4 found). (**b**) Ligation product (**3**) ([M+H$^+$] = 1212.5 calcd, 1212.4 found). (**c**) Reconstructed mass spectrum for $SrtA_{staph}$ adduct (**5**) (19537 calcd, 19538 found). The observed molecular weight is consistent with either the acyl-enzyme intermediate or a by-product in which Ac-K(Dnp)LPET has been ligated to the $SrtA_{staph}$ N-terminus. This material was not characterized further. The $SrtA_{staph}$ adduct co-eluted with free $SrtA_{staph}$ (18760 calcd, 18761 found).

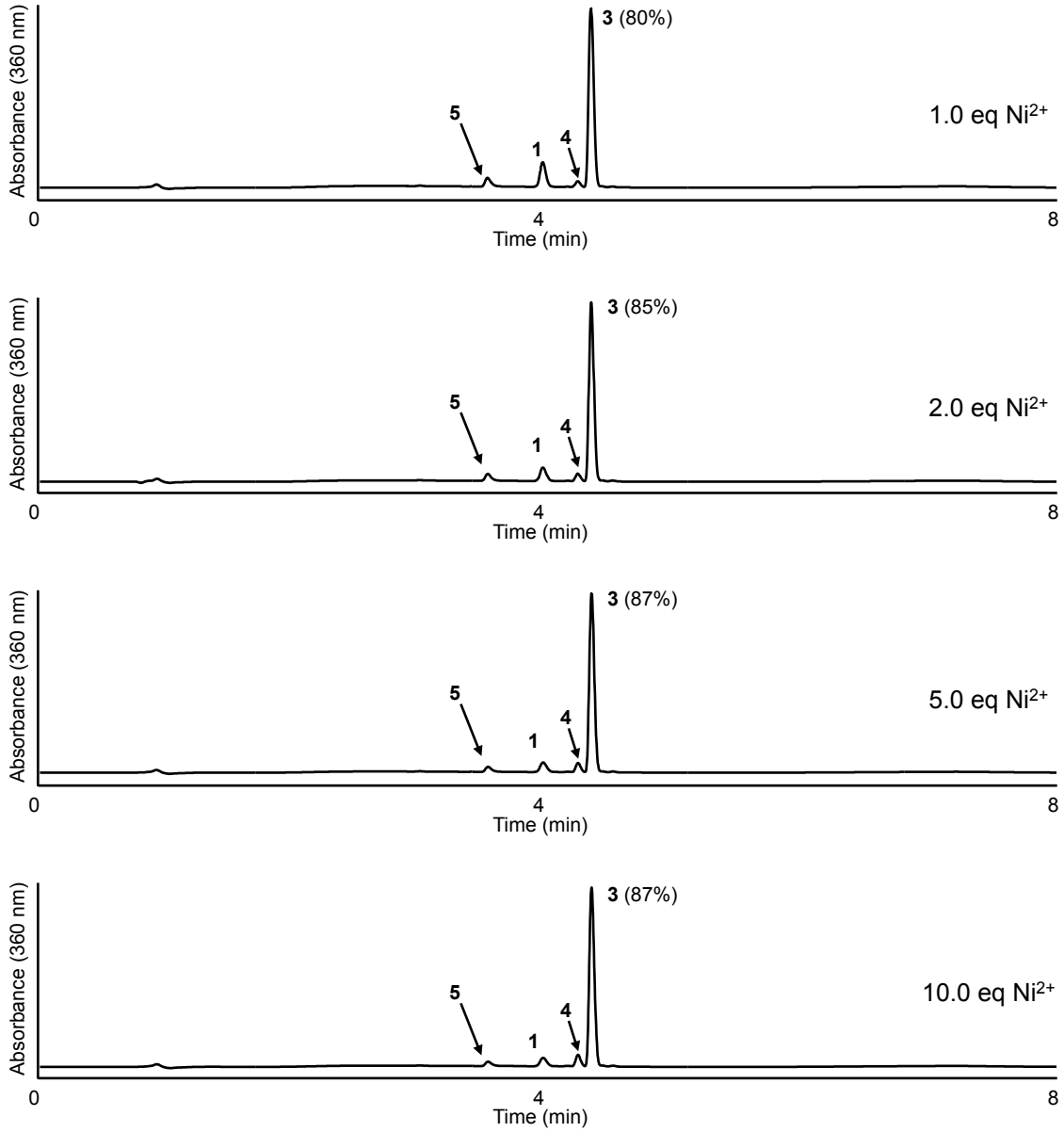

Figure S7. Effect of Ni^{2+} concentration on the efficiency of ligation between Ac-K(DNP)-LPETGGHG (**1**) and GK(Dns) (**2**). Peak labels: **3** = Ac-K(DNP)LPETGK(Dns) ligation product, **4** = Ac-K(DNP)LPET hydrolysis product, **5** = covalent $SrtA_{staph}$ adduct. Reaction conditions: 100 μM **1**, 100 μM **2**, 10 μM $SrtA_{staph}$, 50 mM Tris pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$, 100-1000 μM $NiSO_4$, 0.2% glycerol (v/v), 6% DMSO (v/v), 3 h at 37 ^{o}C. Reactions were analyzed by RP-HPLC (Method A, 360 nm). Extent of product formation was estimated from peak areas derived from the 360 nm chromatogram.

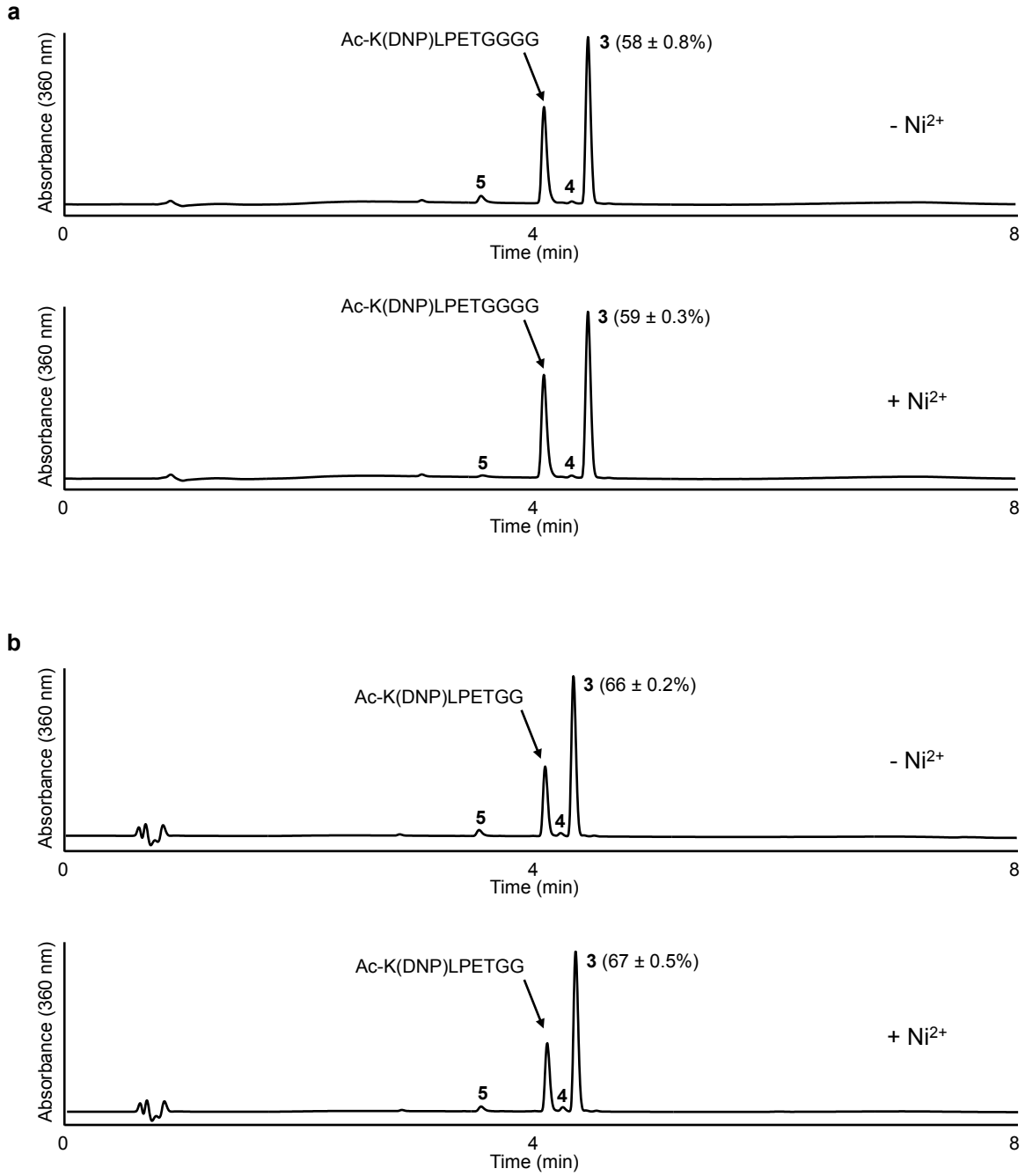

Figure S8. RP-HPLC chromatograms of control reactions utilizing substrates lacking the histidine anchor residue ((**a**) Ac-K(DNP)LPETGGGG and (**b**) Ac-K(DNP)LPETGG). Peak labels: **3** = Ac-K(DNP)LPETGK(Dns) ligation product, **4** = Ac-K(DNP)LPET hydrolysis product, **5** = covalent $SrtA_{staph}$ adduct. For both substrates, Ni^{2+} had no effect on the formation of ligation product **3**. Conditions: 100 μM Ac-K(DNP)LPETGGGG or Ac-K(DNP)LPETGG, 100 μM **2**, 10 μM $SrtA_{staph}$, 50 mM Tris pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$, 0 / 200 μM $NiSO_4$, 0.2% glycerol (v/v), 5-6% DMSO (v/v), 3 h at 37 ^{o}C. For (**a**) Ac-K(DNP)LPETGGGG, values in parentheses represent the average of three independent reactions and are reported as mean ± standard deviation. For (**b**) Ac-K(DNP)LPETGG, values in parentheses represent the average of two independent reactions and are reported as mean ± standard deviation

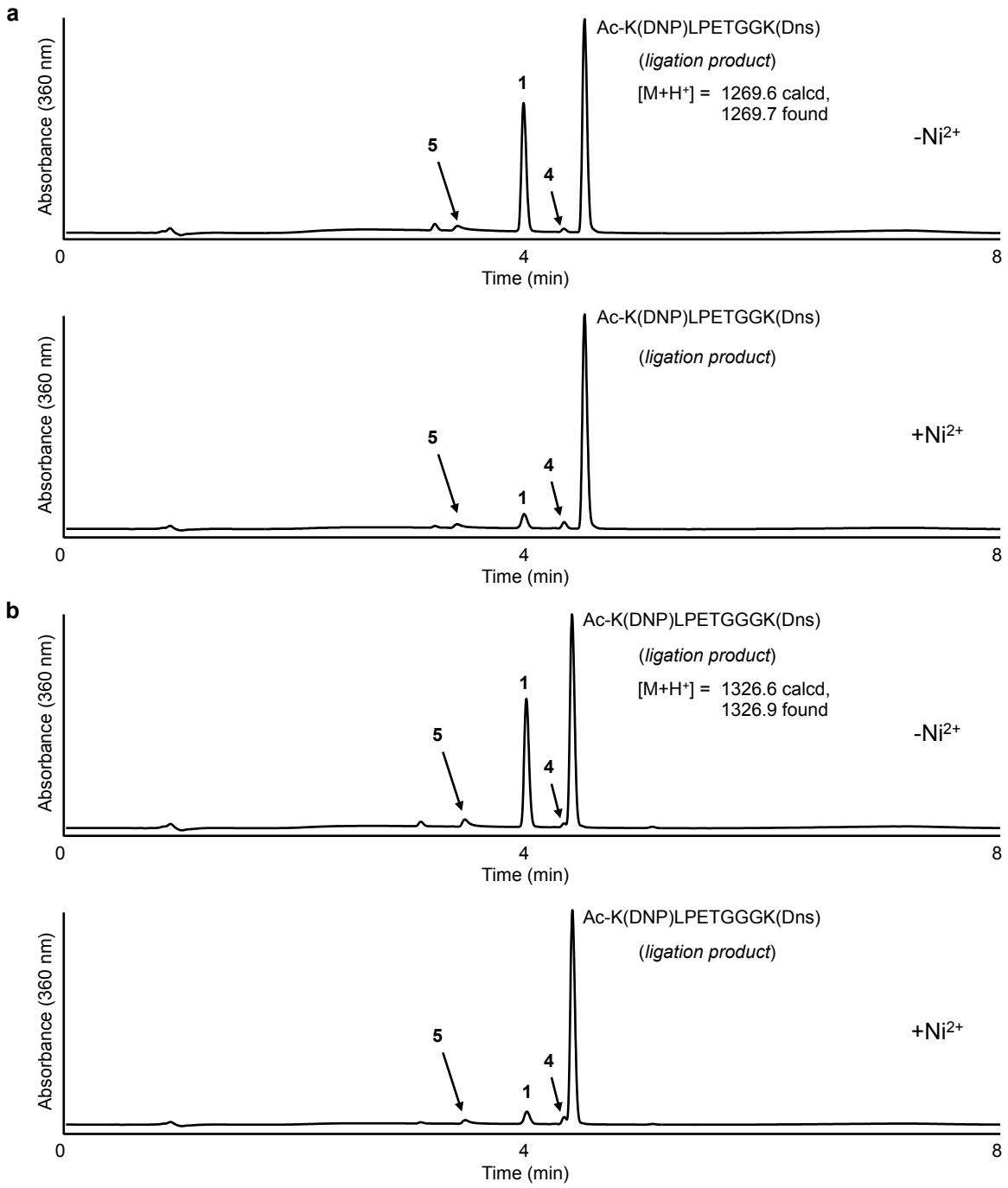

Figure S9. Representative RP-HPLC chromatograms (360 nm) for model ligation reactions between Ac-K(DNP)-LPETGGHG (**1**) and (**a**) GGK(Dns) or (**b**) GGGK(Dns). Peak labels: **4** = Ac-K(DNP)LPET hydrolysis product, **5** = covalent $SrtA_{staph}$ adduct. Reaction conditions: 100 μM **1**, 100 μM GGK(Dns) or GGGK(Dns), 10 μM $SrtA_{staph}$, 50 mM Tris pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$, 0 / 200 μM $NiSO_4$, 0.2% glycerol (v/v), 5% DMSO (v/v), 3 h at 37 ^{o}C. Reactions were analyzed by RP-HPLC (Method A, 360 nm). Extent of product formation was estimated from peak areas derived from the 360 nm chromatogram. The identity of ligation products was confirmed by LC-ESI-MS.

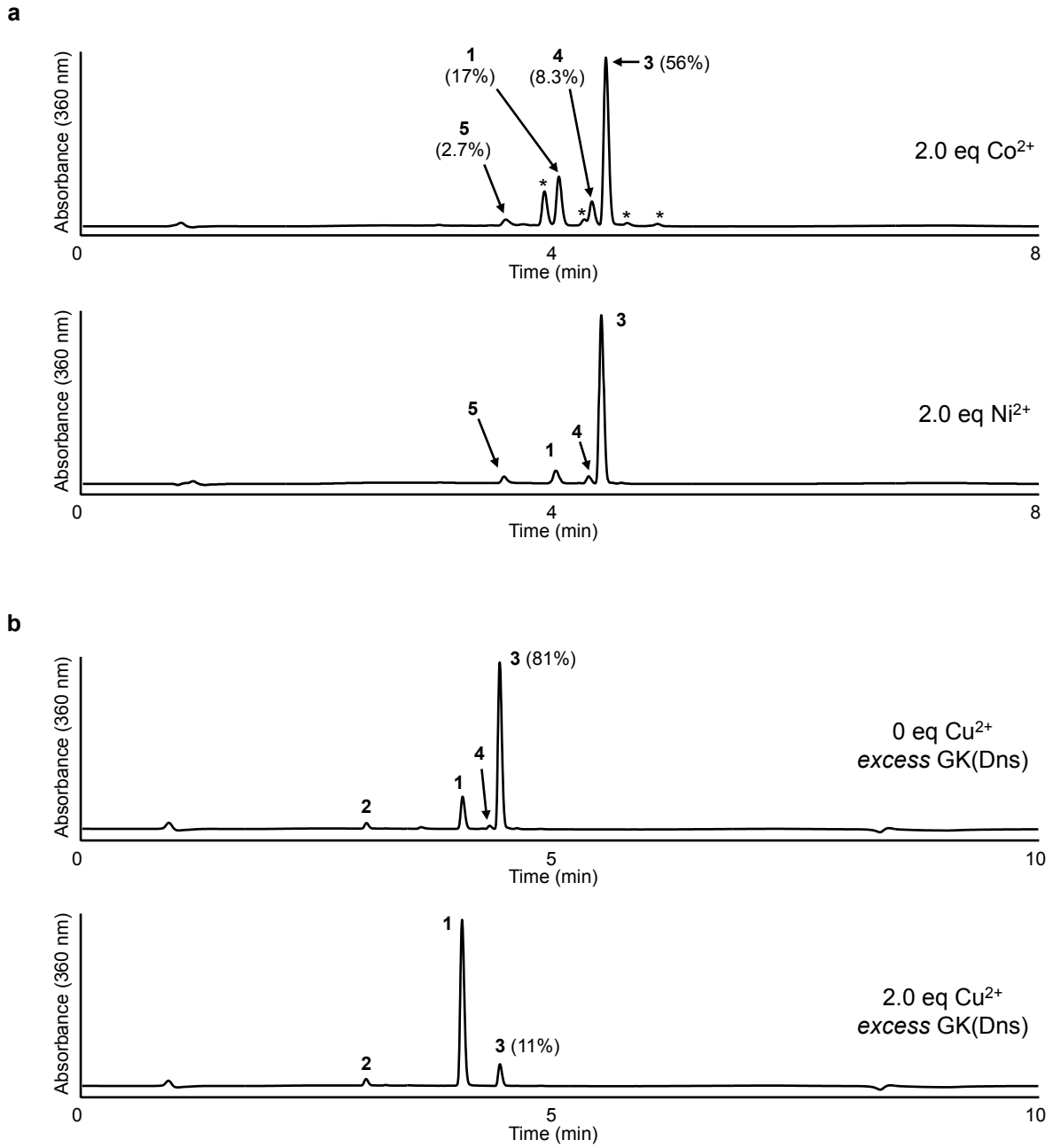

Figure S10. Effect of metal ion identity on the efficiency of ligation between Ac-K(DNP)-LPETGGHG (**1**) and GK(Dns) (**2**). Peak labels: **3** = Ac-K(DNP)LPETGK(Dns) ligation product, **4** = Ac-K(DNP)LPET hydrolysis product, **5** = covalent $SrtA_{staph}$ adduct, ***** = additional, unidentified reaction by-products. (**a**) Comparison of ligation reactions containing Ni^{2+} or Co^{2+}. Reactions containing Co^{2+} generated a more complex product mixture. Reaction conditions: 100 μM **1**, 100 μM **2**, 10 μM $SrtA_{staph}$, 50 mM Tris pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$, 200 μM $NiSO_4$ or $CoCl_2$, 0.2% glycerol (v/v), 6% DMSO (v/v), 3 h at 37 ^oC. Product distribution percentages are based on the total peak area of *all peaks* containing the DNP chromophore. (**b**) Inhibitory effect of Cu^{2+}. Reaction conditions: 100 μM **1**, 300 μM **2**, 10 μM $SrtA_{staph}$, 50 mM Tris pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$, 0 or 200 μM $CuCl_2$, 0.2% glycerol (v/v), 17% DMSO (v/v), 21 h at 37 ^oC, [**Note** = $SrtA_{staph}$ and $CuCl_2$ were incubated for ~2 minutes at room temperature prior to addition of **1** and **2**. All reactions were analyzed by RP-HPLC (Method A, 360 nm). Extent of product formation was estimated from peak areas derived from the 360 nm chromatogram.

Table S2. Product Distribution for Model Peptide Ligation Reactions[a]

Substrate	Nucleophile	2.0 eq Ni^{2+}	Unreacted Substrate	Hydrolysis	$SrtA_{staph}$ Adduct	Ligation Product
1	1.0 eq GK(Dns)	-	37 ± 2.1%	0.7 ± 0.2%	4.0 ± 0.3%	58 ± 2.5%
1	1.2 eq GK(Dns)	-	35 ± 1.3%	0.7 ± 0.1%	3.7 ± 0.4%	61 ± 1.0%
1	1.5 eq GK(Dns)	-	29 ± 0.5%	0.5 ± 0.0%	2.8 ± 0.4%	68 ± 0.5%
1	2.0 eq GK(Dns)	-	23 ± 1.5%	0.0 ± 0.0%	2.1 ± 0.1%	75 ± 1.4%
1	1.0 eq GK(Dns)	+	7.5 ± 0.4%	2.9 ± 0.6%	4.1 ± 0.6%	85 ± 1.5%
1	1.2 eq GK(Dns)	+	7.2 ± 0.7%	1.7 ± 0.5%	2.5 ± 0.1%	89 ± 0.6%
1	1.5 eq GK(Dns)	+	6.6 ± 0.5%	0.9 ± 0.2%	1.6 ± 0.2%	91 ± 0.5%
1	2.0 eq GK(Dns)	+	7.3 ± 1.2%	0.5 ± 0.1%	1.0 ± 0.1%	91 ± 1.3%
1	1.0 eq GGK(Dns)	-	38 ± 0.7%	0.9 ± 0.1%	2.2 ± 0.2%	59 ± 0.8%
1	1.0 eq GGK(Dns)	+	9.0 ± 2.6%	2.6 ± 0.9%	2.2 ± 0.1%	86 ± 2.6%
1	1.0 eq GGGK(Dns)	-	40 ± 2.0%	0.2 ± 0.1%	3.1 ± 0.4%	57 ± 2.4%
1	1.0 eq GGGK(Dns)	+	8.6 ± 3.0%	2.4 ± 0.2%	2.3 ± 0.0%	87 ± 2.8%
Ac-K(DNP)LPETGGGG	1.0 eq GK(Dns)	-	38 ± 0.8%	0.7 ± 0.1%	2.9 ± 0.1%	58 ± 0.8%
Ac-K(DNP)LPETGGGG	1.0 eq GK(Dns)	+	39 ± 0.3%	0.6 ± 0.1%	1.0 ± 0.1%	59 ± 0.3%
Ac-K(DNP)LPETGG[b]	1.0 eq GK(Dns)	-	29 ± 0.5%	1.2 ± 0.1%	3.1 ± 0.1%	66 ± 0.2%
Ac-K(DNP)LPETGG[b]	1.0 eq GK(Dns)	+	30 ± 1.0%	1.4 ± 0.2%	2.2 ± 0.3%	67 ± 0.5%

[a]Reaction Conditions: 100 μM substrate (**1**, Ac-K(DNP)LPETGGGG, or Ac-K(DNP)LPETGG), 1.0-2.0 equivalents of nucleophile, 10 μM $SrtA_{staph}$, 50 mM Tris pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$, 0 / 200 μM $NiSO_4$, 0.2% glycerol (v/v), 5-11% DMSO (v/v), 3 h at 37 ^{o}C. All reactions were analyzed by RP-HPLC (Method A, 360 nm). Product ratios were estimated from peak areas derived from the 360 nm chromatogram. Unless otherwise noted, values represent the average of three independent reactions and are reported as mean ± standard deviation. [b]Reactions run in duplicate, values represent the average of two independent reactions and are reported as mean ± standard deviation

Modification of MBP-LPETGGHG. MBP-LPETGGHG (39.5 µL of a 253 µM stock solution in 20 mM Tris pH 7.5, 150 mM NaCl) was combined with **2** (11.3 µL of a 887 µM stock solution in 1:1 DMSO/H_2O), $NiSO_4$•$6H_2O$ (2.0 µL of a 10 mM stock solution in H_2O), $SrtA_{staph}$ (2.3 µL of a 444 µM stock solution in 90% 50 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol), and 10x sortase reaction buffer (10 µL of 500 mM Tris pH 7.5, 1500 mM NaCl, 100 mM $CaCl_2$). H_2O was then added to bring the final volume to 100 µL. An analogous reaction was set up that lacked the $NiSO_4$ additive. Reactions were incubated at 37 ^oC and analyzed by LC-ESI-MS (main text **Figure 3a** and **Figure S11**).

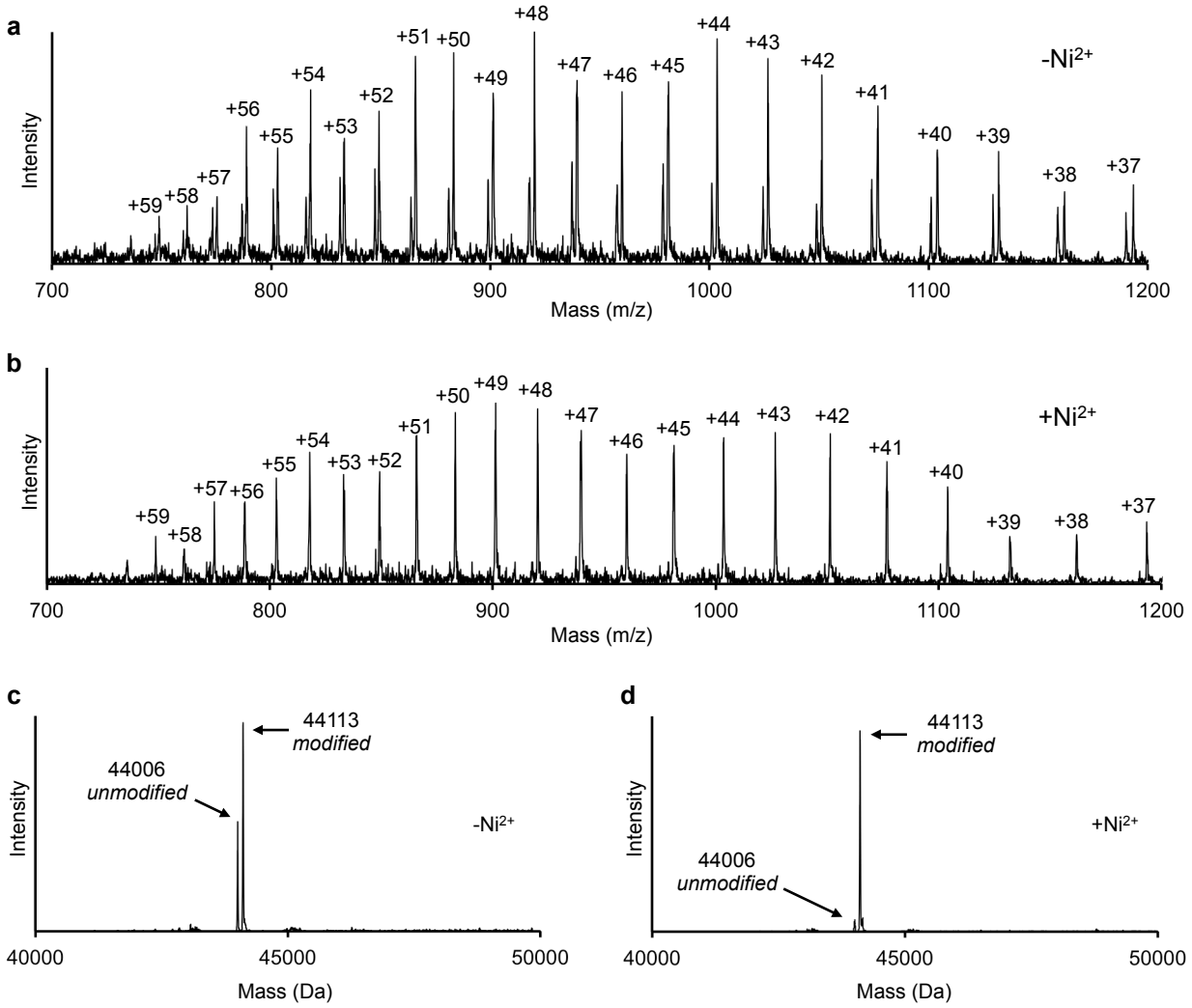

Figure S11. Additional LC-ESI-MS characterization of sortagging reactions between 100 µM MBP-LPETGGHG, 100 µM **2**, and 10 µM $SrtA_{staph}$ in the presence or absence of 200 µM $NiSO_4$ (16 h at 37 ^oC). (**a**) ESI-MS charge series for ligation reaction lacking Ni^{2+}. Ions for both modified and unmodified MBP were evident at each charge state (the lower molecular weight signal in each pair corresponds to unmodified MBP). (**b**) ESI-MS charge series for ligation reaction containing Ni^{2+}. A single charge series corresponding to modified MBP was observed. (**c**) Reconstructed ESI-MS spectrum for ligation reaction lacking Ni^{2+}. (**d**) Reconstructed ESI-MS spectrum for ligation reaction containing Ni^{2+}. [modified MBP (calculated MW = 44112), unmodified MBP (calculated MW = 44003)].

Modification of bovine insulin. A commercially available bovine insulin solution (Sigma-Aldrich I0516, 10 mg/mL in 25 mM HEPES, pH 8.2) was first diluted with Milli-Q water to a final concentration of 1 mM. 10.0 µL of this stock was then combined with 10 equivalents of **7** (10.0 µL of a 10 mM stock solution in DMSO), $NiSO_4 \bullet 6H_2O$ (8.0 µL of a 10 mM stock solution in H_2O), $SrtA_{staph}$ (2.3 µL of a 444 µM stock solution in 90% 50 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol), and 10x sortase reaction buffer (20 µL of 500 mM Tris pH 7.5, 1500 mM NaCl, 100 mM $CaCl_2$). H_2O was then added to bring the final volume to 200 µL. All of the aforementioned stock solutions were used to prepare reactions containing different amounts of $NiSO_4$ and/or **7** (**Figure S12**). All reactions were incubated at 37 ^oC and analyzed by RP-HPLC (Method B, 280 nm) and LC-ESI-MS. Insulin conjugate **8** was isolated by RP-HPLC and its identity was confirmed by LC-ESI-MS (**Figure S13b**).

Purified **8** was also used to assess the site-specificity of the sortase ligation reaction. RP-HPLC fractions containing **8** were first adjusted to pH ~8 using 1 M $NaHCO_3$. DTT was then added to a final concentration of 50 mM and the mixture was incubated for 30 min at room temperature. This resulted in the reduction of all disulfide bonds, and the separation of intact insulin into the corresponding A- and B-chains (**Figure S13a**). This reaction was then analyzed by LC-ESI-MS. As expected, uniformly modified A-chain was observed, along with entirely unmodified B-chain (**Figure S13c,d**). These data are fully consistent with the expected specificity of the $SrtA_{staph}$ ligation reaction, which was predicted to target the N-terminal glycine residue of the insulin A-chain.

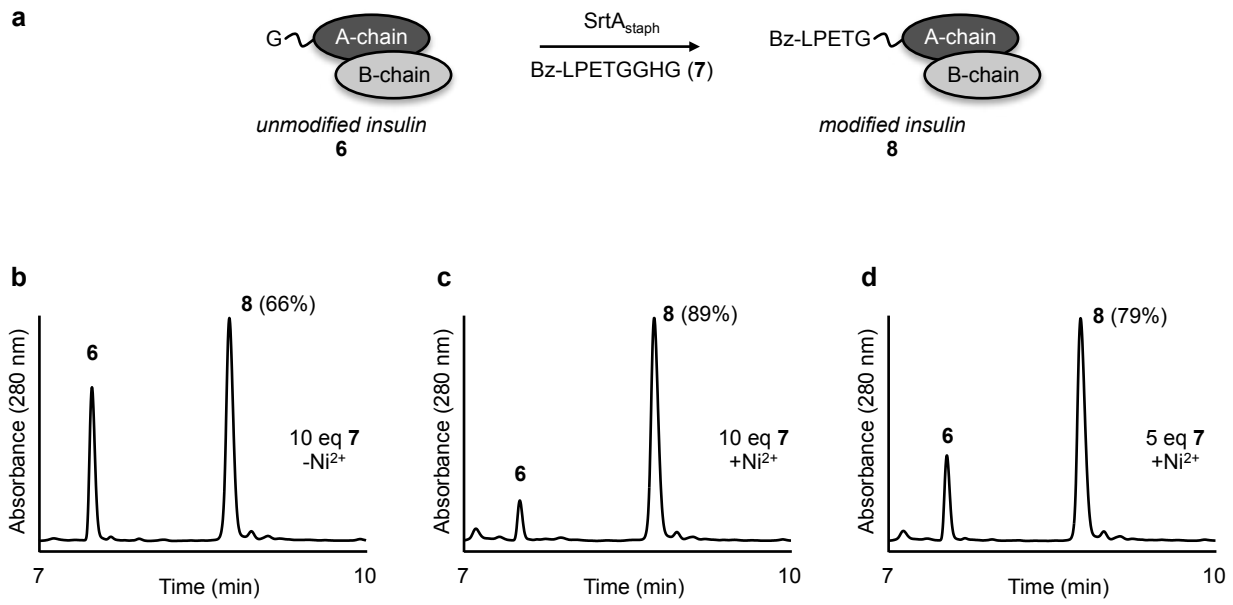

Figure S12. Additional insulin modification data demonstrating relative ligation efficiencies for reactions employing 5 or 10 equivalents of **7** in the presence or absence of Ni^{2+}. (**a**) Overall insulin modification reaction. (**b**-**d**) Representative RP-HPLC chromatograms (280 nm, Method B) illustrating reaction conversion using (**b**) 10 eq **7**, 0 eq $NiSO_4$, 8.5 h at 37 ^oC, (**c**) 10 eq **7**, 8 eq $NiSO_4$, 8 h at 37 ^oC, or (**d**) 5 eq **7**, 4 eq $NiSO_4$, 8 h at 37 ^oC. Values in parentheses represent percent conversion based on the relative peak areas of unmodified insulin (**6**) versus modified insulin (**8**).

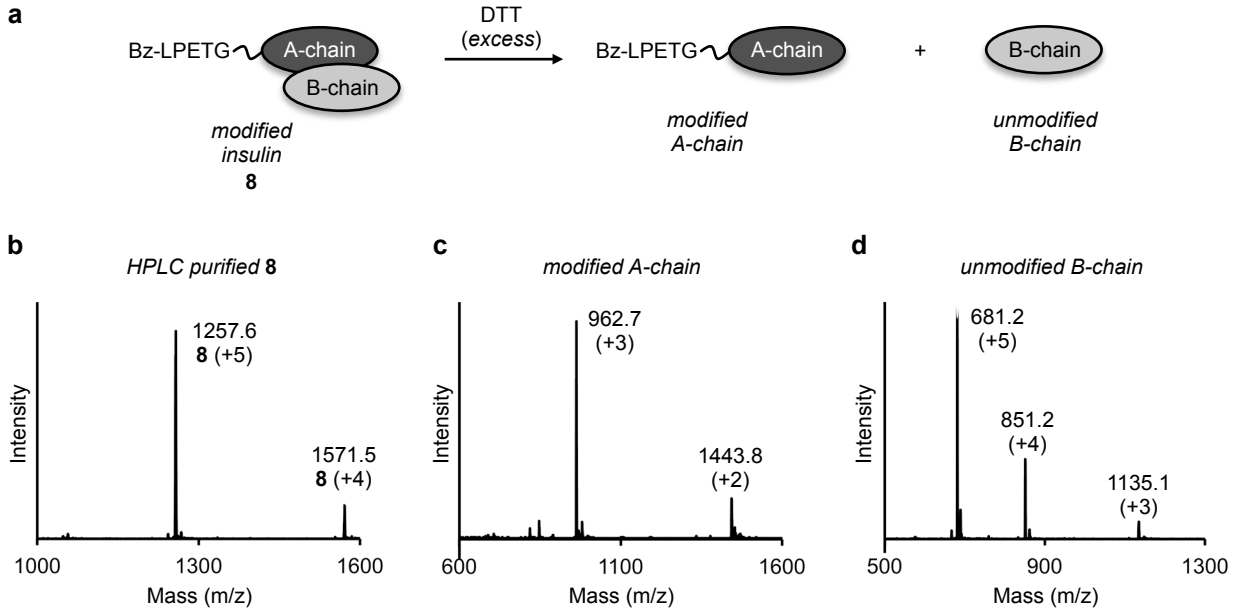

Figure S13. Confirmation of site-specificity in the ligation reaction between bovine insulin and **7**. (**a**) RP-HPLC purified insulin conjugate (**8**) was reduced with excess DTT. As expected, this generated uniformly modified A-chain and unmodified B-chain. (**b**) ESI-MS spectrum of purified **8** prior to DTT reduction. **8**: $[M+5H^+]$ = 1256.6 calcd, 1257.6 found, $[M+4H^+]$ = 1570.5 calcd, 1571.6 found. (**c**) ESI-MS spectrum of modified A-chain following DTT reduction. Modified A-chain: $[M+3H^+]$ = 962.4 calcd, 962.7 found, $[M+2H^+]$ = 1443.1 calcd, 1443.8 found. (**d**) ESI-MS spectrum of unmodified B-chain following DTT reduction. Unmodified B-chain: $[M+5H^+]$ = 681.0 calcd, 681.2 found, $[M+4H^+]$ = 851.0 calcd, 851.2 found, $[M+3H^+]$ = 1134.3 calcd, 1135.1 found.

References

[1] Y. K. Reshetnyak, O. A. Andreev, U. Lehnert, D. M. Engelman, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 6460–6465.

[2] A. Kosoy, C. Möller, D. Perdomo, J. Bubis, *J. Biochem. Mol. Biol.* **2004**, *37*, 260–267.

[3] S. Pritz, Y. Wolf, O. Kraetke, J. Klose, M. Bienert, M. Beyermann, *J. Org. Chem.* **2007**, *72*, 3909–3912.