Supplementary Information

Sublancin is not a lantibiotic but an S-linked glycopeptide

Trent J. Oman, John M. Boettcher, Huan Wang, Xenia N. Okalibe, and Wilfred A. van der Donk*

Department of Chemistry and Howard Hughes Medical Institute,
University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

Telephone: (217) 244 5360
FAX: (217) 244 8533
e-mail vddonk@illinois.edu

Materials
All oligonucleotides were purchased from Integrated DNA Technologies. Restriction endonucleases, DNA polymerases, and T4 DNA ligase were purchased from New England Biolabs or Invitrogen. Factor Xa was purchased from New England Biolabs and chymotrypsin was purchased from Worthington Biochemical Corporation. Media for bacterial culture and chemicals were purchased from Fisher Scientific, Sigma Aldrich, or CalBiochem, unless noted otherwise, and used without further purification.

Strains and Plasmids
*Bacillus subtilis* 168 (sublancin producing organism) was obtained from Bacillus Genetics Stock Center, Ohio State University, Columbus, OH. The indicator strain, *Bacillus subtilis* ATCC 6633 was obtained from American Type Culture Collection, Manassas, VA. *Escherichia coli* DH5α (UIUC Cell Media Facility) was used as host for cloning and plasmid propagation, and *E. coli* BL21 (DE3) and *E. coli* Rosetta 2 (DE3) (Novagen) were used as hosts for protein expression. Cloning vectors (pET15b and pET28b) were obtained from Novagen.

Supplementary Methods

General Methods
All molecular biology manipulations were carried out using standard techniques. Polymerase chain reaction (PCR) amplifications were carried out using an automated thermocycler (C1000, BioRAD). DNA sequencing was performed using appropriate primers by the Biotechnology Center (University of Illinois at Urbana-Champaign, UIUC). GC-MS analysis was performed at the Roy J. Carver Metabolomics Center (UIUC). LC-ESI-Q/TOF MS analyses were conducted using a Synapt MS system equipped with an Acquity UPLC (Waters). MALDI-TOF MS analyses were conducted at the Mass Spectrometry Facility (UIUC) using a Voyager DE-STR (Applied Biosystems). For MALDI-TOF MS analysis of salt-free samples, a 1 μL aliquot of analyte was combined with 1 μL of matrix (saturated α-cyano-4-hydroxy-cinnamic acid matrix in 50% ACN/50% water with 0.1% TFA), and the total volume was spotted onto a MALDI
target and dried under ambient conditions prior to analysis. Salt containing samples were desalted using ZipTip\textsubscript{C18} (Millipore), eluted with 4 \(\mu\)L of matrix, and 2 \(\mu\)L was spotted onto a MALDI target and dried as described above.

**Isolation and purification of sublancin.** A culture of *B. subtilis* 168 was grown in LB media under aerobic conditions at 37 °C for 12-15 h. The overnight culture was used to inoculate (at 1%) 500 mL volumes of Medium A in 2 L flasks. Medium A consisted of 900 mL of Medium A nutrient broth combined with 100 mL of 10X Medium A salts. Medium A nutrient broth was prepared by dissolving 20 g sucrose, 11.7 g citric acid, 4 g Na\textsubscript{2}SO\textsubscript{4}, 4.2 g (NH\textsubscript{4})\textsubscript{2}HPO\textsubscript{4}, and 5 g yeast extract in 900 mL of millipore water. The pH was adjusted to 6.8-6.9 using NaOH and the medium was autoclaved. Medium A salts (10X) were prepared by dissolving 7.62 g KCl, 4.18 g MgCl\textsubscript{2}-6 H\textsubscript{2}O, 0.543 g MnCl\textsubscript{2}-4 H\textsubscript{2}O, 0.49 g FeCl\textsubscript{3}-6 H\textsubscript{2}O, and 0.208 g ZnCl\textsubscript{2} in 1 L of millipore water followed by sterilization via 0.22 \(\mu\)m filtration. The Medium A cultures were grown under aerobic condition at 37 °C for 28-48 h with vigorous agitation. A color change to pinkish-brown was observed and the pH of cultures had lowered to 6-6.5. Sublancin production was consistently observed when these events occurred.

Cultures were acidified to pH 2.5 with concentrated phosphoric acid (85% in water) and centrifuged to remove cells and insoluble material. Sublancin was isolated by reversed-phase solid phase extraction. Briefly, a 100 mL volume of sublancin-containing culture supernatant was loaded onto a Bond Elut C18 solid-phase extraction column (10 g resin, 60 mL volume, Varian, Inc.) previously wetted with 100% methanol and equilibrated in 2% B (solvent A = 0.1% trifluoroacetic acid (TFA) in millipore water; solvent B = 0.0866% TFA in 80/20 acetonitrile/water). Sublancin was eluted using a step gradient with increasing percentage of solvent B in 50 mL volumes that were collected in 10 mL fractions: 2%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 50%, and 100% B. Elution of sublancin from the SPE column was monitored by MALDI-TOF MS. A 1 \(\mu\)L aliquot of elution sample was combined with 1 \(\mu\)L of matrix consisting of saturated \(\alpha\)-cyano-4-hydroxy-cinnamic acid matrix in 50% ACN/50% water with 0.1% TFA, and the total volume was spotted onto a MALDI target and dried under ambient conditions prior to analysis. Sublancin was observed in 50% B fractions. Fractions containing sublancin were combined, lyophilized to dryness, and stored under N\textsubscript{2} at -80 °C until purification by preparative HPLC.

Preparative HPLC was performed using a Waters Delta 600 instrument equipped with a Phenomenex Jupiter Proteo C12 column (10 \(\mu\)m, 90 Å, 250 mm x 15 mm) equilibrated in 2% B (solvent A = 0.1% TFA in water, solvent B = 0.0866% TFA in 80% ACN/20% water). Dry sublancin material was resuspended in 2% B and was applied to the column. Sublancin was eluted by maintaining the mobile phase at 2% B for 1 min, followed by an increase to 100% B over 45 min with a flow rate of 10.0 mL/min. Under these conditions, sublancin eluted at 22.7 min. All fractions were analyzed by MALDI-TOF MS as described above (Supplementary Fig. 1A). Purified sublancin was lyophilized to dryness and stored under N\textsubscript{2} at -80 °C until further use. Typical yields were 10-15 mg sublancin per liter of processed Medium A culture.

**Analysis of the biosynthetic gene cluster of sublancin.** Information on the biosynthetic gene cluster responsible for *in vivo* production of sublancin containing genes *sunI*, *sunA*, *sunT*, *bdbA*, *yolJ* (*sunS*), and *bdbB* was accessed via the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/) with the following accession numbers: RefSeq =
The gene cluster is 5080 nucleotides in length, is of chromosomal origin, and is found between base pairs 2,270,305 – 2,265,225 (complement).

**Iodoacetamide assays for detection of free cysteines.** To detect the presence of free cysteine thiols in native and reduced sublancin, an iodoacetamide assay was used. For the detection of free Cys residues in native sublancin, reactions contained 50 mM Tris (pH 8.3), 10 mM iodoacetamide, and 25 µM sublancin. For the detection of free Cys residues in reduced sublancin, reactions contained 50 mM Tris (pH 8.3), 5 mM TCEP, 10 mM iodoacetamide, and 25 µM sublancin (50 µL total volume). All reactions were incubated for 2.5 h at 25 °C in the dark. At 2.5 h, 10 µL of reaction was quenched with 5% TFA to pH 1-2. Samples were desalted using Zip-Tip C18 prior to MALDI-TOF MS analysis. Samples evaluated by LC-ESI-Q/TOF MS were analyzed without further manipulation (Supplementary Fig. 1). The detection of free cysteines was determined by the presence or absence of carboxyamidomethyl (CAM) thiol modifications.

**LC-ESI-Q/TOF MS and MSMS analyses.** A 5 µL volume of sample was injected on a Waters Acquity UPLC system equipped with a BEH C8 column (1.7 um, 100 mm x 1.0 mm) equilibrated in 3% B (solvent A = 0.1% formic acid in water, solvent B = 0.1% formic in methanol). The sample was fractionated by employing a gradient of 3%-97% B over 12 min and directly subjected to ESI-Q/TOF MS using a Waters Synapt mass spectrometer. Data was acquired in ESI positive mode with the capillary voltage set to 3.0-3.5 kV. Nitrogen was used as cone gas (0 L/h) and desolvation gas (600 L/h). The ionization source and desolvation gas were heated to 120 °C and 300 °C, respectively. The transfer collision energy was set to 4 V for both MS and MSMS analyses. The trap collision energy was set to 6 V (constant) for MS. For MSMS analysis, multiply charged parent ions were selected and fragmented using collision induced dissociation (nitrogen) with a trap collision energy ramp ranging from 20-60 V or a constant value between 25-45 V. The applied collision energies for MSMS were analyte dependent. [Glu^1]-Fibrinopeptide B (Sigma) was directly infused as lock mass with lock spray sampling at minute intervals. The acquired spectra were processed using MaxEnt3 software and analyzed by Protein/Peptide Editor in BioLynx 4.1 (Waters).

**Chymotrypsin digests of sublancin.** For analysis of native sublancin, the peptide was digested under non-reducing conditions in order to maintain the integrity of the in vivo formed disulfides. Sublancin was digested in 100 mM Tris (pH 7.5), 0.05 mg/mL chymotrypsin (Worthington), and 0.5 mg/mL sublancin. For analysis of the reduced peptide, sublancin was digested in 100 mM Tris (pH 7.5), 5 mM TCEP, 0.05 mg/mL chymotrypsin, and 0.5 mg/mL sublancin. All reactions were incubated at 25 °C for 5 h and then quenched with 5% TFA to pH 1-2. Quenched samples were ZipTip C18 desalted prior to analysis by MALDI-TOF MS or analyzed by LC-ESI-Q/TOF MS (Supplementary Fig. 2).

**Determination of the natural sugar modification of sublancin.** The presence of glucose on Cys22 of in vivo isolated sublancin was confirmed via acid-catalyzed hydrolysis and derivatization of the sugar, analysis by gas chromatography-mass spectrometry (GC-MS), and comparison to derivatized hexose standards. Lyophilized sublancin (approximately 0.5-1.0 mg) was dissolved in 1 mL of 0.5 M HCl in water and heated at 75 °C in a high-pressure, sealed vessel for 5 h. The reaction was cooled, transferred to a microcentrifuge tube with the aid of
millipore water, and quenched by addition of aqueous NaOH to pH 5. The total volume was concentrated via vacuum centrifugation. After complete drying, 50 μL of 20 mg/mL O-methylhydroxylamine hydrochloride (Fluka) in pyridine was added to the sample. The sample was vortexed for 1 min and was incubated at 50 °C for 1 h. Next, 50 μL of N-trimethylsilyl-N-methyl trifluoroacetamide (Fluka) was added. Again, the sample was vortexed for 1 min and incubated at 50 °C for 1 h. The sample was cooled at 25 °C for 5 min prior to vortexing for 1 min followed by centrifugation at 15 krpm in a microcentrifuge for 2 min. Hexose sugar standards D-(+)-glucose, D-(+)-galactose, D-(+)-mannose, and D-(−)-fructose were also treated and derivatized using the conditions described above.

The derivatized hexose of sublancin and sugar standards were analyzed individually and in combinations using an Agilent GC-MS system (6890N GC, HP 5973 mass selective detector) equipped with a Phenomenex ZB-1MS column (30 m x 0.32 mm i.d., 0.25 μm df) (Supplementary Fig. 3). Samples were introduced via split injection (3:1 split ratio). The temperature gradient used was 150 °C for 2 min, then 150 °C to 185 °C over 3.5 min (10 °C/min ramp), then 185 °C to 210 °C over 12.5 min (2 °C/min ramp), then 210 °C to 300 °C over 9 min (10 °C/min ramp). The carrier gas was helium and set at a 3.0 mL/min flow rate. The mass selective detector was operated in positive ion scan mode (25-800 m/z). Analyses were performed using electron impact (EI) mode and to preserve the filament of the ionization source, MS records were performed 4 min after injection.

**Molecular cloning of sunA and sunS genes.** Genomic DNA was isolated from an overnight culture of *B. subtilis* 168 using a Microbial DNA Isolation Kit (MO-BIO) according to the manufacturer’s protocol.

The gene *sunA* was PCR amplified by 30 cycles of denaturing (94 °C for 1 min), annealing (50 °C for 1 min), and extending (72 °C for 1 min) using SunA NdeI FP (5’-CGGCAGCCATATGGAAAAAGCTATTTAAAGA -3’) and SunA XhoI RP (5’-GATCCTCGAGTTATCTGCAGAATTGACGATAG -3’) as primers and *B. subtilis* 168 genomic DNA as template (bold indicates restriction sites). The PCR mixture included 1X PCR Buffer (Invitrogen), dNTPs (0.25 mM each), MgCl2 (1.5 mM), Platinum Taq DNA polymerase (0.05 U/μL), and primers (0.5 μM each). Amplifications were confirmed by 2% agarose gel electrophoresis, and the PCR products were purified using a QIAQuick PCR Purification Kit (QIAGEN). The insert DNA fragment and pET15b vector were double digested in separate reactions containing 1X NEBuffer 4 (New England Biolabs) with NdeI and XhoI for 15 h at 37 °C. The *sunA* and pET15b digests were purified by 2% and 1% agarose gel electrophoresis, respectively, and gel extracted using a QIAquick Gel Extraction Kit (QIAGEN). The resulting DNA products were ligated at 25 °C for 5 h in 1X T4 DNA Ligase buffer with T4 DNA Ligase (0.7 U/μL). *E. coli* DH5α cells were transformed with 2.5 μL of the ligation product by heat shock, and cells were plated on LB-ampicillin agar plates and grown for 15 h at 37 °C. Several colonies were picked and used to inoculate separate 5 mL cultures of LB-ampicillin medium. The cultures were grown at 37 °C for 12 h, and plasmids were isolated using a QIAprep Spin Miniprep Kit (QIAGEN). The sequences of the resulting plasmid products were confirmed by DNA sequencing.

The gene *sunS* (previously designated as yolJ) was PCR amplified by 30 cycles of denaturing (94 °C for 1 min), annealing (50 °C for 1 min), and extending (72 °C for 2 min) using SunS NdeI FP (5’-CGGCAGCCATATGGAAAACTGAGTGATATTTA -3’) and SunS XhoI RP (5’-TGCTCGAGTCACTTCAATTTCCCTTTTAC -3’) as primers and *B. subtilis* 168
genomic DNA as template (bold indicates restriction sites). The PCR mixture included 1X PCR Buffer (Invitrogen), dNTPs (0.25 mM each), MgCl₂ (1.5 mM), Platinum Taq DNA polymerase (0.05 U/μL), and primers (0.5 μM each). Amplifications were confirmed by 2% agarose gel electrophoresis, and the PCR products were purified using a QIAquick PCR Purification Kit (QIAGEN). The insert DNA fragment and pET28b vector were double digested in separate reactions containing 1X NEBuffer 4 (New England Biolabs) with NdeI and XhoI for 15 h at 37 °C. The sunS and pET28b digests were purified by 2% and 1% agarose gel electrophoresis, respectively, and gel extracted using a QIAquick Gel Extraction Kit (QIAGEN). The resulting DNA products were ligated at 25 °C for 5 h in 1X T4 DNA Ligase buffer with T4 DNA Ligase (0.7 U/μL). E. coli DH5α cells were transformed with 2.5 μL of the ligation product by heat shock, and cells were plated on LB-kanamycin agar plates and grown for 15 h at 37 °C. Several colonies were picked and used to inoculate separate 5 mL cultures of LB-kanamycin medium. The cultures were grown at 37 °C for 12 h, and plasmids were isolated using a QIAprep Spin Miniprep Kit (QIAGEN). The sequences of the resulting plasmid products were confirmed by DNA sequencing.

Molecular engineering of a Factor Xa cleavage site into the sunA gene. A site-directed mutagenesis approach was used to install a proteolytic cleavage site directly N-terminal to the core peptide of SunA. The primers were designed to contain nucleotide sequences necessary to encode the amino acids IEGR in place of four wild-type peptide residues (QKGS).

The plasmid pET15b sunA Xa was PCR amplified by 30 cycles of denaturing (94 °C for 1 min), annealing (55 °C for 1 min), and extending (72 °C for 6 min) using SunA Xa FP (5'-GAGGAACCTCGAAACATCGAAAGGTCGTGGGATTAGGAAAAGCT -3') and SunA Xa RP (5'-AGCTTTTCCTAATCCACGACCTTCGATGTTCGAGTTCTC -3') as primers and pET15b sunA as template (underline indicates nucleotide change). The PCR mixture included 2X Pfx Amp Buffer (Invitrogen), dNTPs (2.5 mM each), MgSO₄ (1 mM), Platinum Pfx DNA polymerase (0.05 U/μL), and primers (1.5 μL each). Amplifications were confirmed by 1% agarose gel electrophoresis. To digest the methylated template prior to transformation, 1X REACT 4 buffer and 1 μL of DpnI (Invitrogen) was added to the PCR product and the digest was incubated at 37 °C for 3 h. E. coli DH5α cells were transformed with 2.5 μL of the digest product by heat shock, and cells were plated on LB-ampicillin agar plates and grown for 15 h at 37 °C. Several colonies were picked and used to inoculate separate 5 mL cultures of LB-ampicillin medium. The cultures were grown at 37 °C for 12 h, and plasmids were isolated using a QIAprep Spin Miniprep Kit (QIAGEN). The sequences of the resulting plasmid products were confirmed by DNA sequencing.

Mutagenesis of the sunA gene to produce C22S. A site-directed mutagenesis approach was used to convert Cys22 to Ser22 in the SunA peptide. The plasmid pET15b sunA was PCR amplified by 30 cycles of denaturing (94 °C for 1 min), annealing (55 °C for 1 min), and extending (72 °C for 6 min) using SunA C22S FP (5'-GGTACAATTGGTTCTGGGCGGAGCTGTT -3') and SunA C22S RP (5'-AACAGCTCCGCCACCAGAAACAAATTGTACC -3') as primers and pET15b sunA as template (underline indicates nucleotide change). The PCR mixture included 2X Pfx Amp Buffer (Invitrogen), dNTPs (2.5 mM each), MgSO₄ (1 mM), Platinum Pfx DNA polymerase (0.05 U/μL), and primers (1.5 μL each). Amplifications were confirmed by 1% agarose gel electrophoresis. To digest the methylated template prior to transformation, 1X REACT 4 buffer
and 1 μL of DpnI (Invitrogen) was added to the PCR product and the digest was incubated at 37 °C for 3 h. *E. coli* DH5α cells were transformed with 2.5 μL of the digest product by heat shock, and cells were plated on LB-ampicillin agar plates and grown for 15 h at 37 °C. Several colonies were picked and used to inoculate separate 5 mL cultures of LB-ampicillin medium. The cultures were grown at 37 °C for 12 h, and plasmids were isolated using a QIAprep Spin Miniprep Kit (QIAGEN). The sequences of the resulting plasmid products were confirmed by DNA sequencing.

**Mutagenesis of the sunA and sunA Xa genes to produce various mutant precursor peptides.** Site-directed mutagenesis of pET15b sunA and pET15b sunA Xa constructs was performed by multi-step overlap extension PCR. First, the amplification of sunA or sunA Xa was carried out by thirty cycles of denaturing (94 °C for 20 s), annealing (58 °C for 30 s), and extending (72 °C for 20 s) using the SunA NdeI FP and an appropriate mutant reverse primer to yield 5’ fragment of the mutant sunA (or sunA Xa) gene (forward megaprimer (FMP) reaction). The PCR mixtures included 1X FailSafe PreMix G (EPICENTRE Biotechnologies), DMSO (4%), *Phusion* DNA polymerase (Finnzymes) (0.04 U/µL), dNTP (2 mM each) and primers (1 µM each). In parallel, PCR reactions using an appropriate mutant forward primer and the SunA XhoI RP primer was also conducted to produce 3’ fragments of the mutated sunA (or sunA Xa) gene using the same PCR conditions as the FMP reaction (reverse megaprimer (RMP) reaction). The overlapping products from the FMP and RMP reactions were combined in equal amounts and extended by seven cycles of denaturing, annealing and extending using the same PCR conditions. Following the extension, the SunA NdeI FP and SunA XhoI RP primers were added (final concentration 2 µM) and the mixture was subjected to another 25 cycles of denaturing, annealing and extending. Amplification of the final PCR product was confirmed and purified by 2% agarose gel electrophoresis. The resulting DNA inserts and empty pET15b vector were digested with *NdeI* and *XhoI* at 37 °C for 5 h. The digested insert and vector products were purified by 2% and 1% agarose gel electrophoresis, respectively, and gel extracted using a QIAquick Gel Extraction Kit (QIAGEN). The resulting DNA products were ligated at 25 °C for 3 h in 1X T4 DNA Ligase buffer with T4 DNA Ligase (0.7 U/µL). *E. coli* DH5α cells were transformed with 2.5 µL of the ligation product by heat shock, and cells were plated on LB-ampicillin agar plates and grown for 15 h at 37 °C. Several colonies were picked and used to inoculate separate 5 mL cultures of LB-ampicillin medium. The cultures were grown at 37 °C for 12 h, and plasmids were isolated using a QIAprep Spin Miniprep Kit (QIAGEN). The sequences of the resulting plasmid products were confirmed by DNA sequencing.

The table below shows the template and primer sets used for mutagenesis of the sunA gene (underlined font indicates nucleotide change):

| Target Construct | Template | Forward Primer (5’ to 3’) | Reverse Primer (5’ to 3’) |
|------------------|----------|--------------------------|--------------------------|
| sunA G21E | pET15b sunA | SunA G21E FP: AGTGGCGGTACAATGGAAATGTTGGGTGGGAGGACT | SunA G21E RP: AGCTCCGCACCCACATTTGATGACCGCAT |
| sunA G23E | pET15b sunA | SunA G23E FP: GTATGACATCTGGTGAAGCGGAGCTGTGCGT | SunA G23E RP: AGCAACAGCTCCGCTTCACAAACCAATGTGACC |
| sunA G23K | pET15b sunA | SunA G23K FP: GTATGACATCTGGTGAAGCGGAGCTGTGCGT | SunA G23K RP: AGCAACAGCTCCGCTTTCACACCAATGTGACC |
| sunA Xa G21Q/G23A | pET15b sunA Xa | SunA Xa G21Q/G23A FP: AGTGGCGGTACAATGGAAATGTTGGGTGGGAGGACT | SunA Xa G21Q/G23A RP: AGCAACAGCTCCGCTTCACAAACCAATGTGACC |
Overexpression and purification of His\textsubscript{6}-SunA precursor peptides. *E. coli* BL21 (DE3) cells or *E. coli* Rosetta 2 (DE3) cells were transformed via electroporation with a pET15b SunA construct (BL21 used for pET15b SunA; Rosetta 2 was used for all pET15b SunA constructs containing mutations: pET15b SunA Xa, pET15b SunA C22S, pET15b SunA G21E, pET15b SunA G23E, pET15b SunA G23K, and pET15b SunA Xa G21Q/G23A). A single colony transformant was used to inoculate a 30 mL culture of LB supplemented with 100 μg/mL ampicillin (and 25 μg/mL chloramphenicol for Rosetta 2 cells). The culture was grown at 37 °C for 12 h and was used to inoculate 3 L of LB containing 100 μg/mL ampicillin (and 25 μg/mL chloramphenicol for Rosetta 2 cells), and cells were grown at 37 °C to OD\textsubscript{600} = 0.6-0.8. IPTG was added to a final concentration of 1 mM and the culture was incubated at 37 °C for an additional 3 h. Overexpressed His\textsubscript{6}-SunA peptides were expressed as insoluble peptides. Cells were harvested by centrifugation at 12,000 ×g for 15 min at 4 °C, and the pellet was resuspended in 30 mL of start buffer (20 mM NaH\textsubscript{2}PO\textsubscript{4} (pH 7.5), 500 mM NaCl, 0.5 mM imidazole, 20% glycerol) and stored at -80 °C.

The cell paste was suspended in start buffer and the suspension was sonicated on ice for 20 min to lyse the cells. Cell debris was removed by centrifugation at 23,700 ×g for 30 min at 4 °C. The supernatant was discarded and the pellet containing the insoluble peptide was resuspended in the 30 mL of start buffer. The sonication and centrifugation steps were repeated. Again the supernatant was discarded and the pellet was resuspended in 30 mL of buffer 1 (6 M guanidine HCl, 20 mM NaH\textsubscript{2}PO\textsubscript{4} (pH 7.5), 500 mM NaCl, 0.5 mM imidazole). The sample was sonicated and insoluble material was removed by centrifugation at 23,700 ×g for 30 min at 4 °C, followed by filtration of the supernatant through a 0.45 μm filter. The filtered sample was applied to a 5 mL HisTrap HP (GE Healthcare Life Sciences) immobilized metal affinity chromatography (IMAC) column previously charged with NiSO\textsubscript{4} and equilibrated in buffer 1. The column was washed with two column volumes of buffer 1, followed by two column volumes of buffer 2 (4 M guanidine HCl, 20 mM NaH\textsubscript{2}PO\textsubscript{4} (pH 7.5), 500 mM NaCl, 30 mM imidazole). The peptide was eluted with 1-2 column volumes of elution buffer (4 M guanidine HCl, 20 mM NaH\textsubscript{2}PO\textsubscript{4} (pH 7.5), 500 mM NaCl, 1 M imidazole). The fractions were desalted using a ZipTipC\textsubscript{18} and analyzed by MALDI-TOF MS. The fractions containing the desired peptide were pooled and purified using preparative HPLC.

Preparative HPLC was performed using a Waters Delta 600 instrument equipped with a Waters Delta-Pak C4 column (15 μm, 300 Å, 100 mm x 25 mm) equilibrated in 2% B (solvent A = 0.1% TFA in water, solvent B = 0.0866% TFA in 80% ACN/20% water). Peptide containing material was applied to the column and was fractionated using a gradient from 2-100% B over 45 min with a flow rate of 8.0 mL/min. All fractions were analyzed by MALDI-TOF MS as described above. Purified peptide was lyophilized to dryness and stored under N\textsubscript{2} at -80 °C until further use. Typical yields from 3 L of cell culture were 15-20 mg His\textsubscript{6}-SunA, 5-8 mg His\textsubscript{6}-SunA Xa, 3-5 mg of each of His\textsubscript{6}-SunA C22S, His\textsubscript{6}-SunA G21E, His\textsubscript{6}-SunA G23E, His\textsubscript{6}-SunA G23K, and 1-1.5 mg of His\textsubscript{6}-SunA Xa G21Q/G23A.

Overexpression and purification of His\textsubscript{6}-SunS. *E. coli* Rosetta 2 (DE3) cells were transformed with the pET28b SunS construct via electroporation. A single colony transformant was used to inoculate a 30 mL culture of LB supplemented with 50 μg/mL kanamycin. The culture was grown at 37 °C for 12 h and was used to inoculate 3 L of LB containing 50 μg/mL kanamycin, and cells were grown at 37 °C to OD\textsubscript{600} = 0.6. The culture was incubated at 4 °C on ice for 20 min, then IPTG was added to a final concentration of 0.5 mM and the culture was incubated at
Cells were harvested by centrifugation at 12,000 ×g for 15 min at 4 °C, and the pellet was resuspended in 30 mL of start buffer (20 mM Tris (pH 8.0), 500 mM NaCl, 1 mM TCEP, 10% glycerol) and stored at -80 °C.

All protein purification steps were performed at 4 °C. The cell paste was suspended in start buffer and the cells were lysed using a high pressure homogenizer (Avestin, Inc.). Cell debris was pelleted via centrifugation at 23,700 ×g for 20 min at 4 °C. The supernatant was injected via a superloop onto a fast protein liquid chromatography (FPLC) system (ÄKTA, GE Heathcare Life Sciences) equipped with a 5 mL HisTrap HP IMAC column previously charged with Ni²⁺ and equilibrated in start buffer. The column was washed with 50 mL of buffer A (30 mM imidazole, 20 mM Tris, pH 8.0, 1 M NaCl) and the protein was eluted using a linear gradient of 0-100% B (buffer B = 200 mM imidazole, 20 mM Tris, pH 8.0, 1 M NaCl) over 40 min at a 2 mL/min flow rate. UV (280 nm) was monitored and fractions were collected and analyzed by SDS-PAGE (4-20% Tris-glycine READY gel, BioRAD). The fractions containing SunS were combined and concentrated using an Amicon Ultra-15 Centrifugal Filter Unit (10 kDa MWCO, Millipore). Gel filtration purification was used to further purify SunS. The concentrated protein sample was injected onto an FPLC system (ÄKTA) equipped with an XK16 16/60 (GE Healthcare Life Sciences) column packed with SuperDex 75 resin previously equilibrated in 20 mM HEPES (pH 7.5), 100 mM KCl, and 1 mM TCEP. The protein was eluted with a flow rate of 0.9 mL/min. Both UV (280 nm) and conductance were monitored and fractions were collected. Misfolded/aggregated protein was efficiently separated from soluble, correctly folded protein and the desired fractions were combined and concentrated using an Amicon Ultra-15 Centrifugal Filter Unit. The resulting protein sample was stored at -80 °C. Protein concentration was determined using a Bradford Assay Kit (Pierce) and typically yields were 15-18 mg His₆-SunS from 3 L of cell culture (Supplementary Fig. 5).

**Enzymatic in vitro sugar modification of SunA prepeptides by SunS.** His₆-SunA (50 μM) was incubated with His₆-SunS (2 μM) in a reaction buffer containing 50 mM Tris (pH 7.5), 1 mM MgCl₂, 1 mM TCEP, and 5 mM NDP-sugar (final concentrations). Five NDP-sugars were independently evaluated: uridine-5′-diphosphate-α-D-glucose (UDP-Glc, CalBiochem), uridine-5′-diphosphate-α-D-galactose (UDP-Gal, CalBiochem), uridine-5′-diphosphate-α-D-N-acetylglucosamine (UDP-GlcNAc, Sigma), guanosine-5′-diphosphate-α-D-mannose (GDP-Man, CalBiochem), and uridine-5′-diphosphate-α-D-xylene (UDP-Xyl, Complex Carbohydrate Research Center, University of Georgia). The reactions were incubated at 25 °C for various durations of time (1-12 h depending on future use). Reactions were either immediately used for subsequent proteolytic leader peptide removal reactions and oxidative folding or quenched with 5% TFA to pH 1-2 for MALDI-TOF MS analysis (Figure 3A and Supplementary Fig. 7 and 12). For MALDI-TOF MS analysis, samples were desalted using ZipTipC₁₈ and analyzed as described above. For modification of His₆-SunA mutant peptides (His₆-SunA Xa, His₆-SunA C22S, His₆-SunA G21E, His₆-SunA G23E, His₆-SunA G23K, and His₆-SunA Xa G21Q/G23A), reactions were conducted under identical conditions as described above.

**Chymotrypsin digests of sugar modified His₆-SunA peptides.** Modified His₆-SunA (taken directly from the SunS reaction) was digested in 100 mM Tris (pH 7.5), 5 mM TCEP, and 0.1 mg/mL chymotrypsin (Worthington). All reactions were incubated at 25 °C for 5 h and then quenched with 5% TFA to pH 1-2. Quenched samples were desalted with a ZipTipC₁₈ prior to
analysis by MALDI-TOF MS or analyzed by LC-ESI-Q/TOF MS without further manipulation (Supplementary Fig. 6 and 11).

Base-catalyzed β-elimination of glucose from Cys22 of sublancin. Two different strong bases were used to evaluate the ability to β-eliminate the sugar from in vivo isolated sublancin peptide (Supplementary Fig. 8). For elimination with a strong base under reducing conditions, pure sublancin was dissolved to a final concentration of 25 μM in a solution containing 0.3 M NaBH₄ and 0.1 M NaOH and was incubated at 4 °C for 5 h followed by quenching the reaction with 5% TFA to pH 1-2. The sample was desalted using a ZipTipC₁₈ and analyzed by MALDI-TOF and ESI-Q/TOF MS (Supplementary Fig. 8A, C, and D). For elimination with a strong base under non-reducing conditions, pure sublancin was dissolved to a final concentration of 25 μM in a solution containing 0.1 M 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) in n-PrOH and was incubated at 25 °C for 5 h followed by quenching the reaction with 5% TFA to pH 1-2. The sample was desalted using a ZipTipC₁₈ and analyzed by MALDI-TOF MS (Supplementary Fig. 8B). Sublancin samples that were subjected to base-catalyzed β-elimination were digested under reducing conditions with chymotrypsin in a similar manner as described for modified His₆-SunA peptides (Supplementary Fig. 8C and D).

NMR analysis of sublancin. Two samples of sublancin were prepared for NMR analysis. For a sample in 100% D₂O, lyophilized sublancin was dissolved in 100% D₂O (Cambridge Isotope Laboratories) to exchange amide protons and simplify the NMR spectrum in the downfield region. The sample was lyophilized, the procedure was repeated twice, and the final sample was dissolved in D₂O to a final concentration of approximately 2.5 mM. For a sample in 90% H₂O/10% D₂O, lyophilized sublancin was dissolved in 90% H₂O/10% D₂O to a final concentration of approximately 2.5 mM.

Solution NMR spectra were acquired at the NMR Facility (School of Chemical Sciences, University of Illinois at Urbana-Champaign) on a Varian INOVA 500 MHz spectrometer equipped with a 5 mm triple resonance (¹H−¹³C−¹⁵N) triaxial gradient probe, using VNMRJ version 2.1B with the BioPack suite of pulse programs released in early 2006. One and two-dimensional ¹H homonuclear spectra were acquired on both sublancin samples at 25°C. Two-dimensional COSY, TOCSY and NOESY spectra were measured for an average of 2 h per spectrum, digitizing 1024 points in the indirect ¹H dimension (t₁,max = 128 ms) and were utilized to establish correlations among backbone resonances and the conjugated glucose ring of the sublancin sample. An additional COSY spectra with a long pulse delay (9 ms) was acquired with the sublancin sample in 100% D₂O for measurement of 3JH-H couplings. Spectra were processed with NMRPipe (1) and analyzed in Sparky (2) The 3JH-H couplings were measured using the ACME program supplied as a feature of the NMRPipe software (3).

Initially, efforts focused on the region of the spectrum from 6.0 ppm to 7.5 ppm (Supplementary Figure 9a) that was previously reported to contain resonances from the Dha residue, as reported by Paik et al. (4). For this specific investigation the sample in 100% D₂O was used to avoid overlap with amide proton resonances. The previous study relied on one-dimensional ¹H spectra and chemical shift analysis to propose that the resonance at 6.2 ppm belonged to the Dha residue and the remaining resonances were attributable to aromatic side chain protons. Here we use both one-dimensional and two-dimensional data sets to assign the
resonances. Upon initial examination of the two-dimensional COSY spectra it can be seen that all the resonances from ~6.0 ppm to 7.5 ppm are from one distinct spin system (Supplementary Fig. 9c). Furthermore, the 6 resonances in the spin system are not consistent with any of the aromatic residues (W, Y or F) present in the peptide. We tentatively assign these resonances as belonging to the conjugated glucose as it is the only 6-proton spin system identified. Furthermore, all coupling constants are large (>10 Hz), consistent with the axial orientations of the ring protons in a glucose. The only smaller coupling is that observed for the most upfield proton (at 6.25 ppm), consistent with this signal coming from H6. However, we note that for this assignment to be correct the chemical shifts of the sugar would have to be perturbed greatly by the chemical environment of the sugar in the sublancin molecule and since we have not solved the entire structure, this assignment at present must be considered tentative.

Further support for this hypothesis can be seen in the cross peaks to the glucose ring in the two-dimensional NOESY spectrum (Supplementary Fig. 9b). Resonances assigned to proton 1 in the spectra have cross peaks in the 2D-NOESY spectrum to the beta protons of a cysteine residue (possibly Cys22) as well as to the lone isoleucine residue (residue 20). Additionally, the isoleucine delta proton has a strong NOESY cross peak to proton 3 of the glucose. Isoleucine has a unique spin system from other amino acids and isoleucine 20 is the only isoleucine in the sequence; therefore, it serves as a unique identifier in the sequence indicating that the resonances from 6.0 ppm to 7.5 ppm are very close in proximity to the site of proposed conjugation (Cys22). The cross peak intensities in the NOESY spectrum indicate that proton 1 and 3 are within 4-5 angstroms of a Cys and Ile20. This observation is consistent with a conjugated glucose at Cys22 (see Fig. 1c). Efforts to assign the three dimensional structure of sublancin utilizing the additional cross peaks in the NOESY spectrum (Supplementary Fig. 9b and data not shown) are on-going. No spin systems corresponding to Trp, Tyr, and Phe were observed under the conditions used, which is not unusual in protein NMR as these resonances are often weak in intensity under the conditions used. Our continuing efforts are being coupled with atomically resolved structure calculations and we hope to better explain this phenomenon with the full structure in hand.

Provided the assignments of the conjugated glucose are correct, the NMR data can be used to determine the stereochemistry of the glycosylation site. By measuring the $^{3}J_{H-H}$ coupling constant between the anomic proton (proton 1) and the neighboring proton (proton 2) the dihedral angle was calculated using the Karplus relationship (5). A COSY spectrum acquired with a long pulse delay was used to determine this $^{3}J_{H-H}$ coupling constant. Using the ACME software (3) the cross peak of 1-2 (Supplementary Fig. 9c) was best fit with a $^{3}J_{H-H}$ coupling constant of 10.25 Hz. This value results in a dihedral angle of 180° ± 10° based on the Karplus equation (5) and would indicate that the glucose is attached to the peptide through a β-linkage.

References for NMR section:
1. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. 1995 NMRPipe: A multidimensional spectral processing system based on Unix pipes J. Biomol. NMR 6, 277, 293
2. Goddard, T. D., and Kneller, D. G. ( 2005) Sparky; University of California, San Francisco.
3. Delaglio, F., Wu, Z., and Bax, A. Measurement of Homonuclear Proton Couplings from Regular 2D COSY Spectra J. Magn. Reson., 149, 276-281 (2001).
4. Paik, S.H., Chakicherla, A. & Hansen, J.N. J. Biol. Chem. 1998 273, 23134-42.
5. Karplus, M. J. Chem. Phys. 1959. 30, 11.
Kinetics of SunS glycosyltransferase activity using end-point assays. To evaluate the NDP-sugar substrate preference of SunS, His6-SunA was modified with different NDP-sugars at varied NDP-sugar concentrations and products were quantified using HPLC. His6-SunA (5 μM) was incubated with His6-SunS (2 μM) in a 250 μL reaction buffer containing 50 mM Tris (pH 7.5), 1 mM MgCl2, 1 mM TCEP, and a varied concentration of NDP-sugar (final concentrations). NDP-sugar concentrations were 5 mM, 50 μM, or 500 nM. All five NDP-sugars were independently evaluated: UDP-Glc, UDP-Gal, UDP-GlcNAc, GDP-Man, and UDP-Xyl. The reactions were incubated at 25 °C for 1 h. Reactions were quenched with TFA to pH 1-2 and analyzed by MALDI-TOF MS and the extent of glycosylation was quantified by analytical HPLC. For MALDI-TOF MS analysis, samples were desalted using ZipTip C18 and analyzed as described above (Supplementary Fig. 13 upper).

HPLC analysis was performed using a Beckman Coulter System Gold HPLC equipped with a Grace-Vydac Protein C4 column (5 μm, 300 Å, 250 mm x 4.6 mm) equilibrated in 2% solvent B (solvent A = 0.1% TFA in water, solvent B = 0.0866% TFA in 80% ACN/20% water). A 200 μL volume of quenched SunS reaction was applied to the column. Sugar-modified and unmodified His6-SunA wt peptide material were eluted by maintaining the mobile phase at 2% B for 1 min, followed by an increase to 100% B over 45 min with a flow rate of 1.0 mL/min. Under these conditions, sugar-modified His6-SunA wt peptide and unmodified His6-SunA wt peptide eluted at 24.3 and 24.8 min, respectively (Supplementary Fig. 13 lower). All fractions were analyzed by MALDI-TOF MS as described above. The relative amounts of sugar-modified and unmodified His6-SunA peptide was determined by quantifying the peak area corresponding to each form of the peptide by integration. Percent conversion was calculated by dividing the peak area of sugar-modified peptide by the sum of the sugar-modified and unmodified peak areas (Supplementary Table 1).

Proteolysis of SunA Xa precursor and purification of leader and core peptides. To assess whether the leader peptide of SunA is required for SunS-catalyzed core peptide modification, His6-SunA Xa precursor peptide was proteolytically cleaved using Factor Xa protease to generate leader peptide and unmodified core peptide. His6-SunA Xa (1.0 mg/mL) was digested in 50 mM Tris (pH 7.5), 100 mM NaCl, 2 mM CaCl2, and 0.1 mg/mL Factor Xa (New England Biolabs). The reaction was incubated at 25 °C for 2.5 h and an aliquot of the reaction was quenched with 5% TFA to pH 1-2 prior to desalting by ZipTip C18 and analysis by MALDI-TOF MS. At 2.5 h, no starting material was observed by MS (Supplementary Fig. 14B), and the remainder of reaction was quenched with 5% TFA to pH 1-2. The leader and core peptide were then purified using HPLC. HPLC was performed using a Beckman Coulter System Gold HPLC equipped with a Phenomenex Jupiter Proteo C12 column (10 μm, 90 Å, 250 mm x 4.6 mm) equilibrated in 2% solvent B (solvent A = 0.1% TFA in water, solvent B = 0.0866% TFA in 80% ACN/20% water). The quenched proteolysis reaction containing a mixture of core and leader peptide was centrifuged at 15 krpm in a microcentrifuge for 2 min and the resulting supernatant was applied to the column. The core and leader peptides were eluted by maintaining the mobile phase at 2% B for 1 min, followed by an increase to 100% B over 45 min with a flow rate of 1.0 mL/min. Under these conditions, the leader peptide and core peptide eluted at 22.7 and 25.5 min, respectively. All fractions were analyzed by MALDI-TOF MS as described above. Purified material was lyophilized to dryness and stored under N2 at −80 °C until further use. Typical yields were 0.2 mg of leader and 0.15 mg of core peptide per mg of His6-SunA Xa (Supplementary Fig. 14C-E).
Evaluation of the requirement of leader peptide for core peptide modification by SunS.
Purified leader and core peptide were incubated in combination (in trans) and independently in the presence of SunS reaction components. For the in trans reaction, leader peptide (10 μM) and core peptide (10 μM) were incubated with His6-SunS (2 μM) in a reaction buffer containing 50 mM Tris (pH 7.5), 1 mM MgCl2, 1 mM TCEP, and 5 mM UDP-glucose (all final concentrations). For the independent reactions, leader peptide (10 μM) or core peptide (10 μM) was incubated with His6-SunS (2 μM) in a reaction buffer containing 50 mM Tris (pH 7.5), 1 mM MgCl2, 1 mM TCEP, and 5 mM UDP-glucose. All reactions were incubated at 25 °C for 1 h and then quenched with 5% TFA to pH 1-2. Quenched samples were desalted using a ZipTipC18 prior to analysis by MALDI-TOF MS (Supplementary Fig. 14).

In vitro preparation of sublancin and sublancin analogs for bioactivity assays and LC-ESI-Q/TOF MS analysis. Sublancin and sublancin analogs containing non-natural sugars were prepared by modifying His6-SunA Xa peptide with a sugar (SunS reaction), proteolytic removal of the leader peptide (Factor Xa reaction), and oxidative folding of sublancin to afford the disulfide linkages (oxidative folding reaction). Sugar modified His6-SunA Xa was prepared in 250 μL of 50 mM Tris (pH 7.5), 1 mM MgCl2, 1 mM TCEP, 5 mM NDP-sugar, 50 μM His6-SunA Xa, and 2 μM His6-SunS. The reaction was incubated at 25 °C for 12 h. The extent of sugar modification was verified by removing a 5 μL aliquot of the reaction, quenching with 5% TFA to pH 1-2, desalting using a ZipTipC18, and analysis by MALDI-TOF and ESI Q/TOF MS. Following analysis, the leader peptide of sugar modified His6-SunA Xa was proteolytically cleaved by the addition of NaCl and CaCl2 to 100 mM and 2 mM, respectively, and the addition of Factor Xa to 0.075 mg/mL (final concentrations). The reaction was incubated at 25 °C for 4-6 h and the extent of cleavage was monitored by MALDI-TOF MS as stated above. Following analysis, the disulfides of the modified sublancin core peptide were formed by addition of Tris (pH 7.5), oxidized glutathione (GSSG), reduced glutathione (GSH), and EDTA to final concentrations of 50 mM, 2 mM, 2 mM, and 0.1 mM, respectively. The total volume of the oxidative folding reaction was 500 μL and the reaction was incubated at 25 °C for an additional 12 h. The extent of disulfide formation was monitored by removing a 5 μL aliquot of the reaction, quenching with 5% TFA to pH 1-2, desalting using a ZipTipC18, and analyzing by MALDI-TOF MS. Disulfide formation was observed as a peak with a -4 Da mass difference compared with material that was not subjected to oxidative folding.

Evaluation of in vitro prepared sublancin and sublancin analogs. The in vitro preparation of sublancin and sublancin analogs was evaluated in three ways (Figure 3B, Supplementary Fig. 16, 17, and 18). The sugar modification, leader peptide removal, and oxidative folding are all required to obtain bioactivity against the indicator strain B. subtilis ATCC 6633. As such, the in vitro prepared peptides were assessed using the antimicrobial activity assay described below.

In addition to bioactivity assays, sublancin and analogs were structurally characterized via LC-ESI-Q/TOF tandem MS and the resulting fragmentation patterns of the acquired mass spectra were compared to native and reduced authentic sublancin analyzed under identical conditions (Supplementary Fig. 17). For additional confirmation of correct disulfide bond formation of in vitro prepared sublancin and sublancin analogs, the fully modified peptides were independently proteolytically digested with chymotrypsin and the reactions were analyzed by LC-ESI-Q/TOF MS. Briefly, in vitro prepared peptide (SunS treated, Factor Xa treated, and oxidatively folded) was digested under non-reducing conditions in 100 mM Tris (pH 7.5) and
0.05 mg/mL chymotrypsin. All reactions were incubated at 25 °C for 3 h and then quenched with 5% TFA to pH 1-2. Quenched samples were analyzed by LC-ESI-Q/TOF MS without further manipulation (Supplement Fig. 18).

**Antimicrobial activity assays of in vitro prepared sublancin and sublancin analogs.** The 500 μL reactions described above (SunS reaction, Factor Xa reaction, folding reaction) were concentrated to 20 μL via vacuum centrifugation. An overnight culture of *B. subtilis* ATCC 6633 (indicator strain) was grown in LB media under aerobic conditions at 37 °C for 12 h. Ninety-six well agar plates were prepared by combining 20 mL of molten LB medium agar (cooled to 42 °C) with 50 μL of dense overnight culture (approx 10⁸-10⁹ CFU/mL). The seeded agar was poured into a sterile OmniTray (Nunc) and allowed to solidify at 25 °C for 30 min. An additional 30 mL of molten LB medium was cooled to 42 °C, combined with 75 μL of culture, and poured over the lower solidified agar layer. A sterile 96-well PCR plate was placed in the molten agar upper layer and was allowed to solidify at 25 °C for 45 min. After sufficient solidification, the 96-well PCR plate was removed. The total 20 μL volume of each concentrated *in vitro* reaction was dispensed into separate newly formed wells. Authentic sublancin standards were spotted in 15 μL volumes at the concentrations indicated. Plates were left at 25 °C for 15 h and antibacterial activity was qualitatively determined by the presence or absence of growth inhibition (Figure 3B).

**Antimicrobial activity assay with native, reduced, and reduced and alkylated sublancin.** The importance of intact disulfides for the bioactivity of sublancin was assessed by testing the activities of native, reduced, and reduced and alkylated sublancin against *B. subtilis* ATCC 6633. Native sublancin was prepared by dissolving pure sublancin in 50 mM Tris (pH 7.5 or pH 8.3). Reduced sublancin was prepared by dissolving pure sublancin in 50 mM Tris (pH 8.3) with 5 mM TCEP. Reduced and alkylated sublancin was prepared by dissolving pure sublancin in 50 mM Tris (pH 8.3), 5 mM TCEP, and 10 mM iodoacetamide. All reactions contained sublancin at a final concentration of 50 μM and were prepared on a 50 μL scale. All reactions were incubated at 25 °C for 3 h in the dark, then concentrated to 20 μL prior to dispensing into separate wells of a bioactivity plate (prepared as described above) seeded with *B. subtilis* ATCC 6633. Authentic sublancin and nisin standards were spotted in 15 μL volumes at the concentrations indicated. Plates were left at 25 °C for 15 h and antibacterial activity was qualitatively determined by the presence or absence of growth inhibition (Supplementary Fig. 15).

**Antimicrobial activity assay of in vitro prepared sublancin with omission of single steps.** The importance of each step (glycosylation, leader peptide removal, and disulfide formation) in the *in vitro* preparation of sublancin was demonstrated through an antimicrobial activity assay against *B. subtilis* ATCC 6633. Sublancin samples were prepared in parallel whereby one step of the preparation process was omitted.

To prepare non-glucosylated sublancin, leader peptide of unmodified His₆-SunA Xa was first removed by combining 50 mM Tris (pH 7.5), 100 mM NaCl, 2 mM CaCl₂, 0.075 mg/mL Factor Xa, and 50 μM His₆-SunA Xa in a 250 μL reaction. The reaction was incubated at 25 °C for 6 h. The disulfides of the sublancin core peptide were formed by addition of Tris (pH 7.5), oxidized glutathione (GSSG), reduced glutathione (GSH), and EDTA to final concentrations of 50 mM, 2 mM, 2 mM, and 0.1 mM, respectively. The 500 μL total volume reaction was incubated at 25 °C for 12 h.
To prepare sublancin with the leader peptide intact, His6-SunA Xa was first glucosylated by combining 50 mM Tris (pH 7.5), 1 mM MgCl₂, 1 mM TCEP, 5 mM UDP-glucose, 50 μM His6-SunA Xa, and 2 μM His6-SunS in 250 μL. The reaction was incubated at 25 °C for 12 h. The disulfides of the modified His6-SunA Xa peptide were formed by addition of Tris (pH 7.5), oxidized glutathione (GSSG), reduced glutathione (GSH), and EDTA to final concentrations of 50 mM, 2 mM, 2 mM, and 0.1 mM, respectively. The 500 μL total volume reaction was incubated at 25 °C for 12 h.

To prepare unfolded sublancin, His6-SunA Xa was first glucosylated by combining 50 mM Tris (pH 7.5), 1 mM MgCl₂, 1 mM TCEP, 5 mM UDP-glucose, 50 μM His6-SunA Xa, and 2 μM His6-SunS in 250 μL. The reaction was incubated at 25 °C for 12 h. The leader peptide of glucosylated His6-SunA Xa was then cleaved by addition of 50 mM Tris (pH 7.5), 100 mM NaCl, 2 mM CaCl₂, 0.075 mg/mL Factor Xa (final concentrations) to the 250 μL reaction. The reaction was incubated at 25 °C for 6 h.

As a positive control, all three steps were performed in the following order: SunS glucosylation reaction, Factor Xa leader peptide cleavage reaction, and oxidative folding reaction. All of the above reactions were concentrated to 20 μL via vacuum centrifugation and the total 20 μL volumes were separately dispensed into separate wells of a bioactivity plate (prepared as described above) seeded with B. subtilis ATCC 6633. Authentic sublancin standards were spotted in 15 μL volumes at the concentrations indicated. Plates were left at 25 °C for 15 h and antibacterial activity was qualitatively determined by the presence or absence of growth inhibition (Supplementary Fig. 19).

Production of sublancin in xylose containing media. To investigate whether sublancin can undergo growth-medium dependent modifications as described by Dorenbos et al. (J. Biol. Chem. 277, 16682-8, 2002), B. subtilis 168 was cultured in several different culture media and production of sublancin was evaluated by MALDI-TOF MS. Three culture media preparations were used: TY with 1% xylose (1 L scale) contained 10 g tryptone, 5 g yeast extract, and 5 g NaCl. A 20% xylose solution in water was prepared, sterilized via 0.22 μm filtration, and was added to the media to a final concentration of 1%. Medium X was of identical composition as Medium A except that 20 g of xylose was used instead of sucrose. Medium A (described above) was used as a positive control for production of sublancin.

A culture of B. subtilis 168 was grown in LB media under aerobic conditions at 37 °C for 12-15 h. The overnight culture was used to inoculate (at 1%) 500 mL volumes of TY with 1% xylose, Medium X, and Medium A in 2 L flasks. All cultures were grown under aerobic condition at 37 °C for 48 h with vigorous agitation. At various time points during culture growth, a 1 mL aliquot was removed from each culture, acidified to pH 2.5 with concentrated phosphoric acid (85% in water), and centrifuged to remove cells and insoluble material. The production and presence of sublancin, sublancin with additional xylose modifications, subtilosin A, and subtilosin A with additional xylose modifications was assessed by MALDI-TOF MS. The supernatant from each aliquot sample was desalted using ZipTip C₁₈ and was analyzed by MALDI-TOF MS (Supplementary Figure 20).

Attempts to enzymatically install xylose modifications in addition to glucose modification of SunA prepeptide by SunS. His6-SunA (50 μM) was incubated with His6-SunS (2 μM) in a reaction buffer containing 50 mM Tris (pH 7.5), 1 mM MgCl₂, 1 mM TCEP, 5 mM UDP-Glc, and 5 mM UDP-Xyl (final concentrations) (Supplementary Fig. 21). An additional experiment
was performed using purified His\textsubscript{6}-SunA peptide that was previously modified with glucose. This reaction contained His\textsubscript{6}-SunA-Glc (~50 μM), His\textsubscript{6}-SunS (2 μM) in a reaction buffer containing 50 mM Tris (pH 7.5), 1 mM MgCl\textsubscript{2}, 1 mM TCEP, and 5 mM UDP-Xyl (final concentrations). The reactions were incubated at 25 °C for 2.5 h. The reactions were quenched with 5% TFA to pH 1-2 for MALDI-TOF MS analysis (Supplementary Fig. 22).
**Supplementary Results and Figures**

**Supplementary Fig. 1A. ESI-Q/TOF MS analysis of native sublancin.** Pure sublancin isolated from *B. subtilis* 168 without any further manipulation was analyzed by LC-ESI-Q/TOF MS. Expected [M+H]: 3876.74, observed: 3876.7500.

**Supplementary Fig. 1B. ESI-Q/TOF MS analysis of reduced sublancin.** Pure sublancin was reduced using TCEP and analyzed by LC-ESI-Q/TOF MS. The mass of reduced sublancin is 4 mass units greater than native sublancin, suggesting that two disulfide linkages exist in the native structure of sublancin. Expected [M+H]: 3880.77, observed: 3880.7803.
Supplementary Fig. 1C. ESI-Q/TOF MS analysis of reduced and alkylated sublancin. Pure sublancin was reduced and alkylated using TCEP and iodoacetamide and analyzed by LC-ESI-Q/TOF MS. Four of five possible Cys residues were modified with a carboxyamidomethyl (CAM) group, suggesting that a single Cys residue is modified such that its thiol group cannot react with iodoacetamide. Expected [M+H] (for four alkylations): 4107.86, observed: 4107.8340.

Supplementary Fig. 1D. ESI-Q/TOF MS analysis of native, reduced, and reduced and alkylated sublancin. Overlay of mass spectra obtained by LC-ESI-Q/TOF MS analysis of native (blue), reduced (red), and reduced and alkylated (green) sublancin.
Supplementary Fig. 1E. MALDI-TOF MS analysis of native, reduced, and alkylated sublancin. Overlay of mass spectra obtained by MALDI-TOF MS analysis of native (blue), reduced (red), and reduced and alkylated (green) sublancin.

Supplementary Fig. 1F. ESI-Q/TOF MSMS analysis of native, reduced, and alkylated sublancin. Pure sublancin isolated from B. subtilis 168 was analyzed by LC-ESI-Q/TOF MSMS. Fragmentation by collision induced dissociation of the amide bonds of sublancin did not occur when the disulfide bonds remained intact. Rather, the post-translation modification identified as an S-linked glucosylation was lost from the parent ion (-162 Da).
Supplementary Fig. 2A. MALDI-TOF MS analysis of sublancin digested with chymotrypsin under reducing conditions. Sublancin was digested with chymotrypsin under reducing conditions and analyzed by MALDI-TOF MS. The resulting mass spectrum is shown. Chymotrypsin cleavage sites (F, W, and Y residues) are indicated in the peptide sequence above the spectrum. The masses of the observed ions and their corresponding digest fragments are assigned. The results are consistent with Ser16 and Thr19 remaining unmodified and residues Leu12-Tyr32 bearing a +162 Da post-translational modification.
Supplementary Fig. 2B. ESI-Q/TOF MS analysis of sublancin digested with chymotrypsin under non-reducing conditions. Sublancin was digested with chymotrypsin under non-reducing conditions and analyzed using LC-ESI-Q/TOF MS to investigate disulfide bridge connectivity. Chymotrypsin cleavage sites (F, W, and Y residues) and the proposed disulfide bridges are indicated in the peptide sequence above the spectrum. The masses of the observed ions and their corresponding digest fragments are assigned. The results support that a disulfide bond exists between Cys7-Cys36 and between Cys14-Cys29. The results are consistent with Ser16 and Thr19 remaining unmodified and residues Leu12-Tyr32 bearing a +162 Da post-translational modification.
Supplementary Fig. 2C. ESI-Q/TOF MS analysis of sublancin digested with chymotrypsin under reducing conditions. Sublancin was digested with chymotrypsin under reducing conditions and analyzed using LC-ESI-Q/TOF MS. Chymotrypsin cleavage sites (F, W, and Y residues) are indicated in the peptide sequence above the spectrum. The masses of the observed ions and their corresponding digest fragments are assigned. The results support that the disulfide bond existing between Cys7-Cys35 was reduced and peak corresponding to the ion with a mass of 1392.62 Da (native sublancin) was no longer observed. Rather, a peak corresponding to residues Gly1-Trp11 with a mass of 1117.53 Da was observed. In addition, the ion corresponding to Leu12-Tyr32 had a 2 Da mass increase compared to the Leu12-Tyr32 ion found in the native sublancin digest. These results further suggest that a disulfide bond exists between Cys14-Cys29. The results are consistent with Ser16 and Thr19 remaining unmodified and residues Leu12-Tyr32 bearing a +162 Da post-translational modification.
Supplementary Fig. 2D. ESI-Q/TOF MSMS analysis of residues Leu12-Tyr32 of sublancin. Sublancin was digested with chymotrypsin under reducing conditions and analyzed using LC-ESI-Q/TOF MSMS. The digest mixture was fractionated using LC and the multiply charged ion corresponding to residues Leu12-Tyr32 was selected for MSMS analysis. Fragmentation patterns are indicated for the y" ion series (upper) and the b ion series (lower). The results support that Ser16 and Thr19 remain unmodified and Cys22 has a +162 Da post-translational modification.
**Supplementary Fig. 3A. GC-MS analysis of glucose.** D-(+)-Glucose standard was derivatized and analyzed by GC-MS. The TIC is shown with the retention times as indicated (upper). Below the chromatogram are EI-MS mass spectra corresponding to the peaks at 11.017 min (lower left) and 11.399 min (lower right) that correspond to the two isomers of the sugar generated during derivatization. The characteristic ions for TMS derivatized glucose are annotated within the spectra (319, 205, 147, 103, and 73 m/z).
Supplementary Fig. 3B. GC-MS analysis of galactose. D-(+)-Galactose standard was derivatized and analyzed by GC-MS. The TIC is shown with the retention times as indicated (upper). Below the chromatogram EI-MS mass spectra corresponding to the peaks at 10.890 min (lower left) and 11.332 min (lower right) are shown that correspond to the two isomers of the sugar generated during derivatization. The characteristic ions for TMS derivatized galactose are annotated within the spectra (319, 205, 147, 103, and 73 m/z).
Supplementary Fig. 3C. GC-MS analysis of mannose. D-(+)-Mannose standard was derivatized and analyzed by GC-MS. The TIC is shown with the retention times as indicated (upper). Below the chromatogram, EI-MS mass spectra corresponding to the peaks at 10.811 min (lower left) and 11.143 min (lower right) are shown that correspond to the two isomers of the sugar generated during derivatization. The characteristic ions for TMS derivatized mannose are annotated within the spectra (319, 205, 147, 103, and 73 m/z).
Supplementary Fig. 3D. GC-MS analysis of fructose. D-(-)-Fructose standard was derivatized and analyzed by GC-MS. The TIC is shown with the retention times as indicated (upper). Below the chromatogram, EI-MS mass spectra corresponding to the peaks at 10.489 min (lower left) and 10.675 min (lower right) are shown that correspond to the two isomers of the sugar generated during derivatization. The characteristic ions for TMS derivatized fructose are annotated within the spectra (307, 217, 147, 103, and 73 m/z).
Supplementary Fig. 3E. Resolution of derivatized standards by GC-MS. A mixture of derivatized D-(-)-glucose, D(+)-galactose, D(+)-mannose, and D(-)-fructose standards was analyzed by GC-MS. Shown are the TIC for the mixture (upper) and the overlay of the TIC for the mixture and the total ion chromatograms from each individually analyzed sugar (lower). The EI-MS mass spectra corresponding to the individual peaks of the mixture were essentially identical to the spectra obtained when each sugar standard was analyzed individually (Supplementary Fig. 3A-D).
Supplementary Fig. 3F. GC-MS analysis of the sugar attached to sublancin. The S-linked hexose sugar conjugated to Cys22 of in vivo isolated sublancin was chemically released via acid hydrolysis, derivatized, and analyzed by GC-MS. The TIC is shown with the retention times as indicated (upper). EI-MS mass spectra corresponding to the peaks at 11.023 min (lower left) and 11.395 min (lower right) are shown that correspond to the two isomers of the sugar generated during derivatization. The hexose sugar was identified as glucose, as determined by retention time and comparison of the mass spectra with the glucose standard. The characteristic ions for TMS derivatized glucose are annotated within the spectra (319, 205, 147, 103, and 73 m/z).
Supplementary Fig. 3G. Identification of the sugar in sublancin as glucose. An overlay of the total ion chromatograms for the sublancin-derived and silylated sugar sample and for the mixture of silylated sugar standards is shown (upper: full chromatogram, lower: zoom to peaks of interest).
Supplementary Fig. 3H. Confirmation of glucose as the sugar in sublancin by spiking with a standard. To confirm the presence of $S$-linked glucose conjugated to Cys22 of \textit{in vivo} isolated sublancin, the sublancin derived sugar sample was spiked with glucose standard and analyzed by GC-MS. An overlay of the total ion chromatograms for the sublancin derived sugar sample and for the sublancin sample with glucose spike is shown (upper: full chromatogram, lower: zoom to peaks of interest). The EI-MS mass spectra corresponding to the sublancin sample with glucose spike were similar to the spectra obtained when the sublancin derived sugar sample (without added glucose) was analyzed. The characteristic ions for TMS derivatized glucose were observed.
**Supplementary Fig. 4A.** Sequence alignments of SunS with glycosyltransferases sharing high homology. BLAST analysis resulted in several glycosyltransferases (hypothetical and experimentally defined) sharing high sequence homology with SunS and located in a cluster with a putative peptide substrate (see **Supplementary Fig. 10**). Shown is the sequence alignment of SunS and the top hits from the BLAST search in August 2010. The asterisks (*) above the sequences indicate the putative divalent metal (Mg$^{2+}$, Mn$^{2+}$) binding site.
Supplementary Fig. 4B. Representative gene clusters containing both a short open reading frame that may encode a substrate peptide and a glycosyl transferase with homology to SunS. Glycosyl transferase genes shown in green; genes encoding substrate peptide in black; genes encoding transporters shown in red; genes encoding thiol/disulfide isomerases shown in yellow, and genes encoding putative immunity proteins shown in light blue.
Supplementary Fig. 5. Characterization of SunS glycosyltransferase. The gene sunS was cloned and expressed in *Escherichia coli* as an N-terminal fusion protein with a hexa-histidine tag (His<sub>6</sub>-SunS). His<sub>6</sub>-SunS consists of 442 amino acids and has a predicted molecular weight of 51,928 Da. Shown below is an SDS-PAGE (BioRAD Ready Gel 4-20%) analysis of purified SunS protein.
Supplementary Fig. 6A. ESI-Q/TOF MS analysis of glucose modified His$_6$-SunA digested with chymotrypsin under reducing conditions. His$_6$-SunA was modified with glucose, digested with chymotrypsin under reducing conditions, and analyzed by using LC-ESI-Q/TOF MS. Chymotrypsin cleavage sites (F, W, and Y residues) are indicated in the peptide sequence above the spectrum. The masses of the observed ions and their corresponding digest fragments are assigned. The results are consistent with Ser54 and Thr57 remaining unmodified and residues Leu50-Tyr70 bearing a +162 Da post-translational modification. (upper: full spectrum, lower: zoom to peaks of interest).
Supplementary Fig. 6B. ESI-Q/TOF MSMS analysis of residues Leu50-Tyr70 of glucose modified His6-SunA. His6-SunA was modified with glucose, digested with chymotrypsin under reducing conditions, and analyzed by using LC-ESI-Q/TOF MSMS. The digest mixture was fractionated using LC and the multiply charged ion corresponding to residues Leu50-Tyr70 was selected for MSMS analysis. Fragmentation patterns are indicated for the y ion series (upper) and the b ion series (lower). The results confirm that Ser54 and Thr57 remain unmodified and Cys60 has a +162 Da post-translational modification.
**Supplementary Fig. 7A. Modification of SunA C22S with glucose.** Modification of His₆-SunA C22S with glucose was attempted by incubation with His₆-SunS in the presence of UDP-Glc, MgCl₂, and TCEP in Tris (pH 7.5) buffer. The reaction was analyzed by MALDI-TOF MS. Expected [M+H] (with glycosylation): 8162.13. Expected [M+H] (without glycosylation): 7997.98, observed: 7999.85.

![Mass spectrum of modification of SunA C22S with glucose](image1)

**Supplementary Fig. 7B. Modification of SunA G21E with glucose.** Modification of His₆-SunA G21E with glucose was attempted by incubation with His₆-SunS in the presence of UDP-Glc, MgCl₂, and TCEP in Tris (pH 7.5) buffer. The reaction was analyzed by MALDI-TOF MS. Expected [M+H] (with one glycosylation): 8249.25. Expected [M+H] (without glycosylation): 8086.10, observed: 8245.90.

![Mass spectrum of modification of SunA G21E with glucose](image2)
Supplementary Fig. 7C. Modification of SunA G23E with glucose. Modification of His$_6$-SunA G23E with glucose was attempted by incubation with His$_6$-SunS in the presence of UDP-Glc, MgCl$_2$, and TCEP in Tris (pH 7.5) buffer. The reaction was analyzed by MALDI-TOF MS. Expected [M+H] (with one glycosylation): 8249.25. Expected [M+H] (without glycosylation): 8086.10, observed: 8251.70.

Supplementary Fig. 7D. Modification of SunA G23K with glucose. Modification of His$_6$-SunA G23K with glucose was attempted by incubation with His$_6$-SunS in the presence of UDP-Glc, MgCl$_2$, and TCEP in Tris (pH 7.5) buffer. The reaction was analyzed by MALDI-TOF MS. Expected [M+H] (with one glycosylation): 8248.31. Expected [M+H] (without glycosylation): 8085.16, observed: 8243.90 (100% rel. intensity), 8081.60 (50% rel. intensity).
Supplementary Fig. 7E. Modification of SunA Xa G21Q G23A with glucose. Modification of His6-SunA Xa G21Q G23A with glucose was attempted by incubation with His6-SunS in the presence of UDP-Glc, MgCl₂, and TCEP in Tris (pH 7.5) buffer. The reaction was analyzed by MALDI-TOF MS. Expected [M+H] (with one glycosylation): 8317.38. Expected [M+H] (without glycosylation): 8154.23, observed: 8330.27.
Supplementary Fig. 8A. MALDI-TOF MS analysis of base-catalyzed β-elimination of the sugar of sublancin. The sugar of sublancin was β-eliminated by treating native sublancin with NaBH₄ under basic conditions and analyzed by MALDI-TOF MS. Elimination to form an alkene followed by reduction of the alkene was observed. The overlay of mass spectra obtained by MALDI-TOF MS analysis of sublancin before (blue) and after treatment with NaBH₄ (red) is shown.

Supplementary Fig. 8B. Non-reductive elimination of the sugar of sublancin. The sugar of sublancin was β-eliminated by treating native sublancin with DBU and analyzed by MALDI-TOF MS. Elimination to form an alkene was observed. The overlay of mass spectra obtained by MALDI-TOF MS analysis of sublancin before (blue) and after treatment with DBU (magenta) is shown.
Supplementary Fig. 8C. ESI-Q/TOF MS analysis of base-catalyzed β-elimination of the sugar and chymotrypsin digest of sublancin. The sugar of sublancin was β-eliminated by treating native sublancin with NaBH₄ under basic conditions, digested with chymotrypsin under reducing conditions, and analyzed by LC-ESI-Q/TOF MS. The resulting mass spectrum is shown. Chymotrypsin cleavage sites (F, W, and Y residues) are indicated in the peptide sequence above the spectrum. The masses of the observed ions and their corresponding digest fragments are assigned. The results are consistent with Ser54 and Thr57 remaining unmodified and the chemical conversion from Cys22 bearing a +162 Da post-translational modification to form an Ala at position 22.
Supplementary Fig. 8D. ESI-Q/TOF MSMS analysis of residues Leu12-Tyr32 of base-catalyzed β-elimination of the sugar and chymotrypsin digest of sublancin. The sugar of sublancin was β-eliminated by treating native sublancin with NaBH₄ under basic conditions, digested with chymotrypsin under reducing conditions, and analyzed using LC-ESI-Q/TOF MSMS. The digest mixture was fractionated using LC and the multiply charged ion corresponding to residues Leu12-Tyr32 was selected for MSMS analysis. Fragmentation patterns are indicated for the y" ion series (upper) and the b ion series (lower). The results support that Ser16 and Thr19 remain unmodified and Cys22 has lost its post-translational modification and was converted to an Ala residue.
**Supplementary Figure 9.** (a) Expanded region of a one-dimensional $^1$H NMR spectrum of sublancin in 100% D$_2$O. (b) Two-dimensional NOESY spectrum and (c) and two-dimensional COSY spectrum. Shown are the putative resonances of the glucose ring bound to the peptide. Tentative assignments for the proton spin system are labeled in the NOESY and COSY spectra. All spectra were acquired at 500 MHz at 25° C. The spectra were processed with zero filling to double the original size and apodized with a sine bell function.
Supplementary Fig. 10A. Sequence alignments of sublancin precursor peptide with other peptides. Database mining of sequenced genomes revealed several clusters that could potentially harbor genes that encode for glycopeptide precursor peptides. These precursor peptide genes were found clustered near genes encoding for putative glycosyltransferase enzymes and were found among phylogenetically diverse bacteria. Fully conserved residues are in red font. Overall the sequence conservation is low, except for among the four variants in B. cereus where the conservation is moderate in the predicted core peptides (peptide following the predicted GG or GS protease cleavage site). All leader peptides are rich in Glu, as previously noted as a general observation in leader peptides involved in natural product peptide biosynthesis, see Oman and van der Donk, Nat. Chem. Biol. 6, 9-18 (2010). All cysteines are highlighted in green. Like sublancin, the precursor peptides from all organisms except B. cereus and E. faecalis contain five cysteine residues. The cysteine that is glycosylated in sublancin is indicated with an asterisk and is conserved in all peptides. The precursor peptides from B. cereus and E. faecalis have an even number of cysteine residues. If they are involved in two disulfides as in sublancin, then they cannot be glycosylated. The glycosyltransferases in their gene clusters may then function as O-glycosyltransferases.

| Sublancin precursor NP_390031 | 1 | ELKFKEVKLLELENQKG---------------------SGLGKAQCAALWL |
|--------------------------------|---|-----------------------------------------------------|
| B. cereus E33L YP_245531      | 1 | MNKFIRELKEELTKHVG---------------------GVIDGKNETVIHD |
| B. cereus E33L YP_245532      | 1 | MNKFIRELKEELTKHVG---------------------GVIDGKNETVIHD |
| B. cereus E33L YP_245533      | 1 | MNKFIRELKEELTKHVG---------------------GVIDGKNETVIHD |
| B. cereus E33L YP_245534      | 1 | MNKFIRELKEELTKHVG---------------------GVIDGKNETVIHD |
| B. thuringiensis BGSC 4AW1 ZP_04099938 | 1 | KELLIKELNLELTFEGYDGVNYMQHDDGGAGGGSGITACAYFKA |
| B. thuringiensis BGSC 4AW1 ZP_04099938 | 51 | LCYSGGSEWLGGYGSTQ-NNCELARKY |
| Plantaricin ASM1 precursor AB474371 | 1 | SSKLIKTLTVEISKIQTN---------------------GGKFAPDXTLAMS |
| E. faecalis HIP11704 ZP_05566853 | 1 | SSKLIKTLTVEISKIQTN---------------------GGKFAPDXTLAMS |
| Plantaricin ASM1 precursor AB474371 | 32 | CASGGT---------IGG-GGGAVAGMYFQFR-------- |
| E. faecalis HIP11704 ZP_05566853 | 31 | GGGAGPNKL-SLLNQILKNFKAASLFDQVYREWL |
| E. faecalis HIP11704 ZP_05566853 | 31 | GGGAGPNKL-SLLNQILKNFKAASLFDQVYREWL |
| E. faecalis HIP11704 ZP_05566853 | 31 | GGGAGPNKL-SLLNQILKNFKAASLFDQVYREWL |
| E. faecalis HIP11704 ZP_05566853 | 35 | IAASASTFFYGGSGTAMRY-VSEYIRKNS |
| Consensus                      | 51 | GGGAGPNL LGC KCLQFC

S44
Supplementary Fig. 10B. Sequence alignments of SunS with other putative glycosyltransferases. Database mining of sequenced genomes revealed several clusters containing a gene for a putative glycosyltransferase. These glycosyltransferase genes were found clustered near genes encoding for potential for glycopeptide precursor peptides. The putative glycosyltransferase enzymes were found among phylogenetically diverse bacteria. The asterisks (*) above the sequences indicate the putative divalent metal (Mg2+, Mn2+) binding site.
Consensus (301) W YFYARE I SE I LL IKN K

Consensus (351) LL IYIQ L I VLE P C DV YYK LN I

Consensus (401) QIF LQ K YS I SN HF LL LYPY YE AF

Consensus (451) QFF I D M DEIK LN KSNI

Consensus (301) W YFYARE I SE I LL IKN K

Consensus (351) LL IYIQ L I VLE P C DV YYK LN I

Consensus (401) QIF LQ K YS I SN HF LL LYPY YE AF

Consensus (451) QFF I D M DEIK LN KSNI
Supplementary Fig. 11A ESI-Q/TOF MS analysis of N-acetylglucosamine modified His<sub>6</sub>-SunA digested with chymotrypsin under reducing conditions. His<sub>6</sub>-SunA was modified with N-acetylglucosamine, digested with chymotrypsin under reducing conditions, and analyzed by using LC-ESI-Q/TOF MS. Chymotrypsin cleavage sites (F, W, and Y residues) are indicated in the peptide sequence above the spectrum. The masses of the observed ions and their corresponding digest fragments are assigned. The results are consistent with Ser54 and Thr57 remaining unmodified and residues Leu50-Tyr70 bearing a +203 Da post-translational modification. (upper: full spectrum, lower: zoom to peaks of interest).
Supplementary Fig. 11B. ESI-Q/TOF MSMS analysis of residues Leu50-Tyr70 of N-acetylglucosamine modified His6-SunA. His6-SunA was modified with N-acetylglucosamine, digested with chymotrypsin under reducing conditions, and analyzed by using LC-ESI-Q/TOF MSMS. The digest mixture was fractionated using LC and the multiply charged ion corresponding to residues Leu50-Tyr70 was selected for MSMS analysis. Fragmentation patterns are indicated for the \( y '' \) ion series (upper) and the \( b \) ion series (lower). The results confirm that Ser54 and Thr57 remain unmodified and Cys60 has a +203 Da post-translational modification.
**Supplementary Fig. 12A. Modification of SunA with galactose.** His₆-SunA was modified with galactose by incubation with His₆-SunS in the presence of UDP-Gal, MgCl₂, and TCEP in Tris (pH 7.5) buffer. The reaction was analyzed by MALDI-TOF MS. Expected [M+H] (with glycosylation): 8178.19, observed: 8175.28.

**Supplementary Fig. 12B. Modification of SunA with mannose.** His₆-SunA was modified with mannose by incubation with His₆-SunS in the presence of GDP-Man, MgCl₂, and TCEP in Tris (pH 7.5) buffer. The reaction was analyzed by MALDI-TOF MS. Expected [M+H] (with glycosylation): 8178.19, observed: 8176.09.
Supplementary Fig. 12C. Modification of SunA with N-acetyl galactosamine. His$_6$-SunA was modified with N-acetylglucosamine by incubation with His$_6$-SunS in the presence of UDP-GlcNAc, MgCl$_2$, and TCEP in Tris (pH 7.5) buffer. The reaction was analyzed by MALDI-TOF MS. Expected [M+H] (with glycosylation): 8220.23, observed: 8218.75.

Supplementary Fig. 12D. Modification of SunA with xylose. His$_6$-SunA was modified with xylose by incubation with His$_6$-SunS in the presence of UDP-Xyl, MgCl$_2$, and TCEP in Tris (pH 7.5) buffer. The reaction was analyzed by MALDI-TOF MS. Expected [M+H] (with glycosylation): 8148.16, observed: 8147.31.
Supplementary Fig. 12E. As a control reaction, His$_6$-SunA was by incubation with His$_6$-SunS in the presence of MgCl$_2$, and TCEP in Tris (pH 7.5) buffer, but in the absence of any NDP-sugar. The reaction was analyzed by MALDI-TOF MS. Expected [M+H] (without glycosylation): 8015.04, observed: 8015.76.

Supplementary Fig. 12F. As a control reaction without addition of reducing agent, His$_6$-SunA was by incubation with His$_6$-SunS in the presence of UDP-Glc and MgCl$_2$ in Tris (pH 7.5) buffer, but in the absence of TCEP. The reaction was analyzed by MALDI-TOF MS. No glycosylation was observed. Expected [M+H] (with glycosylation): 8178.19. Expected [M+H] (without glycosylation): 8015.04, observed: 8016.25.
Supplementary Fig. 13A-E. Kinetics of SunS glycosyltransferase activity using end-point assays. To identify the NDP-sugar substrate preference of SunS, His₆-SunA (50 µM) was modified with different NDP-sugars at varied NDP-sugar concentrations and products were analyzed by MALDI-TOF MS and quantified using HPLC. The glycosyltransferase reactions were run for 1 h, quenched, and the reaction mixture was analyzed by MALDI-TOF MS. The reaction mixture, containing unmodified starting material and glycosylated product, was fractionated by HPLC and relative amounts of unmodified and modified His₆-SunA peptide were quantified by peak integration. For each NDP-sugar tested, an overlay of MALDI-TOF mass spectra (upper) and an overlay of analytical HPLC chromatograms (lower) are shown.

Supplementary Fig. 13A. Kinetics of SunS glycosyltransferase activity with UDP-glucose.
Supplementary Fig. 13B. Kinetics of SunS glycosyltransferase activity with UDP-galactose.
Supplementary Fig. 13C. Kinetics of SunS glycosyltransferase activity with GDP-mannose.
Supplementary Fig. 13D. Kinetics of SunS glycosyltransferase activity with UDP-N-acetylglucosamine.
Supplementary Fig. 13E. Kinetics of SunS glycosyltransferase activity with UDP-xylose.
Supplementary Table 1. Summary of results for the kinetics of SunS glycosyltransferase activity with NDP-sugars. The relative amounts of sugar-modified and unmodified His<sub>c</sub>-SunA peptide was determined by quantifying the peak area corresponding to each form of the peptide by integration. Percent conversion was calculated by dividing the peak area of sugar-modified peptide by the sum of the sugar-modified and unmodified peak areas. As indicated by the table, the observed conversions suggest that SunS displays preference to NDP-sugars in order from most preferred to least preferred substrate:

UDP-Glc > UDP-Gal > GDP-Man > UDP-GlcNAc > UDP-Xyl

| Sugar  | Concentration of NDP-Sugar | Observed % Conversion |
|--------|---------------------------|-----------------------|
| Glc    | 5 mM                      | 95.2                  |
|        | 50 μM                     | 92.6                  |
|        | 500 nM                    | 11.0                  |
| Gal    | 5 mM                      | 91.6                  |
|        | 50 μM                     | 69.9                  |
|        | 500 nM                    | 6.2                   |
| Man    | 5 mM                      | 90.6                  |
|        | 50 μM                     | 54.9                  |
|        | 500 nM                    | 4.2                   |
| GlcNAc | 5 mM                      | 87.4                  |
|        | 50 μM                     | 57.9                  |
|        | 500 nM                    | 2.5                   |
| Xyl    | 5 mM                      | 69.4                  |
|        | 50 μM                     | 39.4                  |
|        | 500 nM                    | 1.4                   |
**Supplementary Fig. 14A.** Sequences of His₆-SunA and His₆-SunA Xa peptides. His₆ tag (red), leader peptide (blue), and core peptide (black) are indicated. Factor Xa cleavage site is underlined.

His₆-SunA  
\[\text{GSSHHHHHHSGLVPRGSHMEKLFKEVKLEALENGKGSGLGKAQCAALWLCASGGTIGCGGGAVACQNYRQFCR}\]

His₆-SunA Xa  
\[\text{GSSHHHHHHSGLVPRGSHMEKLFKEVKLEALENGKGSGLGKAQCAALWLCASGGTIGCGGGAVACQNYRQFCR}\]

**Supplementary Fig. 14B.** Factor Xa cleavage of SunA Xa precursor and purification of leader and core peptides. His₆-SunA Xa was digested with Factor Xa protease and analyzed by MALDI-TOF MS. The cleavage products (leader and core peptide) were then purified using reversed-phase HPLC and analyzed by MALDI-TOF MS.

The resulting mass spectrum of crude cleavage mixture is shown. Factor Xa cleavage occurred C-terminal to Arg38 resulting in leader peptide (Gly1-Arg38) and core peptide (Gly39-Arg75). Non-specific cleavage occurred C-terminal to Arg71 to produce a truncated core peptide (Gly39-Arg71). The masses of the observed ions and their corresponding digest fragments are assigned.
**Supplementary Fig. 14C.** The MALDI mass spectrum of HPLC purified leader peptide (Gly1-Arg38) is shown. Expected [M+H]: 4367.84, observed: 4369.14.

**Supplementary Fig. 14D.** SunA core peptide was modified with glucose by incubation with His6-SunS in the presence of UDP-Glc, MgCl₂, and TCEP in Tris (pH 7.5) buffer. The leader peptide was not added to the reaction. The reaction was analyzed by MALDI-TOF MS. The overlay of mass spectra obtained by MALDI-TOF MS analysis of the mixture before (black) and after SunS reaction (red) is shown. Expected [M+H] (with glycosylation): 3882.41, observed: 3883.91.
**Supplementary Fig. 14E.** SunA leader peptide was not modified with glucose by incubation with His$_6$-SunS in the presence of UDP-Glc, MgCl$_2$, and TCEP in Tris (pH 7.5) buffer. The core peptide was not added to the reaction. The reaction was analyzed by MALDI-TOF MS. The overlay of mass spectra obtained by MALDI-TOF MS analysis of the mixture before (black) and after SunS reaction (red) is shown. Expected [M+H] (without glycosylation): 4367.84, observed: 4368.81.
Supplementary Fig. 15. Antimicrobial activity assay of native, reduced, and reduced and alkylated sublancin. The importance of intact disulfides for the bioactivity of sublancin was assessed by testing the activities of native, reduced, and reduced and alkylated sublancin against *B. subtilis* ATCC 6633. Native sublancin samples were prepared by incubation in Tris buffer at pH 7.5 and pH 8.3. Reduced sublancin was prepared by incubation in Tris buffer (pH 8.3) in the presence of TCEP. Reduced and alkylated sublancin was prepared by incubation in Tris buffer (pH 8.3) in the presence of TCEP and iodoacetamide. All samples were analyzed by MS (Supplementary Fig. 1). The antimicrobial properties of each sample were assessed by its ability to inhibit the growth of *B. subtilis* ATCC 6633. Authentic sublancin and nisin standards were tested as positive controls. As indicated below, both samples of native sublancin (pH 7.5 and pH 8.3) retained the ability to inhibit the growth of the indicator strain. Both reduced and reduced and alkylated samples did not display antimicrobial activities.
Supplementary Fig. 16A. ESI-Q/TOF MS analysis of *in vitro* prepared sublancin (Glc). His<sub>6</sub>-SunA Xa was modified with glucose, digested with Factor Xa, oxidatively folded, and analyzed by using LC-ESI-Q/TOF MS. Expected [M+H]: 3876.74 Da, observed: 3876.7896 Da.

Supplementary Fig. 16B. ESI-Q/TOF MS analysis of *in vitro* prepared sublancin (Gal). His<sub>6</sub>-SunA Xa was modified with galactose, digested with Factor Xa, oxidatively folded, and analyzed by using LC-ESI-Q/TOF MS. Expected [M+H]: 3876.74 Da, observed: 3876.8430 Da.
Supplementary Fig. 16C. ESI-Q/TOF MS analysis of \textit{in vitro} prepared sublancin (Man). His\textsubscript{6}-SunA Xa was modified with mannose, digested with Factor Xa, oxidatively folded, and analyzed by using LC-ESI-Q/TOF MS. Expected [M+H]: 3876.74 Da, observed: 3876.7253 Da.

Supplementary Fig. 16D. ESI-Q/TOF MS analysis of \textit{in vitro} prepared sublancin (GlcNAc). His\textsubscript{6}-SunA Xa was modified with \textit{N}-acetylglucosamine, digested with Factor Xa, oxidatively folded, and analyzed by using LC-ESI-Q/TOF MS. Expected [M+H]: 3918.76 Da, observed: 3917.8679 Da.
Supplementary Fig. 16E. ESI-Q/TOF MS analysis of in vitro prepared sublancin (aglycon, no sugar). His$_6$-SunA Xa was unmodified, digested with Factor Xa, oxidatively folded, and analyzed by using LC-ESI-Q/TOF MS. Expected [M+H]: 3714.69 Da, observed: 3714.6812 Da.
Supplementary Fig. 17A-F. ESI-Q/TOF MSMS analysis of in vitro prepared sublancin and sublancin analogs. His\textsubscript{6}-SunA Xa was modified with sugar, digested with Factor Xa, oxidatively folded, and analyzed by using LC-ESI-Q/TOF MSMS. Material for all four sugars and aglycon material was analyzed. If disulfide bonds were correctly formed during the oxidative folding process, then fragmentation of the peptide backbone was not expected to occur. Rather, the post-translation modification (S-linked glycosylation) was expected to be lost from the parent ion (loss of 162 Da). As controls, an authentic sublancin sample (isolated and purified from \textit{B. subtilis} 168) and a reduced sublancin sample (authentic sublancin treated with TCEP) were independently analyzed using identical instrument settings as with \textit{in vitro} prepared samples.

Supplementary Fig. 17A. His\textsubscript{6}-SunA Xa was modified with glucose, digested with Factor Xa, oxidatively folded, and analyzed by using LC-ESI-Q/TOF MSMS. Fragmentation of the amide bonds of sublancin (Glc) did not occur, rather, the glucosylation was lost from the parent ion (-162 Da).
Supplementary Fig. 17B. His6-SunA Xa was modified with galactose, digested with Factor Xa, oxidatively folded, and analyzed by using LC-ESI-Q/TOF MSMS. Fragmentation of the amide bonds of sublancin (Gal) did not occur, rather, the galactosylation was lost from the parent ion (-162 Da).

Supplementary Fig. 17C. His6-SunA Xa was modified with mannose, digested with Factor Xa, oxidatively folded, and analyzed by using LC-ESI-Q/TOF MSMS. Fragmentation of the amide bonds of sublancin (Man) did not occur, rather, the mannosylation was lost from the parent ion (-162 Da).
Supplementary Fig. 17D. His<sub>6</sub>-SunA Xa was modified with N-acetylglucosamine, digested with Factor Xa, oxidatively folded, and analyzed by using LC-ESI-Q/TOF MSMS. Fragmentation of the amide bonds of sublancin (GlcNAc) did not occur, rather, the N-acetylglucosamininylation was lost from the parent ion (-203 Da).

Supplementary Fig. 17E. His<sub>6</sub>-SunA Xa was unmodified, digested with Factor Xa, oxidatively folded, and analyzed by using LC-ESI-Q/TOF MSMS. Fragmentation of the amide bonds of sublancin (aglycon) did not occur.
Supplementary Fig. 17F. As control, pure sublancin isolated from *B. subtilis* 168 was analyzed using LC-ESI-Q/TOF MSMS. Fragmentation of the amide bonds of sublancin did not occur, rather, the glycosylation was lost from the parent ion (-162 Da).
Supplementary Fig. 18A-E. ESI-Q/TOF MS analysis of in vitro prepared sublancin and sublancin analogs digested with chymotrypsin under non-reducing conditions. His6-SunA Xa was modified with sugar, digested with Factor Xa, oxidatively folded, and digested with chymotrypsin under non-reducing conditions. The reactions were analyzed by LC-ESI-Q/TOF MS to investigate extent of glycosylation and correct disulfide bridge connectivity. Chymotrypsin cleavage sites (F, W, and Y residues) and the proposed disulfide bridges are indicated in the peptide sequence above the spectrum. The masses of the observed ions and their corresponding digest fragments are shown. The results support that a disulfide bond exists between Cys7-Cys36 and between Cys14-Cys29. The results are consistent with residues Leu12-Tyr32 bearing a glycosylation in reactions which His6-SunA Xa peptide was modified with a sugar. Two masses are evident in all mass spectra of Figures 18A-E (m/z 838.3168 and 1685.8458), but further analysis indicates these are not derived from SunA. They may derive from chymotrypsin-degraded SunS, from chymotrypsin itself, or from Factor Xa or the leader peptide.

Supplementary Fig. 18A. ESI-Q/TOF MS analysis of in vitro prepared sublancin (Glc) digested with chymotrypsin under non-reducing conditions. His6-SunA Xa was modified with glucose, digested with Factor Xa, oxidatively folded, digested with chymotrypsin under non-reducing conditions, and analyzed by using LC-ESI-Q/TOF MS.
Supplementary Fig. 18B. ESI-Q/TOF MS analysis of *in vitro* prepared sublancin (Gal) digested with chymotrypsin under non-reducing conditions. His<sub>S</sub>-SunA Xa was modified with galactose, digested with Factor Xa, oxidatively folded, digested with chymotrypsin under non-reducing conditions, and analyzed by using LC-ESI-Q/TOF MS.
Supplementary Fig. 18C. ESI-Q/TOF MS analysis of *in vitro* prepared sublancin (Man) digested with chymotrypsin under non-reducing conditions. His6-SunA Xa was modified with mannose, digested with Factor Xa, oxidatively folded, digested with chymotrypsin under non-reducing conditions, and analyzed by using LC-ESI-Q/TOF MS.
Supplementary Fig. 18D. ESI-Q/TOF MS analysis of \textit{in vitro} prepared sublancin (GlcNAc) digested with chymotrypsin under non-reducing conditions. His$_e$-SunA Xa was modified with \textit{N}-acetylglucosamine, digested with Factor Xa, oxidatively folded, digested with chymotrypsin under non-reducing conditions, and analyzed by using LC-ESI-Q/TOF MS.
Supplementary Fig. 18E. ESI-Q/TOF MS analysis of in vitro prepared sublancin (aglycon) digested with chymotrypsin under non-reducing conditions. His6-SunA Xa was not modified with a sugar, digested with Factor Xa, oxidatively folded, digested with chymotrypsin under non-reducing conditions, and analyzed by using LC-ESI-Q/TOF MS.

Note: A minor amount of peptide products with different disulfide bond connectivities may form during the oxidative folding of the sublancin core peptide during the final step of the in vitro preparation process. However, these trace quantities were not detected in the tandem MS analyses (Supplementary Fig. 17), which would have likely lead to fragmentation of peptide amide bonds. Moreover, the chymotrypsin digests of each peptide resulting in at least two characteristic digest fragments that support proper disulfide connectivity. In each ESI-MS spectra, an ion is observed corresponding to Gly1-Trp11 and Cys36-Arg37 digest fragments connected via disulfide linkage between Cys7-Cys36. A second ion found in each spectra corresponds to the oxidized form of Leu12-Tyr32 in which a disulfide bond is formed between Cys14-Cys29. Although small amounts of peptide products with different disulfide bond connectivities may be present, it is highly unlikely that such minor products with non-native disulfide bond connectivities would provide the zones of inhibition observed in the antimicrobial activity assays used to evaluate in vitro prepared materials.
Supplementary Fig. 19. Antimicrobial activity assay of in vitro prepared sublancin with omission of single steps. The importance of each step (glycosylation, leader peptide removal, and disulfide formation) in the in vitro preparation of sublancin was demonstrated through an antimicrobial activity assay against B. subtilis ATCC 6633. Sublancin samples were prepared in parallel whereby one step of the preparation process was omitted. Glucose was used as the modifying sugar. The antimicrobial properties of each sample were assessed by its ability to inhibit the growth of B. subtilis ATCC 6633. Authentic sublancin standards were tested as positive controls. The steps that were used are indicated below. The sample in which all three steps (glycosylation, leader peptide removal, and disulfide formation) were used displayed a zone of inhibition (far left). The samples which did not undergo glucosylation (middle left), had the leader peptide intact (middle right), and did not undergo oxidative folding (far right) did not show inhibitory activity.
Supplementary Fig. 20A-C. Attempts to produce xylose modified sublancin and subtilosin

A. Dorenbos et al. reported the production of sublancin and subtilosin in which additional xylose molecules were reported to be attached to the bacteriocins. To determine where the xylose might be attached, *Bacillus subtilis* 168 was cultured in xylose rich media in an effort to produce xylose modified sublancin and subtilosin.

Supplementary Fig. 20A. MALDI-TOF MS analysis of sublancin and subtilosin A production in TY medium with 1% xylose. *B. subtilis* 168 was grown in TY medium with 1% xylose and in Medium X and the culture supernatants were analyzed by MALDI-TOF MS. Medium A (culture medium used to produce sublancin in this study) was used as a positive control. Both sublancin and subtilosin A peptides were produced by *B. subtilis* 168 in TY medium with 1% xylose but no additional xylose modifications were observed. Sublancin expected [M+H]: 3876.74, observed: 3882.68. Subtilosin A expected [M+H]: 3399.90, observed: 3407.36.
Supplementary Fig. 20B. MALDI-TOF MS analysis of sublancin and subtilosin A production in Medium X. Both sublancin and subtilosin A peptides were produced by *B. subtilis* 168 in Medium X but no additional xylose modifications were observed. Sublancin expected [M+H]: 3876.74, observed: 3882.02. Subtilosin A expected [M+H]: 3399.90, observed: 3406.43.

Supplementary Fig. 20C. MALDI-TOF MS analysis of sublancin and subtilosin A production in Medium A. Both sublancin and subtilosin A peptides were produced by *B. subtilis* 168 in Medium A with no additional xylose modifications observed. Sublancin expected [M+H]: 3876.74, observed: 3883.66. Subtilosin A expected [M+H]: 3399.90, observed: 3406.10.
Supplementary Fig. 21. Attempt to enzymatically install glucose and/or xylose modification of SunA prepeptide by SunS. Modification of His6-SunA with glucose and/or xylose was attempted by incubation with His6-SunS in the presence of UDP-Glc, UDP-Xyl, MgCl₂, and TCEP in Tris (pH 7.5) buffer. The reaction was analyzed by MALDI-TOF MS.

Expected [M+H] (with glucosylation only): 8178.19
Expected [M+H] (with xylosylation only): 8148.16
Expected [M+H] (with glucosylation and xylosylation): 8311.31
Observed: 8181.95.

Only a single glucosylation occurred and importantly, no xylose modification was observed in addition to or in substitution of glucose.
Supplementary Fig. 22. Attempt to enzymatically install xylose modification on SunA peptide previously modified with glucose at C22 by SunS. Modification of His6-SunA-Glc (His6-SunA previously modified with a glucose at C22 by SunS and subsequently purified by HPLC) with xylose was attempted by incubation with His6-SunS in the presence of UDP-Xyl, MgCl₂, and TCEP in Tris (pH 7.5) buffer. The reaction was analyzed by MALDI-TOF MS.

Expected [M+H] (with glucosylation only): 8178.19
Expected [M+H] (with glucosylation and xylosylation): 8311.31
Observed: 8174.78