Published in final edited form as:

ACS Chem Biol. 2018 October 19; 13(10): 2966–2972. doi:10.1021/acscchembio.8b00599.

Rapid Discovery of Glycocins through Pathway Refactoring in *Escherichia coli*

Hengqian Ren¹, Subhanip Biswas², Sherri Ho³, Wilfred A. van der Donk^{2,5,6,*}, and Huimin Zhao^{1,2,4,6,*}

¹Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801

²Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801

³Department of Molecular and Cellular Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801

⁴Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801

⁵Howard Hughes Medical Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801

⁶Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801

Abstract

Glycocins (glycosylated bacteriocins) are a family of ribosomally synthesized and post-translationally modified peptides (RiPPs) with antimicrobial activities against pathogens of interest including methicillin-resistant *Staphylococcus aureus* (MRSA), representing a promising source of new antibiotics. Glycocins are still largely underexplored and thus far only six glycocins are known. Here we used genome mining to identify 50 putative glycocin biosynthetic gene clusters and then chose six of them with distinct features for further investigation. Through two rounds of plug-and-play pathway refactoring and expression in *Escherichia coli* BL21(DE3), four systems produced novel glycocins. Further structural characterization revealed that one of them, which belongs to the enterocin 96-type glycocins, was di-glucosylated on a single serine. The other three compounds belong to the SunA/ThuA-type glycocins and exhibit a narrower antimicrobial spectrum than sublancin, the best characterized member in this group, even though they share a similar disulfide topology and glycosylation. Further evaluation of their bioactivities with free glucose at high concentrations suggested that their antimicrobial mechanisms might be both glycocin- and species-specific. These glycocins with distinct features significantly broaden our knowledge on glycocins and may lead to the discovery of new classes of antibiotics.

*To whom correspondence should be addressed. Phone: (217) 333-2631. Fax: (217) 333-5052. zhao5@illinois.edu; wav@illinois.edu.
Author Contributions
H.R., W.A.V., and H.Z. designed the experiments and wrote the manuscript. H.R. performed the pathway refactoring and discovered the new glycocins. H.R. and S.B. contributed to the structure and bioactivity characterization. S.H. helped with molecular cloning and product purification.

ASSOCIATED CONTENT

Supporting Information

This material is available free of charge via the internet at <http://pubs.acs.org>.

Keywords

natural products; synthetic biology; pathway refactoring; glycocin

INTRODUCTION

Ribosomally synthesized and post-translationally modified peptides (RiPPs) constitute an emerging family of natural products with various structural features and bioactivities.¹ Glycocins (glycosylated bacteriocins) are a group of RiPPs that feature sugar moieties installed on serine, threonine, or cysteine residues and two nested disulfide bonds that stabilize their helix-loop-helix structures. A typical biosynthetic pathway for glycocin involves five genes. The gene encoding the precursor peptide is expressed first and the resulting precursor peptide is then glycosylated by a glycosyltransferase. Two disulfide oxidoreductases are responsible for forming the disulfide bonds, and a bifunctional transporter transfers the peptide outside the cell and removes the leader peptide.² Four groups of glycocins with distinct scaffolds have been discovered to date, with the prototypes being glycocin F (GccF-type), sublancin (SunA/ThuA type), enterocin 96 (enterocin 96-type), and enterocin F4–9 (enterocin F4–9-type) (Figure S1b).²

Owing to recent advances in next-generation sequencing and bioinformatics, many putative glycocin biosynthetic gene clusters (BGCs) have been identified through genome mining. However, only six glycocins have been discovered so far: five isolated from native producers and one by chemoenzymatic synthesis.² This dearth of characterized glycocins may be attributed to our limited knowledge of glycocins and technical challenges in product isolation from the native producers, particularly pathogens and extremophiles. All six known glycocins show antimicrobial bioactivities against several Gram-positive pathogens such as *Bacillus cereus*, *Streptococcus pyogenes*, and *Staphylococcus aureus*. One of them, sublancin, was also active against drug resistant strains such as methicillin-resistant *S. aureus* (MRSA) and gentamicin-resistant *Enterococcus faecalis*.^{3–5} Therefore, glycocins appear to be a promising source for drug lead discovery^{3, 6} and it is desirable to discover novel glycocins.

In this work, we sought to use a synthetic biology approach based on the pathway refactoring strategy we previously developed⁷ to rapidly discover new glycocins. RODEO (Rapid ORF Description and Evaluation Online),⁸ a powerful genome mining tool for RiPPs, was used for predicting putative glycocin BGCs. According to their distinct features, six putative glycocin BGCs were chosen for pathway refactoring in *Escherichia coli*. Four of them successfully produced new glycocins, which were named bacillicin CER074, bacillicin BAG20, geocillicin, and listeriocytocin, respectively. These glycocin BGCs originate from various species with most of them either pathogens or extremophiles, but our synthetic biology strategy involving DNA synthesis and pathway refactoring in *E. coli* has made the glycocin discovery process rapid and scalable.

RESULTS AND DISCUSSION

Design of a Modified Pathway Refactoring Workflow

Previously we developed a plug-and-play pathway refactoring workflow for rapid discovery of natural products.⁷ This workflow consists of two-step Golden Gate reactions that first convert biosynthetic genes into individual expression cassettes and then assemble these expression cassettes into one plasmid. In this work, we adapted this workflow for discovery of new glycocins (Figure 1). Briefly, the target glycocin BGC was first partially refactored, which involves only the peptide precursor gene and the glycosyltransferase gene that are required to generate a glycosylated intermediate (first tier).⁹ If a compound corresponding to the expected intermediate was detected, the corresponding BGC was then fully refactored to generate the final product (second tier). *E. coli* BL21(DE3) was used as a heterologous expression host and the BGC genes were codon optimized for *E. coli* and obtained through DNA synthesis. For the refactored partial glycocin BGC, after overnight culture in 1 mL media under various temperatures, the resulting cell pellets were obtained by centrifugation and homogenized in an ultrasonic bath. For the refactored full glycocin BGC, supernatants instead of cell pellets were used for next step analysis. Both the supernatants and the cell lysates were desalted and analyzed by MALDI-TOF MS (Matrix Assisted Laser Desorption/Ionization-Time-of-Flight Mass Spectrometry). While this workflow was designed for the discovery of glycocins, it is also potentially applicable for high-throughput discovery of other RiPP families in the future.

Refactoring of Glycocin BGCs for Heterologous Expression in *E. coli*

Putative glycocin BGCs were predicted by RODEO. The amino acid sequence of the glycosyltransferase for sublancin (SunS) was used as the query for BLAST and the top 5000 hits were used as the input for RODEO, which analyzed the neighboring genes of each input gene in batch through comparison against the Pfam pHMM database. All putative glycosyltransferases with their neighboring genes were then filtered by the typical features of the glycocin BGCs as described in the Materials and Methods. A sequence similarity network was then generated for the putative precursor peptides and six BGC candidates for refactoring were mainly chosen from different groups in this network (Figure S1a). Two out of the six glycocin BGCs which according to the sequence of the precursor peptide belong to the same group as the BGC from *Bacillus megaterium* BHG1.1, were chosen because of their extremophile origin (thermophilic strain, *Geobacillus* sp. 8 and alkaliphilic strain, *Bacillus* sp. JCM19047), which may offer better stability of the product glycocins under certain conditions for future applications (Figure 2 and Figure S1a).^{10–14}

The first tier pathway refactoring was performed as indicated in Figure 1B. Both the precursor peptide and glycosyltransferase were expressed as individual expression cassettes with a T7 promoter and T7 terminator. After the first tier pathway refactoring process, three out of the six partial glycocin BGCs expressed in *E. coli* generated peaks in the MALDI-TOF mass spectrum with molecular weights corresponding to the monoglycosylated precursor peptides (Figure S2a). Based on the mass increase compared to the non-glycosylated precursor peptide (162 Da), all the installed sugars were hexoses. The partial glycocin BGC from *Listeria monocytogenes* generated a new peak in the mass spectrum

consistent with a diglycosylated precursor peptide with an N-terminal truncation. To confirm this result, we introduced site-specific mutations (K7A and M8A) at the putative cleavage sites on its precursor peptide and coexpressed the mutant precursor gene with the glycosyltransferase (Figure 2b). The expected full length mutant precursor peptide with two hexoses installed was indeed observed in the MALDI-TOF mass spectrum (Figure S2b). Given the successful glycosylations, all four glycocin BGCs were then fully refactored with every biosynthetic gene expressed as an individual expression cassette with a T7 promoter and T7 terminator (Figure 1B). Of note, for the glycocin BGC from *Listeria monocytogenes*, the mutant precursor peptide gene was used in pathway refactoring. New peaks in the mass spectrum corresponding to the glycosylated precursor peptides with their leader peptides removed at the expected sites (Figures 2b and S4) were observed from the desalted supernatants. The resulting four glycocins were named bacillicin CER074, bacillicin BAG20, geocillicin, and listeriocytocin. According to the classification by Norris *et al.*,² bacillicin CER074, bacillicin BAG20 and geocillicin are SunA/ThuA-type glycocins,^{10, 13} whereas listeriocytocin is a member of the enterocin 96-type glycocins (Figure S1c).¹²

Isolation and Structural Characterization of the Glycocins

In order to obtain sufficient amounts of glycocins for structure characterization and bioassays (at least 40 µg for each glycocin for both MS based structural characterization and standard agar diffusion assay for bioactivity screening), we further optimized cultivation conditions and purified the glycocins through ammonium sulfate precipitation, solid phase extraction, semiprep HPLC, and analytical HPLC (Figure 3b). The existence of disulfide bonds was first confirmed by an iodoacetamide (IAA) assay that demonstrated the absence of free thiols (Figure S5).⁹ The purified glycocins were then digested by chymotrypsin under non-reducing and reducing conditions and analyzed by MS. Through assignment of the observed fragments, we concluded that all four glycocins shared similar topologies (Figure S6). Although the glycosylated residues for the three SunA/ThuA type glycocins can be easily assigned to the third cysteine by the results of iodoacetamide derivatization and proteolytic digest, MS/MS analysis was necessary to determine the glycosylated residue in listeriocytocin. Through sequence alignment, Ser60 in the precursor peptide of listeriocytocin was predicted to be glycosylated (Figure 2b). MS/MS data unambiguously confirmed this prediction and showed that two successive glycosylations occur on Ser60 (Figure 4, Figure S7, and Figure S8).

To further determine the type of hexoses installed in the four glycocins, the glycosylated precursor peptides were purified from the corresponding partial BGCs, which resulted in a higher yield than refactoring of the full BGCs (Figure S3). Notably, in the first tier pathway refactoring, the precursor genes were cloned into a helper plasmid with an internal N-terminal His-Tag at the cloning site (Figure S11), which enables the purification of the glycosylated intermediates by immobilized metal affinity chromatography (IMAC). These intermediates were then further purified by reversed-phase high-performance liquid chromatography (RP-HPLC) and hydrolyzed. The sugars were derivatized and analyzed using gas chromatography as described elsewhere.^{9, 15} In comparison with various hexose standards, all the installed sugars on these four glycocins were shown to be glucoses (Figure 3D). Previously, two glycosyltransferases for glycocin biosynthesis have been characterized

in vitro and exhibited promiscuous activity towards different nucleotide sugar substrates.^{15, 16} Considering that the concentrations of some potential nucleotide sugar substrates, such as UDP-N-acetylglucosamine, are much higher than UDP-glucose in *E. coli*,^{17, 18} it appears that the glycosyltransferases under investigation here have a preference for UDP-glucose. To further investigate if glycosyltransferases for glycocin biosynthesis exhibit high specificity for their nucleotide sugar substrates during heterologous production in *E. coli*, we also coexpressed the precursor peptide of enterocin F4–9 with its glycosyltransferase in BL21(DE3). A mass representing the peptide modified with two N-acetylglucosamines was clearly observed by MALDI-TOF MS (Figure S12). Hence, glycocins with their native glycosylation can be produced in *E. coli*.

All three SunA/ThuA-type glycocins shared similar topology and sugar modification as sublancin (Figure S6), suggesting these structural features are conserved in this glycocin family. Although enterocin 96 was the first discovered enterocin 96-type glycocin, its structure has yet to be elucidated.¹⁹ Recently the glycosyltransferase of enterocin 96 was reconstituted *in vitro* and the modification was shown to be a disaccharide.¹⁶ As we discovered that the listericytotoxin was also diglycosylated at the same serine residue predicted to be glycosylated in enterocin 96 by sequence alignment, it is very likely that the wild type enterocin 96 shares a similar structure as that produced *in vitro* and the type of sugar modification might be glucose. The diglycosylation might be a distinct feature of enterocin 96-type glycocins from other diglycosylated glycocins, such as enterocin F4–9 and glycocin F.^{20, 21} Listericytotoxin does have some distinct structural features compared to enterocin 96, such as a much smaller loop size (Figure S1c).

Antimicrobial Activity of the Glycocins

The antimicrobial activity of the four glycocins was probed against a panel of Gram-positive and Gram-negative bacteria (*Escherichia coli* JM109, *Pseudomonas aeruginosa* PA01, Methicillin-resistant *Streptococcus aureus* NRS384/USA300, Vancomycin-resistant *Enterococcus faecium* V583 CB807, *Listeria monocytogenes* 4b F2365, *Bacillus subtilis* sp β , *Bacillus halodurans* C-125 and *Bacillus cereus* ATCC14579) by a standard agar diffusion assay (Table S1). Although listericytotoxin lacked antimicrobial activity toward all organisms tested, bacillicin CER074 and bacillicin BAG2O exhibited antimicrobial activity against *Bacillus cereus* ATCC 14579 (minimal inhibitory concentrations: sublancin, 2.5 μ M; bacillicin CER074, 0.156 μ M; bacillicin BAG2O, 0.625 μ M), and geocillicin also exhibited weak bioactivity (Figure 5). Overall, the three newly discovered SunA/ThuA-type glycocins have narrower antimicrobial spectrum than sublancin with all of them inhibiting the growth of *B. cereus* ATCC14579, but not other bacilli tested (Figure S9 and Figure S10).

As the bioactivity of glycocin F, which is a glycocin decorated with two *N*-acetylglucosamine (GlcNAc) moieties, can be reversed by free GlcNAc,²² we tested if free glucose can have a similar effect on the antimicrobial activities of the four new glycocins. While the bioactivities of sublancin and geocillicin could be reversed by free glucose, no significant effects were observed for bacillicin CER074 and bacillicin BAG2O, which means these two glycocins might have different target(s) from sublancin and geocillicin (Figure 5). Previous studies showed that sublancin's activity is mediated by the

phosphoenolpyruvate:sugar phosphotransferase system (PTS) and the large mechanosensitive channel of conductance MscL.^{23, 24} In structure-activity relationship (SAR) studies, Arg33 of sublancin was previously shown to be critical for its bioactivity against *B. subtilis* σ^B and most of the SunA/ThuA-type glycocin are positively charged at the corresponding site in a sequence alignment to sublancin.⁹ However, bioactivity was also observed for bacillicin BAG20 that does not contain a corresponding charged residue, which may imply that bacillicin BAG20 has a different target(s) from sublancin. Therefore, it would be of great interest to further investigate the antimicrobial mechanisms of these new glycocins, and our production platform in *E. coli* provides the opportunity to perform further SAR studies.

CONCLUSION

In this work, we used a rapid pathway refactoring workflow to discover four novel glycocins from various origins, which has significantly expanded the glycocin family. This workflow is potentially scalable, which can be very helpful for RiPP discovery in the future. We provide evidence that di-glycosylation might be a general feature for enterocin 96-type glycocins, and the bioactivity assays demonstrate that factors besides the topology and type of glycosylation may affect the bioactivities of glycocins, which is of great interest for further investigation of structure-function relationships.

METHODS

Materials and Reagents.

All chemicals were purchased from Sigma-Aldrich and Thermo Fisher Scientific. *E. coli* NEB® 10-beta was used for DNA manipulation and *E. coli* BL21(DE3) was used as a host for expression. Genes were obtained as gBlocks® from Integrated DNA Technologies. *Bsa*I, T4 DNA ligase and other restriction enzymes for checking the assembly of plasmids were obtained from New England Biolabs, while the *Bp*I was obtained from Thermo Fisher Scientific and the T7 DNA ligase was obtained from Enzymatics. Plasmids were isolated using QIAGEN Plasmid Mini Kit. Kanamycin was used at 50 µg/ml for all *E. coli* cultures. ZipTip® C18 was purchased from EMD Millipore. All sequencing grade peptidases were obtained from Promega.

Genome Mining by RODEO.

The amino acid sequence of SunS (the glycosyltransferase for the biosynthesis of sublancin) was used as the query for BLASTP, and the first 5000 hits were used as the input for the analysis by RODEO. Each predicted gene cluster was filtered by the existence of all genes which are necessary for glycocin biosynthesis (genes encoding SunA-like precursor, SunS-like glycosyltransferase, disulfide bond formation protein/thioredoxin, SunT-like transporter). The gene encoding a peptide less than 100 aa with at least four cysteines was considered as the putative precursor gene.

Pathway Refactoring through the Plug-and-Play Workflow.

The pathway refactoring process followed the protocol we described before with slight modifications⁷. In the partial pathway refactoring, a new helper plasmid (T7 His Helper 1) with an internal N-terminal His-tag at the cloning site was used in the first tier Golden Gate reaction for cloning the precursor gene, while the glycosyltransferase gene was cloned into the T7 Helper 2/9 plasmid. In the full pathway refactoring, the precursor gene, the glycosyltransferase gene, the disulfide bond formation protein/thioredoxin and the transporter gene were cloned into T7 Helper 1, T7 Helper 2, T7 Helper 3, T7 Helper 4, and T7 Helper 5/9 respectively. For the listeriocytocin BGC, two transporter genes (Figure 2) were cloned into T7 Helper 5 and T7 6/9, respectively.

Detection of the Glycosylated Precursor Peptide from Cell Pellets and Glycocin from Liquid Medium by MALDI-TOF MS.

For the partial refactored BGCs, cells were grown in LB medium and IPTG was added to 0.5 mM when the OD600 reached 0.6–0.8. Cells were grown with shaking at 18 °C, 30 °C and 37 °C in parallel for overnight induction. Cell pellets were harvested from 1 mL of liquid culture by centrifugation and resuspended in 100 µL of cell lysis buffer (8 M urea, 50 mM Tris • HCl, pH 7.5), which were then vigorously vortexed three times (30 s each time) and sonicated in ultrasonic bath (Branson 2800, Danbury, CT) for at least 30 min. The supernatants were then collected by centrifuge, desalted by zip tip and analyzed using MALDI-TOF MS. For the fully refactored BGCs, the supernatants of liquid culture were collected by centrifugation after overnight induction, desalted by zip tip and analyzed using MALDI-TOF MS. As rich media are too complex, which largely affect the product detection from the supernatant possibly due to ion suppression (data not shown), minimal medium M9 was used in the full BGC refactoring.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health (GM077596) (HZ and WAV) and the Howard Hughes Medical Institute (WAV). We thank A. Ulanov in the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign for assistance with the GC-MS analysis. We thank P. Yau in the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign for assistance with the MS/MS analysis.

References

1. Arnison PG, Bibb MJ, Bierbaum G, Bowers AA, Bugni TS, Bulaj G, Camarero JA, Campopiano DJ, Challis GL, Clardy J, Cotter PD, Craik DJ, Dawson M, Dittmann E, Donadio S, Dorrestein PC, Entian KD, Fischbach MA, Garavelli JS, Goransson U, Gruber CW, Haft DH, Hemscheidt TK, Hertweck C, Hill C, Horswill AR, Jaspars M, Kelly WL, Klinman JP, Kuipers OP, Link AJ, Liu W, Marahiel MA, Mitchell DA, Moll GN, Moore BS, Muller R, Nair SK, Nes IF, Norris GE, Olivera BM, Onaka H, Patchett ML, Piel J, Reaney MJT, Rebuffat S, Ross RP, Sahl HG, Schmidt EW, Selsted ME, Severinov K, Shen B, Sivonen K, Smith L, Stein T, Sussmuth RD, Tagg JR, Tang GL, Truman AW, Vederas JC, Walsh CT, Walton JD, Wenzel SC, Willey JM, and van der Donk WA (2013) Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature, *Nat Prod Rep* 30, 1568–1568.

2. Norris GE, and Patchett ML (2016) The glycocins: in a class of their own, *Curr Opin Struc Biol* 40, 112–119.
3. Stein T (2005) *Bacillus subtilis* antibiotics: structures, syntheses and specific functions, *Mol Microbiol* 56, 845–857. [PubMed: 15853875]
4. Wang S, Wang QW, Zeng XF, Ye QH, Huang S, Yu HT, Yang TR, and Qiao SY (2017) Use of the antimicrobial peptide sublancin with combined antibacterial and immunomodulatory activities to protect against methicillin-resistant *Staphylococcus aureus* infection in mice, *J Agr Food Chem* 65, 8595–8605. [PubMed: 28906115]
5. Wang S, Zeng XF, Wang QW, Zhu JL, Peng Q, Hou CL, Thacker P, and Qiao SY (2015) The antimicrobial peptide sublancin ameliorates necrotic enteritis induced by *Clostridium perfringens* in broilers, *J Anim Sci* 93, 4750–4760. [PubMed: 26523568]
6. Paik SH, Chakicherla A, and Hansen JN (1998) Identification and characterization of the structural and transporter genes for, and the chemical and biological properties of, sublancin 168, a novel lantibiotic produced by *Bacillus subtilis* 168, *J Biol Chem* 273, 23134–23142. [PubMed: 9722542]
7. Ren H, Hu P, and Zhao H (2017) A plug-and-play pathway refactoring workflow for natural product research in *Escherichia coli* and *Saccharomyces cerevisiae*, *Biotechnol Bioeng* 114, 1847–1854. [PubMed: 28401530]
8. Tietz JJ, Schwalen CJ, Patel PS, Maxson T, Blair PM, Tai HC, Zakai UI, and Mitchell DA (2017) A new genome-mining tool redefines the lasso peptide biosynthetic landscape, *Nat Chem Biol* 13, 470–478. [PubMed: 28244986]
9. Biswas S, De Gonzalo CVG, Repka LM, and van der Donk WA (2017) Structure-activity relationships of the S-Linked glycocin sublancin, *ACS Chem Biol* 12, 2965–2969. [PubMed: 29112373]
10. Kaunietis A, de Jong A, Pranckute R, Buivydas A, and Kuipers OP (2016) Draft genome sequences of two *Geobacillus* species strains, isolated from oil wells and surface soil above oil pools, *Genome Announc* 4, 01129–01116.
11. Kudo T, Sakamoto K, Akinaga M, Kawauchi A, Nakahara T, Zhang X, Yamada A, Oshima K, Suda W, Kuwahara H, Nakamura N, Nogi Y, Kitamura K, Yuki M, Iida T, Moriya S, Inoue T, Hongoh Y, Hattori M, and Ohkuma M (2014) Draft genome sequences of cyclodextrin-producing alkaliphilic *Bacillus* strains JCM 19045, JCM 19046, and JCM 19047, *Genome Announc* 2, e00211–00214. [PubMed: 24652985]
12. Kuenne C, Billion A, Abu Mraheil M, Strittmatter A, Daniel R, Goesmann A, Barbuddhe S, Hain T, and Chakraborty T (2013) Reassessment of the *Listeria monocytogenes* pan-genome reveals dynamic integration hotspots and mobile genetic elements as major components of the accessory genome, *BMC Genomics* 14, 47. [PubMed: 23339658]
13. Van der Auwera GA, Feldgarden M, Kolter R, and Mahillon J (2013) Whole-genome sequences of 94 environmental isolates of *Bacillus cereus* Ssensu Lato, *Genome Announc* 1, e00380–00313.
14. Wang W, Zheng SS, Sun H, Cao J, Yang F, Wang XL, and Li LX (2016) Draft genome sequence of *Bacillus megaterium* BHG1.1, a strain isolated from bar-headed goose (*Anser indicus*) feces on the Qinghai-Tibet plateau, *Genome Announc* 4, e00317–00316.
15. Oman TJ, Boettcher JM, Wang HA, Okalibe XN, and van der Donk WA (2011) Sublancin is not a lantibiotic but an S-linked glycopeptide, *Nat Chem Biol* 7, 78–80. [PubMed: 21196935]
16. Nagar R, and Rao A (2017) An iterative glycosyltransferase EntS catalyzes transfer and extension of O- and S-linked monosaccharide in enterocin 96, *Glycobiology* 27, 766–776.
17. Guo AC, Jewison T, Wilson M, Liu YF, Knox C, Djoumbou Y, Lo P, Mandal R, Krishnamurthy R, and Wishart DS (2013) ECMDB: the *E. coli* metabolome database, *Nucleic Acids Res* 41, 625–630.
18. Sajed T, Marcu A, Ramirez M, Pon A, Guo AC, Knox C, Wilson M, Grant JR, Djoumbou Y, and Wishart DS (2016) ECMDB 2.0: A richer resource for understanding the biochemistry of *E. coli*, *Nucleic Acids Res* 44, 495–501.
19. Izquierdo E, Wagner C, Marchioni E, Aoude-Werner D, and Ennahar S (2009) Enterocin 96, a novel class II bacteriocin produced by *Enterococcus faecalis* WHE 96, isolated from munster cheese, *Appl Environ Microb* 75, 4273–4276.

20. Maky MA, Ishibashi N, Zendo T, Perez RH, Doud JR, Karmi M, and Sonomoto K (2015) Enterocin F4–9, a novel O-linked glycosylated bacteriocin, *Appl Environ Microb* 81, 4819–4826.
21. Stepper J, Shastri S, Loo TS, Preston JC, Novak P, Man P, Moore CH, Havlicek V, Patchett ML, and Norris GE (2011) Cysteine S-glycosylation, a new post-translational modification found in glycopeptide bacteriocins, *FEBS Lett* 585, 645–650. [PubMed: 21251913]
22. Bisset SW, Yang SH, Amso Z, Harris PWR, Patchett ML, Brimble MA, and Norris GE (2018) Using chemical synthesis to probe structure-activity relationships of the glycoactive bacteriocin glycocin F, *ACS Chem Biol* 13, 1270–1278. [PubMed: 29701461]
23. De Gonzalo CVG, Denham EL, Mars RAT, Stulke J, van der Donk WA, and van Dijk JM (2015) The phosphoenolpyruvate:sugar phosphotransferase system Is involved in sensitivity to the glucosylated bacteriocin sublancin, *Antimicrob Agents Ch* 59, 6844–6854.
24. Kouwen TRHM, Trip EN, Denham EL, Sibbald MJJB, Dubois JYF, and van Dijk JM (2009) The large mechanosensitive channel MscL determines bacterial susceptibility to the bacteriocin sublancin 168, *Antimicrob Agents Ch* 53, 4702–4711.

Figure 1.

The workflow for rapid discovery of glycocins. A) The glycocin BGCs were predicted from genomic sequences that are available in NCBI (National Center for Biotechnology Information) by RODEO, and genes for the glycocin biosynthesis were annotated. B) The candidate BGCs were then refactored using a two-tier strategy and the resulting products were detected using MALDI-TOF MS.

A **B** **Figure 2.**

A) The selected putative glycocin BGCs. B) Alignment of the putative precursor peptides to the precursor peptide of sublancin, a known glycocin previously discovered from *Bacillus subtilis* 168. The red arrow indicates the sites where leader peptides are predicted to be cleaved. Cysteines that are predicted to form disulfide bonds are highlighted in yellow. Residues that are predicted to be glycosylated are highlighted in purple. The mutated residues of the precursor peptide from *Listeria monocytogenes* SLCC2540 BGC to prevent proteolysis are highlighted in green. Both of the residues were mutated to alanines.

Figure 3.

Structures of the identified glycocins. A) Structures of the three SunA/ThuA type glycocins discovered in this work. Disulfide bonds are labeled in yellow while the glycosylated cysteines are labeled in purple. B) The 220 nm trace of the analytical HPLC purification. Peaks containing target glycocins are highlighted in yellow. C) MALDI-TOF MS analysis of the purified glycocins (bacillicin CER074: observed 4584.861 Da, expected 4584.057 Da; geocillicin: observed 4223.228 Da, expected 4222.822 Da; bacillicin BAG20: observed 3650.761 Da, expected 3650.159 Da; listeriocytocin: observed 4667.945 Da, expected 4667.390 Da). D) GC traces of the hydrolyzed and derivatized glycosylated peptides. Traces for hexose standards (galactose, glucose and mannose) were obtained by the same procedure and are shown for comparison.

Figure 4.

Structure of listeriocytocin and MS/MS analysis to determine the site(s) of glycosylation. Cysteine residues that form disulfide bonds are highlighted in yellow and the diglycosylated serine residue is highlighted in purple. The target fragment (underlined with dashed line) for MS/MS analysis was generated by Asp-N endoproteinase under reducing condition. Dehydrated b-ions (−18 Da) are also labeled in the spectrum.

Figure 5.

Activity of glycocins against *Bacillus cereus* ATCC14579. Samples were spotted on LB agar in a volume of 20 μ L a-e: 20 μ M glycocin. f-j: 20 μ M glycocin and 500 mM glucose. Glycocins tested: (a,f) sublancin, (b,g) bacillicin CER074, (c,h) bacillicin BAG2O, (d,i) geocillicin, and (e,j) listeriocytocin. k: 500 mM, glucose. l: water.