Published in final edited form as:

Nat Chem Biol. 2011 February ; 7(2): 78–80. doi:10.1038/nchembio.509.

Sublancin is not a lantibiotic but an S-linked glycopeptide

Trent J. Oman¹, John M. Boettcher¹, Huan Wang¹, Xenia N. Okalibe², and Wilfred A. van der Donk^{1,*}

¹Department of Chemistry and Howard Hughes Medical Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

²HHMI EXROP student, University of Michigan Telephone: (217) 244 5360 FAX: (217) 244 8533

Abstract

Sublancin is shown to be an S-linked glycopeptide containing a glucose attached to a Cys residue, establishing a new post-translational modification. The activity of the S-glycosyl transferase was reconstituted *in vitro* and the enzyme is shown to have relaxed substrate specificity allowing the preparation of analogs of sublancin. Glycosylation is essential for its antimicrobial activity.

Post-translationally modified peptide natural products are a rapidly expanding class of compounds with a wide variety of biological activities¹. One such peptide, sublancin, is produced by *Bacillus subtilis* 168, the laboratory strain that was one of the first bacterial strains for which the genome was sequenced². Sublancin is encoded by the SP β -prophage in strains that are lysogenic for the SP β bacteriophage, and inhibits the growth of non-lysogenic strains³. The reported structure of sublancin contains a dehydroalanine and one of the thioether crosslinks that are characteristic of lantibiotics (Figure 1a)⁴. Like other lantibiotics⁵, sublancin is biosynthesized as a precursor peptide bearing an N-terminal leader peptide of the double glycine-type (19 amino acids, Figure 1b) and a C-terminal core peptide (37 amino acids) that is converted to the mature compound. Interestingly, a search of the *B. subtilis* 168 genome did not reveal genes for any of the four known classes^{5–8} of lantibiotic biosynthetic enzymes. Intrigued by the possibility of a fifth pathway to lanthionine-containing peptides, we investigated the biosynthesis of sublancin. We report here that the compound is not a lantibiotic but rather a very unusual S-linked glycopeptide. We present its structure, report the *in vitro* activity of the glycosyltransferase, and demonstrate its versatility to generate other bioactive S-linked products. A search of the protein databases

Users may view, print, copy, download and text and data- mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

vddonk@illinois.edu.

Author Contributions T.J.O. performed mass spectrometric analyses and all biochemical assays shown, which were designed and analyzed by T.J.O. and W.A.V.. J.M.B. performed NMR analysis. H.W. prepared and performed biochemical assays of the SunA mutant peptides. X.N.O. assisted with purification of sublancin and mass spectrometric analyses. T.J.O. and W.A.V. wrote the manuscript.

Competing financial interests

The authors declare no competing financial interests.

Supplementary information is available online at <http://www.nature.com/naturechemicalbiology/>. Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>. Correspondence and requests for materials should be addressed to W.A.V.

suggests that sublancin is a member of a larger family of novel S-linked glycopeptides, further expanding the already impressive structural diversity of post-translationally modified peptide natural products⁹.

Sublancin was purified as described previously⁴ (Supplementary Methods) and analyzed by mass spectrometry. The mass of the unmodified core peptide predicted from the *sunA* gene sequence is 3717.72 Da whereas the observed mass of purified sublancin is 3875.74 Da. The core peptide contains five Cys residues (Figure 1b), but sublancin does not react with iodoacetamide demonstrating it does not contain free thiols. However, after treatment with TCEP, four out of five cysteines became accessible for alkylation (Supplementary Fig. 1). Sublancin was further analyzed by tandem mass spectrometry. To allow fragmentation, the disulfide bonds were first reduced and the reduced peptide was treated with chymotrypsin, providing a series of peptide fragments (Supplementary Fig. 2). The fragment spanning residues 12-32 of sublancin was analyzed by electrospray ionization-quadrupole/time-of-flight mass spectrometry. The observed fragmentation pattern combined with the sequence of the precursor gene *sunA* indicated that residue 16 is a Ser (Figure 1c) and not a dehydroalanine as previously reported.⁴ Furthermore, the observed fragments indicate that Thr19 is intact and therefore not engaged in a methyllanthionine ring. The data further suggest that Cys22 is carrying a post-translational modification with a mass of 162 Da (Figure 1d). This mass is consistent with a hexose conjugated to the Cys thiol. To determine the identity of the hexose, sublancin was subjected to acidic hydrolysis to release the sugar¹⁰. Subsequent chemical modification to a trimethylsilylated derivative allowed analysis by GC-MS, and comparison with a series of authentic standards demonstrated that the hexose of sublancin is glucose (Supplementary Fig. 3).

The sublancin biosynthetic gene cluster is shown in Figure 2a¹¹. It contains the precursor gene *sunA4* and genes encoding two thiol-disulfide oxidoreductases, BdbA and BdbB¹¹. In addition, it contains two orfs of unknown function, *yolJ* and *yolF*. YofF was recently suggested to be important for immunity of the producing strain and renamed SunI¹²; the function of YolJ has not yet been reported. Bioinformatic analysis showed that YolJ has sequence similarity to the GT-A glycosyl transferases of the GT2 family¹³ (Supplementary Fig. 4), consistent with the observed glucose attached to Cys22. Heretoforth we will refer to the enzyme as SunS, a sublancin biosynthetic enzyme. To confirm its function, *sunS* was cloned and expressed in *Escherichia coli* as an N-terminal fusion protein with a hexahistidine tag (His₆-SunS, Supplementary Fig. 5). Upon purification by immobilized metal affinity chromatography, His₆-SunS was incubated with the purified precursor peptide His₆-SunA, also expressed in *E. coli*. Addition of uridine diphosphate α -D-glucose (UDP-Glc) and Mg²⁺ resulted in conversion of SunA to a product consistent with glucosylation (Figure 2b). Subsequent chymotrypsin digest and ESI-MSⁿ analysis verified that glucose was attached to Cys22 (Supplementary Fig. 6), demonstrating strong regioselectivity of the glycosyltransferase as the SunA substrate contains five Cys residues. Because His₆-SunS glucosylated reduced His₆-SunA, the presence of the disulfides is not required for substrate recognition. Interestingly, the enzyme displays strong chemoselectivity for glucosylation of a thiol because the His₆-SunA mutant C22S was not a substrate (Supplementary Fig. 7a). Additional chemical elimination studies with in vitro glucosylated His₆-SunA confirmed

that the glucose was attached to the sulfur atom of Cys22 (Supplementary Fig. 8). Given the known inversion of stereochemistry of the GT2 family of GT-A glycosyl transferases, a β -linkage is expected but NMR studies were not conclusive to confirm this assumption (Supplementary Fig. 9). The revised structure of sublancin based on our data is shown in Figure 2c. This revised structure also explains the discrepancy noted in the original report between the mass of its proposed structure and the actual mass⁴.

To the best of our knowledge, reports of *S*-linked glycopeptides are extremely rare. No examples of *S*-linked glycoproteins have been reported¹⁴, but an *S*-linked glycopeptide has been found in human urine¹⁰. The conjugation of sugars to Cys results in more stable products than linkage to Ser, both at low and high pH¹⁵⁻¹⁷. As a result, *S*-linked glycopeptides have been the topic of several recent synthetic investigations¹⁸⁻²². Indeed, sublancin is a remarkably stable peptide that tolerates both low and high pH⁴. A search of the databases revealed several gene clusters in diverse host organisms that are similar to the sublancin cluster, including genes for putative precursor peptides and a putative *S*-glycosyl transferase (Supplementary Fig. 10). The precursor peptides have low sequence identity but all are rich in Cys residues in the predicted core peptide. Thus, *S*-linked glycopeptides may be more common than currently appreciated.

To determine the substrate selectivity of SunS, the enzyme was incubated with His₆-SunA and Mg²⁺ in the presence of UDP- α -D-*N*-acetylglucosamine (UDP-GlcNAc) and UDP- α -D-galactose (UDP-Gal) resulting in glycosylation of SunA at Cys22 as determined by tandem MS (Supplementary Fig. 11). Surprisingly, even guanosine diphosphate α -D-mannose (GDP-Man) and UDP- α -D-xylose (UDP-Xyl) proved to be substrates (Supplementary Fig. 12). In the absence of SunS, no activity was detectable. Analysis of the glycosylation reactions with UDP-GlcNAc, UDP-Gal, GDP-Man, and UDP-Xyl demonstrated that these substrates were somewhat less efficient than the corresponding reaction with UDP-glucose but all reactions could be coerced to high conversion (Supplementary Fig. 13 and Supplementary Table 1). Thus, SunS is highly promiscuous with respect to its nucleotide-sugar donor co-substrate and provides a new tool for glycodiversification of peptide natural products^{23,24} adding *S*-linked structures to the repertoire. To also probe the substrate specificity of the peptide, four mutant SunA peptides were prepared. Interestingly, three of the four Cys residues in SunA that are not glucosylated are flanked by Gln and Ala, whereas Cys22 that is glucosylated has two flanking Gly residues. When these two glycines were mutated to Gln and Ala, the resulting SunA-G21Q/G23A peptide was still glucosylated by SunS. Mutation of these flanking glycines to charged residues (Lys or Glu) also did not prevent glucosylation (Supplementary Figure 7b-e) demonstrating that the flanking residues do not provide the basis of substrate recognition, and that the substrate promiscuity extends to the peptide. To evaluate whether the leader peptide guides SunS activity, a Factor Xa cleavage site was engineered into the SunA substrate at the junction of the leader and core peptides. Expression of His₆-SunA-Xa and proteolytic cleavage with Factor Xa allowed purification of the 37-residue core peptide. Incubation with His₆-SunS and UDP-Glc resulted in full glucosylation of Cys22, demonstrating that the leader peptide is not required for enzyme activity (Supplementary Fig. 14).

Reduction of the disulfides of sublancin resulted in complete loss of antimicrobial activity against *B. subtilis* ATCC 6633 in an agar diffusion assay, showing their importance for activity (Supplementary Fig. 15). An *in vitro* oxidative folding protocol was therefore developed to investigate the effect of glycosylation on the antimicrobial activity of sublancin. His₆-SunA-Xa glycosylated by His₆-SunS was treated with Factor Xa, followed by oxidative folding through the addition of a mixture of oxidized and reduced glutathione and EDTA25. Subsequent analysis by antimicrobial assays, proteolytic digests, and ESI-MSⁿ analysis verified the complete *in vitro* reconstitution of sublancin biosynthesis with the correct disulfide connectivities (Figure 2d and Supplementary Figures 16-18). Leaving out any of the three steps (glycosylation, proteolysis or oxidative refolding) resulted in inactive material (Supplementary Fig. 19). These observations show that glycosylation is required for bioactivity and that the leader peptide inhibits bioactivity. We were unable to investigate whether SunS can also glycosylate oxidatively folded SunA because the enzyme required the presence of reducing agent for activity. Interestingly, the SunA core peptides with different sugar moieties attached to Cys22 were also amenable to oxidative folding resulting in various sublancin analogs with the correct disulfide connectivities as determined by proteolytic digests and ESI-MSⁿ analysis (Supplementary Fig. 18). Bioassays demonstrated that the native (glycosylated) sublancin was slightly more active than these analogs (Figure 2d), but we cannot rule out that small differences observed in the inhibition zones may be the result of small amounts of co-eluting peptides with alternative disulfide topologies that may be present in the final material. Regardless, although glycosylation is required for bioactivity, the stereochemistry of the hexose appears to be less critical.

In summary, sublancin is a very unusual S-linked glycopeptide and SunS an unusual member of antibiotic glycosyltransferases with very relaxed substrate specificity. Given the increased chemical stability of S-linked glycopeptides, SunS has great potential for use in preparation of such compounds. Future studies will focus on the mode of action of sublancin.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by the National Institutes of Health (GM58822 to W.A.V.) and a NIH Cellular and Molecular Biology Training Grant (T32 GM007283 to T.J.O.).

References

1. Oman TJ, van der Donk WA. Nat. Chem. Biol. 2010; 6:9–18. [PubMed: 20016494]
2. Kunst F, et al. Nature. 1997; 390:249–56. [PubMed: 9384377]
3. Hemphill HE, Gage I, Zahler SA, Korman RZ. Can. J. Microbiol. 1980; 26:1328–33. [PubMed: 6783280]
4. Paik SH, Chakicherla A, Hansen JN. J. Biol. Chem. 1998; 273:23134–42. [PubMed: 9722542]
5. Willey JM, van der Donk WA. Annu. Rev. Microbiol. 2007; 61:477–501. [PubMed: 17506681]
6. Siezen RJ, Kuipers OP, de Vos WM. Antonie van Leeuwenhoek. 1996; 69:171–84. [PubMed: 8775977]

7. Kodani S, et al. *Proc. Natl. Acad. Sci. USA*. 2004; 101:11448–11453. [PubMed: 15277670]
8. Goto Y, et al. *PLoS Biol*. 2010; 8:e1000339. [PubMed: 20351769]
9. McIntosh JA, Donia MS, Schmidt EW. *Nat. Prod. Rep.* 2009; 26:537–559. [PubMed: 19642421]
10. Lote CJ, Weiss JB. *FEBS Lett*. 1971; 16:81–85. [PubMed: 11945907]
11. Dorenbos R, et al. *J. Biol. Chem.* 2002; 277:16682–8. [PubMed: 11872755]
12. Dubois JY, et al. *Antimicrob. Agents Chemother.* 2009; 53:651–61. [PubMed: 19047653]
13. Lairson LL, Henrissat B, Davies GJ, Withers SG. *Annu. Rev. Biochem.* 2008; 77:521–55. [PubMed: 18518825]
14. Walsh, CT. *Post-translational modification of proteins: expanding nature's inventory*. Roberts & Company, Greenwood Village, CO; 2005.
15. Baran E, Drabarek S. *Pol. J. Chem.* 1978; 52:941–946.
16. Jahn M, Marles J, Warren RA, Withers SG. *Angew. Chem. Int. Ed. Engl.* 2003; 42:352–4. [PubMed: 12548697]
17. Driguez H. *Chembiochem*. 2001; 2:311–8. [PubMed: 11828459]
18. Cohen SB, Halcomb RL. *J. Am. Chem. Soc.* 2002; 124:2534–43. [PubMed: 11890803]
19. Galoni DP, van der Donk WA, Gin DY. *Chem.-Eur. J.* 2003; 24:5997–6006. [PubMed: 14679512]
20. Galoni DP, van der Donk WA, Gin DY. *J. Am. Chem. Soc.* 2004; 127:12–3. [PubMed: 15469231]
21. Galoni DP, Ide ND, van der Donk WA, Gin DY. *J. Am. Chem. Soc.* 2005; 127:7359–69. [PubMed: 15898784]
22. Hili R, Rai V, Yudin AK. *J. Am. Chem. Soc.* 2010; 132:2889–91. [PubMed: 20155938]
23. Blanchard S, Thorson JS. *Curr. Opin. Chem. Biol.* 2006; 10:263–71. [PubMed: 16675288]
24. Thibodeaux CJ, Melancon CE, Liu HW. *Nature*. 2007; 446:1008–16. [PubMed: 17460661]
25. Fuller E, et al. *FEBS J.* 2005; 272:1727–38. [PubMed: 15794759]

Figure 1. Analysis of the structure of sublancin by tandem mass spectrometry

(a) Originally reported structure of sublancin4 proposed to contain a dehydroalanine at position 16 and a methyllanthionine bridge arising from dehydration of Thr19 to Dhb19 and subsequent attack of Cys22 on Dhb19 to form the thioether linkage. (b) Primary sequence of the SunA precursor peptide containing a 19 amino acid N-terminal leader peptide (blue) and a 37 amino acid C-terminal core peptide (black). A double glycine-type proteolytic cleavage site is underlined in red. (c) Fragmentation pattern of the chymotryptic peptide spanning residues 12-32 of sublancin (underlined with dashed line in panel 1b). The fragment ions in the 400-900 m/z range are shown that contain the ion series b_4 through b_{10} . The masses are consistent with Ser16 and Thr19 remaining unmodified. (d) Fragmentation pattern of the chymotryptic peptide spanning residues 12-32 in the 800-1300 m/z range containing the ion series y''_8 through y''_{12} . The masses are consistent with modification of Cys22 with a hexose (+162 Da). Asterisk indicates ion resulting from loss of the hexose from Cys22 during collision induced dissociation.

Figure 2. *In vitro* reconstitution of SunS activity and sublancin antimicrobial activity

(a) The biosynthetic gene cluster of sublancin in *B. subtilis* 168 consists of the glycosyltransferase gene *sunS*, the precursor gene *sunA*, the ABC-transporter gene *sunT*, two thiol-disulfide oxidoreductase genes *bdbA* and *bdbB*, and the immunity protein gene *sunI*. (b) MALDI-TOF MS spectra of His₆-SunA before (black) and after (red) incubation with His₆-SunS, UDP-glucose, and Mg²⁺. (c) The revised structure of sublancin. (d) Antimicrobial activity assays of *in vitro* produced sublancin and sublancin analogs against *B. subtilis* 6633. The type of sugar attached to Cys22 is indicated. Compounds were produced by incubating His₆-SunA-Xa with His₆-SunS and NDP-sugar, followed by proteolysis with Factor Xa to remove the leader peptide, and subsequent oxidative folding to generate the disulfides. Authentic sublancin standards produced and purified from *B. subtilis* 168 were used as positive controls.