ORGANIC LETTERS

2012
Vol. 14, No. 7
1910–1913

Synthesis of the Bacteriocin Glycopeptide Sublancin 168 and *S*-Glycosylated Variants

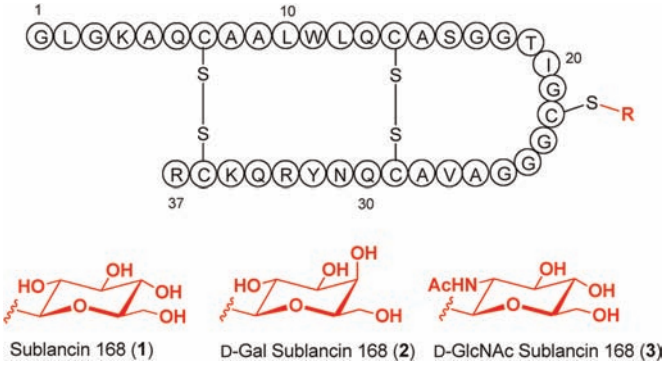
Yves S. Y. Hsieh, Brendan L. Wilkinson, Mitchell R. O'Connell, Joel P. Mackay, Jacqueline M. Matthews, and Richard J. Payne*

School of Chemistry and School of Molecular Bioscience, The University of Sydney, NSW 2006, Australia

richard.payne@sydney.edu.au

Received March 5, 2012

ABSTRACT



The synthesis of sublancin 168, a unique *S*-glucosylated bacteriocin antibiotic, is described. The natural product and two *S*-glycosylated variants were successfully prepared *via* native chemical ligation followed by folding. The synthetic glycopeptides were shown to possess primarily an α-helical secondary structure by CD and NMR studies.

Sublancin 168 (**1**) is a 37 amino acid bacteriocin peptide isolated from the *Bacillus subtilis* 168 strain. The natural product exhibits bactericidal activity against a spectrum of Gram-positive bacteria, including *B. subtilis* 6633, *B. megaterium*, and the pathogenic strains *Staphylococcus aureus* and *Streptococcus pyogenes* with varying degrees of sensitivity.[1]

Sublancin 168 was originally classified as a lantibiotic, since it was thought to possess a characteristic β-methyllanthionine thioether modification.[1] However, a recent report from van der Donk and co-workers has unequivocally demonstrated that sublancin 168 is an unusual *S*-linked glycopeptide with a glucose (Glc) moiety β-linked to cysteine-22 (Cys-22).[2] Although there have been several studies on the synthesis of *S*-linked glycopeptides and glycoproteins as metabolically stable mimics of their *O*-linked counterparts,[3] sublancin 168 was identified as the first naturally occurring bacterially derived *S*-linked glycopeptide. However, Glycocin F[4] (from *Lactobacillus plantarum* KW3) has also since been identified and shown to possess an *S*-linked glycan, in this case *N*-acetylglucosamine (GlcNAc). This may suggest that other bacterial species might produce peptides bearing this intriguing post-translational modification. Although a range of

(1) Paik, S. H.; Chakicherla, A.; Hansen, J. N. *J. Biol. Chem.* **1998**, *273*, 23134.

(2) (a) Oman, T. J.; Boettcher, J. M.; Wang, H.; Okalibe, X. N.; van der Donk, W. A. *Nat. Chem. Biol.* **2011**, *7*, 78. (b) Wang, H.; van der Donk, W. A. *J. Am. Chem. Soc.* **2011**, *133*, 16394.

(3) (a) Gerz, M.; Matter, H.; Kessler, H. *Angew. Chem., Int. Ed.* **1993**, *32*, 269. (b) Zhu, X. M.; Pachamuthu, K.; Schmidt, R. R. *J. Org. Chem.* **2003**, *68*, 5641. (c) Galonic, D. P.; van der Donk, W. A.; Gin, D. Y. *Chem.—Eur. J.* **2003**, *9*, 5997. (d) Zhu, X. M.; Pachamuthu, K.; Schmidt, R. R. *Org. Lett.* **2004**, *6*, 1083. (e) Zhu, X. M.; Schmidt, R. R. *Chem.—Eur. J.* **2004**, *10*, 875. (f) Zhu, X. M.; Haag, T.; Schmidt, R. R. *Org. Biomol. Chem.* **2004**, *2*, 31. (g) Thayer, D. A.; Yu, H. N.; Galan, M. C.; Wong, C.-H. *Angew. Chem., Int. Ed.* **2005**, *44*, 4596. (h) Floyd, N.; Vijayakrishnan, B.; Koeppe, A. R.; Davis, B. G. *Angew. Chem., Int. Ed.* **2009**, *48*, 7798. (i) Chalker, J. M.; Bernardes, G. J. L.; Davis, B. G. *Acc. Chem. Res.* **2011**, *44*, 730.

(4) Stepper, J.; Shastri, S.; Loo, T. S.; Preston, J. C.; Novak, P.; Man, P.; Moore, C. H.; Havlíček, V.; Patchett, M. L.; Norris, G. E. *FEBS Lett.* **2011**, *585*, 645.

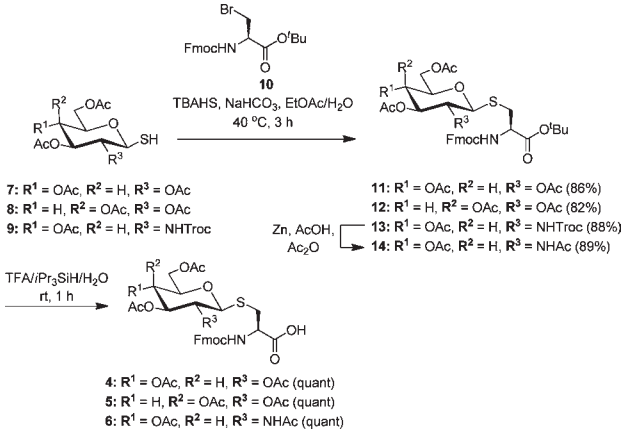
10.1021/ol300557g
Published on Web 03/28/2012

glycan substitutions at Cys-22 (produced *via* the *S*-glycosyltransferase enzyme SunS)[2] and a Cys to Ala mutant of sublancin 168[2b] still possess antimicrobial activity, the precise role of the glycan in determining both the structure and biological function of the natural product remains unclear. Given the potential significance of this new post-translational modification, we were interested in pursuing structural studies on this class of bacteriocin natural products. As such we embarked on the total synthesis of sublancin 168 (**1**) bearing the native D-Glc modification, along with two non-natural analogues **2** and **3** bearing D-galactose (Gal) or D-GalNAc at Cys-22, respectively.

Recently, Hojo and co-workers reported the synthesis of sublancin 168 using a linear Fmoc-strategy solid phase peptide synthesis (Fmoc-SPPS).[5] In our hands, this approach failed to provide sufficient material of a purity that would permit a detailed structural investigation, despite several attempts using different protocols. Therefore, we pursued a native chemical ligation-based strategy[6] for the convergent synthesis of sublancin 168 and non-natural analogues. Herein, we report the successful assembly of sublancin 168 (**1**) and glycosylated variants (**2** and **3**) which has provided sufficient material for the first detailed interrogation of the secondary structure of these glycopeptides using NMR spectroscopy and circular dichroism (CD) spectropolarimetry.

Prior to the synthesis of sublancin 168 and its analogues, we required access to suitably protected thioglycosyl amino acid building blocks **4**–**6** which could be directly incorporated into Fmoc-SPPS (Scheme 1). To this end, 1-thiosugars **7**–**9**[3d,h,7] were each reacted with protected β-bromoalanine **10**[3e] in a mixture of ethyl acetate and water *via* phase-transfer catalysis with tetrabutylammonium hydrogen sulfate (TBAHS) to afford protected thioglycosides **11**–**13** in excellent yields (82–88%).[3e] Reduction of the trichloroethyl (Troc) carbamate in **13** using Zn in acetic acid provided acetamide **14** in good yield (89%). Acidolytic deprotection of the *tert*-butyl esters in **11**, **12**, and **14** using 95:5:5 v/v/v trifluoroacetic acid (TFA)/triisopropylsilane/water then provided thioglycosylamino acids **4**–**6** in quantitative yields.

Scheme 1. Synthesis of Thioglycosylamino Acids **4**–**6**



At the outset, we identified the peptide bond between glutamine (Gln)-13 and Cys-14 as a suitable junction for native chemical ligation. Consequently, 24 amino acid glycopeptides **15**–**17** (bearing *N*-terminal Cys residues) and peptide thioester **18** became the focus for synthesis (Scheme 2). Our approach to *S*-linked glycopeptides **15**–**17** involved the use of a microwave assisted Fmoc-SPPS protocol employing NovaPEG Wang resin immobilized with Fmoc-Arg(Pbf)-OH (Scheme 2 and Supporting Information). Unfortunately, and in agreement with observations made during our previous attempts to prepare the natural product from a single solid phase synthesis, elongation of the glycopeptides under standard conditions proved problematic, largely attributable to aggregation of the peptide *en bloc*. To overcome aggregation, the synthesis was optimized by incorporating the 2,4-dimethoxybenzyl (Dmb) derived dipeptide, Fmoc-Gly-(Dmb)Gly-OH, at Gly-17-Gly-18 and Gly-23-Gly-24 in the sequence. Following successful peptide elongation, glycopeptides were cleaved from the resin with concomitant removal of side chain protecting groups using a mixture of TFA/triisopropylsilane/H_2O (90:5:5, v/v/v). The *O*-acetate protecting groups on the glycans could then be smoothly removed using 20% aqueous hydrazine hydrate. Purification by preparative reversed-phase HPLC afforded the target glycopeptides **15**–**17** in 4–8% yield based on the original resin loading.

The synthesis of the peptide thioester fragment **18** was readily achieved using a previously described side-chain-anchoring strategy (Scheme 2 and Supporting Information).[8] Commercially available Fmoc-Glu-OAllyl was immobilized onto Rink amide resin and elongated using microwave assisted Fmoc-SPPS. Selective demasking of the *C*-terminal carboxylate using $Pd(PPh_3)_4$, followed by thioesterification by treatment with ethyl-3-mercaptopropionate, *N,N*′-diisopropylcarbodiimide (DIC), 1-hydroxybenzotriazole (HOBt), and *N,N*-diisopropylethylamine (DIPEA) as base successfully provided the resin-bound peptide thioester. From here, acidolytic resin cleavage and side-chain deprotection, followed by purification by reversed-phase preparative HPLC, gave the desired peptide thioester **18** in good yield (15% based on the original resin loading).

Having successfully prepared the three requisite glycopeptides **15**–**17** and peptide thioester ligation partner **18**, we next submitted these to native chemical ligation. To this end, *C*-terminal peptide thioester **18** was reacted with **15**, **16**, or **17** in the presence of mercaptophenylacetic acid (MPAA) as a thiol catalyst and tris(2-carboxyethyl)-phosphine (TCEP) as a reductant in 6 M guanidine

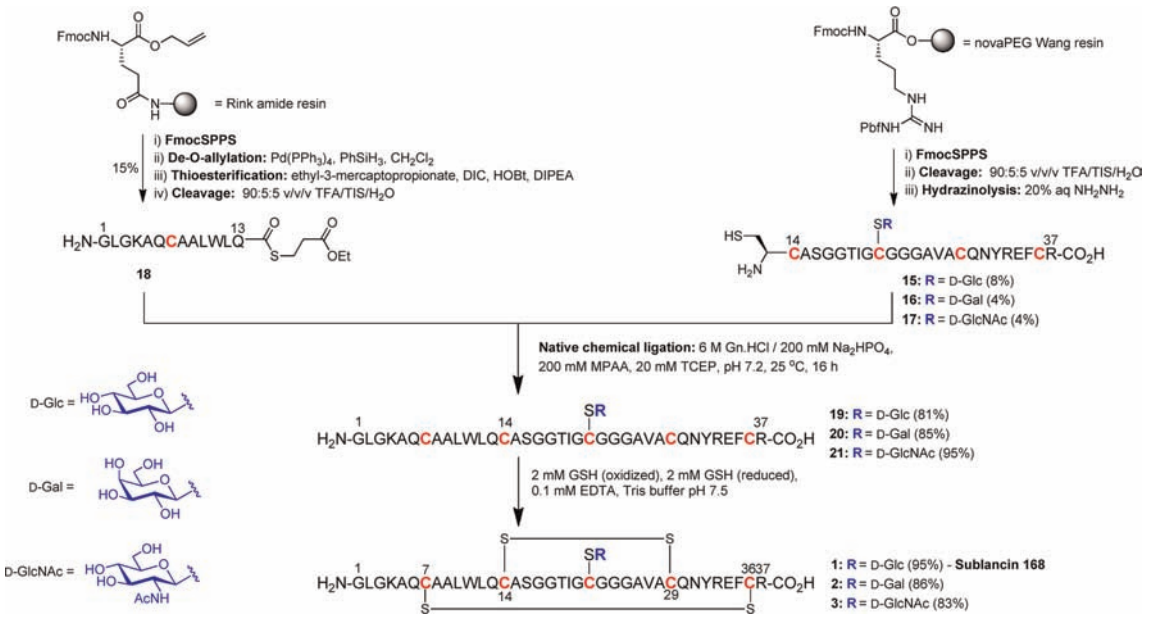
(5) Katayama, H.; Asahina, Y.; Hojo, H. *J. Pept. Sci.* **2011**, *17*, 818.

(6) (a) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776. (b) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10068.

(7) Ferrier, R. J.; Furneaux, R. H. *Carbohydr. Res.* **1977**, *57*, 73.

(8) (a) Ficht, S.; Payne, R. J.; Guy, R. T.; Wong, C.-H. *Chem.—Eur. J.* **2008**, *14*, 3620. (b) Thomas, G. L.; Hsieh, Y. S. Y.; Chun, C. K. Y.; Cai, Z. L.; Reimers, J. R.; Payne, R. J. *Org. Lett.* **2011**, *13*, 4770. (c) Thompson, R. E.; Jolliffe, K. A.; Payne, R. J. *Org. Lett.* **2011**, *13*, 680. (d) Lutsky, M. Y.; Nepomniaschiy, N.; Brik, A. *Chem. Commun.* **2008**, *10*, 1229.

Scheme 2. Synthesis of Sublancin 168 **1** and Glycosylated Variants **2** and **3**



hydrochloride (Gn·HCl)/0.2 M Na_2HPO_4 buffer at pH 7.2 (Scheme 2).[9] The ligation reaction between peptide thioester **18** and all three glycopeptides proceeded to completion within 16 h, as determined by LC-MS (Scheme 2 and Supporting Information). Full-length, unfolded sublancin 168 (**19**) and analogues **20** and **21** were isolated in excellent yields (81−95%) following purification by reversed-phase preparative HPLC. Finally, glycopeptides **19**−**21** were successfully folded using oxidized and reduced glutathione (GSH). The correct disulfide connectivity (Cys7-Cys36 and Cys14-Cys29) was confirmed through digestion with chymotrypsin followed by LC-MS analysis of the resulting peptide and glycopeptide fragments (see Supporting Information).[2a] Gratifyingly, correctly folded glycopeptides **1**−**3** were isolated in 83−95% yields following HPLC purification.

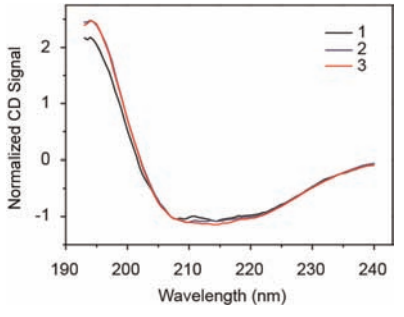


Figure 1. Far-UV circular dichroism (CD) spectra of glycopeptides **1**−**3**.

Having successfully prepared sublancin 168 (**1**) and glycosylated analogues **2** and **3**, we now moved to examine the effect of the different glycans on secondary structure using NMR spectroscopy and CD spectropolarimetry. Far-UV CD spectra of **1**−**3** revealed that all three peptides contain substantial amounts of helical structure, judging from the minima at 208 and 222 nm (Figure 1). Estimates of secondary structure content, made using the method of Raussens et al.,[10] predict that **1**, **2**, and **3** are predominantly helical, with estimates of 47%, 51%, and 48% helix, respectively. Each peptide was also estimated to contain 13% β-turn and 4−7% β-sheet structure. Thermal denaturation curves, obtained by recording the ellipticity at 222 nm as a function of temperature, show that all peptides have similar thermal stability, as evidenced by a gradual loss of signal intensity with increasing temperature (see Supporting Information).

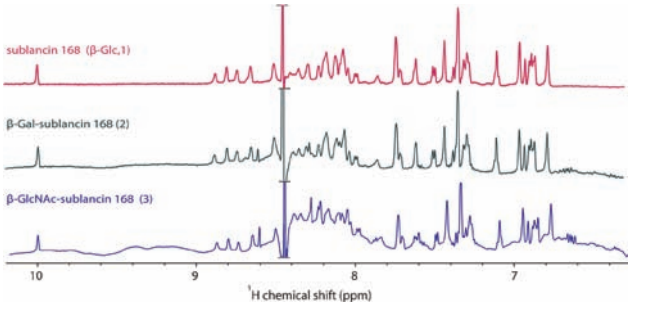


Figure 2. Backbone amide regions of the ^{1}H NMR spectra (recorded at 298 K) of sublancin 168 (**1**) and glycosylated variants **2** and **3**.

(9) Johnson, E. C. B.; Kent, S. B. H. *J. Am. Chem. Soc.* **2006**, *128*, 6640.

The amide regions of the one-dimensional ^{1}H NMR spectra of peptides **1**–**3** (and of an unglycosylated sublancin 168 analogue) were found to exhibit a close resemblance to each other (Figure 2 and Supporting Information). Thus, taken with the CD spectra, these data demonstrate that the overall conformation of sublancin 168 (**1**) is not affected by either the identity or the presence of a glycan moiety. This may explain the prior observation that differentially *S*-glycosylated sublancin 168 analogues all possess antimicrobial activity.[2a] The dispersion and line shapes of the NMR spectra also indicate that sublancin 168 takes up a well-ordered conformation in solution. In order to probe this conformation, we recorded two-dimensional TOCSY, DQF-COSY, and NOESY spectra of **1** and made full chemical shift assignments. The deviations of H^{α} and $C^{\alpha/\beta}$ chemical shifts from their random coil values (Figure 3A and 3B) demonstrate that sublancin 168 (**1**) contains two α-helical segments in the regions 8–16 and 27–34, in agreement with the CD data and the bioinformatic structural prediction using the Protein Structure Prediction Server (PSIPRED) (see Supporting Information). Analysis of medium-range NOE data (see Supporting Information) confirms this conclusion, and furthermore, comparison of the data with the recently reported structure of Glycocin F indicates that the structure of sublancin 168 is likely to be very similar[11] but without the disordered *C*-terminal tail found in the former peptide. Future work in our laboratories will focus on solving the solution structure of sublancin 168.

In summary, we have successfully synthesized sublancin 168 (**1**) and two glycosylated variants (**2** and **3**) using Fmoc-SPPS and a convergent, high yielding assembly using native chemical ligation. Following the synthesis and correct folding of the glycopeptides, we have unambiguously identified the presence of a strong α-helical contribution to the overall structure using CD and 2D NMR spectroscopy. The synthesis of sublancin 168 analogue libraries and other *S*-glycosidically linked bacteriocins with promising biological activity, along with the evaluation of antibacterial activity, will be reported in due course.

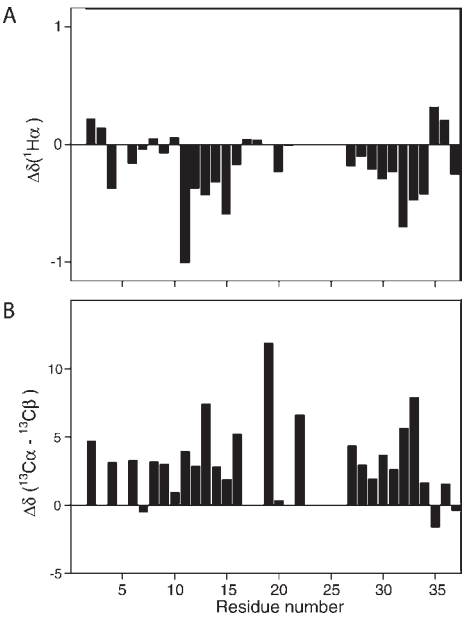


Figure 3. (A) Plot of the deviation of ^{1}Hα chemical shifts from random coil values for sublancin 168 (**1**). An extended series of negative values is indicative of an α-helical region; positive values indicate a *β*-sheet. (B) Plot of the difference ($\Delta^{13}C\alpha - \Delta^{13}C\beta$) for sublancin 168 (**1**), which measures the deviations of $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts from random coil values. An extended series of positive values is indicative of an α-helical region; negative values indicate a *β*-sheet.

Acknowledgment. We acknowledge the University of Sydney, the Australian Postgraduate Award Scheme (Y.H. and M.O.), and the Cancer Institute NSW (M.O.) for funding.

Supporting Information Available. Experimental procedures and characterization of glycosylamino acids and peptide and glycopeptide products. 1D and 2D NMR data and CD spectra of **1**–**3**. This material is available free of charge via the Internet at http://pubs.acs.org.

The authors declare no competing financial interest.

(10) (a) Raussens, V.; Ruysschaert, J.; Goormaghtigh, E. *Anal. Biochem.* **2003**, *319*, 114. (b) Raussens, V.; Ruysschaert, J.-M.; Goormaghtigh, E. *Anal. Biochem.* **2006**, *359*, 150.

(11) Venugopal, H.; Edwards, P. J. B.; Schwalbe, M.; Claridge, J. K.; Libich, D. S.; Stepper, J.; Loo, T.; Patchett, M. L.; Norris, G. E.; Pascal, S. M. *Biochemistry* **2011**, *50*, 2748.