Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

The glycocins: in a class of their own

Gillian E Norris and Mark L Patchett

First reported in 2011, glycocins (**glyco**sylated bacterio**cins**) are bacterial toxins that constitute a subset of ribosomally synthesised and post-translationally modified peptide (RiPP) natural products. Three NMR structures (glycocin F, ASM1 and sublancin 168), two with helix–loop–helix Cs α/α folds, are deposited in the PDB. Each structure contains a monosaccharide β-*S*-linked to a cysteine side chain. Three more glycocins (thurandacin, and enterocins F4-9 and 96) have been biochemically characterised, and others predicted on the basis of bioinformatic analyses. Only glycocin F, ASM1 and enterocin F4-9 are unequivocally glycoactive. This review probes the structure–function relationships of four types of nested disulfide-bonded glycocins.

Address
Institute of Fundamental Sciences, College of Sciences, Massey University, Private Bag 11222, Palmerston North 4442, New Zealand

Corresponding authors: Norris, Gillian E (g.norris@massey.ac.nz) and Patchett, Mark L (m.l.patchett@massey.ac.nz)

Current Opinion in Structural Biology 2016, **40**:112–119

This review comes from a themed issue on **Carbohydrate–protein interactions and glycosylation**

Edited by **Nagasuma Chandra** and **Harry Gilbert**

http://dx.doi.org/10.1016/j.sbi.2016.09.003

Introduction

The synthesis of toxins that harm and/or deter competing, pathogenic, and prey organisms is widespread in all three taxonomic domains. The ribosomally synthesised antimicrobial polypeptides secreted by bacteria are commonly called bacteriocins. Many bacteriocins are small (20–50 amino acids (aa)) with cationic N-terminal regions that bind to negatively charged cell membranes, allowing the hydrophobic C-terminal regions to form pores, killing target cells [1]. A smaller number of bacteriocins are bacteriostatic; that is, they slow or stop the growth of their bacterial targets without killing them.

Classification schemes for lactic acid bacteria bacteriocins have been progressively simplified [2,3•,4•]. The basis for division between unmodified (class II) and modified (class I) bacteriocins is the presence of unusual post-translational modifications (PTMs), for example lanthionine and *S*-glycosylated residues, sulfur to α-carbon linkages, and so on. Bacteriocins containing glycosylated residues, such as sublancin 168 (SunA) and glycocin F (GccF) [5••,6•], comprise the class I group named glycocins [7]. Five (SunA, GccF, ASM1, and enterocins 96 and F4-9) have been purified from bacterial culture, while a sixth, thurandacin, was identified by genomic data mining and chemoenzymatically synthesised [5••,6•,8–10,11•]. Glycocin sugar moieties, so far only one or two monosaccharides per scaffold, are linked to the sidechains of specific cysteine, serine or threonine residues [7,12]. When these moieties are essential for antimicrobial activity, the glycocin can be regarded as being 'glycoactive' (Table 1).

Glycocin F

GccF, secreted by *Lactobacillus plantarum* KW30, was the first glycocin to have its NMR structure determined [13•]. The solution structure has 2 two-turn antiparallel α-helices, joined by an 8-residue loop and constrained by two nested disulfide bonds ($(C–X_6–C)_2$ architecture), and a flexible C-terminal 'tail' (Figures 1a,b and S1). The peptide is decorated by two N-acetylglucosamine (GlcNAc) moieties; one β-*O*-linked through Ser18 in the sequence YD**S**GT (a motif that may contribute to receptor binding) in the interhelical loop, and the other β-*S*-linked to C-terminal Cys43 — a glycosidic linkage rarely described in biology. The amphipathic N-terminal helix (residues 5–12) presents its more hydrophobic face to the solvent and lies at an angle of 25° to the less well ordered C-terminal helix (residues 21–28). The loop between the helices has no obvious interactions with the rest of the structure and is only moderately flexible (Figures 1a,b and S1).

Analysis by CASTp [14] identified a solvent-accessible cavity between the loop and the helices which is ~90 Å^3 in volume and bounded by sidechain atoms of residues 8, 9, 11, 12–21 and 25 (Figure 1d,e) which could be part of a recognition site for receptor docking.

Which of these structural features are involved in the quite extraordinary activity of GccF? Against the *most* susceptible *L. plantarum* strains GccF has an IC_{50} of 4 pM equating to roughly 20 molecules per cell (unpublished). GccF is glycoactive; the *O*-linked GlcNAc is essential for activity and fits neatly into a groove formed by the top 'lip' of the cavity identified by CASTp decreasing the conformational flexibility of the GlcNAc which may facilitate target binding [15] and also protect the β-*O*-glycosidic bond from hydrolysis (Figure 1f). In contrast, the GlcNAc on the

Table 1

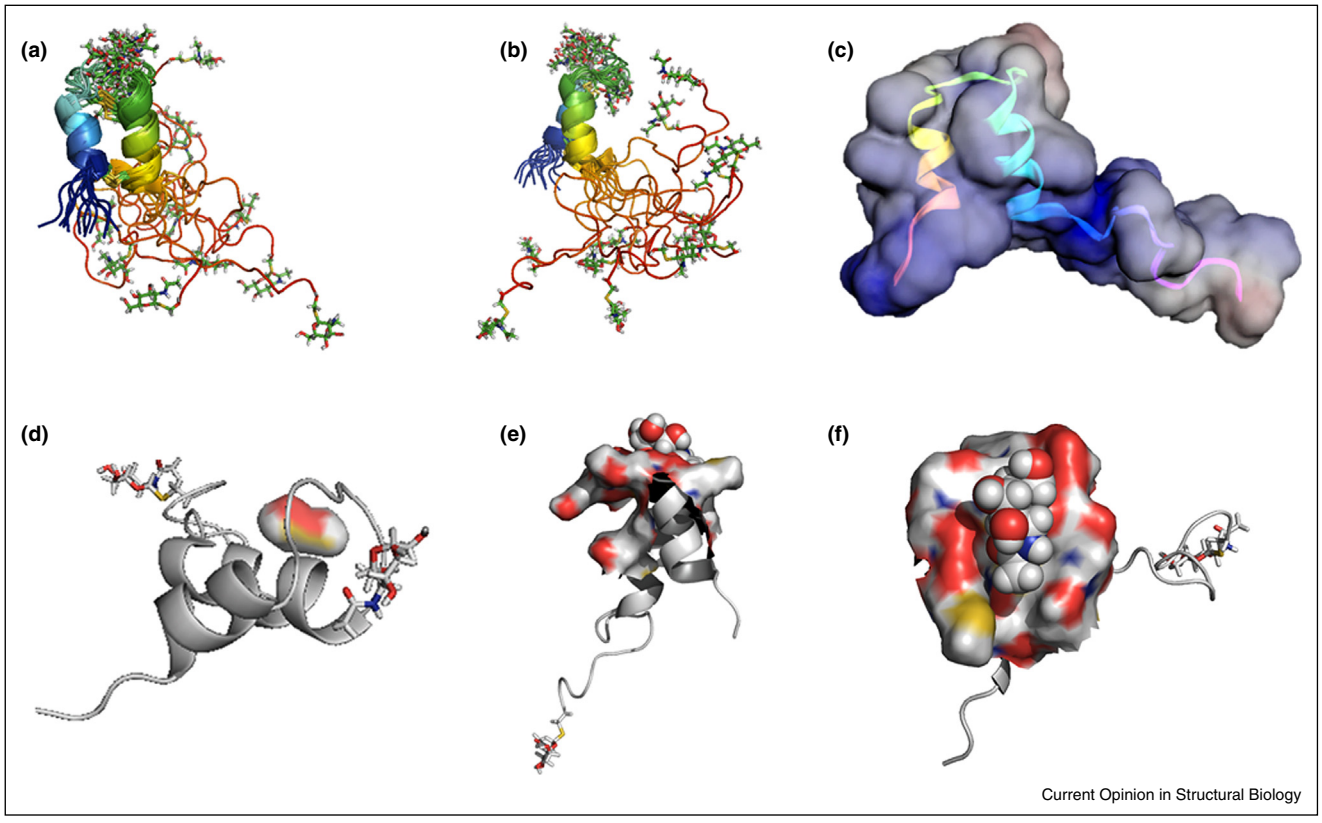
Properties of biochemically characterised glycocins.

Glycocin name[PDB ID]	Peptide scaffold [AAs; S–S (loop)*]	Glycosylation (glycoactive)	Mode of action, predicted p*I*
Glycocin F, GccF[2KUY] ASM1[2MVI]	43; 5–28, 12–21 (8)	S18/C43-GlcNAc[NP, C] (yes) S18/C43-GlcNAc[NP] (yes)	Bacteriostatic, 7.1/7.0
Sublancin, SunA[2MIJ]	37; 7–36, 14–29 (14)	C22-Glc[NP, CE, C] (?)	Bactericidal, 11.0
Thurandacin A, ThuA	42; 7–42, 14–35 (20)	C28-Glc[CE] (?)	Bactericidal, 8.3
Enterocin F4-9	47; 4–47, 32–38 (5)	S37/T46?-GlcNAc [NP] (yes)	Bacteriostatic, 10.7
Enterocin 96	48; Cys6/12/13/39/47	Δ mass ~ +2 hexosamines	Unknown, 9.7
Durancin 61A	Unknown	Glc/Ara? (unknown)	Bactericidal

* Number of residues in the (loop) between the preceding pair of disulfide-bonded cysteine thiols.
NP, natural product; CE, chemoenzymatic synthesis; C, chemical synthesis.

C-terminal 'tail' is unrestrained in its spatial position mainly due to the flexibility of residues 32–43 [13•]. Perhaps unexpectedly, the disordered tail region appears to contribute to GccF activity independently of the C-terminal GlcNAc. Replacing Cys43-β-*S*-GlcNAc with Ser43 (no GlcNAc) increases the IC_{50} from 1 to 100 nM against the standard indicator strain (unpublished), whereas removing the entire tail (residues 33–43) increases the IC_{50} to 350 nM, more than 3-fold higher than the molecule lacking only the C-terminal sugar. In addition, chemically synthesised C-terminally amidated GccF had half the activity of natural product GccF [16••]. Possibly a

Figure 1



Structural analysis of GccF. Structures of GccF generated using PyMOL (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC), and APBS [17,18]. **(a)** Cartoon ensemble of 12 structures of GccF (PDB: 2KUY) coloured N (blue) to C (red). Disulfide bonds are shown as (yellow) sticks and the GlcNAc moieties as sticks. **(b)** Same as (a) but rotated 90°. **(c)** A semi-transparent solvent accessible surface rendering of the GccF peptide, coloured according to electrostatic potential at pH 5: red for negative (−5 kT/e), blue for +ve (+5 kT/e) and white for neutral. A cartoon of the peptide chain is visible through the surface. **(d)** Cavity formed by residues 8, 9, 11–21 and 25 identified by CASTp analysis of GccF shown as a filled shape. **(e)** Side view of the cavity that fits the shape shown in (d) showing the relative position of the *O*-linked GlcNAc (shown as spheres). Atoms are coloured by element: C, grey; O, red; N, blue; S, yellow. **(f)** GccF from the 'top' showing the *O*-linked GlcNAc fitting into a groove on the outward surface of GccF made by the residues that form the cavity identified in (d) and (e).