

Research article

Open Access

The structure of pyogenecin immunity protein, a novel bacteriocin-like immunity protein from *Streptococcus pyogenes*

Changsoo Chang¹, Penny Coggill², Alex Bateman², Robert D Finn², Marcin Cymborowski³, Zbyszek Otwinowski⁴, Wladek Minor³, Lour Volkart¹ and Andrzej Joachimiak*¹

Address: ¹Midwest Center for Structural Genomics and Structural Biology Center, Biosciences Division, Argonne National Laboratory, Argonne, Illinois 60439, USA, ²Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, CB10 1SA, UK, ³Midwest Center for Structural Genomics, Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, VA 22903, USA and ⁴Midwest Center for Structural Genomics, Department of Biochemistry, UT Southwestern Medical Center at Dallas, Dallas, TX 75235, USA

Email: Changsoo Chang - cchang@anl.gov; Penny Coggill - pcc@sanger.ac.uk; Alex Bateman - agb@sanger.ac.uk; Robert D Finn - rdf@sanger.ac.uk; Marcin Cymborowski - mtc8n@virginia.edu; Zbyszek Otwinowski - zbyszek.otwinowski@utsouthwestern.edu; Wladek Minor - wladek@iwonka.med.virginia.edu; Lour Volkart - lvolkart@anl.gov; Andrzej Joachimiak* - andrzejj@anl.gov

* Corresponding author

Published: 17 December 2009

Received: 15 May 2009

BMC Structural Biology 2009, 9:75 doi:10.1186/1472-6807-9-75

Accepted: 17 December 2009

This article is available from: <http://www.biomedcentral.com/1472-6807/9/75>

© 2009 Chang et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Many Gram-positive lactic acid bacteria (LAB) produce anti-bacterial peptides and small proteins called bacteriocins, which enable them to compete against other bacteria in the environment. These peptides fall structurally into three different classes, I, II, III, with class IIa being pediocin-like single entities and class IIb being two-peptide bacteriocins. Self-protective cognate immunity proteins are usually co-transcribed with these toxins. Several examples of cognates for IIa have already been solved structurally. *Streptococcus pyogenes*, closely related to LAB, is one of the most common human pathogens, so knowledge of how it competes against other LAB species is likely to prove invaluable.

Results: We have solved the crystal structure of the gene-product of locus Spy_2152 from *S. pyogenes*, (PDB:2fu2), and found it to comprise an anti-parallel four-helix bundle that is structurally similar to other bacteriocin immunity proteins. Sequence analyses indicate this protein to be a possible immunity protein protective against class IIa or IIb bacteriocins. However, given that *S. pyogenes* appears to lack any IIa pediocin-like proteins but does possess class IIb bacteriocins, we suggest this protein confers immunity to IIb-like peptides.

Conclusions: Combined structural, genomic and proteomic analyses have allowed the identification and *in silico* characterization of a new putative immunity protein from *S. pyogenes*, possibly the first structure of an immunity protein protective against potential class IIb two-peptide bacteriocins. We have named the two pairs of putative bacteriocins found in *S. pyogenes* pyogenecin I, 2, 3 and 4.

Background

Many Gram-positive bacteria produce anti-bacterial peptides and small proteins, called bacteriocins. There are three main classes produced by Gram-positive lactic acid bacteria (LAB): class I bacteriocins are the lantibiotics, small (<4 kDa), post-translationally modified peptides containing unusual amino acids such as lanthionine; class II are small, unmodified, heat-stable bacteriocins (<10 kDa); class III include larger (>30 kDa) heat-labile proteins, such as murein hydrolases [1]. Most bacteriocins are synthesized as precursors, which are matured and secreted, then target a specific bacterium and kill it by increasing its membrane permeability to various small molecules. Class II bacteriocins are subdivided into IIa, pediocin-like unmodified bacteriocins, IIb, two-peptide unmodified bacteriocins, IIc, formerly class V, where the N- and C-termini are covalently linked resulting in a cyclic structure, and class IId, non-pediocin, single, linear peptides [2]. The genetics and biosynthesis of class IIa bacteriocins have been well studied [3], and these constitute one of the most important groups of antimicrobial peptides, due to their useful antibacterial properties. All known IIa bacteriocins are described as being active against *Listeria* and some have already been tested as food preservatives for controlling food-borne pathogens [4]. Structurally, class IIa bacteriocins are related to each other being unstructured in aqueous solution, but with a central amphiphilic alpha-helical region when in lipid micelles or TFE [5-7]; they contain the characteristic conserved N-terminal YGNGVxCxxxxC sequence, though usually not the GxxxG motif(s) characteristic of IIb and IIc bacteriocins [7].

The class IIb two-peptide unmodified bacteriocins, for example plantaricin E/F [8], need the complementary action of both peptides to be active [9]. These bacteriocins contain long amphiphilic alpha-helical stretches, and the two complementary peptides interact when exposed to membrane-like entities. The GxxxG motif is conserved in many two-peptide bacteriocins, and it is postulated that the two complementary peptides dimerize via a helix-helix interaction that involves this motif, to form the functionally active heterodimer [10]. The dimer functions by creating a pore within the membrane through which small molecules leak out, and, typically, the genes encoding the two peptides are found adjacent to each other on the same operon [10].

The cyclic class IIc bacteriocins are characterized by being tryptophan-rich and lacking any GG, GxxxG or YGNGVxCxxxxC motifs, [11]; the class IId bacteriocins have none of these features but some have recently been found also to be circular, rather than linear, with conserved AxxhhN and AhhW/F motifs [12-14].

In order to neutralize the toxic effect of the peptide on the "producing" cell the genes encoding bacteriocins are generally co-transcribed with a cognate immunity protein. These small proteins (typically 88-115 amino acids) interact very tightly with a specific bacteriocin or pair of bacteriocins and protect the "producing" microbe from the toxic effect of its own bacteriocin [15,16]. The immunity proteins usually show high specificity for their cognate bacteriocins [17,18]. For each class IIa bacteriocin encoded in a genome there is a one to one relationship between bacteriocin and cognate immunity protein; whereas, in contrast, for each pair of class IIb two-peptide bacteriocins there is a single cognate immunity protein encoded in a genome [9]. Structures of five immunity proteins have already been solved: ImB2 [19], EntA-im [20], PedB [21], PisI [22], and Mun-Im [23], all protective against IIa bacteriocins. As yet, no structures for immunity proteins protective against IIb or IIc bacteriocins have been solved.

The sequencing of bacterial genomes, including those from human pathogens, has revealed a number of genes which might potentially code for new bacteriocins and immunity proteins, suggesting that the use of these antimicrobial peptides is more widespread than previously thought and that bacteria might be targeting several different bacterial species using these toxins. Understanding which specific immunity protein neutralizes which bacteriocin toxin is important if these peptides are to be used as antimicrobials.

The Gram-positive bacterium *Streptococcus pyogenes*, closely related to LAB, is one of the most common human pathogens. It causes a wide range of both minor diseases such as pharyngitis, erysipelas and pyodermas, that are readily controlled by antibiotics, as well as major, often lethal, conditions such as acute rheumatic fever, necrotizing fasciitis and streptococcal toxic shock syndrome, in developing countries and in the western world [24]. The search for new antibacterial agents effective against this species is thus urgent.

Other streptococcal species have been shown to secrete bacteriocin-like toxins, as there are reports of *S. salivarius* producing a variety of bacteriocin-like inhibitory substances showing *in vitro* inhibitory activity against *S. pyogenes*, including Salivaricin A [25,26]. Such observations suggest that antibacterial toxins are playing a very important role in controlling the level of the *S. pyogenes* population in human microbiomes. Bacteriocin-like toxins and antitoxins may well have an impact on the development of new antibacterial strategies and treatments. A thorough understanding of the biology of the bacteriocins in combination with their immunity proteins is important for any possible therapeutic use of bacteriocins. The full

sequences of LAB genomes provide opportunities to scan for the presence of toxins and their corresponding immunity proteins. However, the sequence alone may not be sufficient to identify these proteins.

Here we present the first structure of the protein from locus *Sp*_{py}_2152 (gene names taken from *S. pyogenes* M1 GAS), named pyogenecin immunity protein Sp-PIP, determined at 2.15 Å resolution. We provide structural and sequence analyses, and identify the putative corresponding bacteriocin-like toxins in the *S. pyogenes* genome, which are found to belong to the class IIb two-peptide bacteriocins.

Results

Structure determination

We have expressed the pyogenecin immunity protein from *S. pyogenes* in *Escherichia coli* and purified it to homogeneity. The crystal structure of Sp-PIP was solved by the single wavelength anomalous diffraction (SAD) method using selenomethionine (SeMet)-substituted protein. The model was refined to 2.15 Å with an R-factor of 15.7% and an R-free of 25.4%. The crystal asymmetric unit contains 78 (out of the 102) well ordered residues and 74 water molecules. The first methionine and the 23 C-terminal residues are disordered and cannot be identified in the electron density maps. The structure is high quality; all

non-glycine and non-proline residues of the model lie either in the most favorable region (95.9% of residues) or in the additionally allowed region (4.1% of residues) of the Ramachandran plot. Detailed refinement statistics and crystallographic data are shown in Table 1. The crystal structure of Sp-PIP, shown in Figure 1, is a typical anti-parallel four-helix bundle. The four long helices (H1, residues 3-16; H3, 23-41; H4, 43-61; H5, 66-79) pack tightly together around a well-defined hydrophobic core. One additional short helix (H2, residues 17-22) connects helices H1 and H3. All long helices are between 14 and 17 residues in length. Helix-pairs H1/H4 and H2/H3 are parallel to each other and helix-pairs H1/H2 and H3/H4 cross over each other at an angle of approximately 30°. In the structure, the N- and C-termini are very close together (7 Å). Overall, the protein is slightly acidic (calculated pI = 6.06) with two large acidic patches near each end of the bundle and one large positively charged patch near the N- and C-termini.

Structural analysis

Structural comparisons, using the DALI server [27], reveal that the structure of Sp-PIP matches a range of proteins all having a four helix bundle topology, with a Z-score of 8.7 or less. In this range of Z-scores, it is not clear whether the matches are due to homology or simply to structural similarity. However, structural homology combined with

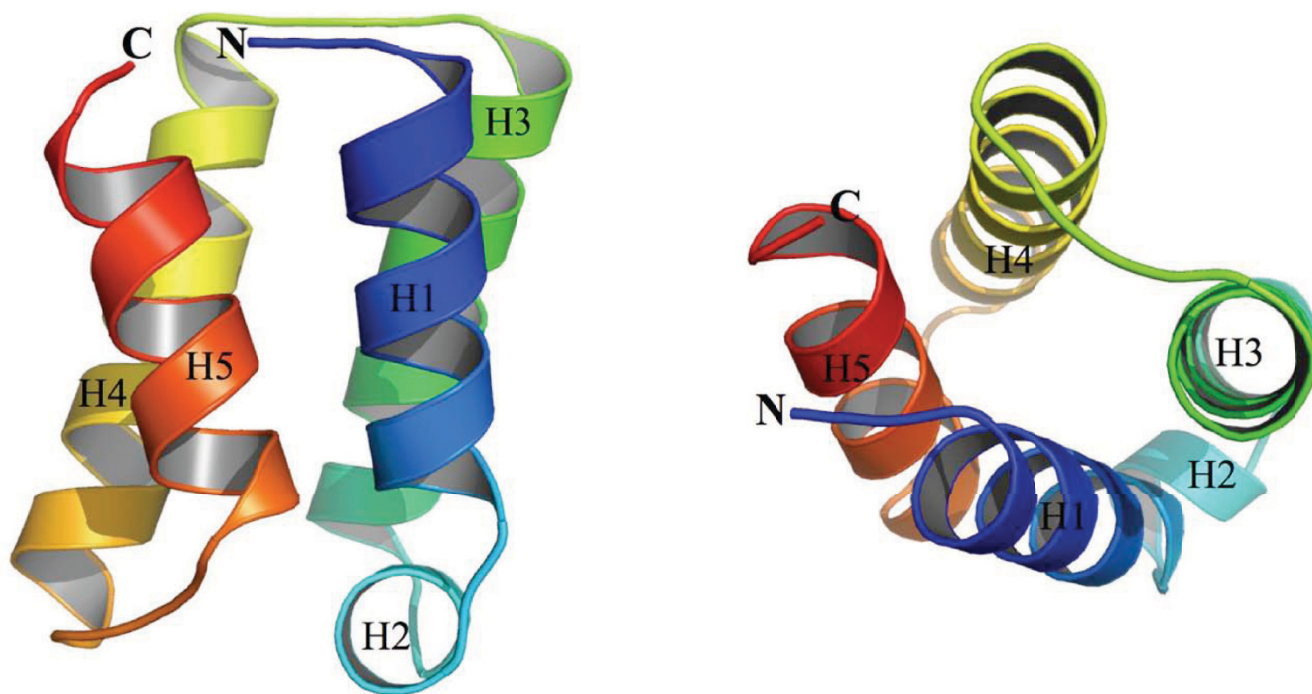


Figure 1

Structure of 2fu2. The structure of Sp-PIP - PDB ID: [2fu2](#) - shown as a four helix bundle in two orthogonal views. Both N- and C-termini of proteins are labeled and the protein is colored from blue (N-terminus) to red (C-terminus).

Table 1: Crystallographic results. The crystallographic data and refinement statistics for the structure 2fu2.

Data collection statistics	
Wavelength (Å)	0.9791
Space group	C2
Cell parameters(Å, °)	$a = 76.09$ $b = 30.28$ $c = 35.86$ $\beta = 113.26$
Resolution (Å)	50-2.15
Total reflections	82,574
Unique reflections	4,042
Completeness (%)	96.3(83.0)
I/sigma	34.3(7.4)
R _{merge} (%)	8.1(22.9)
Refinement statistics	
Resolution range (Å)	50-2.15
R _{cryst} (%)	15.7
R _{free} (%)	25.4
R.m.s. deviations from ideality	
Bond lengths (Å)	0.026
Angles (°)	1.81
No. of protein atoms	627
No. of solvent atoms	67
Average B-factor	
main-chain atoms (Å ²)	33.7
Side-chain atoms (Å ²)	35.4
Protein atoms (Å ²)	34.6
Solvent atoms (Å ²)	48.0

sequence analysis, enabled us to relate this protein to structures that ranked 2nd (1tdp), 10th (2bl8) and 23rd (2k19) in the list of DALI matches, and the results are shown in Table 2. These homologous proteins are all pediocin-like immunity proteins. Sp-PIP shares a RMSD of 2.4 Å with ImB2, 1tdp, over 75 residues, and, using the structure comparison service SSM at European Bioinformatics Institute <http://www.ebi.ac.uk/msd-srv/ssm>[28], the RMSD is found to be 2.11 Å.

Sequence analysis

Guided by the structure of Sp-PIP, whose sequence had not been assigned to any known protein family, we sought to expand our knowledge of its sequence relatives using the Pfam database [29]. Having identified that it was structurally similar to known members of the Pfam EntA_Immun family (Pfam:PF08951) we carried out an analysis with profile Hidden Markov models (HMMs) to determine whether Sp-PIP should also belong to this family or not. Taking the EntA_Immun family from Pfam release 22.0 as a starting point we carried out iterative searches using the HMMER package (v2.3.2). After multiple rounds of searching of the HMM against the sequence database (UniProt version 12.5) using an E-value threshold of 0.04 along with careful manual inspection of the resulting matches we were able to detect 172 sequences compared to the 19 sequences found in Pfam release 22.0. The family has been updated in the current release of Pfam - 24.0 - and the alignment is now available on the Pfam web-site at: <http://pfam.sanger.ac.uk/family/Pf08951>.

Domain analysis

We investigated whether any other protein domains were to be found on any immunity proteins in the EntA_Immun family. Some members of the family do also carry associated domains. One protein contains an N-terminal helix-turn-helix transcriptional regulator (UniProt:Q88Y45), and a group of methionine sulfoxide reductases from *Enterococcus faecium* (UniProt:Q3Y319) and closely related species carry a C-terminal PMSR domain - peptide methionine sulfoxide reductase (Pfam:PF01625) - as well as the N-terminal immunity protein domain.

Proteome analysis

Sp-PIP belongs to a large family of putative bacteriocin immunity proteins. To investigate the possible targets of the presumed immunity protein we carried out an analysis of the proteomes of several strains of *S. pyogenes*, including M1 GAS, MGAS10750, and MGAS8232, to try to identify any complementary bacteriocins and immunity proteins. We searched the proteomes with HMMs of both the known class I, antibacterial18 (Pfam:PF08130), and class II bacteriocins, bacteriocin_II (Pfam:PF01721),

Table 2: DALI comparison of Sp-PIP with other immunity proteins.

Structure with PDB ID	Z-score	RMSD Å	# residues
<u>1tdp</u> -- Carnobacteriocin immunity protein, ImB2	8.3	2.4	75
<u>2bl8</u> -- Enterocin A immunity protein, EntA-im	7.3	3.0	72
<u>2k19</u> -- Piscicolin 126 immunity protein, PisI	6.9	3.0	75

DALI scores for structures representative of groups A, B and C immunity proteins against Sp-PIP.

bacteriocin_IIC (Pfam:PF10439) as well as with the EntA_Immun families, from the Pfam database (release 23.0). We have identified a small genomic region containing four genes encoding putative bacteriocin proteins of class IIb (Spy_0479, Spy_0480, Spy_0484 and Spy_0486), and have called the proteins Pyogenecin 1, 2, 3, and 4, respectively (UniProt:Q9A139, Q9A138, Q9A137, Q9A136). The organization of this region is shown in Figure 2a. We have also identified two separate gene loci encoding immunity proteins in the *S. pyogenes* genome: Spy_2152 (Sp-PIP) and Spy_1675. Both encoded proteins were found in all the *S. pyogenes* strains examined, and are identified as belonging to the EntA_Immun family (Pfam:PF08951), a family which contains nearly 200 different members from LAB and closely related species.

Discussion

The four-helix bundle structure of Sp-PIP was found to be most closely related to those of a number of pediocin-like IIA immunity proteins. Three subgroups of pediocin-like IIA immunity proteins have been defined on the basis of common sequence motifs and phylogenetic analysis [17], being denoted as groups A, B and C, as shown in Figure 3:

- Group A: EntA-im, the enterocin A immunity protein, [20]; PDB ID: 2bl8
- Group B: PisI, the piscicolin 126 immunity protein [22]; PDB ID: 2k19
- Group C: ImB2, the carnobacteriocin B2 immunity protein [19]; PDB ID: 1tdp

The structural comparison results suggest that Sp-PIP is most closely related to Imb2, from subgroup C. These results suggest that Sp-PIP is highly likely to function as a bacteriocin immunity protein with similarities to proteins from group C.

Sequence analysis reveals that Sp-PIP belongs to the now expanded Pfam family of immunity proteins EntA_Immun that also includes pediocin-like immunity proteins from groups A, B and C as detailed above, thus confirming its evolutionary link to known bacteriocin immunity proteins. A search for other immunity proteins of a similar nature in the *S. pyogenes* genome identified another gene locus encoding a protein matching the



Figure 2
Alignment of representative immunity proteins. Multiple sequence alignment of representative immunity proteins from groups A, B and C colored using CHROMA software (Goodstadt and Ponting 2001).

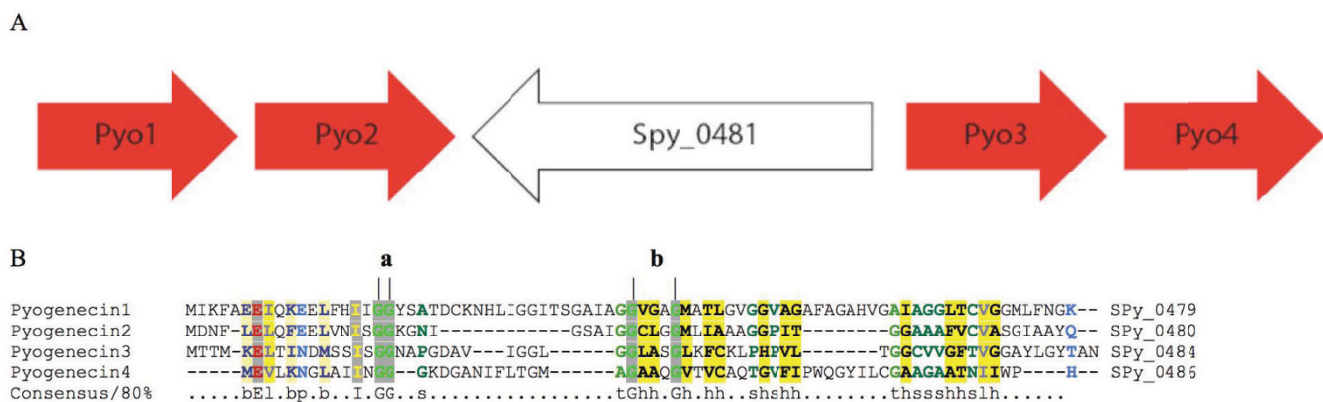


Figure 3
Streptococcus pyogenes bacteriocin cluster. (A) Organization of the bacteriocin locus of *S. pyogenes*, containing four peptides shown by red arrows, Pyogenecin I-4 as Pyo I-4; (B) Multiple sequence alignment of putative bacteriocins using T-Coffee (Notredame *et al.*, 2000), colored using CHROMA software (Goodstadt and Ponting 2001). The GG "double-glycine" leader peptide motif is indicated at (a) and the GxxxG helix-interacting motif at (b).

EntA_Immun family. Both of these loci were seen in all strains examined. Unlike the case in many other species, in *S. pyogenes* these bacteriocin immunity proteins were not found closely linked to their bacteriocins, a condition that might be explained by the remnants of a transposon identified between them.

Functions for proteins may often be inferred through consideration of the combinations of the functional domains found on them. A number of domain-pairings were identified on some of the immunity family members, such as an N-terminal helix-turn-helix transcriptional regulator on one protein, which might be suggestive of immunity proteins acting as transcriptional regulators or of interacting with transcription regulators and modulating their function. A group of methionine sulfoxide reductases are found to carry both a C-terminal PMSR domain and the immunity domain. The PMSR domain is involved in reversing protein inactivation by oxidation of methionine residues. However, it is not clear what, if any, functional significance these domain-pairings might have.

The search for potential bacteriocin targets for this putative immunity protein took in classes I, II and III. Our analyses identified that *S. pyogenes* contains a class I lantibiotic system composed of one unique lantibiotic type A sequence called *srtA* and just one unique immunity protein *srtI*. These two genes are closely associated on the genome, but are separated from the proteins of the class II system described below as well as from the presumed immunity proteins. This suggests that the Sp-PIP is unlikely to act as an immunity protein for class I bacteriocins. Class II bacteriocins are subdivided into classes IIa, IIb, IIc and IID. We did not find any bacteriocins of class IIa encoded by the *S. pyogenes* genome, which would have

the well-conserved N-terminal YGNGVxCxxxC motif [30,31]. Searches were not made for bacteriocins of types IIc or IID since the immunity proteins complementary for these two groups identified so far, gassericin A [32] and circularin A, [33] do not share sequence or size similarity with those of the pediocin-like immunity proteins. The four putative class IIb proteins we identified from a short genomic region, and named pyogenecins 1-4, are found as two tandem pairs of genes separated by a gene in the antisense orientation that encodes a transposase pseudogene. The presence of this pseudogene suggests that this region might once have been part of a mobile genetic element. Finding bacteriocins in pairs is characteristic of the class IIb bacteriocins [34]. As with other class IIb bacteriocins, these pyogenecins have a conserved GG leader peptide, as well as the conserved GXXXG motif necessary for helix-helix interaction between the two proteins, as illustrated in Figure 2b. All four putative pyogenecins are found to fall into the Pfam Bacteriocin_IIc family (Pfam:PF10439), a family of bacteriocins secreted by streptococcal and other LAB species.

Based on the sequence similarity to other two-peptide bacteriocins, the genomic arrangement of the genes found, and the identification of two immunity proteins, we hypothesize that the putative pyogenecins 1 and 2 associate to form one active bacteriocin and pyogenecins 3 and 4 associate to form a second active bacteriocin possibly targeting a different bacterial species. These peptides need now to be synthesized to allow testing of their biological activity in a suitable bacteriocin-assay. Given that we find two immunity proteins we suggest that each one is specific for one of the putative bacteriocin pairs, Pyogenecin1/Pyogenecin2 and Pyogenecin3/Pyogenecin4.

There is also the possibility that one or both of these immunity proteins might be an orphan immunity protein [17], so-called because no cognate bacteriocin in the genome has been identified; and it has been pointed out that most other immunity proteins so far determined for Ila two-peptide bacteriocins are small trans-membrane proteins closely coupled transcriptionally with their bacteriocins [9]. However, it would appear that the sequences of only a handful of two-peptide immunity proteins have been deduced and neither are there defined structures for any of these nor are most of them more than putative immunity proteins, the designation being based only on the presence of predicted transmembrane helices [35-37], so comparison with the system we propose here is difficult. The exact method by which immunity proteins protect their host cell or interact with the bacteriocins is also unclear, but it may be that the orphan immunity proteins do confer resistance to other bacteriocins in addition to their cognate bacteriocin [17,9]. It is possible of course that many more two-peptide-like systems remain still to be characterized, as there are a large number of additional LAB peptides in the Pfam family Bacteriocin_IIC PF10439 which are "pairs" in the sense that they are expressed from closely adjacent genes and for which a cognate immunity protein has not yet been identified.

Conclusions

The structure of the putative immunity protein Sp-PIP protective against potential class IIB two-peptide bacteriocins could be the first structural representative of this class. A combined structural, genomic and proteomic analyses has allowed the identification and *in silico* characterization of a new putative immunity protein from *S. pyogenes*. Structurally similar proteins seem to provide immunity protection for single peptide Ila class bacteriocins. Further biological and biochemical studies are needed to verify the antibacterial activity of the putative pyogenecins and to determine the degree of resistance and cross-resistance provided by the proposed immunity protein.

Methods

Protein cloning, expression and purification

The ORF from the Spy_2152 locus was amplified from genomic *S. pyogenes* M1 DNA with KOD DNA polymerase using conditions and reagents provided by the vendor (Novagen, Madison, WI). The gene was cloned into the pMCSG7 vector using a modified ligation-independent cloning protocol [38] and over-expressed in *E. coli* BL21 (DE3) - Gold (Stratagene) harboring an extra plasmid encoding three rare tRNAs (AGG and AGA for Arg, ATA for Ile). The pMCSG7 vector bearing a tobacco etch virus (TEV) protease cleavage site creates a construct with a cleavable His₆-tag fused onto the N-terminus of the target protein and adds three artificial residues (Ser-Asn-Ala) on

that end. The cells were grown using SeMet-containing enriched M9 medium and conditions known to inhibit methionine biosynthesis. The cells were grown at 37 °C to an OD₆₀₀ of ~0.6 and protein expression induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG). After induction, the cells were grown overnight with shaking at 20 °C. The harvested cells were re-suspended in lysis buffer (50 mM HEPES pH 8.0, 500 mM NaCl, 10 mM imidazole, 10 mM β-mercaptoethanol, and 5% v/v glycerol) in the presence of lysozyme (1 mg/mL) and of protease inhibitor cocktail (100 μL per 2 g of wet cells), and then kept on ice for 20 min before sonication. Then the cells were lysed using sonication. The lysate was clarified by centrifugation at 30,000 × g (RC5C-Plus centrifuge, Sorval) for 20 min, followed by filtration through 0.45 μm and 0.22 μm in-line filters (Gelman). The lysate was then applied to a 5 mL HiTrap Ni-NTA column on the AKTExpress (GE Health Systems). His₆-tagged protein was eluted using buffer containing a higher concentration of imidazole (500 mM NaCl, 5% glycerol, 50 mM HEPES, pH 8.0, 250 mM imidazole, 10 mM 2-mercaptoethanol), and the His₆-tag was cleaved from the protein by treatment with recombinant His₆-tagged TEV protease. A second Ni-NTA affinity chromatography was performed manually to remove the His₆-tag and His₆-tagged TEV protease. A total of 90 mg of protein was purified for finding crystallization conditions. Then the protein was dialyzed in 20 mM Tris-HCl pH 7.1, 50 mM NaCl, 2 mM dithiothreitol (DTT) and concentrated to 110 mg/mL using a Centricon Plus-20 Centrifugal Concentrator (Millipore). The standard purification protocol has been thoroughly described previously [39].

Protein crystallization

The initial crystallization conditions were searched using the sitting drops vapor diffusion method at 18 °C with the help of the Mosquito crystallization workstation (TTP Labtech) with Index (Hampton Research) and Wizard I and II (Emerald Biostructures) crystallization screens. Crystals suitable for X-ray experiment were obtained at the initial stage of crystal screening, the crystallization condition number 23 of Wizard II within several days. Crystals were flash-frozen in liquid nitrogen with crystallization solution containing 25% (v/v) glycerol as cryoprotectant prior to data collection.

Data collection and structure determination

Anomalous diffraction data were collected at the selenium peak from crystals of a SeMet-substituted protein. The data sets were collected on an ADSC quantum Q315 charge-coupled device-detector at 100 K on the Structural Biology Center beamline 19ID at the Advanced Photon Source, Argonne National Laboratory. The space group was found to be C2, with cell parameters of $a = 76.09 \text{ \AA}$, $b = 30.28 \text{ \AA}$, $c = 35.86 \text{ \AA}$, $\alpha = 90.00^\circ$, $\beta = 113.26^\circ$, and $\gamma =$

90.00°. The diffraction data were processed using the HKL3000 suite of programs [40].

All procedures for single wavelength anomalous dispersion (SAD), phasing, phase improvement by density modification, and initial protein model building were done using the structure module of the HKL3000 software package. The mean figure of merit of the phase-set was 0.207 for 50-2.25 Å data and improved to 0.542 after density modification (DM). The autotracing Arp/wArp module [41] in HKL3000 built 77 out of 102 residues with fitted sequence. The initial model was rebuilt with the program COOT [42] using electron density maps based on DM-phased reflection file. After each cycle of rebuilding, the model was refined using REFMAC5 [43] from the CCP4 suite with TLS refinement. The stereochemistry of the structure was checked with PROCHECK [44]. Atomic coordinates and experimental structure factors of Sp-PIP have been deposited with the PDB database and are accessible under the code [2fu2](#).

Authors' contributions

CC, LV, MC performed the crystallography part of the research; AB, PC and RF performed the genome sequence searches; PC, AB, ZO, WM and AJ analyzed data; AJ designed the research; PC, AB and AJ wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors wish to thank members of The Structural Biology Center at Argonne National Laboratory for their help with data collection at the 19ID beamline. We also thank Lindsey Butler for help in preparation of this manuscript. This work was supported by National Institutes of Health grant GM074942 and by the U. S. Department of Energy, Office of Biological and Environmental Research, under contract DE-AC02-06CH11357. PC, AB and RDF are funded by the Wellcome Trust and this work was supported by the Wellcome Trust [grant number WT077044/Z/05/Z]. The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory ("Argonne"). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.

References

1. Drider D, Fimland G, et al.: "The continuing story of class IIa bacteriocins". *Microbiol Mol Biol Rev* 2006, **70(2)**:564-82.
2. Cotter PD, Hill C, et al.: "Bacteriocins: developing innate immunity for food". *Nat Rev Microbiol* 2005, **3(10)**:777-88.
3. Ennahar S, Sashihara T, et al.: "Class IIa bacteriocins: biosynthesis, structure and activity". *FEMS Microbiol Rev* 2000, **24(1)**:85-106.
4. Galvez A, Abriouel H, et al.: "Bacteriocin-based strategies for food biopreservation". *Int J Food Microbiol* 2007, **120(1-2)**:51-70.
5. Tichaczek PS, Vogel RF, et al.: "Cloning and sequencing of curA encoding curvacin A, the bacteriocin produced by *Lactobacillus curvatus* LTH1174". *Arch Microbiol* 1993, **160(4)**:279-83.
6. Uteng M, Hauge HH, et al.: "Three-dimensional structure in lipid micelles of the pediocin-like antimicrobial peptide sakacin P and a sakacin P variant that is structurally stabilized by an inserted C-terminal disulfide bridge". *Biochemistry* 2003, **42(39)**:11417-26.
7. Haugen HS, Fimland G, et al.: "Three-dimensional structure in lipid micelles of the pediocin-like antimicrobial peptide curvacin A". *Biochemistry* 2005, **44(49)**:16149-57.
8. Fimland N, Rogne P, et al.: "Three-dimensional structure of the two peptides that constitute the two-peptide bacteriocin plantaricin EF". *Biochim Biophys Acta* 2008, **1784(11)**:1711-9.
9. Oppegard C, Rogne P, et al.: "The two-peptide class II bacteriocins: structure, production, and mode of action". *J Mol Microbiol Biotechnol* 2007, **13(4)**:210-9.
10. Rogne P, Fimland G, et al.: "Three-dimensional structure of the two peptides that constitute the two-peptide bacteriocin lactococcin G". *Biochim Biophys Acta* 2008, **1784(3)**:543-54.
11. Leer RJ, Vossen JM van der, et al.: "Genetic analysis of acidocin B, a novel bacteriocin produced by *Lactobacillus acidophilus*". *Microbiology* 1995, **141(Pt 7)**:1629-35.
12. Kemperman R, Kuipers A, et al.: "Identification and characterization of two novel clostridial bacteriocins, circularin A and closticin 574". *Appl Environ Microbiol* 2003, **69(3)**:1589-97.
13. Wirawan RE, Swanson KM, et al.: "Uberolysin: a novel cyclic bacteriocin produced by *Streptococcus uberis*". *Microbiology* 2007, **153(Pt 5)**:1619-30.
14. Maqueda M, Sanchez-Hidalgo M, et al.: "Genetic features of circular bacteriocins produced by Gram-positive bacteria". *FEMS Microbiol Rev* 2008, **32(1)**:2-22.
15. Axelsson L, Holck A: "The genes involved in production of and immunity to sakacin A, a bacteriocin from *Lactobacillus sake* Lb706". *J Bacteriol* 1995, **177(8)**:2125-37.
16. Quadri LE, Kleerebezem M, et al.: "Characterization of a locus from *Carnobacterium piscicola* LV17B involved in bacteriocin production and immunity: evidence for global inducer-mediated transcriptional regulation". *J Bacteriol* 1997, **179(19)**:6163-71.
17. Fimland G, Eijsink VG, et al.: "Comparative studies of immunity proteins of pediocin-like bacteriocins". *Microbiology* 2002, **148(Pt 11)**:3661-70.
18. Johnsen L, Fimland G, et al.: "The C-terminal domain of pediocin-like antimicrobial peptides (class IIa bacteriocins) is involved in specific recognition of the C-terminal part of cognate immunity proteins and in determining the antimicrobial spectrum". *J Biol Chem* 2005, **280(10)**:9243-50.
19. Sprules T, Kawulka KE, et al.: "NMR solution structure of ImB2, a protein conferring immunity to antimicrobial activity of the type IIa bacteriocin, carnobacteriocin B2". *Biochemistry* 2004, **43(37)**:11740-9.
20. Johnsen L, Dalhus B, et al.: "1.6-Angstroms crystal structure of EntA-im. A bacterial immunity protein conferring immunity to the antimicrobial activity of the pediocin-like bacteriocin enterocin A". *J Biol Chem* 2005, **280(19)**:19045-50.
21. Kim IK, Kim MK, et al.: "High resolution crystal structure of PedB: a structural basis for the classification of pediocin-like immunity proteins". *BMC Struct Biol* 2007, **7**:35.
22. Martin-Visscher LA, Sprules T, et al.: "Nuclear magnetic resonance solution structure of Pisl, a group B immunity protein that provides protection against the type IIa bacteriocin piscicolin 126, PisA". *Biochemistry* 2008, **47(24)**:6427-36.
23. Jeon HJ, Noda M, et al.: "Crystal structure and mutagenic analysis of a bacteriocin immunity protein, Mun-im". *Biochem Biophys Res Commun* 2009, **378(3)**:574-8.
24. Cunningham MW: "Pathogenesis of group A streptococcal infections". *Clin Microbiol Rev* 2000, **13(3)**:470-511.
25. Robson CL, Wescombe PA, et al.: "Isolation and partial characterization of the *Streptococcus mutans* type AII lantibiotic mutacin K8". *Microbiology* 2007, **153(Pt 5)**:1631-41.
26. Lee SV, Mitchell DA, et al.: "Discovery of a widely distributed toxin biosynthetic gene cluster". *Proc Natl Acad Sci USA* 2008, **105(15)**:5879-84.
27. Holm L, Park J: "DaliLite workbench for protein structure comparison". *Bioinformatics* 2000, **16(6)**:566-7.
28. Krissinel E, Henrick K: "Detection of protein assemblies in crystals". In *Computational Life Sciences: Proceedings of the First Interna-*

- tional Symposium, *Complife* Berlin Heidelberg: Springer-Verlag; 2005:163-174.
29. Finn RD, Tate J, et al.: "**The Pfam protein families database.**". *Nucleic Acids Res* 2008:D281-8.
 30. Cintas LM, Casaus P, et al.: "**Biochemical and genetic characterization of enterocin P, a novel sec-dependent bacteriocin from *Enterococcus faecium* PI3 with a broad antimicrobial spectrum.**". *Appl Environ Microbiol* 1997, **63(11)**:4321-30.
 31. Bennik MH, Vanloo B, et al.: "**A novel bacteriocin with a YGNGV motif from vegetable-associated *Enterococcus mundtii*: full characterization and interaction with target organisms.**". *Biochim Biophys Acta* 1998, **1373(1)**:47-58.
 32. Kawai Y, Kusnadi J, et al.: "**DNA sequencing and homologous expression of a small peptide conferring immunity to gassericin A, a circular bacteriocin produced by *Lactobacillus gasseri* LA39.**". *Appl Environ Microbiol* 2009, **75(5)**:1324-30.
 33. Kemperman R, Jonker M, et al.: "**Functional analysis of the gene cluster involved in production of the bacteriocin circularin A by *Clostridium beijerinckii* ATCC 25752.**". *Appl Environ Microbiol* 2003, **69(10)**:5839-48.
 34. Martinez B, Fernandez M, et al.: "**Synthesis of lactococcin 972, a bacteriocin produced by *Lactococcus lactis* IPLA 972, depends on the expression of a plasmid-encoded bicistronic operon.**". *Microbiology* 1999, **145(Pt 11)**:3155-61.
 35. Marciset O, Jeronimus-Stratingh MC, et al.: "**Thermophilin 13, a nontypical antilisterial poration complex bacteriocin, that functions without a receptor.**". *J Biol Chem* 1997, **272(22)**:14277-84.
 36. McCormick JK, Poon A, et al.: "**Genetic characterization and heterologous expression of brochocin-C, an antibotulinal, two-peptide bacteriocin produced by *Brochothrix campestris* ATCC 43754.**". *Appl Environ Microbiol* 1998, **64(12)**:4757-66.
 37. Stephens SK, Floriano B, et al.: "**Molecular analysis of the locus responsible for production of plantaricin S, a two-peptide bacteriocin produced by *Lactobacillus plantarum* LPCO10.**". *Appl Environ Microbiol* 1998, **64(5)**:1871-7.
 38. Dieckman L, Gu M, et al.: "**High-throughput methods for gene cloning and expression.**". *Protein Expr Purif* 2002, **25(1)**:1-7.
 39. Kim Y, Dementieva I, et al.: "**Automation of protein purification for structural genomics.**". *J Struct Funct Genomics* 2004, **5(1-2)**:111-8.
 40. Minor W, Cymborowski M, et al.: "**HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes.**". *Acta Crystallogr D Biol Crystallogr* 2006, **62(Pt 8)**:859-66.
 41. Perrakis A, Morris R, et al.: "**Automated protein model building combined with iterative structure refinement.**". *Nat Struct Biol* 1999, **6(5)**:458-63.
 42. Emsley P, Cowtan K: "**Coot: Model-Building Tools for Molecular Graphics.**". *Acta Crystallographica Section D - Biological Crystallography* 2004, **60(Pt 12 Pt 1)**:2126-32.
 43. Murshudov GN, Vagin AA, et al.: "**Refinement of macromolecular structures by the maximum-likelihood method.**". *Acta Crystallogr D Biol Crystallogr* 1997, **53(Pt 3)**:240-55.
 44. Laskowski RA, MacArthur MW, et al.: "**PROCHECK: a program to check the stereochemical quality of protein structures.**". *J Apple Cryst* 1993, **26**:283-291.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

