Antonie van Leeuwenhoek **78:** 379–384, 2000.

Teicoplanin biosynthesis genes in *Actinoplanes teichomyceticus*

Margherita Sosio*, Alessandra Bianchi, Elena Bossi & Stefano Donadio
*Biosearch Italia, via R. Lepetit 34, 21040 Gerenzano (VA), Italy. (*Author for correspondence; E-mail: msosio@biosearch.it)*

Key words: antibiotics, glycopeptides, peptide synthetases, resistance, vancomycin

Abstract

The genetic determinants for the biosynthesis of the glycopeptide antibiotic teicoplanin were identified. In order to isolate the corresponding gene cluster, oligonucleotides derived from highly conserved motifs in peptide synthetases were used. These synthetic probes, and gene fragments derived from the balhimycin gene cluster of *Amycolatopsis mediterranei*, led to the identification of the likely teicoplanin gene cluster centered on a region of ca. 110 kb from the genome of *Actinoplanes teichomyceticus*, the teicoplanin producer. Partial nucleotide sequences identified partial ORFs likely to encode two glycosyltransferases, three P-450 monooxygenases and one ABC transporter. The corresponding genes have been found in other glycopeptide gene clusters. Furthermore, upstream to the peptide synthetase region a segment was identified with a remarkable similarity to the *vanHAX* operon, conferring resistance to glycopeptides in enterococci. Thus, in contrast to the other glycopeptide producers thus far analyzed, in *A. teichomyceticus* the genes for teicoplanin biosynthesis are closely linked to homologs of glycopeptide resistance commonly found in vancomycin-resistant enterococci.

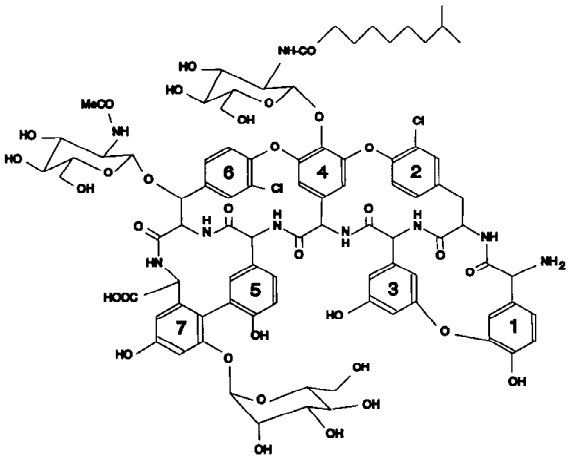

Figure 1. Structure of teicoplanin. The figure reports the structure of factor A2 (Lancini 1989).

Introduction

Teicoplanin (Figure 1) is a glycopeptide antibiotic produced by *Actinoplanes teichomyceticus*. It belongs to the therapeutically important antibiotics of the vancomycin group with which it shares the mechanism of action and the chemical class (Lancini 1989). A heptapeptide backbone, with five common amino acids, is the conserved feature to all members of this group of antibiotics. Differences are frequently found at residues 1 and 3 and in the number and position of various substituents: teicoplanin presents a *p*-hydroxyphenylglycine and a 3,5-dihydroxyphenylglycine at positions 1 and 3; a chlorine atom on each of the tyrosine residues; and three sugar moieties, a D-mannose, an *N*-acetyl-β-D-glucosamine and an *N*-acyl-β-D-glucosamine (Lancini & Cavalleri 1997).

As for vancomycin, the mechanism of action of teicoplanin consists of inhibiting cell-wall assembly by binding to the terminal D-Ala-D-Ala of the muramyl pentapeptide intermediate of peptidoglycan synthesis (Lancini & Cavalleri 1997). Teicoplanin and vancomycin are active against Gram-positive bacteria, including the major pathogens such as enterococci, streptococci and staphylococci; they are often 'last choice' drugs for life-threatening infections caused by pathogens resistant to most other antibiotics. During the last decade, the occurrence of glycopeptide resistance in enteroccoci, the so-called VRE strains

for vancomycin-resistant enterococci, was observed and has been increasing. Recently, some isolates of methicillin resistant *Staphylococcus aureus* with decreased susceptibility to glycopeptides have also been observed (John 1999).

The increasing emergence of resistance to glycopeptides has prompted the search for more effective derivatives of this class of antibiotics. The identification of the genes responsible for the biosynthesis of the glycopeptides chloroeremomycin and balhimycin (Solenberg et al. 1997; van Wageningen et al. 1998; Pelzer et al. 1999) and the selective expression or disruption of some of these genes has been successfully exploited to produce hybrid glycopeptides (Solenberg et al. 1997) or pathway intermediates (Pelzer et al. 1999). These results are particularly significant since no pathway intermediates had been previously isolated from glycopeptide producing strains (Lancini & Cavalleri 1997). Thus, a contribution to the generation of improved glycopeptides can also derive from the genetic manipulation of the biosynthesis genes. The isolation and preliminary characterization of the teicoplanin gene cluster is presented here.

Identification of the teicoplanin gene cluster

The biosynthesis of many peptide antibiotics occurs on a protein template through the action of non-ribosomal peptide synthetases (NRPS). These enzymes present a modular organization, where a different module is used for recognition and activation, covalent binding and amide bond formation for each amino acid incorporated into the peptide. Three distinct domains within each module carry out these reactions (Marahiel 1997). The nature and number of the amino acids present in the teicoplanin peptide skeleton suggests that this antibiotic is synthesized by NRPS(s) consisting of seven modules. Recent genetic evidence from the chloroeremomycin (*cep*; van Wageningen et al. 1998) and balhimycin (*bal*; Pelzer et al. 1999; W. Wohlleben, pers. comm.) clusters indicates that glycopeptide formation occurs by the non-ribosomal route.

In order to identify the teicoplanin biosynthesis gene cluster, we took advantage of two oligo probes, Pep6 and Pep8, derived from highly conserved motifs present in NRPSs (Marahiel 1997) and designed for the identification of the corresponding genes in actinomycetes (Sosio et al. 2000). From an *A.teichomyceticus* library, 37 cosmids positive to both probes were identified. Restriction mapping and hybridization profiling allowed the grouping of ten of these cosmids into three different, unlinked contigs, encoding likely NRPS modules (data not shown). The other positive cosmids have not been organized into contigs yet. As seen with other actinomycetes (Sosio et al. 2000), *A. teichomyceticus* contains many more DNA fragments hybridizing to Pep6 and Pep8 than one would expect if this strain had the genetic potential to make only teicoplanin via the non-ribosomal system.

In order to identify the teicoplanin biosynthesis genes among the multiple NRPS clusters, we used a series of probes derived from the *bal* and *cep* clusters. The *bgtf*A gene, derived from the *bal* cluster of *Amycolatopsis mediterranei* and involved in one of the glycosylation steps (Pelzer et al. 1999), identified four positive cosmids previously grouped in a single contig. All together, these cosmids, denominated pAT9, pAT17, pAT32 and pAT42, contained two distinct fragments hybridizing to *bgtf*A. They covered a 68-kb region of the *A. teichomyceticus* chromosome, with each of the *bgtfA*-hybridizing fragments located on either side of a tight sequence of seven fragments hybridizing to Pep6 and Pep8, for a total of seven presumptive NRPS modules (Figure 2).

One *bgfA*-positive fragment was also detected in each of three additional, unrelated, Pep6- and Pep8-positive cosmids (pAT5, pAT13 and pAT38). However, the signals observed with *bgtfA* from the fragments of the contig of Figure 2 were stronger than those from pAT5, pAT13 and pAT38 (data not shown). Furthermore, none of these latter cosmids contained a sufficient number of Pep8-hybridizing fragments to be a likely candidate for teicoplanin biosynthesis. On the basis of the number of putative NRPS modules and the close relatedness to *bgtfA*, the contig of Figure 2 appeared as the desired cluster. Interestingly, if the positive signals observed with *bgtfA* in pAT5, pAT13 and pAT38 do correspond to *bona fide* glycosyltransferase genes, then *A. teichomyceticus* would possess the genetic information for the synthesis of other glycosylated peptides.

Organization of the tcp *genes*

The contig delimited by pAT9, pAT17, pAT32 and pAT42 spanned a 68 kb segment of the *A. teichomyceticus* genome, a size that may not be sufficient to contain all the genes required for teicoplanin biosynthesis (van Wageningen et al. 1998). By genomic walking, this contig was extended by four more cos-

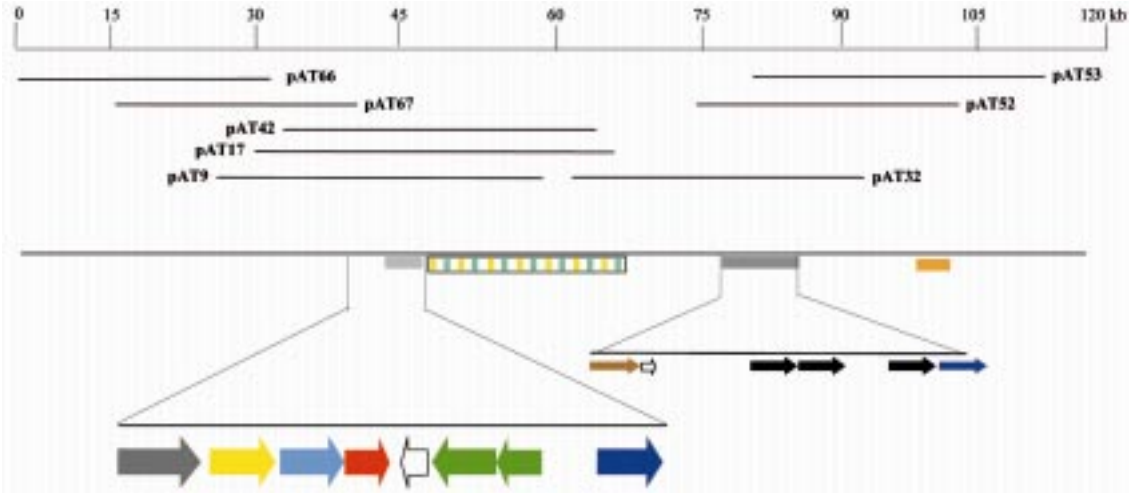

Figure 2. Organization of the *tcp* cluster. The region of the *A. teichomyceticus* genome containing the teicoplanin biosynthesis genes is shown with the cosmids encompassing it above. The putative NRPS region is indicated by an open box with with yellow and green stripes representing Pep6 and Pep8-hybridizing segments, respectively. The fragments hybridizing with *bgtf*A are drawn as shaded bars. The *csg* hybridizing portion is shown as an orange bar. In the two enlargments, thick and thin arrows represent complete and partial ORFs, respectively. Color codes are: brown, ABC transporter; black, P450-monoxygenase; dark blue, glycosyltransferase; grey, MurF; yellow,VanH; light blue, VanA; red, VanX; white, unknown; green, sensor kinase and response regulator.

mids for a total of 114 kb (Figure 2). Next, we made use of a series of probes derived from other genes involved in glycopeptide biosynthesis. These cosmids were hybridized with ORF7 from the *cep* cluster (van Wageningen et al. 1998). This gene is a homologue of *oxyA*, whose gene products, a P-450 monooxygenase, is likely involved in cross-linking of aromatic side chains during balhymicin formation (Pelzer et al. 1999). Three ORF7-related genes exist in the *cep* and *bal* clusters. This probe identified three different fragments, as shown in Figure 2, as expected from the similar cross-linking reactions needed for teicoplanin biosynthesis.

Another hybridization experiment involved a gene (‘*csg*’) from the *bal* cluster, encoding a polypeptide resembling a chalcone synthase (Wohlleben, pers.comm.). Chalcone synthases are involved in the formation of polyketides in plants (Hopwood 1997). A similar sequence is also present in the *cep* cluster (van Wageningen et al. 1998). Although its exact role has not been demonstrated, the presence of this gene in two glycopeptide clusters suggests a role in the formation of the polyketide chain precursor of 3,5-dihydroxyphenylglycine. This amino acid residue is present in all glycopeptides and is made through the condensation of acetate units by a hypothetical polyketide pathway (Lancini & Cavalleri 1997). Hybridization with a *csg* fragment led to the identification of a positive segment from cosmids pAT52 and pAT53 (Figure 2).

Although we do not have any direct proof yet, the finding of seven presumptive NRPS modules, two *bgtfA*-like sequences, three *cep*ORF7 signals and one *csg*-positive fragment, strongly suggests the involvement of the *A. teichomyceticus* segment of Figure 2 in teicoplanin biosynthesis. From now on, this contig is referred as the teicoplanin (*tcp*) cluster. The sequence information described below strengthens this assumption.

In order to confirm the hybridization data, we obtained partial nucleotide sequences from the ends of a few subclones from the *tcp* cluster. The sequences generated from a 10 kb portion to the right of the NRPS segment revealed the presence of incomplete ORFs with significant similarities to the following sequences (left to right in Figure 2): multi-drug resistant determinants belonging to the ABC transporter superfamily; a small polypeptide of unknown function present in the *cep* cluster; three different cytochrome P450-related monooxygenases; and a glycosyltransferase. In all cases, the best matching sequences were from the other two glycopeptide clusters. Sequence information from the NRPS-proximal end of the 4-kb fragment that hybridized to *bgtfA*, identified a second glycosyltransferase-like sequence.

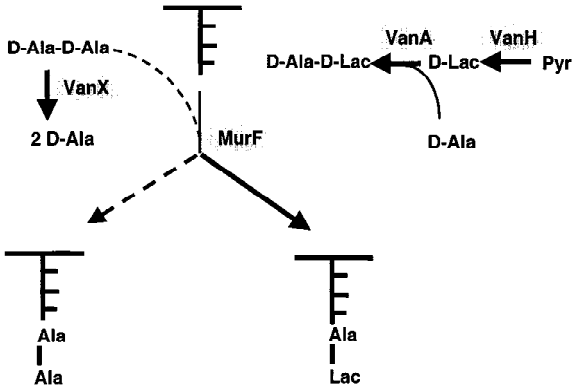

Figure 3. Synthesis of a modified cell wall in VRE strains. The muramyl-tripeptide is schematized by the branched symbol. Solid arrows represent the pathway operating in resistant strains, while dotted arrows the steps occurring in sensitive strains. The names of the enzymes involved are circled, while substrates are abbreviated as: D-Ala, D-alanine; D-Lac, D-lactate; Pyr, pyruvate. Adapted from Walsh et al. 1996.

Likely resistance determinants in the tcp *cluster*

At the other end of the 4 kb fragment mentioned above, we identified the 3′ end of an ORF which showed similarity to *vanX*, a gene involved in glycopeptide resistance in VRE strains (Arthur et al. 1996; Walsh et al. 1996). This result suggested that other resistance genes might be present in the cluster next to this segment. Therefore the sequence of an 11-kb segment centered on the *vanX*-homologue was determined. This region contains eight complete ORFs (designated ORF 1 through 8). ORF8 corresponds to the glycosyltransferase gene mentioned above. The deduced gene products of ORFs 1–4 exhibit remarkable similarity to MurF, VanH, VanA and VanX, respectively. These enzymes participate in the resistance mechanism in VRE strains by redirecting a portion of the peptidoglycan pathway, substituting the terminal D-alanyl-D-alanine (D-Ala-D-Ala) residue in the pentapeptide with a D-alanyl-D-lactate (D-Ala-D-Lac) moiety (Arthur et al. 1996; Walsh et al. 1996). As schematized in Figure 3, VanH is an α-ketoacid reductase that converts pyruvate into D-Lac; VanA condenses D-Ala and D-Lac to yield the depsipeptide D-Ala-D-Lac; VanX is a specific d,d-dipeptidase that cleaves D-Ala-D-Ala (made by a housekeeping D-Ala-D-Ala ligase). Although *murF* is not physically linked to the *vanHAX* operon in VRE strains, its gene product participates in the synthesis of peptidoglycan by adding the dipeptide D-Ala-D-Ala or the depsipeptide D-Ala-D-Lac to the muramyl-tripeptide in vancomycin-sensitive and -resistant strains, respectively (Figure 3). The overall result of the action of three specific resistance determinants and one broad-specificity enzyme in VRE strains is a peptidoglycan lacking the target where glycopeptides bind with high affinity.

Of the other three ORFs detected in this sequenced segment from the *tcp* cluster, ORF6 resembles sensor kinases and ORF7 response regulators. The best matches were with putative two-component signal transduction systems from the *S. coelicolor* genome sequence (data not shown). ORF5 does not find any significant match in the database.

Conclusions

The information thus far generated on the *tcp* cluster is summarized in Figure 2. In terms of biosynthesis genes, the overall organization of the *tcp* cluster appears to be very similar to that of the *cep* cluster (van Wageningen et al. 1998). However, the relative positions of one of the glycosyltransferase genes and of the ABC transporter gene are inverted between the *cep* and *tcp* clusters. The available portion of the *bal* cluster shares an organization identical to that of *cep*. Both *bal* and *cep* clusters originate from *Amycolatopsis* species.

The uniqueness of the *tcp* cluster lies in the presence of a likely glycopeptide resistance cassette. The *vanHAX* determinants have been studied in great detail in enterococci (Arthur et al. 1996; Walsh et al. 1996). Thus far, homologs of the *vanHAX* genes have also been found in glycopeptide producers (Marshall et al. 1998). In addition, a *vanHAX*-like sequence can also be detected from the ongoing sequence of the *Streptomyces coelicolor* genome (EMBL accession No. AL079348). The organization of the *vanHAX*-like sequences found in other actinomycetes, including *A. teichomyceticus*, are summarized in Figure 4, and compared to the VRE strains. In *Streptomyces toyocaensis* and *Amycolatopsis orientalis*, producers of A47934 and vancomycin, respectively, the available evidence indicates that the *van*-like genes are not linked to glycopeptide biosynthesis determinants (Marshall et al. 1998). In the chloroeremomycin producer *A. orientalis*, the sequence of a ca. 100 kb segment is available (van Wageningen et al. 1998; EMBL accession No. AL078635), and *van*-like genes are not present. While there are some variations in gene organization within actinomycetes (Figure 4),

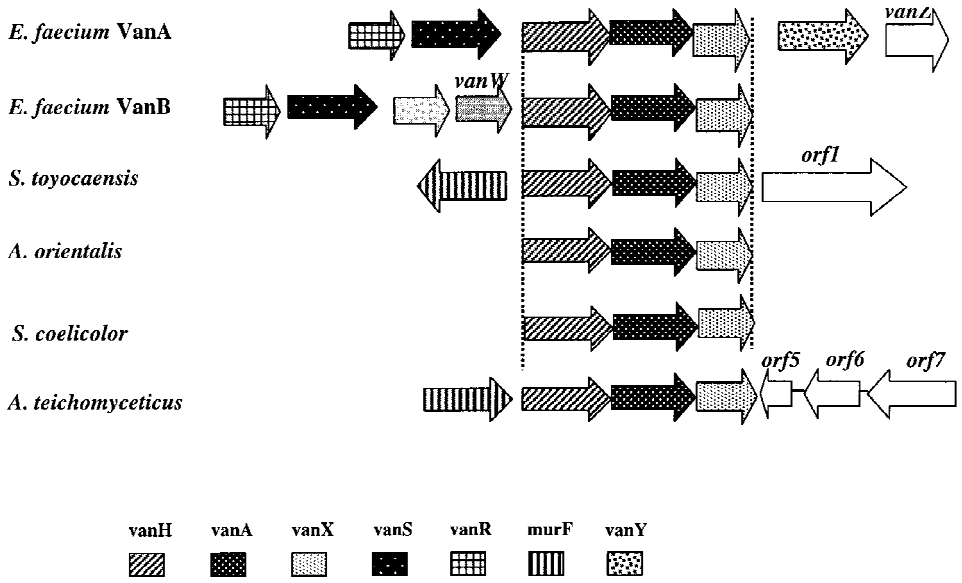

Figure 4. Organization of *van*HAX region in actinomycetes. The genes unique to each strain are labelled with their function, if known.

one feature is unique to *A. teichomyceticus*: the close linkage of the putative *vanHAX* cassette to the biosynthesis genes. The existence in *A. teichomyceticus* of a *murF* homologue closely linked to *vanHAX* (not found in VRE strains, but required for synthesis of the modified peptidoglycan; see Figure 3) suggests that this gene may also participate in resistance. Interestingly, a *murF* homologue is closely linked to *vanHAX*-like genes in *S. toyocaensis* (Figure 4). However, the location of glycopeptide biosynthesis genes in this strain is not currently known (Marshall et al. 1998).

Some glycopeptide producers possess both D-Ala-D-Lac and D-Ala-D-Ala ligases (Marshall et al. 1997; Marshall & Wright 1997 1998). It has also been speculated that the close linkage of resistance and biosynthesis genes found in actinomycetes might allow a close coordination of their expression, thereby ensuring that antibiotic synthesis occurs when the producer strain has become resistant to the antibiotic (Chater & Hopwood 1993). The results presented here suggest that synthesis of a glycopeptide-insensitive peptidoglycan might be closely coordinated with teicoplanin biosynthesis in *A. teichomyceticus*. A comparison with the other strains thus suggests that different glycopeptide producers may adopt different strategies to escape to the toxic action of their own products.

Acknowledgments

We are indebted to Stefan Pelzer, John Robinson, Wim Vrijbloed and Wolfgang Wohlleben for sharing unpublished results and valuable probes. We are grateful to Giancarlo Lancini for continuous encouragement and fruiful discussions. This work was supported by the Italian CNR, *PF Biotecnologie*.

References

Arthur M, Reynolds P & Courvalin P (1996) Glycopeptide resistance in enterococci. Trends Microbiol. 4: 401–407

Chater KF & Hopwood DA (1993) *Streptomyces*,. In: Sonenshein AL, Hoch JA& Losick R (eds) *Bacillus subtilis* and other gram-positive bacteria. Biochemistry, Physiology and Molecular Genetics (pp 83–99). American Society for Microbiology, Washington, DC

John CC (1999) Vancomycin resistance in *Staphylococcus aureus*. N. Engl. J. Med. 341: 207–208

Hopwood D (1997) Genetic contribution to understanding polyketide synthases. Chem. Rev. 97: 2465–2497

Lancini G (1989) Fermentation and biosynthesis of glycopeptide antibiotics. In: Bushell ME and Grafe U (eds) Bioactive Metabolites from Microrganisms (pp 283–296). Elsevier Science Publishers B.V., Amsterdam

Lancini G & Cavalleri B (1997) Glycopeptide antibiotics (Dalbaaheptides). In: Kleinkauf H and von Dohren H (eds) Biotechnology, Second Edition, vol 7 (pp 369–396). VHC, Weinheim

Marahiel MA (1997) Protein templates for the biosynthesis of peptide antibiotics. Chem. Biol. 4: 561–567

Marshall CG, Lessard IAD, Park I-S & Wright GD (1998) Glycopeptide antibiotic resistance genes in glycopeptide-producing organisms. Antimicrob. Agents Chemother. 42: 2215–2220

Marshall CG & Wright GD (1997) The glycopeptide antibiotic producer *Streptomyces toyocaensis* NRRL 15009 has both D-alanyl-D-alanine and D-alanyl-D-lactate ligases. FEMS Microbiol. Lett. 157: 295–299

Marshall CG & Wright GD (1998) DdlN from vancomycin-producing *Amycolatopsis orientalis* C329.2 is a VanA homologue with D-alanyl-D-lactate ligase activity. J. Bacteriol. 180: 5792–5795

Pelzer S, Sussmuth R, Heckmann D, Recktenwald J, Huber P, Jung G & Wohlleben W (1999) Identification and analysis of the balhimycin biosynthetic gene cluster and its use for manipulating glycopeptide biosynthesis in *Amycolatopsis mediterranei* DSM5908. Antimicrob. Agents Chemother. 43: 1565–1573

Solenberg PJ, Matsushima P, Stack DR, Wilkie SC, Thompson RC & Baltz RH (1997) Production of hybrid glycopeptide antibiotics *in vitro* and in *Streptomyces toyocaensis*. Chem. Biol. 4: 195–202

Sosio M, Bossi E, Bianchi A & Donadio S (2000) Multiple peptide synthetase gene clusters in actinomycetes. (Mol. Gem. Genet. in press)

van Wageningen AMA, Kirkpatrick PN, Williams DH, Harris RB, Kershaw JK, Lennard NJ, Jones M, Jones SJM & Solenberg PJ (1998) Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. Chem. Biol. 5: 155–162

Walsh CT, Fisher SL, Park IS, Prahalad M & Wu Z (1996) Bacterial resistance to vancomycin: five genes and one missing hydrogen bond tell the story. Chem. Biol. 3: 21–28