Chemistry & Biology, Vol. 11, 107–119, January, 2004, ©2004 Elsevier Science Ltd. All rights reserved. DOI 10.1016/j.chembiol.2004.01.001

Biosynthetic Gene Cluster of the Glycopeptide Antibiotic Teicoplanin: Characterization of Two Glycosyltransferases and the Key Acyltransferase

Tsung-Lin Li,[1] Fanglu Huang,[1] Stephen F. Haydock,[2,*] Tatiana Mironenko,[2] Peter F. Leadlay,[2] and Jonathan B. Spencer[1,*]
[1]Department of Chemistry
University of Cambridge
Lensfield Road
Cambridge CB2 1EW
[2]Department of Biochemistry
University of Cambridge
Tennis Court Road
Cambridge CB2 1GA
United Kingdom

Summary

The gene cluster encoding biosynthesis of the clinically important glycopeptide antibiotic teicoplanin has been cloned from *Actinoplanes teichomyceticus*. Forty-nine putative open reading frames (ORFs) were identified within an 89 kbp genetic locus and assigned roles in teicoplanin biosynthesis, export, resistance, and regulation. Two ORFs, designated *orf*s 1 and 10*, showed significant homology to known glycosyltransferases. When heterologously expressed in *Escherichia coli*, these glycosyltransferases were shown to catalyze the transfer of UDP-(*N*-acetyl)-glucosamine onto, respectively, 3-chloro-β-hydroxytyrosine-6 (3-Cl-6βHty) and 4-hydroxyphenylglycine-4 (4Hpg) of the teicoplanin heptapeptide aglycone. The product of another ORF, *orf*11*, was demonstrated in vitro to transfer *n*-acetyl-, *n*-butyryl-, and *n*-octanoyl-groups from acyl-CoA donors either to a free UDP-aminosugar or to an aminosugar moiety in the teicoplanin pseudoaglycone, thus identifying Orf11* as the key acyltransferase in teicoplanin maturation. These findings should accelerate the combinatorial engineering of new and improved glycopeptide drugs.

Introduction

Teicoplanin is one of only two glycopeptide antimicrobials currently licensed for use in man. It is reserved for the treatment of life-threatening infections caused by gram-positive organisms, including methicillin-resistant *Staphylococcus aureus* (MRSA) [1]. The rapidly escalating problem of antibiotic drug resistance has driven a search for newer agents with superior efficacy against drug-resistant strains. Teicoplanin, however, remains an extremely important agent for the treatment of severe sepsis. It is considerably more potent than vancomycin against gram-positive bacteria, has a lower toxicity, and a superior pharmacokinetic profile [2]. Teicoplanin for clinical administration is a mixture of five teicoplanins (A2-1 to A2-5) produced by *Actinoplanes teichomyceticus*, which are representative of a subgroup (type IV) of glycopeptides characterized by the presence of a unique long aliphatic chain attached to a sugar moiety (Figure 1) [3]. This fatty acyl side chain is thought to be derived from β-oxidation of fatty acids [4]. In addition, four minor components referred to as *r*elated substances 1 to 4 (RS1-4) have been recently isolated [5], which differ in the nature of the *N*-acyl chain. The heptapeptide backbone of the aglycone core is composed of five amino acids that are common to all members of this group. Teicoplanin specifically contains 4-hydroxyphenylglycine and 3,5-dihydroxyphenylglycine residues at positions 1 and 3; a chlorine atom substituted on each of the tyrosine residues; and three sugar moieties, *N*-fatty acyl-β-D-glucosamine, *N*-acetyl-β-D-glucosamine, and D-mannose at positions 4, 6, and 7, respectively.

Particular attention has been drawn to the *N*-fatty acyl-β-D-glucosamine substituent by the recent finding that a semisynthetic glycopeptide, LY264826, has remarkably increased activity owing to the chemical introduction of a hydrophobic moiety onto a sugar residue of chloroeremomycin [6]. Alternative approaches to such modification of glycopeptide antibiotics employ biotransformation techniques or genetically modified producer organisms to alter the structure of the target molecule, as is commonplace for polyketide biosynthesis [7, 8]. Our institute previously reported the first gene cluster governing the formation of a glycopeptide, that for chloroeremomycin [9]. This and subsequent reports relating to the biosynthesis of balhimycin, complestatin, A47934, and A40926 have greatly expanded our knowledge of the biosynthesis of this important group of compounds [10, 11, 12, 13]. Gene disruption or in vitro characterization of genes heterologously expressed in *E. coli* have further elucidated the steps involved in the biosynthesis of these agents, and this has enabled the generation of novel chimeric glycopeptides either through the manipulation of biosynthetic pathways or through in vitro glycosylation of glycopeptide aglycones or pseudoaglycones [14, 15, 16, 17]. These initial successes have spurred us to obtain detailed information relating to the biosynthesis of the clinically important glycopeptide teicoplanin. Preliminary studies reported by others [18] had demonstrated that the genes for teicoplanin biosynthesis were clustered and lay within a 100 kbp locus in *A. teichomyceticus*. As might be expected, the biosynthetic genes appeared to show homology to genes responsible for glycopeptide biosynthesis in other published clusters.

We describe here the entire DNA sequence and identification of 49 genes involved in the biosynthesis of teicoplanin. As discussed below, the biosynthetic cluster apparently includes all the genes needed for teicoplanin biosynthesis, regulation of biosynthesis, export, and resistance. We also describe the heterologous expression and enzymatic assay of two glycosyltransferases and of a candidate gene for the key fatty acyl transferase. The expressed gene products were found to be active, enabling us to identify their respective roles in teicoplanin maturation. The substrate flexibility shown by

*Correspondence: jbs20@cam.ac.uk (J.B.S.); sfh12@mole.bio.cam.ac.uk (S.F.H.)

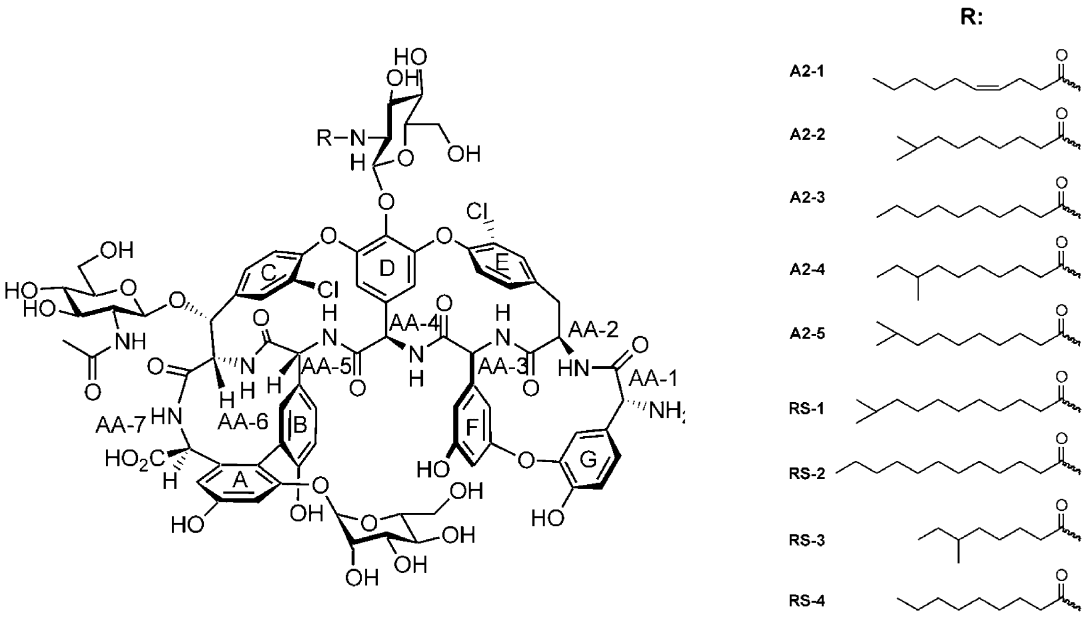


Figure 1. Structures of Teicoplanins

Teicoplanin for clinical administration is a mixture of five major teicoplanins (A2-1 to A2-5). Four minor components referred to as related substances 1 to 4 (RS1–4) have also been isolated. The seven amino acids of teicoplanin from the N-terminal end to C-terminal end are designated as AA-1 to AA-7, and the seven aromatic rings are lettered A thorugh G [3].

the fatty acyltransferase suggests real utility for this enzyme in future combinatorial engineering of glycopeptide drugs.

Results and Discussion

Cloning and Sequencing of the *tei* Gene Cluster

Approximately 500 cosmids from a genomic library of *A. teichomyceticus* genomic DNA were randomly end sequenced. The resulting DNA sequences were used to search published protein databases using BLAST. Five cosmids were identified with a high degree of homology to genes previously identified in glycopeptide gene clusters. Four cosmids were sequenced in their entirety and found to encode genes consistent with biosynthesis of teicoplanin. The identification of genes involved in carbohydrate metabolism confirmed that the upstream limit of the cluster had been identified. The downstream limit could not be identified. After PCR-based screening of the library, a further clone was obtained that extended downstream of the sequenced locus. This clone was also sequenced in its entirety and found to contain the downstream limit of the cluster, as shown by the presence of insertion sequences and genes involved in manganese transport. In total, 120 kbp of the teicoplanin cluster and of the flanking sequences were isolated and sequenced from *A. teichomyceticus*.

Organization of the Teicoplanin Cluster

The biosynthetic gene cluster of teicoplanin (*tei*) spans approximately 89 kbp and includes 49 putative *orfs* (Figure 2A). Likely functions for most of these could be assigned by comparison of the predicted protein translation with gene products deposited in the published sequence databases. These 49 genes are predicted to be responsible for assembly of the antibiotic, resistance, export, and regulation of synthesis. 35 *orf*s find homologs in other glycopeptide gene clusters (Table 1).

Biosynthesis of Nonproteinogenic Amino Acids

Tyrosine and three nonproteinogenic amino acids, (*S*)-4-hydroxyphenylglycine (Hpg), 3,5-dihydroxyphenylglycine (Dpg), and β-hydroxytyrosine (βHty), are used as building blocks in the teicoplanin group of glycopeptide antibiotics. Chlorination of Tyr and βHty at amino acid positions 2 and 6 is proposed to occur at an early stage of synthesis prior to phenolic oxidative coupling, possibly with tyrosine or βHty as the substrate [19]. Four enzyme-catalyzed steps have been identified in the biosynthesis of HPG in the chloroeremomycin producer *A. orientalis*. These are designated prephenate dehydrogenase (Pdh), 4-hydroxymandelic acid synthase (HmaS), 4-hydroxymandelic acid oxidase (HmO), and 4-hydroxyphenylglycine aminotransferase (HpgT) [20, 21, 22]. Pdh is commonly associated with the shikimate primary metabolic pathway and converts prephenate to 4-hydroxyphenylpyruvate, the substrate for HmaS [23]. In addition to a Pdh homolog designated Orf24* in the *tei* cluster, Orf 14* is a homolog of a dual-function enzyme that possesses both 7-phospho-2-dehydro-3-deoxy-D-arabinoheptulosonate synthetase (DAHPS) and chorismate mutase activities [23]. The presence of these enzymes in the *tei* cluster may reflect the high demand for tyrosine (five out of seven amino acids in teicoplanin originate from tyrosine). Combining information from the *tei* and *cep* clusters permits a detailed proposal to be made for

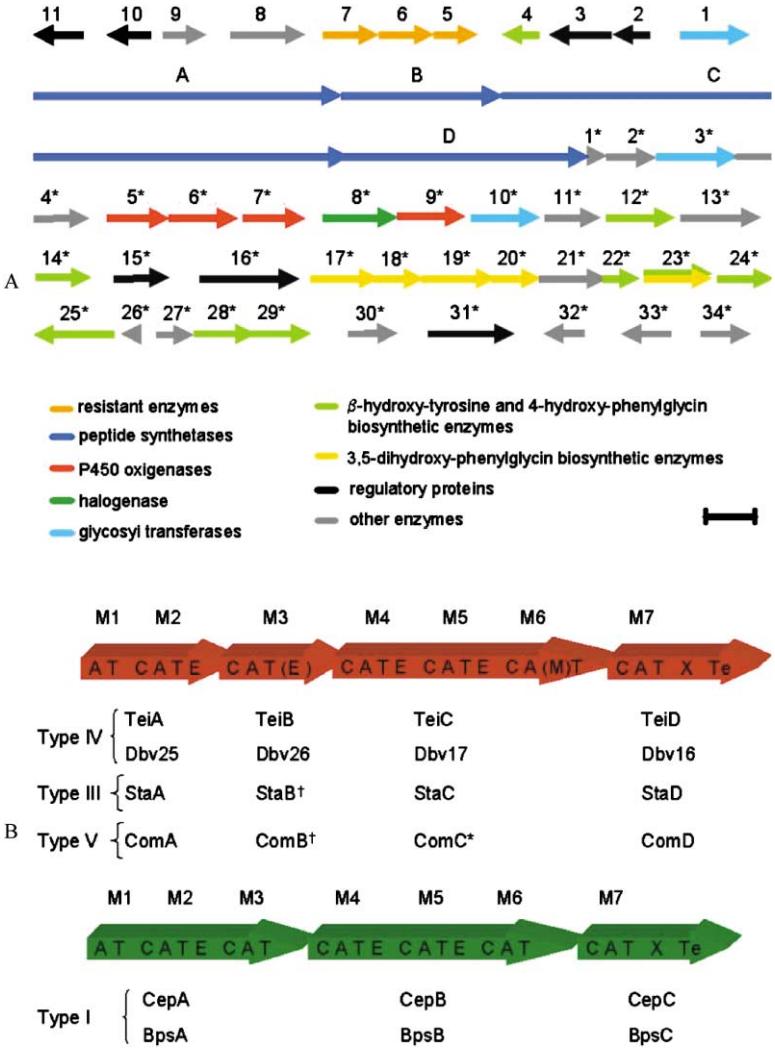


Figure 2. Genetic Arrangement and Proposed Gene Functions in the *tei* Cluster

(A) The designation of the *tei* cluster follows conventional assignment for this class of antibiotics, in which *orf** indicates genes that appear after NRPS. Each gene is represented as an arrow bar that also indicates the relative size of the gene and its transcriptional direction.

(B) The four NRPSs in the *tei* cluster (TeiA–D) house two, one, three, and one extension modules, respectively. Comparison with other previously described glycopeptide clusters reveals two different arrangements in which seven modules are distributed between three NRPSs in *cep* and *bps* clusters and between four NRPSs in *dbv*, *sta*, *com*, and *tei*. †, module 3 in ComB and StaB includes an extra epimerization domain, but the domain in StaB is proposed to be inactive; *, module 6 in ComC includes an extra methylation domain that is proposed to *N*-methylate residue 6 of complestatin; X, all six clusters contain a degenerate C or E domain adjacent to the end TE domain.

the biosynthetic pathway to Hpg in *A. teichomyceticus*. The first step involves the action of Orf14* followed by other shikimate pathway enzymes to produce chorismate. Orf14* converts the chorismate into prephenate, which undergoes a subsequent decarboxylation by Orf24* (Pdh) to generate 4-hydroxyphenylpyruvate. The remaining steps are likely to be identical to those biochemically characterized in the *cep* cluster. Orf28* (HmaS) catalyzes the formation of (*S*)-4-hydroxymandelate from 4-hydroxyphenylpyruvate. The (*S*)-4-hydroxymandelate is subsequently oxidized by Orf29* (HmO, a flavin-dependent oxidase) to 4-hydroxyphenylglyoxylate, followed by transamination catalyzed by Orf23* (HpgT, a pyridoxal phosphate-dependent transaminase) using tyrosine as the amino donor to produce Hpg. The deaminated Tyr can directly serve as a substrate for HmaS in the next catalytic cycle.

Three ORFs have inferred products with homology, respectively, to phenylalanine hydroxylase (PAH, a nonhaem dioxygenase) (Orf25*), dihydrofolate reductase (Orf4), and GTP cyclohydrolase I (Orf22*). These three enzymes might represent an alternative route of tyrosine biosynthesis in which phenylalanine is hydroxylated to Tyr in the presence of Fe^{2+}. This reaction requires one equivalent of the cofactor tetrahydrobiopterin, which is produced from dihydrobiopterin by dihydrofolate reductase at the expense of one equivalent of NADPH [24]. In Tyr synthesis, a pivotal enzyme, GTP cyclohydrolase I, hydrolyzes GTP into 7,8-dihydroneopterin triphosphate, the precursor to tetrahydrobiopterin. Therefore, we suggest that Orf25*, Orf22*, and Orf4 together help to meet the high demand for Tyr, which serves not only as an amino donor but also as a substrate for Tyr β-hydroxylase (see below).

Dpg biosynthesis has also been studied in the glycopeptide producers *A. orientalis* and *A. mediterranei* [25, 26, 27]. Four ORFs (DpgA–D) organized into a putative transcriptional unit together with one additional ORF (HpgT) governing a transamination step have been shown to be involved in DPG production by the in vitro characterization of biological activities [28]. The equivalent genes encoding these enzymes in *A. teichomyceticus* are designated Orfs 17*–20* and are organized as found in other published glycopeptide clusters. The synthesis of β-Hty from Tyr in balhimycin and chloroeremomycin biosynthesis appears to follow a novel mecha-

Table 1. Summary of Genes Identified from the *tei* Cluster

	Related Clusters[a]					Best Match[b]		
tei orfs	*cep*	*bal*	*com*	*sta*	*dbv*	Source[d]	Entry[e]	Proposed Function[c]
11[R]	-	-	-	-	-	*Streptomyces coelicolor*	Q9XA73	Transcriptional regulator
10[R]	-	-	-	-	-	*Streptomyces avermitilis*	Q82PM0	Transcriptional regulator
9	-	-	-	-	-	*Streptomyces coelicolor*	Q9L0G1	Short chain dehydrogenase
8	?	-	-	MurF	-	*Streptomyces toyocaensis*	O33804	MurF; D-Ala- D-Ala adding enzyme
7	?	-	-	$VanH_{st}$	-	*Amycolatopsis orientalis*	Q9Z6H3	D-lactate dehydrogenase
6	?	-	-	$VanA_{st}$	-	*Streptomyces coelicolor*	Q9XAK7	D-Ala- D-Lac ligase
5	?	-	-	$VanX_{st}$	-	*Streptomyces toyocaensis*	Q9ZI81	D-Ala- D-Ala dipeptidase
4[R]	-	-	-	-	-	*Vibrio vulnificus*	Q8D590	Dihydrofolate reductase
3[R]	?	-	-	$VanS_{st}$	-	*Streptomyces toyocaensis*	Q8KLL0	Transcriptional regulator
2[R]	?	-	-	$VanR_{st}$	-	*Streptomyces coelicolor*	Q8CJW1	Two component response regulator
1	GtfA	BgtfA	-	-	9	*Amycolatopsis orientalis*	P96559	Glycosyltransferase
A	A	BpsA	A	A	25	*Streptomyces toyocaensis*	Q8KLL3	Peptide synthetase; module 1-2
B	A	BpsA	B	B	26	*Streptomyces toyocaensis*	Q8KLL4	Peptide synthetase; module 3
C	B	BpsB	C	C	17	*Streptomyces toyocaensis*	Q8KLL5	Peptide synthetase; module 4-6
D	C	BpsC	D	D	16	*Amycolatopsis orientalis*	O52821	Peptide synthetase; module 7
1*	6	1	E	E	15	*Nonomuraea*, ATCC39727	CAD91210	Unknown
2*	15	2	-	?	21	*Amycolatopsis orientalis*	O52804	Hypothetical protein
3*	-	-	-	-	20	*Nonomuraea*, ATCC39727	CAD91215	Mannossyltransferase
4*	2	?	L	U	24	*Amycolatopsis orientalis*	O52818	ABC transporter
5*	7	OxyA	I	F	14	*Streptomyces toyocaensis*	Q8KLL7	P450 monooxygenase
6*	-	-	-	G	13	*Streptomyces toyocaensis*	Q8KLL8	P450 monooxygenase
7*	8	OxyB	J	H	12	*Nonomuraea*, ATCC39727	CAD91207	P450 monooxygenase
8*	10	BhaA	-	I	10	*Streptomyces toyocaensis*	Q8KLM0	Halogenase
9*	9	OxyC	-	J	11	*Nonomuraea*, ATCC39727	CAD91206	P450 monooxygenase
10*	GtfB	Bgtf	-	-	9	*Amycolatopsis orientalis*	P96559	Glycosyltransferase
11*	-	-	-	-	8	*Nonomuraea*, ATCC39727	CAD91203	Acyltransferase
12*	-	-	-	M	28	*Nonomuraea*, ATCC39727	CAD91223	β-hydroxylase
13*	-	-	-	-	?	*Streptomyces avermitilis*	Q82H77	Acyl-CoA synthase
14*	31	?	-	?	?	*Thermotoga maritima*	Q9WYH8	DAHPS/Chorismate mutase
15*	-	-	-	Q	4	*Streptomyces toyocaensis*	Q8KLK0	Transcriptional regulator
16*	-	-	-	?	3	*Streptomyces avermitilis*	Q93HJ7	LuxR regulatory protein
17*	DpgA	DpgA	-	DpgA	31	*Amycolatopsis mediterranei*	Q939X3	DpgA
18*	DpgB	DpgB	-	DpgB	32	*Amycolatopsis mediterranei*	Q939X2	DpgB
19*	DpgC	DpgC	-	DpgC	33	*Streptomyces toyocaensis*	Q8KLK7	DpgC
20*	DpgD	DpgD	-	DpgD	34	*Streptomyces toyocaensis*	Q8KLK8	DpgD
21*	?	7	-	?	35	*Amycolatopsis mediterranei*	Q939X8	Na^+-H^+ antiporter
22*	-	-	-	?	?	*Synechococcus elongatus*	Q8DJB8	GTP cyclohydrolase I
23*	HpgT	PgaT	HpgT	HpgT	37	*Streptomyces toyocaensis*	Q8KLK4	HpgT
24*	Pdh	?	Pdh	Pdh	5	*Streptomyces toyocaensis*	Q8KLK3	Prephenate dehydrogenase
25*[R]	-	-	-	M	28	*Streptomyces toyocaensis*	Q8KLM4	Nonhaem iron dioxygenase
26*[R]	-	-	-	-	-	*Mycobacterium leprae*	Q9CD74	Hypothetic protein
27*	-	-	-	-	-	*Streptomyces coelicolor*	Q9RJ78	Hypothetic protein
28*	HmaS	5	HmaS	HmaS	2	*Amycolatopsis orientalis*	O52791	HmaS
29*	HmaO	6	HmaO	HmaO	1	*Nonomuraea*, ATCC39727	CAD91196	HmO
30*	-	-	-	-	36	*Streptomyces fradiae*	Q54145	Type II thioesterase
31*	-	-	-	-	-	*Streptomyces avermitilis*	Q82GU2	Transcriptional regulator
32*[R]	-	-	-	-	-	*Streptomyces avermitilis*	Q82RX4	Transposase
33*[R]	-	-	-	-	-	*Streptomyces coelicolor*	Q9RK03	Hypothetical protein
34*	-	-	-	-	-	*Rhizobium loti*	Q983I7	Manganese ABC transporter

R, gene on opposite strand; -, clearly missing in cluster; ?, perhaps present (may not have been sequenced or published yet).
[a] Existence of a homolog in other related glycopeptide clusters. A dash denotes its absence.
[b] Best relevant match found by BLAST searches in Swiss-Prot/TrEMBL.
[c] Proposed function in the teicoplanin biosynthetic gene cluster.
[d] Organism origin of the best matching sequence.
[e] Accession number of the best matching sequence.

nism proposed for the biosynthesis of novobiocin and nikkomycin, in which three enzyme-catalyzed steps are required [29, 30]. Tyr is initially activated by the adenylation domain of a di-domain NRPS and is covalently tethered to an adjacent PCP domain of the same NRPS. A P450 hydroxylase subsequently oxidizes the tethered Tyr at its β position. The modified Tyr is then released from the NRPS by a thioesterase. However, no counterparts of these genes are found in the *tei*, *dbv*, and *sta* gene clusters. Analysis of the *tei* gene cluster reveals a putative gene product (Orf12*) that contains motifs typical of nonhaem iron dioxygenases and has a homolog in both the *dbv* and *sta* clusters. As previously suggested [13], this enzyme might hydroxylate the C3 position of Tyr. Another homolog of this enzyme is found in the chloramphenicol cluster, in which β-hydroxylated Phe residue is present in the peptide backbone [31].

Assembly of the Heptapeptide

Synthesis of the teicoplanin heptapeptide backbone is catalyzed by four nonribosomal peptide synthetases

(NRPSs) designated TeiA–D. NRPSs are multifunctional enzymes with a characteristic modular organization. Each module harbors the catalytic domains necessary for activation, covalent tethering, and functional modification of each monomer as well as for catalysis of peptide-bond formation. The number and order of these domains within the enzyme is normally consistent with the sequence and structure of the peptide product [32, 33, 34]. Organization of the domains for adenylation (A), thiolation (T), condensation (C), epimerization (E), and thioesterase (TE) within the four proteins was identified through comparison with consensus sequences from the published databases. Extension modules are distributed 2-1-3-1 between the four NRPSs, as in the A40926 (*dbv*) (type IV), A47934 (*sta*) (type III), and complestatin (*com*) (type V) biosynthetic clusters and in contrast to the 3-3-1 arrangement in the chloroeremomycin (*cep*) and balhimycin (*bal*) biosynthetic clusters (type I) (Figure 2B).

The assignment of modules is also fully consistent with the predicted amino acid specificities of the corresponding A domains [35, 36]. As in all four previously reported sequences for glycopeptide NRPSs, module 1 has no E domain, although the configuration of the amino acid in this position is D rather than L (non-amino acid 3,5 dichloro-4-hydroxybenzoylformate is the first building block in *com*) [9, 10, 11, 12, 13]. A further stereochemical inconsistency is seen in the sequence analysis of the *sta* cluster, where an extra E domain found in module 3 is claimed to be functionally inactive [12]. All six glycopeptide NRPSs contain a degenerate domain (either C or E) of obscure function flanking the terminal TE domain.

A small but highly conserved gene of unknown function, designated *orf*1* in the *tei* cluster, is found immediately downstream of the terminal TE domain of the NRPS in many NRPS gene clusters [9, 10, 11, 13]. Orf30*, a homolog of type II thioesterases (TE II), is found only in the *tei* and *dbv* clusters but not others. The gene product may be involved in reducing the mispriming of NRPSs [37]. A similar role has been proposed for TE II enzymes in polyketide synthases [38].

Cyclisation of the Linear Heptapeptide

The characteristic rigid concave shape of teicoplanin and other glycopeptide antibiotics [39] is conferred by intramolecular crosslinks generated by phenolic oxidative coupling reactions. It has been possible, for several of the glycopeptide antibiotics, to relate each crosslink to a corresponding gene encoding a cytochrome P450 oxygenase. A series of gene disruption studies in the balhimycin-producing strain have revealed the order of the regioselective couplings. OxyB is suggested to prime the first reaction, coupling residues 4 and 6, OxyA then couples residues 2 and 4, and OxyC then forms a C-C bond between residues 5 and 7 [40, 41, 42]. The *tei* cluster contains four such genes, namely *orf*s 5*, 6*, 7*, and 9*, as found in the *sta* and *dbv* clusters, consistent with the four crosslinks that are found in the corresponding glycopeptides. From sequence comparisons, we propose that Orfs 5*, 7*, and 9* are related to OxyA, OxyB, and OxyC (similarity/identity: 90/78, 80/72, and 84/74), respectively, and hence responsible for crosslinking residues 2–4, 4–6, and 5–7. The remaining P450 enzyme, Orf 6*, is therefore a candidate to catalyze crosslinking of residues 1–3.

Halogenation

Teicoplanin has two chlorinated amino acids at positions 2 (3-Cl-Tyr) and 6 (3-Cl-β-Hty), and the putative halogenase Orf8* is proposed to catalyze halogenation of both of these amino acids. Indeed, *orf*8* is homologous to *bha*A in the *bal* cluster, which has been shown to chlorinate both Tyr/β-Hty residues 2 and 6 [19]. The specificity and the timing of halogenation remain to be determined in each case.

Glycosylations and Acylation of Teicoplanin Aglycone

The pattern of glycosylation of teicoplanin antibiotics is central to their biological activity [43]. It has been shown that specific glycosylation occurs after the formation of heptapeptide aglycone [44]. By analogy to previously characterized clusters, three separate glycosyl transferases should be required for the glycosylation of the teicoplanin aglycone. The translated amino acid sequences from *orf*1 and *orf*10* in the *tei* cluster show significant sequence similarity to GtfA and GtfB encoded by the *cep* cluster, which glycosylate residues 6 and 4 of the chloroeremomycin aglycone, respectively. We suspected therefore that Orf1 and Orf10* fulfilled analogous functions in the *tei* cluster. The roles of these two putative glycosyltransferases were experimentally tested as described below. A third glycosyltransferase with specificity for mannose is required to complete the glycosylation of teicoplanin. Orf3* in the *tei* cluster shows significant sequence similarity to many putative integral membrane proteins. A gene encoding a similar protein has been identified in the *dbv* cluster (*orf20*). Orf20 (*dbv*) has been convincingly shown [13] to contain sequence motifs characteristic of authentic mannosyltransferases, and these are shared by *tei* Orf3* (data not shown).

Resistance

The organization of genes conferring teicoplanin resistance in the *tei* cluster is similar to that of glycopeptide resistance in vancomycin-resistant *Enterococci* (VRE) [45, 46]. *orf*7* is homologous to *van*H (encoding a D-lactate dehydrogenase), *orf*6* is homologous to *van*A (encoding D-Ala-D-lactate ligase), and *orf*5* to *van*X (coded for D-Ala-D-Ala dipeptidase). Activation of the resistance genes in VRE requires a two-component regulatory system. A transmembrane sensory histidine protein kinase, VanS, senses and responds to environmental stimuli and activates an intracellular DNA binding response regulator designated VanR [43, 45, 46]. The *tei* cluster contains homologs of *van*R and *van*S designated *orf*2* and *orf*3*. These two genes are adjacent to the *van*H, A, and X cluster, forming the most intact set of resistant genes in a glycopeptide-producing strain that parallels those found in VRE, but differing from those of *S. toyocaensis*, where the genes are separated by 20 kb. This appears to provide additional direct evidence of resistance dissemination between species. Two extra *orf*s are also probably linked with the resistance genes: *orf*8*, encoding a D-Ala-D-Ala ligase (MurF in the *sat* cluster), and *orf*9*, a putative short chain dehydrogenase, are also probably involved in cell wall biosynthesis. The *dbv* and *bal* clusters utilize a different resistance mechanism in which a carboxypeptidases, VanY, is responsible (VanY hydrolyzes the terminal D-Ala residue of late peptidoglycan precursors) [43, 47].

Genes for Export and Regulation

Biosynthesis of secondary metabolites is highly regulated, is associated with specific stages of the *Streptomyces* growth cycle, and particularly occurs during stages of nutrient limitation and growth of aerial hyphae. Genes can be identified in the teicoplanin cluster that are likely to perform important regulatory functions.

As previously described, *orfs* 2 and 3 encode components of a bacterial two-component sensor kinase systems probably involved in regulation of expression of the resistance determinants.

A second class of regulators, designated the *luxR* family, is activated when bound to autoinducer molecules such as *N*-(3-oxohexanoyl)-L-homoserine lactone (OHHL). The "helix-turn-helix" DNA binding motif of these proteins is located in the C-terminal section of the sequence [48]. *orf16** encodes a protein of 818 amino acids with end-to-end homology to many *luxR* class regulators, particularly to those encountered in secondary metabolite biosynthetic clusters. In addition, *orfs* 15* and 31* show similarities to gene products implicated in a regulatory role in bacterial species. *orfs* 10 and 11 also show a high similarity to known regulators. Their location at the edges of the cluster and lack of any homologs to these genes in other clusters means that a regulatory role for these genes awaits further study.

*orf*s 1*, 2*, 26*, and 27* are similar to putative open reading frames in published databases. No functional data are available to characterize these genes, and their function remains speculative.

orf 4* encodes a membrane-spanning ATP-dependent export pump with close matches to other transporters found in glycopeptide gene clusters. It is therefore likely that this is responsible for the export of mature teicoplanin from the cell. Orf21* is similar to a number of ion transporters, and similar genes have been identified in other glycopeptide gene clusters. Its role is uncertain.

Heterologous Expression, Purification, and Assay of Putative Glycosyltransferases Orf1 and Orf10*

By heterologous expression of these two predicted glycosyltransferases in *E. coli*, we aimed both to verify their involvement in the production of teicoplanin and to assign them individual catalytic roles. In addition, knowledge of their substrate specificity would increase their usefulness in future attempts to engineer hybrid maturation pathways to obtain novel glycopeptides [15, 16, 17]. Each enzyme was expressed in soluble form in *E. coli* and purified by Ni-NTA affinity chromatography, as outlined in Experimental Procedures.

The enzymatic activity of Orf1 and Orf10* was monitored by incubation of the enzyme with teicoplanin (pseudo)aglycones and analysis of the products by LC-MS. Newly-formed peak(s) were compared with LC-MS of reference samples of partially hydrolyzed teicoplanin metabolites (Figures 3A and Figure 4, LC trace A). The identity of the peaks in the LC spectrum of the partially hydrolyzed teicoplanins was deduced as follows. The peak at 13.8 min with m/z 1401 $[M + H]^+$ must be *N*-acetyl-glucosaminoyl-3-Cl-6βHty pseudoaglycone (**2**), since only one *N*-acetyl-glucosamine moiety is present in teicoplanin. The two peaks both with m/z 1359 $[M+H]^+$ appearing at 10.8 and 13 min must arise from either glucosaminoyl-3-Cl-6βHty pseudoaglycone (**5**), which may undergo deacylation from *N*-acetyl-glucosaminoyl-3-Cl-6βHty or glucosaminoyl-4Hpg pseudoaglycone (**4**), which may undergo deacylation from *N*-acyl-glucosaminoyl-4Hpg. The fourth peak shown at 15 min with m/z 1360 $[M + H]^+$ must be mannosyl-7Dpg pseudoaglycone (**6**).

After incubation of purified recombinant Orf1 with UDP-*N*-acetyl-glucosamine and teicoplanin aglycone, the LC/MS trace of the assay mixture showed a new prominent peak at 13.9 min (Figure 4, LC trace C). The mass spectrum revealed that this peak has exactly the expected mass $[M + H]^+ = 1401$ corresponding to *N*-acetyl-glucosaminoyl heptapeptide pseudoaglycone and the unique isotope profile of a compound containing two chlorine atoms (Figure 4, **2**). Incubation with UDP-glucosamine as an alternative substrate did not produce any further new peaks. The *orf*1 gene product is therefore revealed as a glycosyltransferase that transfers UDP-*N*-acetyl-glucosamine to the C3 hydroxy of 3-Cl-6βHty in the teicoplanin heptapeptide aglycone to form (**2**).

The second GtfB-like glycosyltransferase in the *tei* cluster, Orf10*, was therefore expected to transfer a sugar to the 4Hpg residue of the teicoplanin aglycone. In agreement with this, when purified recombinant Orf10* was incubated with UDP-*N*-acetyl-glucosamine and teicoplanin aglycone, it resulted in the appearance of a new peak at 11.5 min in the LC/MS trace, again with m/z 1401, which did not match any of the reference compounds, consistent with it being structure **3** (Figure 4, LC trace D). When instead recombinant Orf10* was assayed with UDP-glucosamine, a different new peak was observed at 10.8 min in the LC/MS trace with m/z 1359, which is in excellent agreement with the reference compound glucosaminoyl-4Hpg pseudoaglycone (**4**) at 10.7 min (data not shown). Therefore, Orf10* is a glycosyltransferase transferring either UDP-glucosamine or UDP-*N*-acetyl-glucosamine specifically onto the 4HPG of the teicoplanin aglycone. Whether UDP-glucosamine or UDP-*N*-acetyl-glucosamine is the preferred substrate in vivo remains to be determined (see also below).

The fact that Orf1 and Orf10* glycosyltransferases can react independently with the teicoplanin aglycone distinguishes this system from chloroeremomycin biosynthesis, where glycoyltransfer appears to be ordered, with GtfB known to act first with residue 4 [15, 16]. Our analysis suggests that Orf1 has higher activity than Orf10* against the aglycone, meaning that glycosylation at residue 6 is faster than at residue 4, but detailed kinetic studies will be needed to confirm this (the conversion of teicoplanin aglycone to corresponding monosugar aglycone catalyzed by Orf1 and Orf10 with UDP-*N*-Acetyl-glucosamine present in the typical condition were 50% and 20%, respectively). When incubations contained both Orf1 and Orf10*, a new peak corresponding to the disaccharide pseudoaglycone (**7**) appeared in the LC trace, unambiguously confirming the role of these two enzymes in the glycosylation of teicoplanin (Figure 4, LC trace E). The proposed reaction sequences catalyzed by Orf1 and 10* are summarized in Figure 3B.

Orf1 and Orf10* were also tested using the vancomycin aglycone as substrate. However, after incubation no

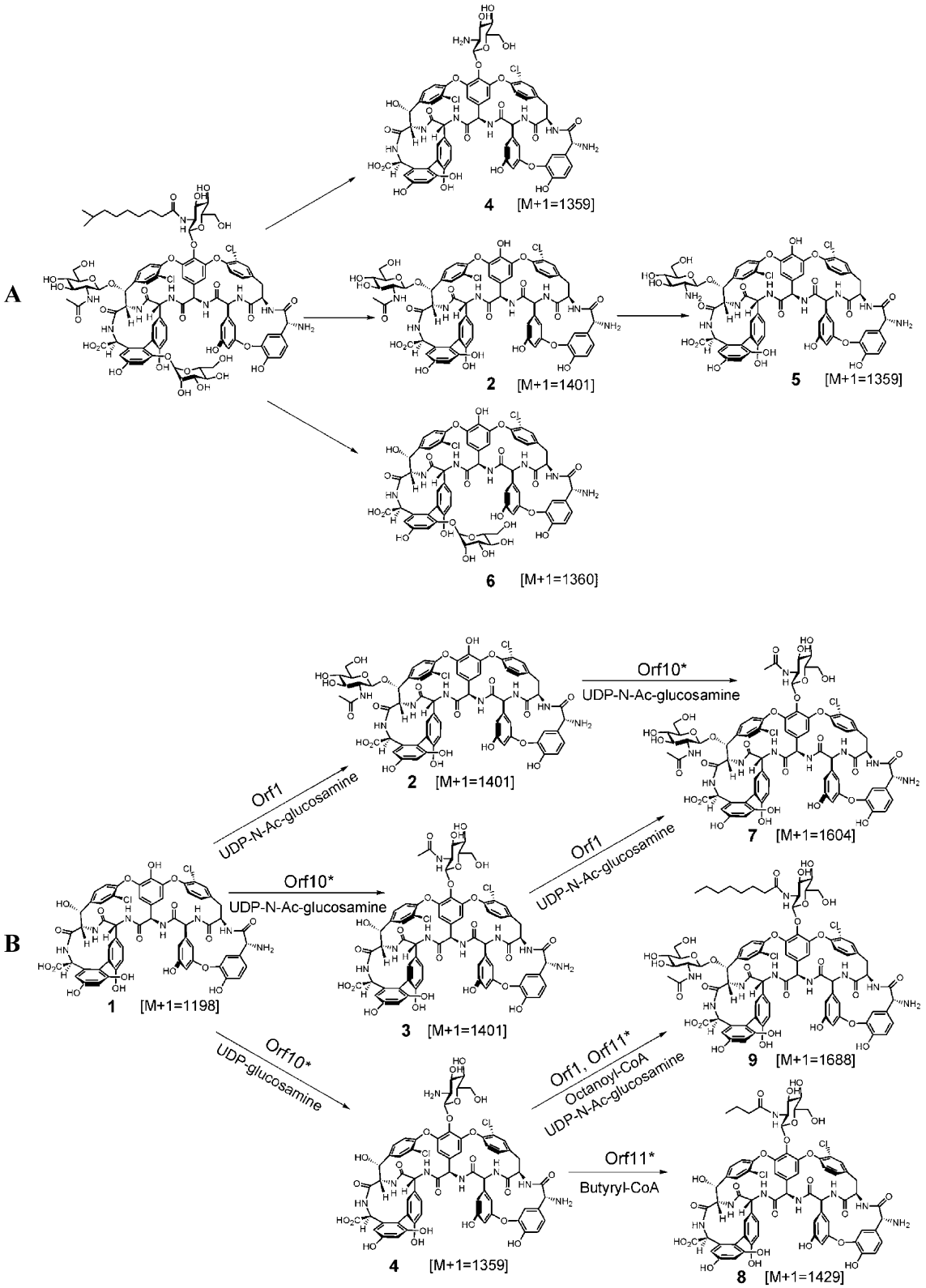
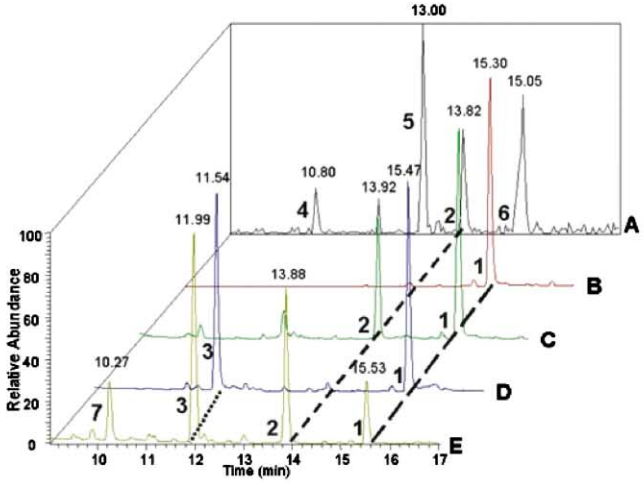


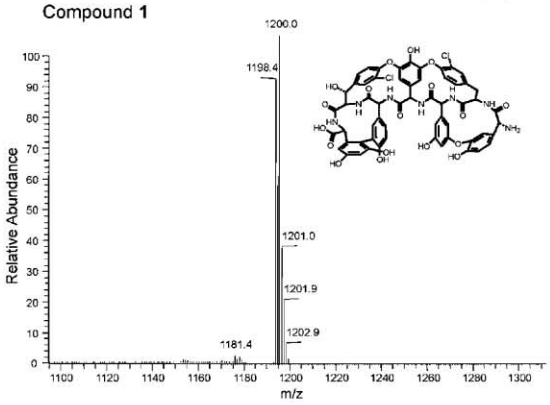
Figure 3. Structures of Reference Teicoplanins and of the Products of Coupled Enzymatic Reactions of Orf1, 10*, and 11*

(A) Products of partial hydrolysis of teicoplanin referred to in this study.

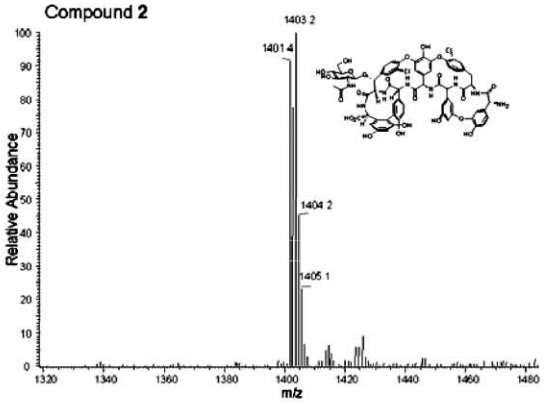
(B) The sequential enzymatic reactions of Orf1, 10*, and 11* are shown. Orf1 and Orf10* independently catalyze conversion of teicoplanin aglycone into the corresponding pseudoaglycones (**2** and **3**), which can each be further converted by the second glycosyltransferase to **7**. Orf10* transfers a UDP-glucosamine to teicoplanin aglycone, forming **4**, which can be acylated by Orf11* into **8** using butyryl-CoA or can be converted by coupled action of Orf1 and Orf11* into **9** using octanoyl-CoA and UDP-*N*-Ac-glucosamine.



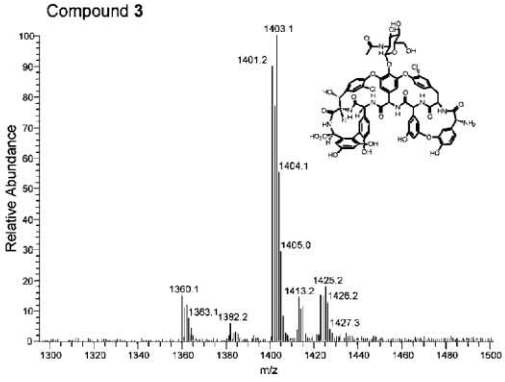
Compound **1**



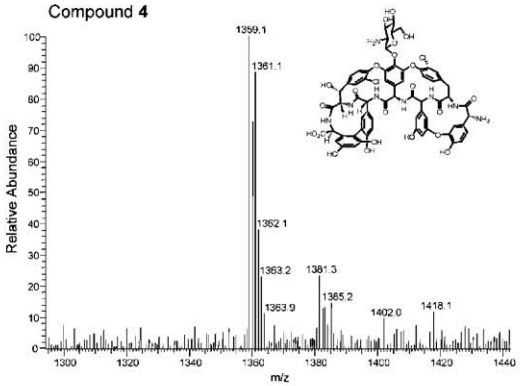
Compound **2**



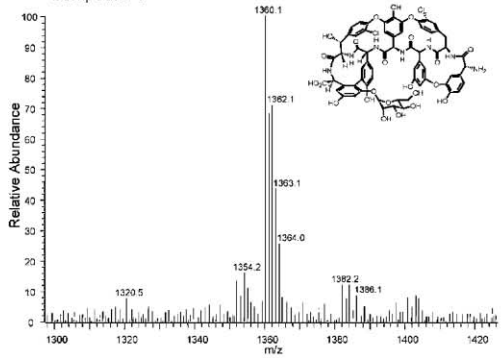
Compound **3**



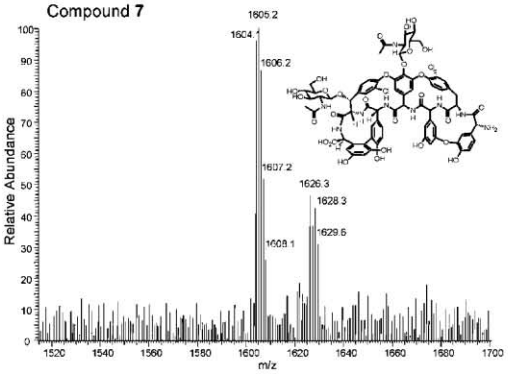
Compound **4**



Compound **6**



Compound **7**



new products were observed. This contrasts with the activity of GtfB in the *cep* cluster and GtfE in the vancomycin cluster, in which both enzymes can transfer activated sugar either to vancomycin or teicoplanin aglycone, but GtfE with much greater efficiency than GtfB [15, 16, 17]. Sugar donor specificity appears to be stricter for Orf1, which can only utilize UDP-*N*-acetyl-glucosamine, while Orf10* catalyzes sugar transfer from UDP-glucosamine, UDP-*N*-acetyl-glucosamine, and UDP-glucose. GDP-mannose was also used to test whether either enzyme could transfer mannosyl groups to the teicoplanin (pseudo)aglycone, but (as expected) no reaction was observed. The lower substrate specificity of Orf10* suggests its potential use to add other NDP-activated sugars to produce novel teicoplanin derivatives.

In the present work, mannosyl transfer was not examined directly, but our results are consistent with the view that mannosyl transfer is likely to be a later step in biosynthesis [49].

Heterologous Expression, Purification, and Assay of Putative Acyltransferase Orf11*

The superior pharmacokinetic profile of teicoplanin has been plausibly attributed to the hydrophobic aliphatic side chain on the aminosugar attached to residue 4. An acyltransferase (Orf23) found in the *dbv* cluster has been proposed to carry out this reaction [13], but there is no homolog of this gene in the *tei* cluster. On the other hand, *tei orf*11*, which is homologous to unassigned gene *orf*8 in the *dbv* cluster, is adjacent to glycosyltransferase gene *orf*10* and possibly cotranscribed with it, so we considered Orf11* an attractive candidate for the acyltransferase catalyzing the key *N*-acylation of glucosaminoyl-4Hpg pseudoaglycone.

Orf11* was expressed in *E. coli* as soluble protein, and the recombinant enzyme was purified by Ni-NTA affinity chromatography as detailed in Experimental Procedures. Its ability to transfer an acyl group to a glucosaminoyl-4Hpg pseudoaglycone was assayed by incubation together with teicoplanin aglycone, UDP-glucosamine, and butyryl-CoA in the presence of Orf10*. A new peak was observed with m/z 1429 $[M + H]^+$, and a characteristic isotope profile was revealed in the LC trace corresponding to compound **8** (Figure 5A, left panel). Formation of this compound was not seen in the absence of Orf11* (only compound **4** was formed instead). Orf11* was thus revealed as an acyltransferase with the specificity expected of the long-sought *N*-acyltransferase.

Incubation of Orf11* in the presence of both Orf1 and Orf10* with teicoplanin aglycone and a different acyl donor, octanoyl-CoA, confirmed this conclusion. A new peak was observed that eluted after teicoplanin aglycone (indicating increased hydrophobicity) and with the m/z 1688 $[M + H]^+$ and characteristic isotope profile for compound **9** (Figure 5A, right panel). Again, a control experiment without Orf11* did not lead to formation of this product.

To address whether the acylation occurs before or after sugar attachment to the teicoplanin aglycone, we added Orf11* to an assay containing only UDP-glucosamine and octanoyl-CoA as substrates. LC/MS analysis clearly revealed a new peak (absent from a control lacking Orf11*) with m/z 692 $[M + H]^+$, which corresponds to UDP-*N*-octanoyl-glucosamine (Figure 5B, LC trace a, **10**). This result showed that the acylation can occur with a free activated sugar. Next, the enzymatically synthesized UDP-*N*-octanoyl-glucosamine was used to test the glycosylation activity of Orf10* on this *N*-acylated substrate. In parallel, the acyltransferase activity of Orf11* was assayed using glucosaminoyl-4Hpg pseudoaglycone (prepared using Orf10*) and octanoyl-CoA as substrates. Remarkably, *N*-octanoyl-glucosaminoyl-4Hpg pseudoaglycone (with m/z 1484 $[M + H]^+$) was produced in both assays (Figure 5C, **11**). Evidently, the acylation catalyzed by Orf11* may occur either before or after the attachment of the aminosugar to the teicoplanin aglycone, and glycosyltransferase Orf10* can utilize as acyl donors either UDP-*N*-octanoyl-, UDP-*N*-butyryl-, or UDP-*N*-acetyl-glucosamine.

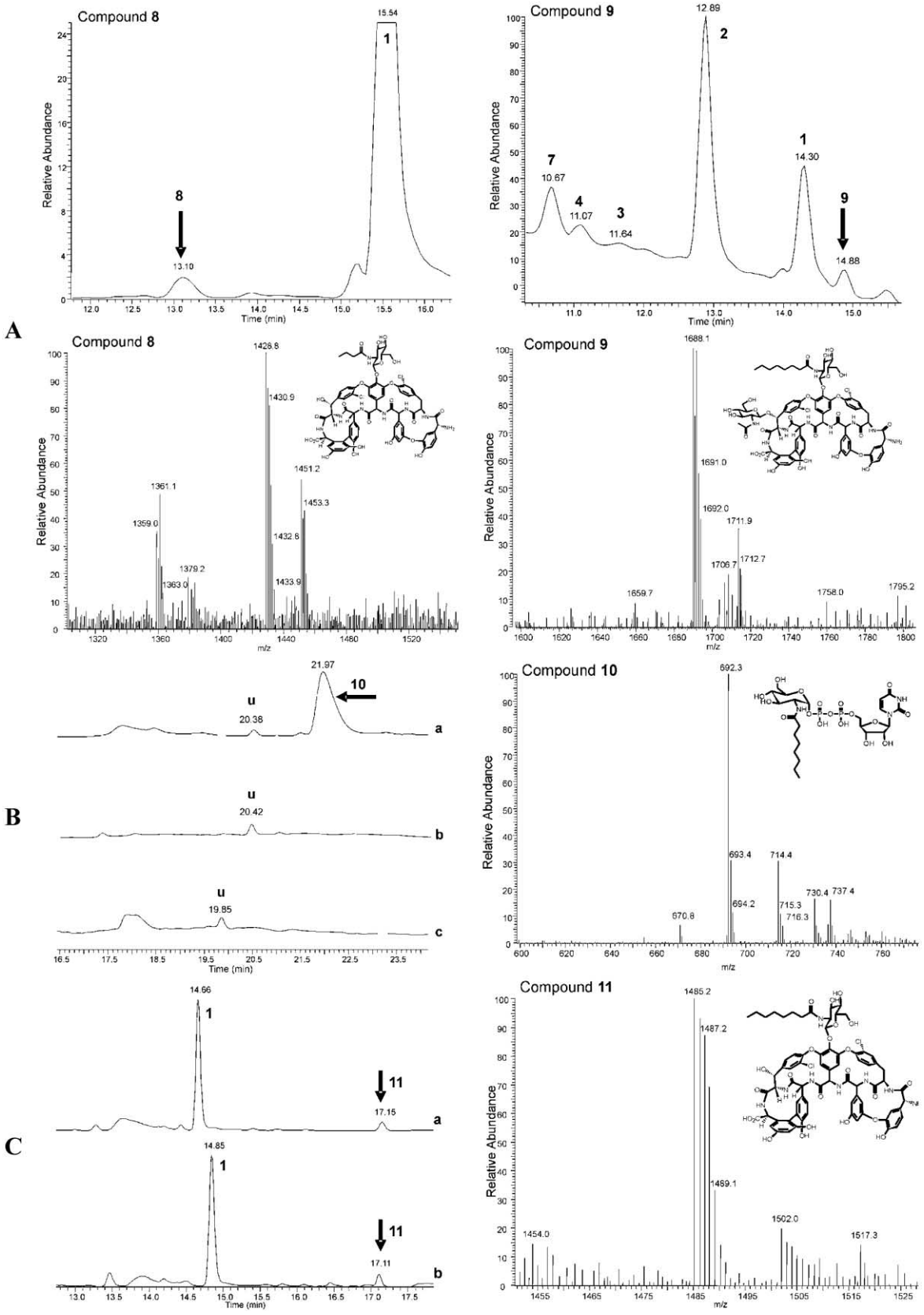
Our demonstration that acyl-CoA species can serve in vitro as acyl donors for the *N*-acylation does not exclude the possibility that in vivo acyl-ACP species could be alternative or even preferred donors for this reaction. We note that *orf*13* in the *tei* gene cluster, which has no counterparts in other related gene clusters, appears from sequence comparisons to encode an acyl-CoA synthase that may be involved in the ATP-dependent ligation of a free long fatty acid chain with free CoASH [50], which lends some support to the notion that acyl-CoA species are utilized as donors.

If UDP-*N*-acetyl-glucosamine is the preferred substrate for Orf10* in vivo, then removal of the acetyl group must precede attachment of the fatty acyl chain. We considered the possibility that Orf11* might behave as a bifunctional enzyme catalyzing this net transacylation, since there appear to be no other deacylases present in the *tei* cluster. However, there was no observable deacylation activity (data not shown) when Orf11* was incubated either with free UDP-*N*-acetyl-glucosamine or with *N*-acetyl-glucosaminoyl 4Hpg teicoplanin pseudoaglycone (prepared from enzymatic product of Orf10*). Therefore, if in vivo acylation follows aminosugar attachment, UDP-glucosamine may be the preferred substrate for Orf10*.

Whichever of these parallel pathways is followed, the observed relaxed substrate specificities of the newly identified Orf11* acyltransferase enzyme and of the Orf10* glycosyltransferase certainly suggest that they should prove useful in future engineering of novel glycopeptide production.

Figure 4. LC Traces of Enzymatic Assays and Mass Spectra of Relevant Compounds

The LC traces and peaks shown are as follows. Trace A, partially hydrolyzed reference compounds; trace B, teicoplanin aglycone; trace C, assay of Orf1; trace D, assay of Orf 10*; trace E, coupled assay of Orf1 and 10*. Compounds: **1**, teicoplanin aglycone; **2**, *N*-Ac-glucosaminoyl-3-Cl-6βHty pseudoaglycone; **3**, *N*-Ac-glucosaminoyl 4Hpg pseudoaglycone; **4**, glucosaminoyl-4Hpg pseudoaglycone; **5**, glucosaminoyl-3-Cl-6βHty pseudoaglycone; **6**, mannosyl-7Dpg pseudoaglycone; **7**, disaccharide pseudoaglycone.


A
B
C
Compound 8
Compound 9
Compound 10
Compound 11
Relative Abundance
Time (min)
m/z
15.54
1
8
13.10
12.89
2
7
10.67
4
11.07
3
11.64
1
14.30
9
14.88
1428.8
1430.9
1451.2
1453.3
1361.1
1359.0
1379.2
1363.0
1432.8
1433.9
1688.1
1691.0
1692.0
1711.9
1706.7
1712.7
1659.7
1758.0
1795.2
21.97
10
u
20.38
a
u
20.42
b
u
19.85
c
692.3
693.4
714.4
694.2
730.4
737.4
715.3
716.3
670.8
14.66
1
11
17.15
a
14.85
1
11
17.11
b
1485.2
1487.2
1489.1
1502.0
1454.0
1517.3

Significance

Teicoplanin is one of only two glycopeptide antibiotics approved for clinical use. The *tei* cluster required for its biosynthesis is the most comprehensive yet reported for glycopeptide formation, with 49 ORFs, including several unique features, an organization of the resistance genes that is wholly different from those found in other glycopeptide-producing organisms but closely related to the resistance genes in vancomycin-resistant enterococci (VRE), providing the clearest yet evidence of resistance transfer between species. Purified recombinant enzymes from the *tei* cluster have been used to obtain direct evidence for the identity and role of Orfs 1 and 10* as glycosyltransferases catalyzing the transfer of UDP-(*N*-acetyl)-glucosamine onto 3-Cl-6βHty and 4Hpg of teicoplanin heptapeptide aglycone, respectively. Another gene product designated Orf11* has been identified by in vitro assay as the key *N*-acyltransferase that transfers a long fatty acid chain to a sugar moiety of teicoplainin pseudoaglycone, a modification which appears to markedly enhance glycopeptide activity. These and other tailoring genes and enzymes of the *tei* cluster can now be deployed for the generation of new and improved glycopeptide drugs.

Experimental Procedures

Isolation and Sequencing of the *tei* Cluster

Genomic DNA was isolated from *Actinoplanes teichomyceticus* DSM 43866 by "procedure B" of Hopwood et al [51]. Preparation of a cosmid library of *A. teichomyceticus* chromosomal DNA in the vector Supercos 1 was carried out according to manufacturer's instructions (Stratagene). Briefly, total DNA was partially digested with BamHI, dephosphorylated with shrimp intestinal alkaline phosphatase, and purified prior to ligation into Supercos 1 vector without size fractionation. The ligation mixture was packaged with Gigapack III Gold Packaging Extract (Stratagene) followed by transfection into *E. coli*. XL1 Blue MR. Four hundred eighty individual colonies were inoculated into 1 ml of 2 × LB medium with selection for ampicillin and kanamycin resistance and incubated at 37°C overnight with shaking in 5 × 96 well pates (ABgene). A 100 μl aliquot of the grown culture was preserved in 40% glycerol in 5 × 96 well plates (Orange Scientific) and stored at −80°C. Cosmid DNA was prepared using High Yield Protocol 1 for plasmid miniprep using MultiScreen (Millipore) and end sequenced with T3 and T7 primers. All sequencing was performed using Big Dye Terminator kit in an ABI 373A sequencer according to the manufacturer's protocols (ABI) in the DNA Sequencing Facility of the Department of Biochemistry, University of Cambridge. Cosmids containing sequences with significant similarity to related known biosynthetic genes of glycopeptide antibiotics were identified using the BLAST algorithm and amplified in 50 ml cultures from the stored glycerol deeps and DNA prepared using Qiagen midiprep kits. These clones were subjected to partial digestion with Sau3AI, subcloned into pSHG397 plasmid [52], and sequenced utilizing pUC forward and reverse primers. The complete sequence of the cosmid was obtained from assembly of the overlapping Sau3*A*I fragments. The sequence was edited with SeqEd v1.0.3. The GAP 4 program (Staden) was employed for sequence assembly. NRPS *tei* genes were discovered in three overlapping cosmids spanning ~110 kb. Detailed sequence analysis confirmed that the boundary of the cluster 5′ to the NRPS modules had been cloned but not the 3′ boundary. It was therefore necessary to obtain a further cosmid that extended further downstream of the cloned sequence. Based on the newly identified sequence, three primers, TE1: 5′-AAG ACG ACG CGG ACG AAG GCG TCG GTG CA-3′, TE2: 5′-GAT GGC GAT GGA TCG GTC GGA GAT TAG CA-3′, and TE3: 5′-CTC TTC CGG GGC CAT CGT GAT CTC CCA TG-3′, were designed, synthesized (QIAGEN), and used in PCR screening of the 6 × 96 colonies of cosmid library. One cosmid that gave a positive PCR product was further subcloned and sequenced as above and was confirmed to extend beyond the downstream limit of the cluster A. The probable boundaries of the *tei* cluster were established by comparison with other glycopeptide gene clusters and from the deduced functions of the *tei* and flanking gene products.

Characterization of Orf1, Orf10*, and Orf11*

The *orf1* was amplified from cosmid p3G10, and the *orf10** and *11** genes were amplified from cosmid p4A7 by using PCR with the following primers: *orf1*-p1, 5′-GAA AAG GAT GTG **CAT ATG** CGC GTG CTG TTT-3′; *orf1*-p2, 5′-GCG GGC CCG GCG **GAA TTC** ACG CGG GAA CCG-3′; *orf10**-p1, 5′-TGG GGA TGC **CAT ATG** CGT GTG TTG TTG TCG-3′; *orf10**-p2, 5′-AGC CGG TG**G AAT TC**A CGC GGA GAC CGA CGA-3′; *orf11**-p1, 5′-CGT GAA **CAT ATG** ATG ATG GAT CCC GAG ACC-3′; *orf11**-p2, 5′-CTC CGC **GAA TTC** TTA CGG CAG CTT CAC GAA-3′. The products were each ligated according to the manufacturer's instructions (Novagene) into expression vector pET-28a(+), which provides an *N*-terminal His_6-tagged protein. *E. coli* cells were grown at 37°C in 1 liter of LB medium until an A_{600} of about 0.7 was reached. Protein expression was induced with 0.2 mM IPTG at 16°C, and after 12 hr cells were harvested by centrifugation, resuspended in 10 mM imidazole-HCl buffer (pH 7.8) (binding buffer), and ruptured by sonification. The cell-free extract was applied to a column of Ni^{2+}-NTA resin, which was then washed successively with binding buffer and wash buffer (30 mM imidazole-HCl) before eluting the target protein with 100 mM imidazole-HCl. The buffer was exchanged for 50 mM HEPES buffer (pH 7.2) using Millipore centrifugal filters. The relative molecular masses of purified proteins Orf1, Orf10*, and Orf11* were determined by electrospray mass spectrometry (ESI-MS) to be 43,227, 44,355, and 39,017 Da, respectively. The recombinant proteins are in good agreement with the predicted calculated masses, taking the His_6 tag into consideration.

Typical assays for the enzymes (10 μg each) discussed in the text contained HEPES buffer (50 mM, pH 7.2), UDP-activated sugar (2 mM), heptapeptide aglycone (2 mM), and coenzyme A derivative (2 mM) in the presence of 1 mM $MgCl_2$ and 1 mM DTT and were incubated (150 μl final assay volume) for 2.5 hr at 25°C. After incubation, mixtures were centrifuged at 16,000 × g for 5 min (Heraeus Biofuge Pico) and then spun in an ultracentrifugal filter unit fitted with a 5 kDa cut-off membrane (Millipore). The resulting filtrate was directly subjected to HPLC-ESI/MS analysis with a gradient solvent system of increasing acetonitrile (0.1% TFA) in water (0.1% TFA) from 0 to 60% over 30 min. On-line LC-ESI/MS spectra were obtained from a Hewlett-Packard HPLC 1100 series instrument coupled to a Finnigan MAT LCQ ion trap mass spectrometer fitted with a positive mode ESI source.

Figure 5. LC Traces of Assays Using Acyltransferase Orf11* and Mass Spectra of the Products

(A) LC traces and MS analysis of acylation products. Compounds: **8**, *N*-Bu-glucosaminoyl-4Hpg pseudoaglycone; **9**, *N*-octanoyl-glucosaminoyl-4Hpg, *N*-acetyl-glucosaminoyl-3-Cl-6βHty pseudoaglycone.

(B) LC traces and MS analysis of UDP-glucosamine acylation catalyzed by Orf11*, in which traces a and b are assays in the presence or absence, respectively, of Orf11* with UDP-glucosamine and octanoyl-CoA as substrates. Trace c is as for a, but with UDP-*N*-Ac-glucosamine instead of UDP-glucosamine. Compounds: **10**, UDP-*N*-octanoyl-glucosamine; u, unknown peak.

(C) LC traces and MS analysis of enzymatic assays catalyzed by Orf10* and Orf11*, respectively, in which traces a and b are assays using as substrates UDP-*N*-octanoyl-glucosamine and *N*-Ac-glucosamine-4Hpg pseudoaglycone, respectively. Compound **11**, *N*-octanoyl-glucosamine-4Hpg pseudoaglycone, was produced in both assays.

Acknowledgments

We thank Professors Dudley Williams (Department of Chemistry, University of Cambridge) and Robert Field (School of Chemical Science & Pharmacy, University of East Anglia) for kind gifts of glycopeptide aglycones and of UDP-glucosamine. This work was supported by project grants from the BBSRC to J.B.S., S.F.H., and P.F.L. T.-L.L. was supported by a Fellowship from the Ministry of Education, Taiwan.

Received: September 8, 2003
Revised: October 24, 2003
Accepted: November 6, 2003
Published: January 23, 2004

References

1. Khare, M., and Keady, D. (2003). Antimicrobial therapy of methicillin resistant *Staphylococcus aureus* infection. Expert Opin. Pharmacother. *4*, 165–177.
2. Nagarajan, R. (1993). Structure-activity relationships of vancomycin-type glycopeptide antibiotics. J. Antibiot. (Tokyo) *46*, 1181–1195.
3. Nicolaou, K.C., Boddy, C.N.C., Bräse, S., and Winssinger, N. (1999). Chemistry, biology, and medicine of the glycopeptide antibiotics. Angew. Chem. Int. Ed. Engl. *38*, 2096–2152.
4. Borghi, A., Edwards, D., Zerilli, L.F., and Lancini, G. (1991). Factors affecting the normal and branched-chain acyl moieties of teicoplanin components produced by *Actinoplanes teichomyceticus*. J. Gen. Microbiol. *137*, 587–592.
5. Borghi, A., Ferrari, P., Gallo, G.G., Zanol, M., Zerilli, L.F., and Lancini, G. (1991). Microbial de-mannosylation and mannosylation of teicoplanin derivatives. J. Antibiot. (Tokyo) *44*, 1444–1451.
6. Allen, N.E., and Nicas, T.I. (2003). Mechanism of action of oritavancin and related glycopeptide antibiotics. FEMS Microbiol. Rev. *26*, 511–532.
7. Marsden, A.F.A., Wilkinson, B., Cortés, J., Dunster, N.J., Staunton, J., and Leadlay, P.F. (1998). Engineering broader specificity into an antibiotic-producing polyketide synthase. Science *279*, 199–202.
8. Jacobsen, J.R., Hutchinson, C.R., Cane, D.E., and Khosla, C. (1997). Precursor-directed biosynthesis of erythromycin analogs by an engineered polyketides synthase. Science *277*, 367–369.
9. van Wageningen, A.M., Kirkpatrick, P.N., Williams, D.H., Harris, B.R., Kershaw, J.K., Lennard, N.J., Jones, M., Jones, S.J., and Solenberg, P.J. (1998). Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. Chem. Biol. *5*, 155–162.
10. Pelzer, S., Süßmuth, R., Heckmann, D., Recktenwald, J., Huber, P., Jung, G., and Wohlleben, W. (1999). Identification and analysis of the balhimycin biosynthetic gene cluster and its use for manipulating glycopeptide biosynthesis in *Amycolaptopsis mediterranei* DSM5908. Antimicrob. Agents Chemother. *43*, 1565–1573.
11. Chiu, H.T., Hubbard, B.K., Shah, A.N., Eide, J., Fredenburg, R.A., Walsh, C.T., and Khosla, C. (2001). Molecular cloning and sequence analysis of the complestatin biosynthetic gene cluster. Proc. Natl. Acad. Sci. USA *98*, 8548–8553.
12. Pootoolal, J., Thomas, M.G., Marshall, C.G., Neu, J.M., Hubbard, B.K., Walsh, C.T., and Wright, G.D. (2002). Assembling the glycopeptide antibiotic scaffold: The biosynthesis of A47934 from *Streptomyces toyocaensis* NRRL15009. Proc. Natl. Acad. Sci. USA *99*, 8962–8967.
13. Sosio, M., Stinchi, S., Beltrametti, F., Lazzarini, A., and Donadio, S. (2003). The gene cluster for the biosynthesis of the glycopeptide antibiotic A40926 by *Nonomuraea* species. Chem. Biol. *5*, 155–162.
14. Weist, S., Bister, B., Puk, O., Bischoff, D., Pelzer, S., Nicholson, G.J., Wohlleben, W., Jung, G., and Süssmuth, R.D. (2002). Fluorobalhimycin: a new chapter in glycopeptide antibiotic research. Angew. Chem. Int. Ed. Engl. *41*, 3383–3385.
15. Solenberg, P.J., Matsushima, P., Stack, D.R., Wilkie, S.C., Thompson, R.C., and Baltz, R.H. (1997). Production of hybrid glycopeptide antibiotics *in vitro* and in *Streptomyces toyocaensis*. Chem. Biol. *4*, 195–202.
16. Losey, H.C., Peczuh, M.W., Chen, Z., Eggert, U.S., Dong, S.D., Pelczer, I., Kahne, D., and Walsh, C.T. (2001). Tandem action of glycosyltransferases in the maturation of vancomycin and teicoplanin aglycones: novel glycopeptides. Biochemistry *40*, 4745–4755.
17. Losey, H.C., Jiang, J., Biggins, J.B., Oberthür, M., Ye, X.Y., Dong, S.D., Kahne, D., Thorson, J.S., and Walsh, C.T. (2002). Incorporation of glucose analogs by GtfE and GtfD from the vancomycin biosynthetic pathway to generate variant glycopeptides. Chem. Biol. *9*, 1305–1314.
18. Sosio, M., Bianchi, A., Bossi, E., and Donadio, S. (2000). Teicoplanin biosynthesis genes in *Actinoplanes teichomyceticus*. Antonie Van Leeuwenhoek *78*, 379–384.
19. Puk, O., Huber, P., Bischoff, D., Recktenwald, J., Jung, G., Süßmuth, R., van Pée, K., Wohlleben, W., and Pelzer, S. (2002). Glycopeptide biosynthesis in *Amycolaptopsis mediterranei* DSM5908: Function of a halogenase and a haloperoxidase/perhydrolase. Chem. Biol. *9*, 225–235.
20. Choroba, O.W., Williams, D.H., and Spencer, J.B. (2000). Biosynthesis of the vancomycin group of antibiotics: Involvement of an unusual dioxygenase in the pathway to (*S*)-4-hydroxyphenylglycine. J. Am. Chem. Soc. *122*, 5389–5390.
21. Li, T.L., Choroba, O.W., Charles, E.H., Sandercock, A.M., Williams, D.H., and Spencer, J.B. (2001). Characterisation of a hydroxymandelate oxidase involved in the biosynthesis of two unusual amino acids occurring in the vancomycin group of antibiotics. Chem. Commun., 1752–1753.
22. Hubbard, B.K., Thomas, M.G., and Walsh, C.T. (2000). Biosynthesis of L-*p*-hydroxyphenylglycine, a non-proteinogenic amino acid constituent of peptide antibiotics. Chem. Biol. *7*, 931–942.
23. Knaggs, A.R. (2003). The biosynthesis of shikimate metabolites. Nat. Prod. Rep. *20*, 119–136.
24. Bassan, A., Blomberg, M.R.A., and Siegbahn, P.E.M. (2003). Mechanism of dioxygen cleavage in tetrahydrobiopterin-dependent amino acid hydroxylases. Chemistry *9*, 106–115.
25. Pfeifer, V., Nicholson, G.J., Ries, J., Recktenwald, J., Schefer, A.B., Shawky, R.M., Schröder, J., Wohlleben, W., and Pelzer, S. (2001). A polyketide synthase in glycopeptide biosynthesis – The biosynthesis of the non-proteinogenic amino acid (*S*)-3,5-dihydroxyphenylglycine. J. Biol. Chem. *276*, 38370–38377.
26. Li, T.L., Choroba, O.W., Hong, H., Williams, D.H., and Spencer, J.B. (2001). Biosynthesis of the vancomycin group of antibiotics: characterisation of a type III polyketide synthase in the pathway to (*S*)-3,5-dihydroxyphenylglycine. Chem. Commun., 2156–2157.
27. Chen, H., Tseng, C.C., Hubbard, B.K., and Walsh, C.T. (2001). Glycopeptide antibiotic biosynthesis: enzymatic assembly of the dedicated amino acid monomer (*S*)-3,5-dihydroxyphenylglycine. Proc. Natl. Acad. Sci. USA *98*, 14901–14906.
28. Sandercock, A.M., Charles, E.H., Scaife, W., Kirkpatrick, P.N., O'Brien, S., Papageorgiou, E.A., Spencer, J.B., and Williams, D.H. (2001). Biosynthesis of the di-meta-hydroxyphenylglycine constituent of the vancomycin-group antibiotic chloroeremomycin. Chem. Commun., 1700–1701.
29. Chen, H., Hubbard, B.K., O'Connor, S.E., and Walsh, C.T. (2002). Formation of beta-hydroxy histidine in the biosynthesis of nikkomycin antibiotics. Chem. Biol. *9*, 103–112.
30. Chen, H., and Walsh, C.T. (2001). Coumarin formation in novobiocin biosynthesis: beta-hydroxylation of the aminoacyl enzyme tyrosyl-S-NovH by a cytochrome P450 NovI. Chem. Biol. *8*, 301–312.
31. He, J., Magarvey, N., Piraee, M., and Vining, L.C. (2001). The gene cluster for chloramphenicol biosynthesis in Streptomyces venezuelae ISP5230 includes novel shikimate pathway homologues and a monomodular non-ribosomal peptide synthetase gene. Microbiol. *147*, 2817–2829.
32. Schwarzer, D., Finking, R., and Marahiel, M.A. (2003). Nonribosomal peptides: from genes to products. Nat. Prod. Rep. *20*, 275–287.
33. Mootz, H.D., Schwarzer, D., and Marahiel, M.A. (2002). Ways of

assembling complex natural products on modular nonribosomal peptide synthetases. Chembiochem *3*, 490–504.
34. Konz, D., and Marahiel, M.A. (1999). How do peptide synthetases generate structural diversity? Chem. Biol. *6*, R39–R48.
35. Stachelhaus, T., Mootz, H.D., and Marahiel, M.A. (1999). The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. Chem. Biol. *6*, 493–505.
36. Challis, G.L., Ravel, J., and Townsend, C.A. (2000). Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem. Biol. *7*, 211–224.
37. Schwarzer, D., Mootz, H.D., Linne, U., and Marahiel, M.A. (2002). Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases. Proc. Natl. Acad. Sci. USA *99*, 14083–14088.
38. Heathcote, M.L., Staunton, J., and Leadlay, P.F. (2001). Role of type II thioesterases: evidence for removal of short acyl chains produced by aberrant decarboxylation of chain extender units. Chem. Biol. *8*, 207–220.
39. Williams, D.H., and Bardsley, B. (1999). The vancomycin group of antibiotics and the fight against resistant bacteria. Angew. Chem. Int. Ed. Engl. *38*, 1173–1193.
40. Bischoff, D., Pelzer, S., Bister, B., Nicholson, G.J., Stockert, S., Schirle, M., Wohlleben, W., Jung, G., and Süssmuth, R.D. (2001). The biosynthesis of vancomycin-type glycopeptide antibiotics: the order of the cyclization steps. Angew. Chem. Int. Ed. Engl. *40*, 4688–4691.
41. Bischoff, D., Pelzer, S., Holtzel, A., Nicholson, G.J., Stockert, S., Wohlleben, W., Jung, G., and Süssmuth, R.D. (2001). The biosynthesis of vancomycin-type glycopeptide antibiotics: new insights into the cyclization steps. Angew. Chem. Int. Ed. Engl. *40*, 1693–1696.
42. Süssmuth, R.D., Pelzer, S., Nicholson, G., Walk, T., Wohlleben, W., and Jung, G. (1999). New advances in the biosynthesis of glycopeptide antibiotics of the vancomycin type from *Amycolaptopsis mediterranei*. Angew. Chem. Int. Ed. Engl. *38*, 1976–1979.
43. Boneca, I.G., and Chiosis, G. (2003). Vancomycin resistance: occurrence, mechanisms and strategies to combat it. Expert Opin. Ther. Targets *7*, 311–328.
44. Kaplan, J., Korty, B.D., Axelsen, P.H., and Loll, P.J. (2001). The role of sugar residues in molecular recognition by vancomycin. J. Med. Chem. *44*, 1837–1840.
45. Cetinkaya, Y., Falk, P., and Mayhall, C.G. (2000). Vancomycin-resistant enterococci. Clin. Microbiol. Rev. *13*, 686–707.
46. Murray, B.E. (1997). Vancomycin-resistant enterococci. Am. J. Med. *102*, 284–293.
47. Reynolds, P.E. (1998). Control of peptidoglycan synthesis in vancomycin-resistant enterococci: D,D-peptidases and D,D-carboxypeptidases. Cell. Mol. Life Sci. *54*, 325–331.
48. Fuqua, C., Winans, S.C., and Greenberg, E.P. (1996). Census and consensus in bacterial ecosystems: the LuxR-LuxI family of quorum-sensing transcriptional regulators. Annu. Rev. Microbiol. *50*, 727–751.
49. Smith, T.K., Cottaz, S., Brimacombe, J.S., and Ferguson, M.A. (1996). Substrate specificity of the dolichol phosphate mannose: glucosaminyl phosphatidylinositol alpha1–4-mannosyltransferase of the glycosylphosphatidylinositol biosynthetic pathway of African trypanosomes. J. Biol. Chem. *271*, 6476–6482.
50. Schnurr, J.A., Shockey, J.M., de Boer, G.J., and Browse, J.A. (2002). Fatty acid export from the chloroplast. Molecular characterization of a major plastidial acyl-coenzyme A synthetase from Arabidopsis. Plant Physiol. *129*, 1700–1709.
51. Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., and Hopwood, D.A. (2000). Practical Streptomyces Genetics (Norwich, United Kingdom: Crowes).
52. Hashimoto-Gotoh, T., Mizuno, T., Ogasahara, Y., and Nakagawa, M. (1995). An oligodeoxyribonucleotide-directed dual amber method for site-directed mutagenesis. Gene *152*, 271–275.