J Ind Microbiol Biotechnol (2006) 33: 129–140
DOI 10.1007/s10295-005-0028-5

ORIGINAL PAPER-JMBM

**Vivian Miao · Renee Brost · Joanne Chapple
Kevin She · Marie-François Coëffet-Le Gal
Richard H. Baltz**

# The lipopeptide antibiotic A54145 biosynthetic gene cluster from *Streptomyces fradiae*

Received: 13 April 2005 / Accepted: 1 August 2005 / Published online: 6 October 2005


**Abstract** $Ca^{2+}$-dependent cyclic lipodepsipeptides are an emerging class of antibiotics for the treatment of infections caused by Gram-positive pathogens. These compounds are synthesized by nonribosomal peptide synthetase (NRPS) complexes encoded by large gene clusters. The gene cluster encoding biosynthetic pathway enzymes for the *Streptomyces fradiae* A54145 NRP was cloned from a cosmid library and characterized. Four NRPS-encoding genes, responsible for subunits of the synthetase, as well as genes for accessory functions such as acylation, methylation and hydroxylation, were identified by sequence analysis in a 127 kb region of DNA that appears to be located subterminally in the bacterial chromosome. Deduced epimerase domain-encoding sequences within the NRPS genes indicated a D-stereochemistry for Glu, Lys and Asn residues, as observed for positionally analogous residues in two related compounds, daptomycin, and the calcium-dependent antibiotic (CDA) produced by *Streptomyces roseosporus* and *Streptomyces coelicolor*, respectively. A comparison of the structure and the biosynthetic gene cluster of A54145 with those of the related peptides showed many similarities. This information may contribute to the design of experiments to address both fundamental and applied questions in lipopeptide biosynthesis, engineering and drug development.

*Note*: A patent application on the sequence of the A54145 gene cluster was published (PCT Int. Appl., W003060127A2) during the preparation of this manuscript. A small deletion of 67 nucleotides (nucleotides 56952–57018) was observed relative to the sequence assembled in this study. The present sequence, obtained from a library clone, was confirmed by sequencing products from replicate PCR amplifications of the region in question from two stocks of *S. fradiae* DNA.

V. Miao · R. Brost · J. Chapple · K. She · M.-F. C.-L. Gal
R. H. Baltz (✉)
Cubist Pharmaceuticals Inc., 65 Hayden Av.,
Lexington, MA 02421, USA
E-mail: rbaltz@cubist.com
Tel.: +1-781-8608444
Fax: +1-781-8611164

**Keywords** Biosynthetic genes · Lipopeptide · Nonribosomal peptide synthetase · *Streptomyces fradiae*

## Introduction

A54145 is a family of related acidic, cyclic lipodepsipeptide antibiotics produced by *Streptomyces fradiae* [5]. A54145 contains a 13 amino acid peptide core that is cyclized by an ester bond to give a ten-membered ring with a three amino acid tail (Fig. 1). The core includes four unusual amino acids: 3-hydroxyasparagine ($hAsn_3$), 3-*O*-methyl-aspartic acid ($OmAsp_9$), 3-methyl-glutamic acid ($3mGlu_{12}$) and sarcosine ($Sar_5$). Natural variants are formed in the ring when Glu is incorporated instead of 3mGlu at position 12, and when Val is added instead of Ile at position 13. Additional members of the family derive from variation in the lipid: $Trp_1$ of the tail may be coupled to *n*-decanoyl, 8-methylnonanoyl, or 8-methyldecanoyl lipid side chains [15]. Eight factors are normally produced in substantial amounts by fermentation [5], but their ratios can be modulated by feeding amino acids such as L-Val, L-Ile or L-Glu. New acyl derivatives can also be generated by feeding various medium-chain fatty acids such as hexanoic acid, caprylic acid, or nonanoic acid [6].

A54145 has antibacterial activity against strains of *Staphylococcus, Streptococcus, Clostridium* and enterococci [12]. The compounds with $3mGlu_{12}$ are generally more potent and more toxic than those with $Glu_{12}$, and compounds with longer acyl side chains are more active than those with shorter side chains. Fukuda et al. [14] removed the lipid side chains of the *N*-Lys-(*tert*-BOC) protected A54145 factors by incubation with *Actinoplanes utahensis* which secretes a deacylase. Three of the deacylated A54145 factors were reacylated with different fatty acids, and the semisynthetic *N*-acyl compounds were tested for antibacterial activity. Several $Glu_{12}$-containing analogs had efficacy against *Staphylococcus aureus* and *Streptococcus pyogenes* in mice [12].

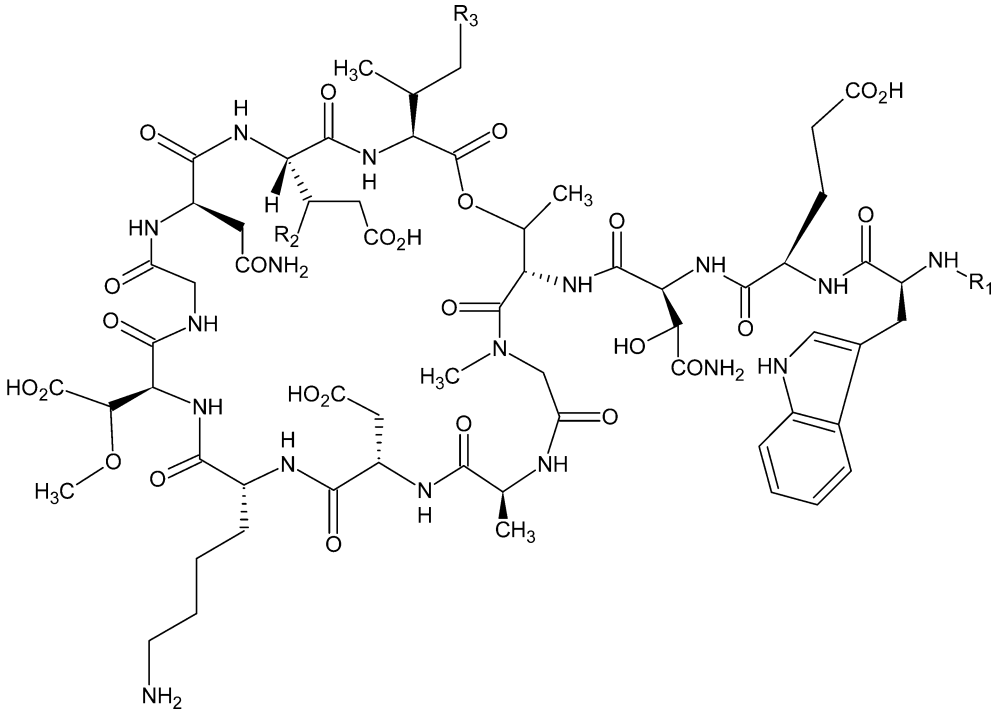

**Fig. 1** Structure and proposed stereochemistry of A54145. $R_1$ = *n*C10, *i*C10 or *a*C11 fatty acid side chain; $R_2$ = H or $CH_3$; $R_3$ = H or $CH_3$. Amino acid residues as discussed in this paper are numbered starting from *N*-terminal, exocyclic Trp attached to the R group and proceeding clockwise

A key aspect of the biological activity of A54145 is that it is $Ca^{2+}$-dependent [12] like two other cyclic lipodepsipeptides: daptomycin produced by *Streptomyces roseosporus* [2, 13], and the calcium-dependent antibiotic (CDA) produced by *Streptomyces coelicolor* [18, 22]. Daptomycin is the active ingredient in Cubicin® (daptomycin-for-injection), and the interaction with calcium is a key component of its bactericidal activity. Other similarities shared by A54145, daptomycin and CDA include the ten-membered depsipeptide ring that is formed by an ester linkage through the hydroxyl group of Thr, and the presence of Asp (or a modified Asp) and 3mGlu (or Glu) residues at the same positions in the ring (Fig. 2). Daptomycin and CDA have D- or achiral residues at equivalent positions as well, but the published literature on A54145 does not address amino acid stereochemistry. The stereochemistry at position 2 in the tail of daptomycin was recently assigned during analysis of the daptomycin biosynthetic gene cluster [27]. The undoubted relationship between the structural similarities and the mode of action of this class of compounds can be explored productively by analoging chemically or biosynthetically.

Like CDA and daptomycin, A54145 is synthesized by a nonribosomal peptide synthetase (NRPS). The NRPS enzymes synthesize peptides, including many secondary metabolites that may include nonproteinogenic and D-amino acids, on large multisubunit, multimodular enzymes [30]. A typical module in an NRPS is comprised of a C-domain for condensing the nascent peptide and a new amino acid, an A-domain for activating a specific amino acid, and a carrier (thiolation or T-) domain following the A-domain; accessory domains for methylation (M-) or epimerization (E-) that tailor individual amino acids may also be present. The modules, grouped often in protein subunits, act in an ordered manner to build the final peptide, and a terminal thioesterase (Te) domain participates in the release of the completed peptide from the enzyme. Earlier reports [2, 3] briefly describe obtaining portions of the A54145 biosynthetic gene cluster from an *S. fradiae* cosmid library by heterologous hybridization using an NRPS probe sequence. The A54145 gene sequences were not published, although sequence sampling resulted in identification of *cps-1*, a 4.2 kb fragment encoding an NRPS module that was used for disrupting A54145 production by targeted insertional inactivation. Rare restriction sites in the inserted plasmid allowed the mapping of the A54145 biosynthetic gene cluster to a locus 150–200 kb from one end of the linear chromosome of *S. fradiae*.

We report here the construction of an independent library of *S. fradiae* DNA, and the cloning and characterization of the complete A54145 biosynthetic pathway (*lpt*) from *S. fradiae*. We deduced the chirality for amino acids in A54145, and compared the features of the *lpt* gene cluster with those for the biosynthesis of the daptomycin (*dpt*) and CDA (*cda*) [16, 27].

## Materials and methods

### Strains

*S. fradiae* NRRL18158 was used as the source of the library and maintained on trypticase soy agar. *Escherichia coli* XL1-Blue-MR (Stratagene, La Jolla, Calif.) was used for propagating cosmid clones.

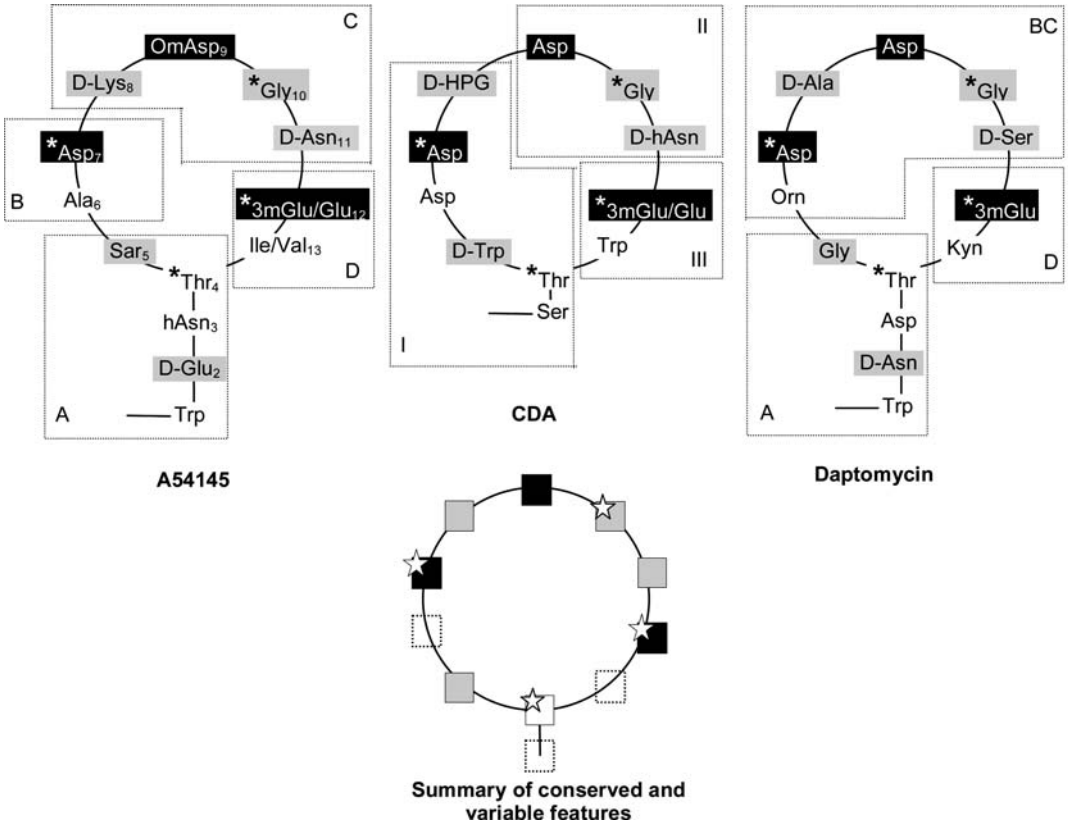


**Fig. 2** Peptide core of cyclic lipopeptides. Non-proteinogenic amino acids: $hAsn_3$ hydroxyasparagine, $OmAsp_9$ 3-*O*-methyl-aspartic acid, $3mGlu_{12}$ 3-methyl-glutamic acid, $Sar_5$ sarcosine, *HPG* hydroxyphenyl glycine, *Orn* ornithine, *Kyn* kynurenine. *Top*: *dotted lines* demarcate groups of amino acids in A54145, CDA and daptomycin that are incorporated by the indicated subunits: LptA, LptB, LptC and LptD in A54145; CdaPSI, CdaPSII and CdaPSIII in CDA; and DptA, DptBC and DptD in daptomycin. Conserved features among the compounds: *asterisks* identical residues at positions 4, 7, 10; *black highlight* acidic residues at 7, 9, 12; *gray background* achiral or D-residues at 2, 5, 8, 10 and 11 (some residues belong to more than one feature category). *Bottom*: composite of conserved features include identical (*asterisks*), acidic (*black*), D- or

## Cosmid library construction and screening

Biomass for DNA extraction was grown from *S. fradiae* spores and mycelium inoculated into 25 ml F10A medium and shaken overnight at 200 rpm at 30°C. Cells were harvested and processed to release genomic DNA into agar plugs as described previously [27]. DNA was partially digested with *Bam*HI, and fragments of 30–60 kb were purified by gel electrophoresis, ligated to a *Bam*HI-digested cosmid vector, and packaged with Gigapack III Gold packaging extract (Stratagene). The vector, pStreptoCos X, is derived from SuperCos (Stratagene) and was modified to carry markers for resistance to ampicillin and apramycin, as well as sequences for conjugative transfer (*oriT*) and site-specific integration (*att*P/*int*$^{\phi C31}$) [21]. Packaged DNA was introduced into *E. coli* XL1-Blue-MR and a library of 1900 arrayed clones was screened by PCR. A first set of 11 cosmids carrying genes from the A54145 biosynthetic pathway were identified using primers P138 (5′-CCGTGCTCTTCGACACCTCC-3′) and P139 (5′-CGTGTCGGTCTTCGACACCCT-3′) derived from *cps-1* (Genbank AF016696); a second set of flanking clones was then found by cosmid walking.

## DNA sequencing

End-sequencing of cosmid DNAs using vector primers P196 (5′-CGACGGCCAGTGAGCGCGCG-3′) and P197 (5′-GTCCGTGGAATGAACAATGGAAG-3′) and restriction mapping was used to identify the best candidates for sequencing. A 45 kb contig, represented by pCV14 and pCV15, was obtained by primer walking and then confirmed and extended into pCV16 and pCV17 by shot gun sequencing (SeqWright, Houston, Tx.). Primer walking was initiated from multiple sites by using *cps-1* as well as sequences derived from the ends of the 11 overlapping cosmids or from amplifying expected landmark regions, and ordering these fragments for an optimized walking strategy using multiple starting points. Landmarks were established using degenerate primers to conserved NRPS features. For example, comparison of *dpt* and *cda* sequences lead to design of primers P140 and P141 (5′-ACSSWSGGSGTSSCCTTCATGAA-3′, 5′-ATGGTGTTCGAGAACTAYCC-3′) targeting the 3mGlu-incorporating module, and P144 and P145 (5′-SCSCTSCAGGAGGGSHTSSTSTTCC-3′, 5′-CCGAASACSACGTCGTCSCGSCC-3′) targeting a $C^{II}$ type of condensation domain [27]. Each

walking step was conducted on two or more independent library clones. Flanking sequences were obtained from pCV16 and pCV17 and contiguity of the cloned DNA was verified by PCR across the internal *Bam*HI junctions. Contiguity between pCV14 and pCV16, which do not overlap but adjoin at the same *Bam*HI site, was validated by PCR using oligonucleotides that primed on each side of the junction, as well as by comparing sequences of these PCR products with the same region from overlapping clones. The final contig was deposited in Genbank as DQ118863.

Data were assembled and analyzed using MacVector Version 7.0 (Accelrys, San Diego, Calif.) and comparisons to publicly available genes were made using BLAST [1]. Protein alignments were performed initially in Clustal W, and adjusted slightly by hand. A dendrogram comparing A-domains was constructed by the Neighbor-Joining method (MacVector 7.0, default values for user definable parameters) using aligned protein sequences between motifs A4 and A5 [25] from the public domain. Amino acid binding pocket residues were determined by alignment with *grsA* PheA domain [33].

## Results

### Isolation of cosmid clones carrying the A54145 biosynthetic genes

A cosmid library of *S. fradiae* NRRL18158 with an average insert size of $34 \pm 7$ kb was screened by PCR using primers derived from the ends of the 4.2 kb *cps-1* fragment. Analysis of *cps-1* indicated that it corresponded to the module for incorporating $Ala_6$: this fragment would mark the center of the pathway, provided there was colinearity between the NRPS modules and the A54145 peptide sequence. Screening consisted of using *cps-1* to first find clones that included large portions of the NRPS genes, and then cosmid walking to find flanking DNA. The use of NRPS landmark regions obtained through degenerate PCR as additional starting points for sequencing accelerated data gathering and also permitted early examination of regions of particular interest. The final assembled contig (Fig. 3) encompassed a 127 kb region, with 61 open reading frames (ORFs) (Table 1).

### A54145 NRPS biosynthetic genes

The region represented by pCV14 and pCV15 contained 14 ORFs, including four terminally overlapping genes (*lptA, lptB, lptC* and *lptD*) that encode subunits of the A54145 synthetase (Fig. 4, Table 1). Detailed analyses of the deduced LptA, LptB, LptC and LptD proteins showed all of the signature motifs and domains indicative of an NRPS [25] (Fig. 4). Together, they comprised 13 modules responsible for incorporating the 13 amino acids in A54145. Biochemical studies [35] on partially purified immunoreactive *S. fradiae* enzyme fractions previously suggested that A54145 synthetase was composed of three subunits, but the DNA sequence data support a four subunit model in which the subunits incorporate 5, 2, 4 and 2 amino acids, respectively. The highly similar daptomycin NRPS has three subunits, with the largest, DptBC, being analogous to a fusion of LptB and LptC [27]. The subunit junction sequences suggest that the LptB to LptC junction may be unusual (below).

The A-domain of each module has a binding pocket region responsible for specifying the selection of amino acids [9, 33]. Comparisons of predicted Lpt A-domains and their specificity conferring residues with those from characterized NRPSs were generally consistent with the amino acid content and order of residues expected for A54145 (Figs. 4 and 5); for example, the domains activating $Thr_4$, $Ala_6$, as well as $Gly_{10}$ and $Sar_5$ (*N*-methyl-$Gly_5$) were clearly identifiable by clustering with corresponding sequences from known modules (Fig. 5).

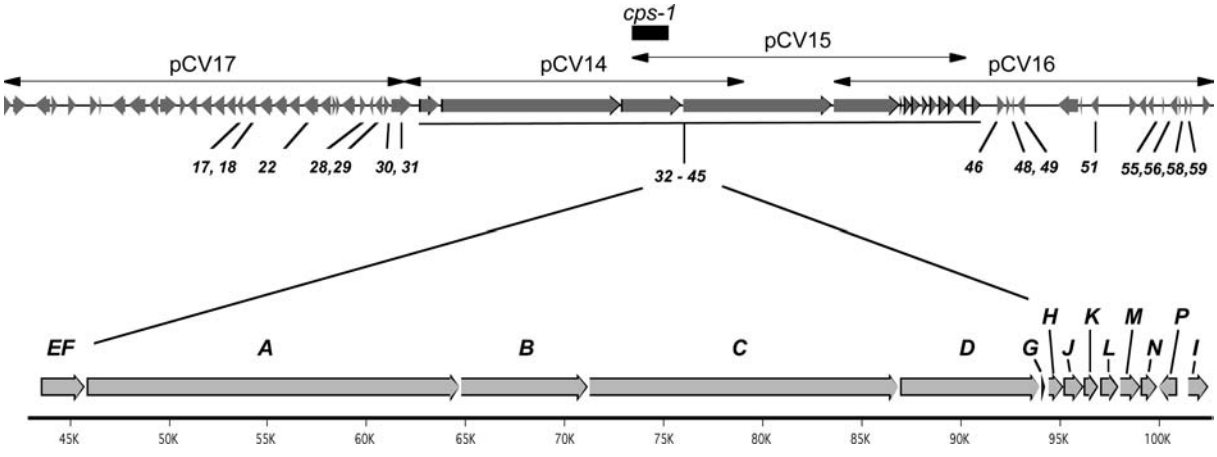


**Fig. 3** A54145 biosynthetic gene cluster and surrounding DNA. The four main cosmids, *cps-1* [4], and ORFs in the 127 kb region of *S. fradiae* chromosomal DNA are shown. The ORF32 to ORF45 region is expanded below to show the named *lpt* genes

**Table 1** *S. fradiae* open reading frames (ORFs) in the *lpt* region. ORFs are located as shown in Fig. 3. A predicted protein size (number of amino acids) is provided. Functions were hypothesized on the basis of BLAST analyses where only E-values $< 10$ (probability of random hit) were considered significant. NCBI references for the best hit at time of search are given. *Dash* no significant hits (ORF 14 and ORF 25); *asterisk* similarity to known sequences (ORFs 11, 52 and 53), but no associated predicted protein size owing to a possible frame shift in each (they were not reinvestigated as they were distant from the main *lpt* genes)

| ORF number | Predicted size (aa) | Gene name | Proposed function/similarity to | Reference | Source | *E*-value |
|---|---|---|---|---|---|---|
| 1 | 282 | | Possible amidohydrolase | NP_691226 | *Oceanobacillus iheyensis* | $e^{-48}$ |
| 2 | 541 | | Putative *p*-aminobenzoyl-glutamate transporter | NP_736715 | *Corynebacterium glutamicum* | $e^{-132}$ |
| 3 | 511 | | Putative histidine ammonia-lyase | NP_824501 | *S. avermitilis* | 0 |
| 4 | 255 | | Possible transcriptional regulator | YP_063002 | *Leifsonia xyli* subsp. *xyli str* | $e^{-69}$ |
| 5 | 298 | | Putative hydrolase | NP_631261 | *S. coelicolor* | $e^{-117}$ |
| 6 | 352 | | Putative oxidoreductases | YP_051382 | *Erwinia carotovora* subsp. *atroseptica* | $e^{-82}$ |
| 7 | 146 | | Hypothetical protein | NP_628834 | *S. coelicolor* | $e^{-26}$ |
| 8 | 499 | | Putative sugar kinase | NP_625114 | *S. coelicolor* | $e^{-168}$ |
| 9 | 679 | | Putative dehydrogenase | NP_828590 | *S. avermitilis* | 0 |
| 10 | 389 | | Putative sugar isomerase | NP_625112 | *S. coelicolor* | 0 |
| 11 | * | | (Probable glycosidase) | (AAM88355) | (*Streptomyces narbonensis*) | ($e^{-67}$) |
| 12 | 265 | | Hypothetical protein | NP_822661 | *S. avermitilis* | $e^{-33}$ |
| 13 | 347 | | Possible hydrolase | NP_631489 | *S. coelicolor* | $e^{-27}$ |
| 14 | 361 | | – | | | – |
| 15 | 430 | | Putative integral membrane protein | NP_627949 | *S. coelicolor* | $e^{-152}$ |
| 16 | 385 | | Putative secreted protein | NP_627950 | *S. coelicolor* | $e^{-92}$ |
| 17 | 227 | | Putative response regulator | NP_627946 | *S. coelicolor* | $e^{-100}$ |
| 18 | 465 | | Probable sensor kinase | NP_627947 | *S. coelicolor* | $e^{-123}$ |
| 19 | 550 | | Probable D-aminoacylase | ZP_00329501 | *Moorella thermoacetica* | $e^{-102}$ |
| 20 | 479 | | Possible permease | ZP_00271523 | *Ralstonia metallidurans* | $e^{-58}$ |
| 21 | 470 | | Putative decarboxylase | NP_624642 | *S. coelicolor* | $e^{-153}$ |
| 22 | 561 | | Putative DNA-binding protein | NP_624643 | *S. coelicolor* | $e^{-90}$ |
| 23 | 433 | | Hypothetical protein | NP_736825 | *Corynebacterium efficiens* | $e^{-25}$ |
| 24 | 192 | | Putative non-heme chloroperoxidase | NP_102663 | *Mesorhizobium loti* | $e^{-16}$ |
| 25 | 158 | | Hypothetical protein | YP_119431 | *Nocardia farcinica* | $e^{-22}$ |
| 26 | 473 | | – | | | – |
| 27 | 273 | | Hypothetical protein | AAQ93598 | *Streptomyces clavuligerus pSCL2* | $e^{-22}$ |
| 28 | 241 | | Possible response regulator | AAC45136 | *Streptomyces lividans* | $e^{-76}$ |
| 29 | 282 | | Probable sensor kinase | NP_825857 | *S. avermitilis* | $e^{-54}$ |
| 30 | 236 | | Probable ABC transporter, ATP-binding protein | NP_923438 | *Gloeobacter violaceus* | $e^{-51}$ |
| 31 | 769 | | Putative ABC transporter, permease | NP_827479 | *S. avermitilis* | $e^{-35}$ |
| 32 | 732 | *lptEF* | Acyl-CoA ligase | CAE53367 | *Actinoplanes teichomyceticus* | $e^{-118}$ |
| 33 | 6292 | *lptA* | Peptide synthetase | NP_627443 | *S. coelicolor* | 0 |
| 34 | 2143 | *lptB* | Peptide synthetase | NP_627443 | *S. coelicolor* | 0 |
| 35 | 5246 | *lptC* | Peptide synthetase | NP_627444 | *S. coelicolor* | 0 |
| 36 | 2384 | *lptD* | Peptide synthetase | NP_733597 | *S. coelicolor* | 0 |
| 37 | 80 | *lptG* | Possible esterase | NP_627432 | *S. coelicolor* | $e^{-21}$ |
| 38 | 264 | *lptH* | Probable thioesterase | NP_627445 | *S. coelicolor* | $e^{-73}$ |
| 39 | 331 | *lptJ* | Possible regulator | NP_824815 | *S. avermitilis* | $e^{-110}$ |
| 40 | 262 | *lptK* | Probable *O*-methyltransferase | AAC44360 | *Streptomyces hygroscopicus* | $e^{-36}$ |
| 41 | 319 | *lptL* | Probable hydrolase | NP_627448 | *S. coelicolor* | $e^{-65}$ |
| 42 | 353 | *lptM* | Probable ABC transporter | CAC22118 | *Streptomyces griseus* | $e^{-81}$ |
| 43 | 282 | *lptN* | Probable ABC transporter, permease | CAC22119 | *S. griseus* | $e^{-49}$ |
| 44 | 206 | *lptP* | Hypothetical protein (membrane associated) | CAD60532 | *Streptomyces cinnamoneus* | $e^{-15}$ |
| 45 | 352 | *lptI* | Probable methyltransferase | NP_627429 | *S. coelicolor* | $e^{-29}$ |
| 46 | 312 | | Putative regulator (*merR*-like) | NP_821492 | *S. avermitilis* | $e^{-95}$ |
| 47 | 136 | | Hypothetical protein | YP_121226 | *N. farcinica* | $e^{-15}$ |
| 48 | 61 | | Transposase IS1647-like | BAC75265 | *S. avermitilis* | $e^{-19}$ |
| 49 | 285 | | Putative transposase IS6 family | NP_828721 | *S. avermitilis* | $e^{-58}$ |
| 50 | 703 | | Possible protein kinase | NP_630931 | *S. coelicolor* | $e^{-62}$ |
| 51 | 139 | | Hypothetical protein | NP_630932 | *S. coelicolor* | $e^{-10}$ |
| 52 | * | | (Putative transposase) | (NP_631829) | (*S. coelicolor*) | ($e^{-66}$) |
| 53 | * | | (Putative DNA helicase) | (NP_628097) | (*S. coelicolor*) | ($e^{-87}$) |
| 54 | 342 | | Hypothetical protein | NP_822305 | *S. avermitilis* | $e^{-139}$ |
| 55 | 221 | | Possible undecaprenyl diphosphate synthase | NP_854770 | *Mycobacterium bovis* | $e^{-40}$ |
| 56 | 132 | | Putative Tra3-protein | NP_828781 | *S. avermitilis SAP1* | $e^{-21}$ |
| 57 | 291 | | Putative transposase | CAG15044 | *A. teichomyceticus* | $e^{-111}$ |
| 58 | 104 | | Possible carboxy peptidase | CAB37344 | *S. fradiae* | $e^{-19}$ |
| 59 | 156 | | Putative transposase IS1647-like | NP_821220 | *S. avermitilis* | $e^{-64}$ |
| 60 | 124 | | Putative transposase IS5 family | NP_821219 | *S. avermitilis* | $e^{-55}$ |
| 61 | 315 | | Hypothetical protein | NP_336863 | *Mycobacterium tuberculosis* | $e^{-76}$ |

| Subunit | Modules and domains |
|---|---|
| LptA | $C^{III}A_{Trp}T \bullet CA_{Glu}\mathbf{TE} \bullet C^{II}A_{hAsn}T \bullet CA_{Thr}T \bullet CA_{Gly}\mathbf{MT}$ |
| LptB | $C^{II}{}_{Ala}T \bullet CA_{Asp}T$ |
| LptC | $CA_{Lys}\mathbf{TE} \bullet C^{II}A_{OmAsp}T \bullet CA_{Gly}T \bullet CA_{Asn}\mathbf{TE}$ |
| LptD | $C^{II}A_{3mGlu/Glu}T \bullet CA_{Ile/Val}T\ \mathbf{Te}$ |

| Mod. | Residue | A-domain | Pocket | C-domain | T domain | $E^1$ or $M^2$-domain |
|---|---|---|---|---|---|---|
| LptA1 | $Trp_1$ | VHTSGSTGRPKG | DVALVGVVQ | HHLALDGY | LGGDS | |
| LptA2 | $Glu_2$ | ILTSGSTGRPKA | DLVKVASVN | HHVAADGA | LGGDS | FGFNYLG[1] |
| LptA3 | $hAsn_3$ | LYTSGSTGRPKG | DLTKVGDVN | HHLVLDGW($C^{II}$) | LGGHS | |
| LptA4 | $Thr_4$ | IYTSGSTGRPKG | DFWSVGMVH | HHIAADGW | LGGHS | |
| LptA5 | $Sar_5$ | LHTSGSTGRPKG | DILQLGVIW | HHIAGDGW | LGGDS | VLEIGVGTG[2] |
| LptB1 | $Ala_6$ | LHTSGSSGRPKG | DVFNLALVF | HHLVLDGW($C^{II}$) | LGGHS | |
| LptB2 | $Asp_7$ | LYTSGSTGRPKG | DLTKVGAVN | HHIAADGW | LGGHS | |
| LptC1 | $Lys_8$ | IHTSGSTGRPKG | DAWDAGTVD | HHIASDGW | RGGDS | FGFNYLG[1] |
| LptC2 | $OmAsp_9$ | LYTSGSTGRPKG | DLTKIGAVN | HHTIVDGW($C^{II}$) | LGGHS | |
| LptC3 | $Gly_{10}$ | IYTSGSTGRPKG | DILQLGLVW | HHIATDGW | LGGHS | |
| LptC4 | $Asn_{11}$ | VYTSGSTGTPKG | DLTKVGDVS | HHIATDAW | LGGDS | YGFNYLG[1] |
| LptD1 | $3mGlu/Glu_{12}$ | LYTSGSTGRPKG | DLGKTGVVN | HHILLDGW($C^{II}$) | LGGHS | |
| LptD2 | $Ile/Val_{13}$ | VYTSGSTGRPKG | DGLFVGIAV | HHIASDGW | NGGHS | |

**Fig. 4** Organization of the A54145 NRPS subunits. *Top*: each module (separated by *large dots*) has a set of core catalytic domains: condensation ($C^I$, $C^{II}$ or $C^{III}$), adenylation ($A_X$, where the subscript indicates the amino acid in A54145) and thiolation (*T*); additional domains, epimerization (*E*) or methylation (*M*), are associated with modules 2, 5, 8, 11. A terminal thioesterase (*Te*) is responsible for cyclization and release of the nascent peptide from the NRPS. *Bottom*: representatives of the sets of motifs [25] indicative of adenylation, condensation, thiolation, epimerization or methylation domains (A3, C3, T and E7 or M1 motifs, respectively). Amino acids are represented by their one letter code. The "pocket" of specificity conferring residues in the A-domains corresponds to the set of residues at positions 235, 236, 239, 278, 299, 300, 301, 322 and 330 in the *GrsA* PheA protein, as deduced by sequence alignment [33]. C3 motifs from $C^{II}$ type domains are indicated parenthetically. The Asp in the T-domain LGGD motif is typical of a T-domain in a module incorporating a D-amino acid; here, it is also found in the T-domain of the first module

A-domain reference sequences for acidic amino acids from streptomycetes have not been well represented until recently, but the *lpt, dpt* and *cda* NRPS pathways provide a group of 16 domains recognizing Glu, Asp, Asn and their analogs. This set clustered distinctly from domains recognizing other amino acids, and formed Asp, Asn and Glu/3mGlu subgroups that segregated further by source pathway or organism. The binding pocket residues were highly conserved, with the signature $K_{278}$ preserved in all sequences (Table 2), but $T_{239}$ clearly distinguished Asp and Asn activating domains from Glu and Gln domains, and notably, $A_{322}$ (equivalent of the conserved $H_{322}$ residue in *Bacillus* NRPS) of Asp-binding domains was replaced by an acidic residue in Asn-binding domains. The latter is consistent with studies in *Bacillus* where introduction of an acidic residue converts the binding specificity from Asp to Asn for a module of surfactin synthetase [33]. The clustering of hAsn domains with Asp-binding domains is surprising in this regard, but may afford a clue for consideration of the evolutionary derivation of acidic and polar substrate-binding domains and possibly on the mechanism of biochemical modification. The $Glu_2$ A-domain groups with, yet is distinct from, the $3mGlu_{12}$ A-domains: a pair of reciprocal changes in the binding pockets, $V_{299}$ to $T_{299}$ and $S_{322}$ to $V_{322}$, might be compensatory.

The A-domain of the $Lys_8$-incorporating module is novel among actinomycete NRPSs, and resembles a *dpt* pathway module for incorporating a similar amino acid, ornithine (Orn). Smaller side-chains among some of the specificity-determining residues ($A_{236}$, $A_{299}$ and $T_{322}$) for Lys relative to those ($T_{236}$, $M_{299}$ and $Y_{322}$) for Orn may facilitate accommodation of the bulkier substrate in A54145. $D_{278}$ and $T_{322}$ (or $Y_{322}$ for Orn) appear as conservative substitutions for *Bacillus* GrsB residues $E_{279}$ and $S_{322}$ that have been suggested to be involved in salt-bridge formation and H-bonding with the amino group of Lys [9].

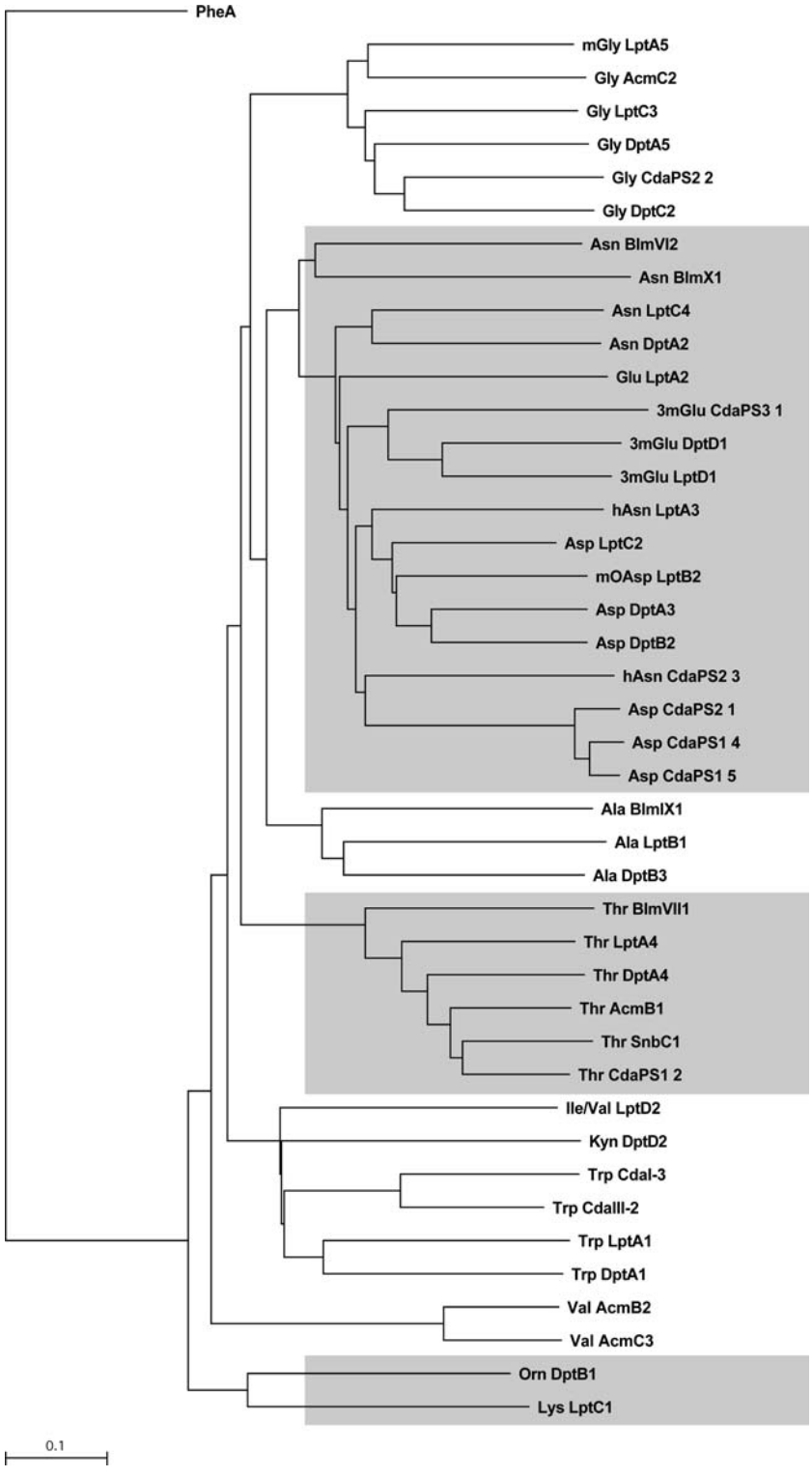


**Fig. 5** Dendrogram comparing approximately 110 amino acids inside the region defined by motifs A4 and A5 of A-domains of A54145 and other streptomycete NRPS. Source pathways are designated by protein name, subunit name or number, and module number: daptomycin (*Dpt*, *S. roseosporus*), calcium-dependent antibiotic (*Cda*, *S. coelicolor*), actinomycin (*Acm*, *Streptomyces chrysomallus*) and bleomycin (*Blm*, *Streptomyces verticillus*). PheA was used as an outgroup

Ile is usually incorporated in position 13 of A54145, but a small percent of Val-containing analogs is also found, particularly late in a fermentation [5]. Modules with relaxed specificity—particularly for hydrophobic amino acids—are not uncommon: a number have been reported in *Bacillus* spp. The Ile/$Val_{13}$ A-domain was notable in that it grouped with the Trp and Kyn activating A-domains from *cda* and *dpt* clusters. This observation raises a possibility that specificity for Ile/Val might have evolved in place, in an ancestral module or subunit related to *dptD* and *cdaPS3*, which is also adapted for chain termination.

Comparison of C-domain protein sequences has revealed the presence of three classes, $C^{I}$, $C^{II}$ and $C^{III}$ as

**Table 2** Specificity residues for Asp, Asn, Glu and analogs

| Amino acid | Source | Residue number[a] | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 235 | 236 | 239 | 278 | 299 | 301 | 322 | 330 | 331 |
| Asp | LptB2 | D | L | **T** | **K** | V | G | A | V | N |
| | DptA3 | D | L | **T** | **K** | L | G | A | V | N |
| | DptB2 | D | L | **T** | **K** | L | G | A | V | N |
| | DptC1 | D | L | **T** | **K** | L | G | A | V | N |
| | CdaPS1-4 | D | L | **T** | **K** | I | G | A | V | N |
| | CdaPS1-5 | D | L | **T** | **K** | I | G | A | V | N |
| | CdaPS2-1 | D | L | **T** | **K** | I | G | A | V | N |
| OmAsp | LptC2 | D | L | **T** | **K** | I | G | A | V | N |
| Asn | LptC4 | D | L | **T** | **K** | V | G | **D** | V | S |
| | DptA2 | D | L | **T** | **K** | L | G | **D** | V | N |
| | CepA3 | D | L | **T** | **K** | L | G | **E** | V | G |
| | BlmX1 | D | L | **T** | **K** | V | G | **E** | V | G |
| | BlmVI2 | D | L | **T** | **K** | V | G | **E** | V | G |
| HAsn | LptA3 | D | L | **T** | **K** | V | G | **D** | V | N |
| | CdaPS2-3 | D | L | **T** | **K** | V | G | **E** | V | N |
| Glu | LptA2 | D | L | V | **K** | V | A | S | V | N |
| 3mGlu | LptD1 | D | L | G | **K** | T | G | V | V | N |
| | DptD1 | D | L | G | **K** | T | G | V | I | N |
| | CdaPS3-1 | D | Q | G | **K** | T | G | V | G | H |

[a]Refers to the corresponding position in the PheA sequence [33]

previously proposed [27]. $C^{I}$-domains, those that follow L-amino acid activating modules and typically found with internal modules, are most abundant. C-domains that have specificity for a D-amino acid upstream have been demonstrated in *Bacillus* [11]. $C^{II}$-domains in the daptomycin and CDA biosynthetic pathways are distinguishable by sequence, located downstream of modules that include epimerase domains, and are often subunit starters; the four $C^{II}$-domains in Lpt proteins are consistent with this (Fig. 6). The C-domain starting LptC is notable in that it is of the $C^{I}$ type, instead of the $C^{II}$ type, and the LptB to LptC junction resembles that of an intrasubunit junction, rather than other intersubunit junctions with an extended linker region (Fig. 6). The positionally analogous junction in the *dpt* pathway (DptBC2–DptBC3), also has a $C^{I}$-domain, but in daptomycin synthetase, the module is internal. This raises the speculation that genes for two subunits might have become fused to form a larger subunit, or vice versa, that *dptBC* might have become split, if the *dpt* and *lpt* pathways shared ancestry. The remaining C-domain at the beginning of LptA may belong to $C^{III}$, a small heterogeneous group proposed to interact with an acyl rather than a peptidyl substrate, and proposed here for coupling the *N*-terminal fatty acid starter to $Trp_1$.

Domains responsible for amino acid modification were present in some modules. E-domains in modules 2, 8 and 11 of A54145 indicate that $Glu_2$, $Lys_8$ and $Asn_{11}$ should be in the D-configuration, as observed for the residues in those positions in daptomycin. An approximately 400 residue *N*-methylation domain with two conserved motifs in module 5 (Fig. 4), for which A-domain analyses indicated as likely to incorporate Gly, accounts for the presence of $Sar_5$. Previous results with partially purified A54145 enzyme that activated Gly, but not Sar, have also suggested that a methylation domain would be found on the multienzyme complex [2, 35]. No methylation domain was associated with the other methylated amino acid, $3mGlu_{12}$, suggesting that free 3mGlu is either incorporated by the penultimate module in A54145, or that methylation is accomplished by a separate enzyme after Glu is incorporated. In the first instance, specificity of the A-domain may be relaxed, and Glu or 3mGlu may be incorporated. Comparison of the protein sequences in the three-lipopeptide pathways revealed no obvious distinction among A-domains from modules incorporating only Glu (Lpt), only 3mGlu (Dpt), or both (LptD and Cda).

The thioesterase (Te) domain at the end of LptD is similar to domains at the end of the Cda and Dpt NRPSs and has the conserved motif GXSXG; this domain accepts either Ile or Val for the cyclization reaction that ultimately releases the peptide from the enzyme [7, 20].

## Accessory genes

The ORFs immediately flanking the *lpt* NRPS genes are likely to be involved in the biosynthesis of A54145 (Table 1). The *N*-terminus of the *lptEF* protein has similarity to acyl-CoA ligases, while its *C*-terminus has the PP-binding, phosphopantetheine attachment site of an acyl carrier protein (ACP). LptEF, together with the $C^{III}$-domain in LptA, is probably responsible for adding the fatty acid. The *N*-terminal domain of LptEF could form activated acyl-CoA thioesters of medium-chain fatty acids which could be then transferred by the *C*-terminal domain to the Trp-incorporating module. In the *dpt* gene cluster, analogous roles

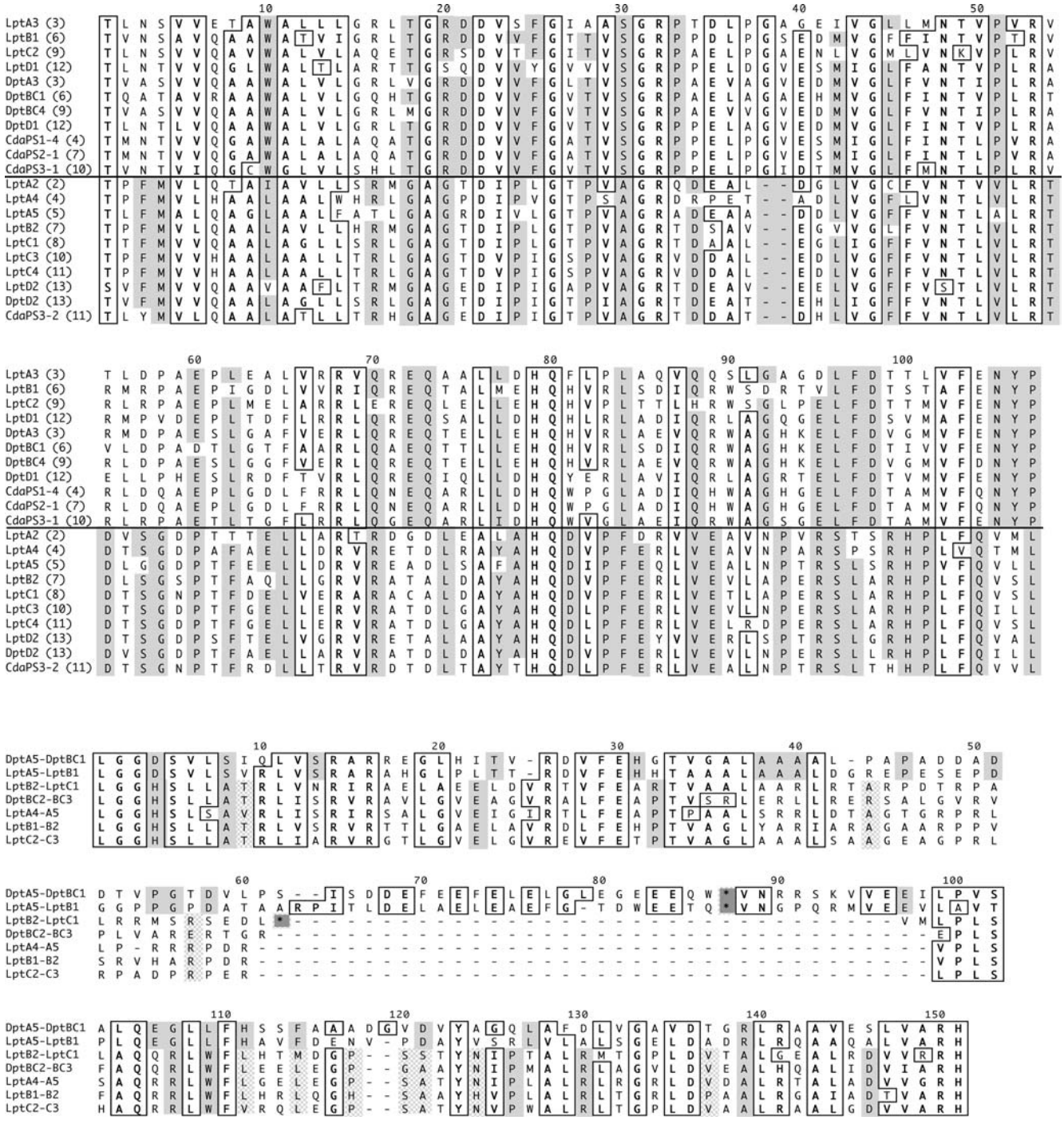

**Fig. 6** $C^{II}$ condensation domains and intersubunit junctions. *Top*: example of sequences distinguishing $C^{II}$ (top 11 lines) and $C^{I}$ domains (the rest). Residues conserved for C-domains in general are *boxed*; additional sequences conserved by domain subtypes are shaded. The subunit and the module are specified: e.g. *LptA4* designates sequences from *LptA*, module 4; *CdaPS1-2* designates sequences from *CdaPS1*, module 2. The associated amino acid position is shown parenthetically. *Bottom*: LptB2–LptC1 intersubunit junction resembles intrasubunit junctions. Intersubunit (top 2 lines) and intrasubunit (lowest 4 lines) junction regions starting in the middle of the T-domain of the upstream module, and ending at the C2 motif in the C-domain of the downstream module are shown. *Asterisks* (position 87 in lines 1 and 2, position 62 in line 3) mark the ends of subunits

are attributed to proteins encoded by *dptE* and *dptF*. Straight and branched $C_{10\text{-}11}$ fatty acids are typically found in A54145, but feeding experiments have shown that the enzymatic step involved in coupling the fatty acid is indiscriminant with respect to substrates with chain lengths of $C_6$ to $C_{11}$ [6].

Orthologs of *lptG*, encoding a small esterase-like conserved protein, and *lptH*, encoding a Te, occur in

many polyketide and NRPS biosynthetic gene clusters, including also *dpt* and *cda*. There has been no experimental evidence for the function of *lptG*-like genes in any pathway, but it is conceivable that they may have a role in secondary metabolism, given their genomic context [23]. For example, one of two possible start codons for *lptG* overlaps the end of *lptD* in *S. fradiae*, and a possible ribosome binding site associated with the second start codon lies within the end of *lptD*. In *Streptomyces tendae* Tü901, an LptG-like domain resides within the NRPS encoded by *NikP1*. Bioinformatic analyses of LptG-like domains (MtbH-like proteins) in *S. coelicolor* have also led to a suggestion that these genes function in export or regulation [36]. Enzymes such as those encoded by *lptH* can unblock NRPSs that have been misprimed with an acylated (instead of a free) 4′ phosphopantethiene prosthetic group and therefore cannot sustain chain elongation [30]; a hydrolase could remove the acyl group to regenerate the active holoenzyme and re-enable biosynthetic activity. These enzymes have been shown in polyketide biosynthetic pathways to be important in maintaining a high level of antibiotic production [8].

## Genes encoding tailoring enzymes

The *lptK*, *lptL* and *lptI* genes most likely encode modification enzymes for $OmAsp_9$, $hAsn_3$ and $3mGlu_{12}$. As with the $Glu/3mGlu_{12}$ incorporating module, the A-domains and the amino acid binding pockets of the modules for Asp and Asn are not noticeably different between those associated with natural and modified residues. This leaves open the possibility that modification occurs on an amino acid tethered to a carrier domain in the multimodular enzyme, a model similar to that of a free-standing module with A and T domains that participate in the building of a precursor that is then used by the incorporating module in the biosynthetic pathway [10]. Alternatively, these amino acids may be modified after the peptide is assembled. The predicted product of *lptK*, one of two methyltransferases, could be responsible for elaborating OmAsp by methylation of a hydroxylated intermediate. The rationale is twofold: LptK most resembles an *O*-methyltransferase encoded by *fkbM* in the FK506 PKS pathway [28], and the second methyltransferase gene could be reasonably assigned to the methylation of Glu (below). About 60 nucleotides downstream, the hydrolase predicted from *lptL* might participate in biosynthesis of OmAsp; however, it is very similar to that of an oxygenase (AnsO) encoded by *S. coelicolor asnO* (SCO3236), which has been proposed to be involved in converting Asn to hAsn during CDA biosynthesis [16]. Thus, an alternate role for LptL may be the formation of hAsn in A54145. The *lptI* gene has orthologs in the *cda* and *dpt* gene clusters (*glmT* and *dptI*, respectively) and sequence analyses suggest that they all encode methyltransferases. As both daptomycin and CDA also contain 3mGlu (but not OmAsp), one can deduce that *lptI* encodes the methyltransferase for 3mGlu. LptI is the only ORF in the entire *lpt* region that uses a TTA codon. This occurs at the first Leu, residue 17, in the beginning of the deduced protein. TTA is rarely used in streptomycetes, and is observed only in genes expressed late in development, including some secondary metabolic genes [19, 24]. The accumulation of 3mGlu-containing factors only late in production [5] could possibly be attributed to delayed onset of translation of mRNA from this gene.

## Regulatory, precursor supply and resistance genes

Several candidates for regulatory genes are in or near the *lpt* gene cluster. The predicted *lptJ* gene product is similar to that of *syrP*, a proposed regulator of syringomycin synthesis in *Pseudomonas syringae*. Inactivation of *syrP* abolishes production of syringomycin in broth, but causes overproduction on agar, and is accompanied by reduced sensitivity to phosphate [37]. The *syrP* gene does not fall into any known class of regulatory genes, so the interpretation of *lptJ* function must be left open at this time. Upstream of the NRPS genes are two pairs of ORFs, ORF28 and ORF29, and ORF17 and ORF18, each of which appears to encode prokaryotic 2-component regulatory systems comprising a response regulator and a sensor kinase. Similar proteins encoded by *absA1*/*absA2* in the *cda* gene cluster negatively regulate biosynthesis of several antibiotics in *S. coelicolor* [29]. ORF21 and ORF22 are two other genes of potential interest: ORF22 may encode a DNA-binding transcriptional regulator, while ORF21, encoding diaminopimelic acid decarboxylase, an enzyme in lysine biosynthesis, might be involved in enhancement of available lysine levels, a role analogous to that for the Trp biosynthetic genes near the *cda* cluster [16]. Downstream, ORF46, encoding a MerR-like repressor protein, is the only candidate regulatory gene.

Resistance genes are often found associated with antibiotic biosynthetic pathways [26, 31], and putative ABC transporters encoded by ORF30 and ORF31, or *lptM* and *lptN*, may have a role in self-resistance and/or transport to A54145. The *lptM* and *lptN* genes are next to *lptP*, the last a conserved protein of unknown function notable for several deduced transmembrane domains and a strong basic charge at the *C*-terminus. An ortholog of *lptP* is present in the *dpt* pathway (*dptP*), and in the same physical arrangement relative to two ABC transporters (*dpt* ORF27 and *dpt* ORF28) in *S. roseosporus*.

## Distal flanking regions

The upstream cosmid pCV16 contains a relatively high density of ORFs and the border of the *lpt* cluster is

unclear. It probably extends beyond *lptEF*, to ORF30 and ORF31, but possibly up to ORF17 and ORF18. ORFs further upstream appear to encode housekeeping genes (e.g. β-glucoside and sugar isomerase). Downstream, *lptI*, or possibly ORF46, would appear to mark the right border of the A54145 biosynthetic gene cluster. The sequence beyond this had a deficiency of clearly identifiable ORFs. Potential coding regions were dispersed through only about half of the 20 kb, even though the overall 67% GC content seems reasonably high (compared to the *lpt* genes, at 71%, and to the NRPS genes in particular, at 76%). Except for a 2.1 kb ORF likely to encode a Ser/Thr protein kinase homolog (ORF50), the rest are small and one-third appear to be insertion sequence (IS) element-related or transposase derivatives (Fig. 3, Table 2). ORF55, a 400 bp region with 80% GC content, contains similarity to TRA3, a gene found in some streptomycete plasmids that has been implicated in DNA transfer [32, 38]. The localization of A54145 biosynthetic genes in a region containing IS and transposase sequences is consistent with its mapping by pulsed field gel electrophoresis to one end of the linear chromosome [4], as such sequences are often found near the ends of the linear chromosomes of *Streptomyces avermitilis* and *S. coelicolor* also [17].

## Discussion

The cyclic lipodepsipeptides are emerging as an important new class of antibiotics. Daptomycin, under the name Cubicin® (daptomycin-for-injection), was approved recently for clinical treatment of skin and skin structure infections caused by Gram-positive pathogens. Previous work has demonstrated that certain A54145 derivatives are as potent as daptomycin in vitro, but are less efficacious and more toxic in mice [2]. Both the peptide composition and the fatty acid side-chain affect antibacterial activity and toxicity of lipopeptides. Appropriate combinations of amino acid modifications and fatty acid side-chains might lead to improvements in the therapeutic index or pharmacological properties of lipopeptides.

Consideration of A54145, along with CDA and daptomycin, as a family of naturally occurring compounds with $Ca^{2+}$-dependent biological activity leads to the development of a framework to explore structure-activity relationships within this class of antibiotics. Comparison of A54145, CDA and daptomycin would allow discrimination, at the first level, between features that are conserved and therefore likely to be intrinsic to the basic mode of action, and those that are mutable and that might be manipulated to augment potency or improve pharmacological properties while still retaining the primary mode of action. The former may include the ring size, the multiple acidic positions represented by Asp and analogs [34], and the conservation of achiral or D-amino acids at particular positions in the ring or the tail [27]. The amino acids that are not conserved would then represent locations to begin investigations for tailoring desirable changes.

The integration of the sequence information and clones of the A54145 biosynthetic gene cluster with those of the *dpt* and *cda* NRPS gene clusters provide opportunities for fundamental genetic and biochemical studies [35]. *S. fradiae, S. roseosporus* and *S. coelicolor* are all hosts that are amenable to molecular genetic manipulation [3, 16] and hypotheses generated by comparative analyses in silico can be tested in vivo using components of all three systems. These range from studies of intersubunit communications for similar and dissimilar subunits (DptD, CdaPS3 and LptD, or DptA and LptA versus CdaPS1) to dissection of position versus specificity determinants in the NRPS, for instance, using the set of Trp-incorporating modules represented by the three pathways. Heterologous expression of orthologous genes for tailoring enzymes may help refine our understanding of their unique features and teach us how new substrate specificities may be gained. Adding the A54145 system to the extant knowledge of the cellular machinery behind the CDA and daptomycin biosynthetic pathways should greatly accelerate design and engineering of useful new compounds from natural products (Baltz et al. and Miao et al., submitted for publication).

**Acknowledgements** The authors would like to thank J. McMenanim and L. Thurston for technical assistance, and S.K. Wrigley for reviewing and discussing the manuscript.

## References

1. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410
2. Baltz RH (1997) Lipopeptide antibiotics produced by *Streptomyces roseosporus* and *Streptomyces fradiae*. In: Strohl WR (ed) Biotechnology of antibiotics. Marcel Dekker, New York, pp 415–435
3. Baltz RH (1998) Genetic manipulation of antibiotic-producing *Streptomyces*. Trends Microbiol 6:76–83
4. Baltz RH, McHenney MA, Hosted TJ (1997) Genetics of lipopeptide antibiotic biosynthesis in *Streptomyces fradiae* A54145 and *Streptomyces roseosporus* A21978C. In: Baltz RH, Hegeman GD, Skatrud PL (eds) Developments in industrial microbiology. Society for Industrial Microbiology, Fairfax, VA, pp 93–98
5. Boeck LD, Papiska HR, Wetzel RW, Mynderse JS, Fukuda DS, Mertz FP, Berry DM (1990) A54145, a new lipopeptide antibiotic complex: discovery, taxonomy, fermentation and HPLC. J Antibiot (Tokyo) 43:587–593
6. Boeck LD, Wetzel RW (1990) A54145, a new lipopeptide antibiotic complex: factor control through precursor directed biosynthesis. J Antibiot (Tokyo) 43:607–615
7. Bruner SD, Weber T, Kohli RM, Schwarzer D, Marahiel MA, Walsh CT, Stubbs MT (2002) Structural basis for the cyclization of the lipopeptide antibiotic surfactin by the thioesterase domain SrfTE. Structure (Camb) 10:301–310
8. Butler AR, Bate N, Cundliffe E (1999) Impact of thioesterase activity on tylosin biosynthesis in *Streptomyces fradiae*. Chem Biol 6:287–292
9. Challis GL, Ravel J, Townsend CA (2000) Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem Biol 7:211–224

10. Chen H, Hubbard BK, O'Connor SE, Walsh CT (2002) Formation of beta-hydroxy histidine in the biosynthesis of nikkomycin antibiotics. Chem Biol 9:103–112
11. Clugston SL, Sieber SA, Marahiel MA, Walsh CT (2003) Chirality of peptide bond-forming condensation domains in nonribosomal peptide synthetases: the C5 domain of tyrocidine synthetase is a (D)C(L) catalyst. Biochemistry 42:12095–12104
12. Counter FT, Allen NE, Fukuda DS, Hobbs JN, Ott J, Ensminger PW, Mynderse JS, Preston DA, Wu CY (1990) A54145, a new lipopeptide antibiotic complex: microbiological evaluation. J Antibiot 43:616–622
13. Debono M, Barnhart M, Carrell CB, Hoffmann JA, Occolowitz JL, Abbott BJ, Fukuda DS, Hamill RL, Biemann K, Herlihy WC (1987) A21978C, a complex of new acidic peptide antibiotics: isolation, chemistry, and mass spectral structure elucidation. J Antibiot 40:761–777
14. Fukuda DS, Debono M, Molloy RM, Mynderse JS (1990) A54145, a new lipopeptide antibiotic complex: microbial and chemical modification. J Antibiot 43:601–606
15. Fukuda DS, Du Bus RH, Baker PJ, Berry DM, Mynderse JS (1990) A54145, a new lipopeptide antibiotic complex: isolation and characterization. J Antibiot 43:594–600
16. Hojati Z, Milne C, Harvey B, Gordon L, Borg M, Flett F, Wilkinson B, Sidebottom PJ, Rudd BA, Hayes MA, Smith CP, Micklefield J (2002) Structure, biosynthetic origin, and engineered biosynthesis of calcium-dependent antibiotics from *Streptomyces coelicolor*. Chem Biol 9:1175–1187
17. Ikeda H, Ishikawa J, Hanamoto A, Shinose M, Kikuchi H, Shiba T, Sakaki Y, Hattori M, Õmura S (2003) Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. Nat Biotechnol 21:526–531
18. Kempter C, Kaiser D, Haag S, Nicholson G, Gnau V, Walk T, Gierling KH, Decker H, Zähner H, Jung G, Metzger JW (1997) CDA: calcium-dependent peptide antibiotics from *Streptomyces coelicolor* A3(2) containing unusual residues. Angew Chem Int Ed Engl 36:498–501
19. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) Practical *Streptomyces* genetics. John Innes Foundation, Norwich
20. Kohli RM, Walsh CT (2003) Enzymology of acyl chain macrocyclization in natural product biosynthesis. Chem Commun (Camb) 3:297–307
21. Kuhstoss S, Rao RN (1991) Analysis of the integration function of the streptomycete bacteriophage phi C31. J Mol Biol 222:897–908
22. Lakey JH, Lea EJ, Rudd BA, Wright HM, Hopwood DA (1983) A new channel-forming antibiotic from *Streptomyces coelicolor A3(2)* which requires calcium for its activity. J Gen Microbiol 129:3565–3573
23. Lauer B, Russwurm R, Schwarz W, Kalmanczhelyi A, Bruntner C, Rosemeier A, Bormann C (2001) Molecular characterization of co-transcribed genes from *Streptomyces tendae* Tu901 involved in the biosynthesis of the peptidyl moiety and assembly of the peptidyl nucleoside antibiotic nikkomycin. Mol Gen Genet 264:662–673
24. Leskiw BK, Bibb MJ, Chater KF (1991) The use of a rare codon specifically during development? Mol Microbiol 5:2861–2867
25. Marahiel MA, Stachelhaus T, Mootz HD (1997) Modular peptide synthetases involved in nonribosomal peptide synthesis. Chem Rev 97:2651–2674
26. Mendez C, Salas JA (1998) ABC transporters in antibiotic-producing actinomycetes. FEMS Microbiol Lett 158:1–8
27. Miao V, Coeffet-Legal MF, Brian P, Brost R, Penn J, Whiting A, Martin S, Ford R, Parr IB, Bouchard M, Silva CJ, Wrigley SK, Baltz RH (2005) Daptomycin biosynthesis in *Streptomyces roseosporus*: cloning and analysis of the gene cluster and revision of peptide stereochemistry. Microbiology 151(Pt 5):1507–1523
28. Motamedi H, Shafiee A, Cai SJ, Streicher SL, Arison BH, Miller RR (1996) Characterization of methyltransferase and hydroxylase genes involved in the biosynthesis of the immunosuppressants FK506 and FK520. J Bacteriol 178:5243–5248
29. Ryding NJ, Anderson TB, Champness WC (2002) Regulation of the *Streptomyces coelicolor* calcium-dependent antibiotic by *absA*, encoding a cluster-linked two-component system. J Bacteriol 184:794–805
30. Schwarzer D, Finking R, Marahiel MA (2003) Nonribosomal peptides: from genes to products. Nat Prod Rep 20:275–287
31. Seno E, Baltz RH (1989) Structural organization and regulation of antibiotic biosynthesis and resistance genes in actinomycetes. In: Shipiro S (ed) Regulation of secondary metabolism in actinomycetes. CRC Press, Boca Raton, pp 1–48
32. Spatz K, Kohn H, Redenbach M (2002) Characterization of the *Streptomyces violaceoruber* SANK95570 plasmids pSV1 and pSV2. FEMS Microbiol Lett 213:87–92
33. Stachelhaus T, Mootz HD, Marahiel MA (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. Chem Biol 6:493–505
34. Uguru GC, Milne C, Borg M, Flett F, Smith CP, Micklefield J (2004) Active-site modifications of adenylation domains lead to hydrolysis of upstream nonribosomal peptidyl thioester intermediates. J Am Chem Soc 126:5032–5033
35. Wessels P, von Döhren H, Kleinkauf H (1996) Biosynthesis of acylpeptidolactones of the daptomycin type. A comparative analysis of peptide synthetases forming A21978C and A54145. Eur J Biochem 242:665–673
36. Yeats C, Bentley S, Bateman A (2003) New knowledge from old: in silico discovery of novel protein domains in *Streptomyces coelicolor*. BMC Microbiol 3:3–20
37. Zhang JH, Quigley NB, Gross DC (1997) Analysis of the *syrP* gene, which regulates syringomycin synthesis by *Pseudomonas syringae pv. syringae*. Appl Environ Microbiol 63:2771–2778
38. Zotchev SB, Schrempf H (1994) The linear *Streptomyces* plasmid pBL1: analyses of transfer functions. Mol Gen Genet 242:374–382