Non-ribosomal peptide synthetase module fusions to produce derivatives of daptomycin in *Streptomyces roseosporus*

Sascha Doekel, Marie-Françoise Coëffet-Le Gal, Jian-Qiao Gu, Min Chu, Richard H. Baltz and Paul Brian

Cubist Pharmaceuticals Inc., 65 Hayden Avenue, Lexington, MA 02421, USA

Genetic engineering has been applied to reprogramme non-ribosomal peptide synthetases (NRPSs) to produce novel antibiotics, but little is known about what determines the efficiency of production. We explored module exchanges at nucleotide sequences encoding interpeptide linkers in *dptD*, a gene encoding a di-modular NRPS subunit that incorporates 3-methylglutamic acid (3mGlu$_{12}$) and kynurenine (Kyn$_{13}$) into daptomycin. Mutations causing amino acid substitutions, deletions or insertions in the inter-module linker had no negative effects on lipopeptide yields. Hybrid DptD subunits were generated by fusing the 3mGlu$_{12}$ module to terminal modules from calcium-dependent antibiotic (CDA) or A54145 NRPSs, and recombinants produced daptomycin analogues with Trp$_{13}$ or Ile$_{13}$ at high efficiencies. A recombinant expressing DptD with a hybrid Kyn$_{13}$ module containing a di-domain from a D-hAsn module caused the production of a new daptomycin analogue containing Asn$_{13}$.

**INTRODUCTION**

*Streptomyces roseosporus* NRRL 15998 produces the A21978C complex of acidic lipopeptides of the daptomycin family, which consist of at least six factors with different fatty acids attached to the amino group of L-Trp$_1$ of the peptide (Baltz et al., 2005; Debono et al., 1987). (The subscripts refer to positions of amino acids in the peptide, starting with L-Trp$_1$, as discussed below and in Fig. 1b.) The three major fermentation components of the A21978C complex have 11-, 12- and 13-carbon branched-chain fatty acids, whereas the minor factor daptomycin, which can be enriched by feeding decanoic acid during fermentation (Baltz et al., 2005; Huber et al., 1988), has a 10-carbon straight chain (Fig. 1a). Daptomycin has been approved as Cubicin® (daptomycin-for-injection) for treatment of skin and skin structure infections caused by Gram-positive pathogens (Arbeit et al., 2004), and for bacteremia and endocarditis caused by *Staphylococcus aureus*, including strains resistant to meticillin (MRSA) (Fowler et al., 2006).

The daptomycin biosynthetic gene cluster was cloned and sequenced, and the non-ribosomal peptide synthetase (NRPS) and other genes identified (Miao et al., 2005; Fig. 1b). The cyclic, branched, 13-membered lipopeptide backbone of daptomycin is produced by three giant NRPS multi-enzymes, DptA, DptBC and DptD, encoded by the *dptA*, *dptBC* and *dptD* genes, respectively. NRPSs are composed of conserved repeating units (modules) each with an average size of 120 kDa (Sieber & Marahiel, 2005). Each module catalyses the incorporation of one specific amino acid in a sequence that is collinear with the modular organization in the DNA coding sequences. A typical module has three catalytic domains, responsible for peptide bond formation (condensation domain, C; Stachelhaus et al., 1998), recognition and activation of substrate amino acids (adenylation domain, A; Stachelhaus et al., 1999), and translocation (thiolation domain, T; or peptidyl carrier protein, PCP; Stachelhaus et al., 1996). Modules may also contain accessory domains, such as epimerization (E) domains that catalyse the conversion of L-amino acids to D-isomers (Stachelhaus & Walsh, 2000). The terminal modules of NRPSs often have thioesterase (Te) domains involved in cleavage of the nascent peptide, releasing linear or cyclic peptides, or cyclic depsipeptides (Kohli & Walsh, 2003). Such modules are arranged as C-A-T-Te, and individual enzyme domains are connected by linker peptides.

Because of their collinear gene/protein module organization, NRPSs present promising targets for combinatorial
biosynthesis (Baltz, 2008; Baltz et al., 2005; Fischbach & Walsh, 2006; Sieber & Marahiel, 2005). Module exchanges in NRPS model systems have proven to be useful to generate peptides of a targeted sequence (Doekel & Marahiel, 2000; Mootz et al., 2000), and several novel derivatives of surfactin have been produced by the targeted replacement of modules (Mootz et al., 2002; Stachelhaus et al., 1995; Yakimov et al., 2000). In other in vivo studies, the specificity-conferring codes in the active sites of A domains (Stachelhaus et al., 1999; Challis et al., 2000) have been exploited to redirect surfactin biosynthesis by site-specific mutagenesis (Eppelmann et al., 2002). Biochemical investigations on NRPSs have helped define rules of substrate specificity and promiscuity, and to define inter-domain linkers and inter-peptide docking domains, thereby providing tools to guide the engineering of NRPS biosynthetic genes to produce novel peptides in vivo (Fischbach & Walsh, 2006; Sieber & Marahiel, 2005).

The enzymic domains of NRPSs are connected by peptide linkers that have no apparent sequence conservation. The linkers have been used in vitro (Doekel & Marahiel, 2000; Mootz et al., 2000) and in vivo (Stachelhaus et al., 1995; Nguyen et al., 2006b) to engineer modular NRPSs by module and domain exchanges. Production levels of novel lipopeptides related to daptomycin, engineered by single module exchanges, ranged from 17 to 50% of controls (Nguyen et al., 2006b).

The functions of NRPS inter-domain linkers have not been extensively explored, although there has been speculation about a role in initiation of protein–protein interactions (Shen et al., 2002). The 18 aa linker connecting the T (or PCP) and downstream C domains in a T-C di-domain of the tyrocidine synthetase TycC apparently exhibits considerable conformational flexibility, and the final seven amino acids lack defined interactions with either domain and are mobile in the PCP-C crystal structure (Samel et al., 2007). The flexibility in this linker reflects the requirement for the T (PCP) to interact with the A domain and the preceding and following C domains to function in peptide assembly.

![Fig. 1. Structures of lipopeptide antibiotics and organization of NRPS genes. (a) Chemical structures of daptomycin, A54145 (factor B1) and CDA. (b) Organization of the NRPS gene clusters of daptomycin, A54145 and CDA. The dptA, dptBC and dptD genes encode multi-modular NRPS enzymes, DptA, DptBC and DptD, that sequentially incorporate L-Trp1-D-Asn2-L-Asp3-L-Thr4-Gly5, L-Orn6-L-Asp7-L-Asp8-D-Ala9-L-Asp10-Gly11-L-Kyn12-L-Thr13, respectively. The lptA, lptB, lptC and lptD genes encode LptA, LptB, LptC and DptD that incorporate L-Trp1-D-Glu2-L-Asn3-L-Asp4-Sar5, L-Ala6-L-Asp7, D-Lys8-L-mOAsp9-L-Gly10-D-Asn11, and L-3mGlu12-L-Thr13, respectively. The cdaPS1, cdaPS2 and cdaPS3 genes encode CdaPS1, CdaPS2 and CdaPS3 that incorporate L-Ser1-L-Thr2-D-Trp3-L-Asp4-D-hPhg5, L-Asp6-Gly7-D-hAsn8, and L-3mGlu10-L-Trp11, respectively. (See Baltz, 2008; Baltz et al., 2005.)](http://mic.sgmjournals.org)
Daptomycin production can be restored in a dptD deletion mutant of *S. roseosporus* (ΔdptD) with dptD expressed under the transcriptional control of the ermEp* promoter by trans-complementation at an ectopic site in the chromosome (Coéffet-Le Gal et al., 2006; Miao et al., 2006b). This ΔdptD mutation can also be complemented by the genes lptD and cdaPS3, encoding the terminal subunits from the distinctly related lipopeptide biosynthetic pathways of *A54145* and calcium-dependent antibiotic (CDA), respectively, to produce novel hybrid lipopeptides (Baltz, 2008; Coéffet-Le Gal et al., 2006; Miao et al., 2006b; Nguyen et al., 2006b). The analogues contained Trp or Ile/Val in place of Kyn1, when subunits CdaPS3 (Hojati et al., 2002) or LptD (Miao et al., 2006a) were exchanged for the native DptD (Fig. 1a). This work indicated that the Te domains from the CdaPS3 and LptD subunits were able to cyclize heterologous peptides differing substantially from the native CDA and *A54145* peptides, consistent with *in vitro* chemoenzymatic studies with excised T-Te di-domains that generated lipopeptides related to daptomycin (Grünwald et al., 2004; Kopp et al., 2006). In the *in vivo* studies, the recombinants produced daptomycin derivatives containing Trp12 or Ile12/Val13 in yields of about 50% and 25% of control, respectively (Coéffet-Le Gal et al., 2006; Miao et al., 2006b). The reduction in yields might be due to suboptimal interpeptide docking, substrate recognition or ring closure.

This heterologous whole-subunit trans-complementation system for combinatorial biosynthesis of daptomycin analogues provides a means to test parameters important for high productivity of engineered NRPSs. In particular, it can be used to explore the efficiency of producing hybrid molecules by fusing modules or exchanging domains from di-modular DptD, LptD and CdaPS3 subunits to engineer positions 12 and 13. This also has a practical advantage in that positions 12 and 13 are important determinants for antibacterial activity (Coéffet-Le Gal et al., 2006; Miao et al., 2006b; Nguyen et al., 2006a).

In the present study, we generated module fusions at T-C linkers, and showed that the linker region can be modified by amino acid substitutions, and by deletion or insertion of four amino acids, without disrupting lipopeptide production. We also generated a functional domain exchange by fusing a heterologous C-A di-domain to the T-Te di-domain of DptD at the A-T linker, and the recombinant produced a novel derivative of daptomycin with Asn13.

**METHODS**

**Strains, growth conditions and conjugal transfer of plasmids.** Strains and plasmids are listed in Table 1. *S. roseosporus* was grown on trypticase soy (TS) agar or in TS broth at 30 °C. As-1 agar and fermentation medium F10A were described previously (Miao et al., 2006b). Plasmids were transferred from *Escherichia coli* ML22 into *S. roseosporus* UA376 (ΔdptD) by conjugation (Miao et al., 2006b). Exconjugants were selected for hygromycin resistance at 50 μg ml−1 on As-1 agar.

**Antibiotic production, purification and assay.** *S. roseosporus* strains containing recombinant plasmids were fermented in triplicate for 5 days as described previously (Miao et al., 2006b) and supernatants of fermentation broths were tested for antibacterial activity. Samples (100 μl) were added to wells (5 mm diameter) embedded in AS-1 agar plates supplemented with 10 mM CaCl2 and *Staphylococcus aureus* 42 (10⁵ cf.u. ml⁻¹). After incubation at 37 °C overnight, the plates were inspected for zones of inhibition, and MICs were determined (Miao et al., 2006b). Lipopeptides were quantified from broth cultures by HPLC and LC-MS using daptomycin as the standard, and were purified as described by Nguyen et al. (2006a).

**Plasmid constructions.** Expression plasmids for engineered dptD genes were constructed based on pRB04, a pHM11a-derived plasmid in which the dptD gene is expressed from the strong, constitutive *ermEp* promoter (Miao et al., 2006b). The coding sequence of dptD was cloned by inserting a *BsrGI/BamHI* (blunted) fragment (primers CB312 and CB313). Primer CB313 (see Supplementary Table S1, available with the online version of this paper) contained artificial AvrII and Pmel sites that were positioned in the A-site of the T-C linker of dptD as shown in Fig. 2(b), giving rise to plasmid pCB301. This cloning step was repeated using primer pairs CB312/CB314 and CB312/CB315, respectively, to shift the AvrII site downstream to the B and C sites of the T-C linker (plasmids pCB302 and pCB303, respectively). Plasmid pCB302 was used to ligate in-frame an AvrII/HpaI fragment comprising the terminal modules (CATTe) for dptD2 (Kyn), LptD2 (Ile) and cdaPS3-2 (Trp) (primer pairs CB323/CB325, CB327/CB329 and CB330/CB331) generating plasmids pCB304, pCB305 and pCB306, respectively. Plasmids pCB301 and pCB303 were used for cloning a dptD2 fragment with varying 5′ terminal ends to accommodate A and C sites of the T-C linker, generating plasmids pCB307 and pCB308 (using primer pairs CB322/CB325 for the A site and CB324/CB325 for the C site, respectively). Plasmids pCB309 and pCB310 were constructed as above, but by switching the inserts to generate a deletion or insertion of 12 nt in the linker coding sequences (Table 1). C-A di-domains for dptD2 and lptC4 were amplified using primer pairs CB323/CB354 and CB355/CB356, respectively, and cloned into pCB302. CAT modules and C-A di-domains were generated using λ Red-mediated recombination (Datsenko & Wanner, 2000; Nguyen et al., 2006b; Zhang et al., 1998) in *E. coli* DH10B(pKD78) using the following primers: CB346/CB347 for dptD2(CAT), CB346/CB357 for dptD2(CA), CB352/CB353 for lptC4(CAT) and CB352/CB358 for lptC4(CA). The Te domain from dptD2 was amplified using the primers CB332 and CB333 and subcloned along with spec (amplified with CB357/CB358) using Pmel and HindIII to form the Te/spec cassette. This cassette was excised and cloned into Hpal/HindIII sites in conjunction with the CAT modules of Kyn and Asn, respectively, to give pCB401 (Kyn) and pCB407 (Asn). The T-Te di-domain from dptD2 was amplified using the primers CB359 and CB333 and subcloned along with spec (amplified with CB357/CB358) using Pmel and HindIII to form the TTe/spec cassette. This cassette was excised and cloned into Hpal/HindIII sites in conjunction with the C-A di-domains of Kyn and Asn, respectively, to give pCB409 (Kyn) and pCB410 (Asn). Cloning junctions were sequenced for confirmation.

**RESULTS**

**Module fusions and domain exchanges in NRPS subunit DptD**

Previous successful whole-subunit trans-complementation studies, using the di-modular DptD, and distantly related LptC and CdaPS3 (Baltz, 2008; Coéffet-Le Gal et al., 2006;
Miao et al., 2006b), suggested a relatively simple test system to explore the requirements for functional module and domain exchanges. The di-modular dptD, lptD and cdaPS3 genes can be split into individual modules, and ligated in different combinations to produce hybrid di-modules; these can be tested for efficiency of trans-complementation of a S. roseosporus DdptD mutant. In this way, the interpeptide docking sequences can be kept constant by using the 3mGlu12 module from dptD, and fusing it in different ways to the Kyn13, Ile13 and Trp11 modules from the dptD, lptD and cdaPS3 genes, respectively. The efficiencies of lipopeptide production can be compared to the results of whole-subunit exchanges carried out previously. In describing the current experiments, we represent inter-domain linker regions as hyphens (e.g. C-A-T-C-A-T-Te for DptD and related di-modules), and use :: to represent fusion sites (e.g. C-A-T::C-A-T-Te or CAT::CATTe for fusion at the T-C linker).

Table 1. Strains and plasmids

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant characteristics</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S. roseosporus</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UA343</td>
<td>A21978C producer</td>
<td>NRRL 15998</td>
</tr>
<tr>
<td>UA117</td>
<td>rpsL (Sm')</td>
<td>Miao et al. (2006b)</td>
</tr>
<tr>
<td>UA378</td>
<td>ΔdptD::ermE rpsL7</td>
<td>Miao et al. (2006b)</td>
</tr>
<tr>
<td>SD301</td>
<td>UA378::pCB301</td>
<td>This study</td>
</tr>
<tr>
<td>SD302</td>
<td>UA378::pCB302</td>
<td>This study</td>
</tr>
<tr>
<td>SD303</td>
<td>UA378::pCB303</td>
<td>This study</td>
</tr>
<tr>
<td>SD304</td>
<td>UA378::pCB304</td>
<td>This study</td>
</tr>
<tr>
<td>SD305</td>
<td>UA378::pCB305</td>
<td>This study</td>
</tr>
<tr>
<td>SD306</td>
<td>UA378::pCB306</td>
<td>This study</td>
</tr>
<tr>
<td>SD307</td>
<td>UA378::pCB307</td>
<td>This study</td>
</tr>
<tr>
<td>SD308</td>
<td>UA378::pCB308</td>
<td>This study</td>
</tr>
<tr>
<td>SD309</td>
<td>UA378::pCB309</td>
<td>This study</td>
</tr>
<tr>
<td>SD310</td>
<td>UA378::pCB310</td>
<td>This study</td>
</tr>
<tr>
<td>SD401</td>
<td>UA378::pCB401</td>
<td>This study</td>
</tr>
<tr>
<td>SD407</td>
<td>UA378::pCB407</td>
<td>This study</td>
</tr>
<tr>
<td>SD409</td>
<td>UA378::pCB409</td>
<td>This study</td>
</tr>
<tr>
<td>SD410</td>
<td>UA378::pCB410</td>
<td>This study</td>
</tr>
<tr>
<td><strong>E. coli</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH10B</td>
<td>F' mcrA Δ(mrr-hsdRMS-mcrBC) ∆801lacZΔM15 ∆lacX74 recA1 endA1 araD139 (ara, leu)7697 galU galK Δ rpsL(SmR) nupG</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>ML22</td>
<td>DH10B::pUZ8002</td>
<td></td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BW25113</td>
<td>lacI plmB14 ∆lacZ016 hsdR514 ΔaraBADΔ133 ΔrhaBADΔ2378</td>
<td>Datsenko &amp; Wanner (2000)</td>
</tr>
<tr>
<td>pRH5338</td>
<td>rpsI rep Am't</td>
<td>Hosted &amp; Baltz (1997)</td>
</tr>
<tr>
<td>pHM11a</td>
<td>Hm' oriT ermEp' att/int&lt;sub&gt;&lt;sup&gt;81177&lt;/sup&gt;&lt;/sub&gt;</td>
<td>Motamedi et al. (1995)</td>
</tr>
<tr>
<td>PR04</td>
<td>pHM11a::dptD</td>
<td>Miao et al. (2006b)</td>
</tr>
<tr>
<td>pStreptoBAC V</td>
<td>BAC vector' oriT att/int&lt;sub&gt;&lt;sup&gt;88217&lt;/sup&gt;&lt;/sub&gt; Am't</td>
<td>Miao et al. (2005)</td>
</tr>
<tr>
<td>pUZ8002</td>
<td>RK2 derivative defective in oriT</td>
<td>Coeffet-Le Gal et al. (2006)</td>
</tr>
<tr>
<td>pCB301</td>
<td>pHM11a::dptD1 AvrII 81172 (A site) ΔdptD2</td>
<td>This study</td>
</tr>
<tr>
<td>pCB302</td>
<td>pHM11a::dptD1 AvrII 81178 (B site) ΔdptD2</td>
<td>This study</td>
</tr>
<tr>
<td>pCB303</td>
<td>pHM11a::dptD1 AvrII 81184 (C site) ΔdptD2</td>
<td>This study</td>
</tr>
<tr>
<td>pCB304</td>
<td>pCB302 (ΔdptD2 B site) :: dptD2(CATTe)</td>
<td>This study</td>
</tr>
<tr>
<td>pCB305</td>
<td>pCB302 (ΔdptD2 B site) :: lptD2(CATTe)</td>
<td>This study</td>
</tr>
<tr>
<td>pCB306</td>
<td>pCB302 (ΔdptD2 B site) :: dcaPS3-2(CATTe)</td>
<td>This study</td>
</tr>
<tr>
<td>pCB307</td>
<td>pCB301 (ΔdptD2 A site) :: dptD2(CATTe)</td>
<td>This study</td>
</tr>
<tr>
<td>pCB308</td>
<td>pCB303 (ΔdptD2 C site) :: dptD2(CATTe)</td>
<td>This study</td>
</tr>
<tr>
<td>pCB309</td>
<td>pCB301 (ΔdptD2 C site) :: dptD2(CATTe) Δ81172–81184</td>
<td>This study</td>
</tr>
<tr>
<td>pCB310</td>
<td>pCB303 (ΔdptD2 A site) :: dptD2(CATTe) + 81172–81184</td>
<td>This study</td>
</tr>
<tr>
<td>pCB401</td>
<td>pCB302::dptD2(CAT) :: dptD2(Te)</td>
<td>This study</td>
</tr>
<tr>
<td>pCB407</td>
<td>pCB302::lptC4(CAT) :: dptD2(Te)</td>
<td>This study</td>
</tr>
<tr>
<td>pCB409</td>
<td>pCB302::dptD2(CA) :: dptD2(Te)</td>
<td>This study</td>
</tr>
<tr>
<td>pCB410</td>
<td>pCB302::lptC4(CA) :: dptD2(Te)</td>
<td>This study</td>
</tr>
</tbody>
</table>

†BAC, bacterial artificial chromosome.
The first experiments explored the simple fusion of the two
*dptD* module coding sequences by inserting restriction
endonuclease cleavage sites into the T-C linker. From
sequence comparisons between several T-C linker regions
in the daptomycin NRPS proteins, LptD and CdaPS3, we
noted a stretch of about 21 aa with little sequence
conservation (Fig. 2a). The first six amino acids showed
minimal conservation, so we predicted that they were at
least part of the T-C linker. An *AvrII* restriction site was
introduced in various positions of this proposed T-C linker
coding sequence to generate three different new sequences
(A, B and C; Fig. 2b), resulting in three sequential 2 aa
substitutions in the protein. The recombinant plasmids
were introduced into *S. roseosporus* UA378 (*dptD*)
by conjugation and integrated at the IS117 *attB*
site, and all recombinants produced daptomycin at similar high levels
(Table 2). The results indicate that the 6 aa peptide is
indeed part of a functional linker, and that the precise
amino acid sequence is not critical for DptD function
during daptomycin biosynthesis.

To further explore the flexibility of the linker region, we engineered linkers that were contracted or expanded by 4 aa
(linkers D and E; Fig. 2b). Recombinants expressing DptD with these modified linkers produced the same yields of lipopeptides as those containing 2 aa substitutions (Table
2), and the same relative yields of individual lipopeptides
(not shown). The results suggest that the T-C linker region
in DptD has no obvious enzymic and structural function,
other than connecting T and C domains.

**BLAST** analysis (http://www.ncbi.nlm.nih.gov/) of DptD
against the TycC linker from *Bacillus brevis* ATCC 8185
indicated that the 6 aa segment of the T-C linker in DptD
lies within the TycC linker, starting at position 4. The TycC
linker ends in the conserved proline also observed 3 aa
from the right end in the sequences shown in Fig. 2(a)
(Samel *et al.*, 2007). Thus the DptD and other streptomy-

![Fig. 2. T-C linker sequences in daptomycin
NRPSs, CdaPS3 and LptD, and structures of
engineered T-C linkers. (a) Alignment of the T-
C linker regions of lipopeptide NRPSs. The
boxed area depicts the proposed T-C linker
region used for module fusions. (b) Introduction of an *AvrII* site in the A, B, C, D
and E sites of the proposed T-C linker region
in DptD in strains SD307, SD304, SD308,
SD309 and SD310, respectively. The shaded
area depicts the linker region. Contraction and
expansion of the T-C linkers D and E in strains
SD309 and 310 are also shown.]
cete lipopeptide linkers in Fig. 2(a) may span 20–22 aa, providing additional sites for module fusions.

**Heterologous module fusions in DptD**

Module fusions were generated to link the DptD module for 3mGlu$_{12}$ to the terminal modules for Trp$_{11}$ and Ile$_{13}$ from CDA and A54145 NRPSs, respectively (Fig. 1). The substrate specificities of these modules redirect the synthesis of daptomycin to incorporate Trp and Ile/Val, respectively, in place of Kyn$_{13}$. Both modules were introduced with their C-terminal Te domains in the CAT::CATTe fusions using the splice site B in the T-C linker (Fig. 2b and Fig. 3). These hybrid subunits restored lipopeptide biosynthesis in strain UA378 deleted for dptD at yields similar to those produced by recombinants reconstructed to produce daptomycin (Table 2). LC-MS analysis of fermentation broths detected the mass ions indicative of hybrid compounds with Trp$_{13}$ (1630, 1644 and 1658) and Ile$_{13}$ (1543, 1557, 1571 and 1585), respectively (Table 3). The mass ions (1529, 1543, 1557 and 1571) indicative of compounds containing Val$_{13}$ were also detected in lower abundance, consistent with the fact that LptD can also incorporate Val at lower efficiency than Ile (Baltz et al., 2005; Miao et al., 2006b). Similar results were obtained when the splice site C was used to generate hybrid DptD subunits (not shown).

**Domain exchanges in DptD and the role of native TTe linkage**

Although exchanges of the complete CATTe module of the daptomycin NRPS with similar CATTe modules from the CDA and A54145 pathways were successful, the availability of modules containing Te domains is limited. It would be useful to be able to generate hybrid CATTe modules by excising CA or CAT from CAT or CATE modules, and ligating at A-T or T-Te linkers (Fig. 3). To test this, we chose the CATE module for D-Asn$_{11}$ from lptC (Miao et al., 2006a; Fig. 1) to generate C-A-T::Te and C-A::TTe fusions (Table 3, Fig. 3). As controls, similar homologous fusions were made to reconstruct the Kyn$_{13}$ module (strains SD401 and SD409; Table 1). These plasmids containing the hybrid dptD genes were introduced into S. roseosporus UA378 and fermentations of recombinant strains were carried out. The strains expressing the reconstructed Kyn$_{13}$ CAT::Te and CA::TTe fusions produced the predicted lipopeptides. The recombinant containing the Asn$_{13}$ CAT::Te fusion produced no detectable lipopeptides, but the strain with the CA::TTe fusion produced lipopeptides with ESI mass ions of 1558, 1572 and 1586 and UV maximal absorptions at 220 and 281 nm, as expected for the three novel compounds containing A21978C$_{1-3}$(Asn$_{13}$), at yields of about 36–60 mg l$^{-1}$. As a representative example, the structure of A21978C$_{1-3}$(Asn$_{13}$) was supported by its HR-MS data at $m/z$ 1558.7042 [M + H]$^+$. Furthermore, analysis of LC-MS/MS data revealed that the Asn residue was located in the depsipeptide ring, as expected (Fig. 4).

**Biological activities of hybrid peptides**

While most of the daptomycin derivatives were very active against *Staph. aureus* 42, they were slightly less potent than daptomycin. The MIC of daptomycin against *Staph. aureus* 42 is 0.5–1 mg l$^{-1}$, whereas the MICs of the daptomycin...
derivatives were 1 μg ml⁻¹ for the Trp₁₃ derivative and 2–4 μg ml⁻¹ for the Ile₁₃ derivative. This is consistent with results obtained previously with heterologous whole-subunit exchanges (Nguyen et al., 2006a; Miao et al., 2006b). All three Asn₁₃ derivatives had MICs of 128 μg ml⁻¹, further emphasizing the critical importance of position 13 for antibacterial activity (Nguyen et al., 2006b; Miao et al., 2006b).

### Table 3. Production of lipopeptides by recombinants containing reconstructed or hybrid dptD genes

The module and domain fusions of daptomycin NRPS using domains from dptD (Kyn), lptD (Ile), cdaPS3 (Trp) and lptC (Asn) were constructed as described in Methods and in Table 1. The fusion sites are marked by ::. Fermentation broths of recombinants carrying plasmids encoding hybrid subunits were analysed by LC-MS. The heterologous domains from the A54145 and CDA pathways are italicized.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Fusion scheme</th>
<th>Predicted mass ions</th>
<th>Mass ions detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD304</td>
<td>CAT::CA₅₅Te</td>
<td>1634, 1648, 1662</td>
<td>+</td>
</tr>
<tr>
<td>SD305</td>
<td>CAT::CA₅₅Te</td>
<td>1543, 1557, 1571, 1585</td>
<td>+</td>
</tr>
<tr>
<td>SD306</td>
<td>CAT::CA₅₅Te</td>
<td>1630, 1644, 1658</td>
<td>+</td>
</tr>
<tr>
<td>SD401</td>
<td>CAT::CA₅₅Te</td>
<td>1634, 1648, 1662</td>
<td>+</td>
</tr>
<tr>
<td>SD407</td>
<td>CAT::CA₅₅Te</td>
<td>1558, 1572, 1586</td>
<td>-</td>
</tr>
<tr>
<td>SD409</td>
<td>CAT::CA₅₅Te</td>
<td>1634, 1648, 1662</td>
<td>+</td>
</tr>
<tr>
<td>SD410</td>
<td>CAT::CA₅₅Te</td>
<td>1558, 1572, 1586</td>
<td>+</td>
</tr>
</tbody>
</table>

**Fig. 4.** LC-MS/MS analysis of A21978C₁⁻ (Asn₁₃) produced by strain SD410. (a) Chemical structure of A21978C₁⁻ (Asn₁₃) and six characteristic product ions b₁–b₃ and y₁₀–y₁₂. (b) MS/MS spectrum of A21978C₁⁻ (Asn₁₃) and assignment of key product ions.
DISCUSSION

It is now feasible to reprogramme the amino acid specificity of NRPS multi-enzymes, but the factors determining high-efficiency in vivo productivity from hybrid enzymes are poorly understood. This needs to be addressed by linking the application of combinatorial biosynthesis to industrial-scale fermentation production. In the present work, we addressed the efficiency of producing hybrid lipopeptides related to daptomycin by fusing complete modules at T-C linkers in the DptD subunit. We found that neither the sequence nor the length (within ±4 aa) of a segment of the inter-module T-C linker is important for catalytic efficiency of DptD. This linker may simply act as a tether which ensures that domains are aligned and in close proximity.

The flexible T-C linker nucleotide sequences containing inserted restriction endonuclease sites were exploited to construct hybrid genes to redirect lipopeptide biosynthesis to produce hybrid compounds containing Trp13 or Ile13/Val13 at levels comparable to those produced by recombinants producing the native daptomycin factors. This compares favourably with yields obtained using whole-subunit exchanges to produce the same hybrid molecules at about 50% and 25% of controls, respectively (Miao et al., 2006b; Coëffet-Le Gal et al., 2006). The high productivity of recombinant S. roseosporus strains containing heterologous module exchanges may be due in part to the use of homologous inter-peptide docking sites and promiscuous Te domains, but also must be due in part to the inherent flexibility of the inter-module linker. Thus two advantages of using module exchange, as opposed to subunit exchange, are that it conserves the natural inter-peptide docking sites required for efficient communication between DptBC and DptD, and it provides a mechanism to expand the number of substitutions at position 13, which is important for antibacterial activity (Nguyen et al., 2006a; Miao et al., 2006b).

To further explore the engineering of the daptomycin biosynthetic pathway at the terminal module, we made fusions at T-C linkers, keeping the 3mGlu12 module intact, coupled with fusions at A-T or T-Te linkers to exchange the CA or CAT at position 13. Of these, the CAT::CA::TTe double fusion, using the C-A di-domain from the d-Asn1 (CATte) module of A54145 was successful, whereas the CAT::CAT::Te double fusion did not yield product. This is consistent with previous work (Zhou et al., 2006) that demonstrated a functional interaction between T and Te in T-Te di-domains. T domains from typical CAT modules might not be expected to carry out this function efficiently, if at all. Thus maintaining the integrity of the T-Te di-domain may be an important consideration in engineering terminal modules containing Te domains.

The fermentation yield of the recombinant containing the CAT::CA::TTe double fusion in dptD to insert Asn13 was about 30–50% of the yields obtained with strains containing single module fusions. This promising approach provides a route to insert other amino acids at the important position 13. Suboptimal fermentation yields of recombinants may be addressed by focused mutagenesis, as has been demonstrated in the engineering of the mixed polyketide-peptide andrimid (Fischbach et al., 2007).

ACKNOWLEDGEMENTS

We thank N. Cotroneo for MIC determinations. This work was supported by Cubist Pharmaceuticals Inc. and all authors were employed by Cubist during the execution of this work.

REFERENCES


http://mic.sgmjournals.org


Edited by: L. Heide