

ORIGINAL ARTICLE

Structural characterization of a lipopeptide antibiotic A54145E(Asn₃Asp₉) produced by a genetically engineered strain of *Streptomyces fradiae*

Jian-Qiao Gu, Dylan C Alexander, Jessica Rock, Paul Brian¹, Min Chu and Richard H Baltz²

A potent new lipopeptide antibiotic, A54145E(Asn₃Asp₉), was isolated from the fermentation broth of *Streptomyces fradiae* DA1489 engineered to delete genes encoding enzymes involved in hydroxylation of Asn₃ and methoxylation of Asp₉. The chemical structure predicted from the genetic changes in the biosynthetic pathway was determined by analyses of chemical transformations, D, L-amino acid quantitation by enantiomer labeling, tandem LC-MS/MS and 2D NMR techniques. These studies confirmed the primary amino acid sequence of A54145E(Asn₃Asp₉) predicted from the genetic engineering strategy, and also confirmed the structure and locations of three D-amino acids predicted from bioinformatic studies.

The Journal of Antibiotics (2011) 64, 111–116; doi:10.1038/ja.2010.140; published online 24 November 2010

Keywords: A54145E(Asn₃Asp₉); antibiotic; lipopeptide; structure elucidation; *Staphylococcus aureus*; *Streptomyces fradiae*

INTRODUCTION

Daptomycin is a cyclic 13-membered lipopeptide antibiotic approved for the treatment of complicated skin and skin structure infections caused by Gram-positive pathogens¹ and bloodstream infections (bacteremia/endocarditis) caused by *Staphylococcus aureus*, including methicillin-resistant strains.² However, daptomycin is not indicated for community-acquired pneumonia.³ The reduction of daptomycin antibacterial activity by pulmonary surfactant may explain the failure of daptomycin in clinical trials for community-acquired pneumonia.⁴ Many derivatives of daptomycin have been generated by combinatorial biosynthesis,^{5–7} but none had sufficient potency in the presence of bovine surfactant to pursue clinically.⁸

As part of our studies on generating derivatives of daptomycin by combinatorial biosynthesis, we cloned and sequenced the A54145 biosynthetic gene cluster to provide nonribosomal peptide synthetase (NRPS) genes, modules and domains to exchange with those of the daptomycin NRPSs.^{9,10} A54145 is a family of cyclic tridecapeptide antibiotics related to daptomycin and produced by *Streptomyces fradiae*.^{11–15} The A54145 lipopeptides contain several nonproteinogenic amino acids, including hydroxyasparagine (hAsn₃), sarcosine (Sar₅), methoxyaspartic acid (moAsp₉) and 3-methylglutamic acid (3mGlu₁₂).¹² Four different cyclic peptide nuclei vary in the ring amino acid residues Glu/3mGlu at position 12 and Ile/Val at position 13.^{12,14} In addition, bioinformatic analysis predicted that A54145

factors would contain D-amino acids at position 2, 8 and 11 (D-Glu₂, D-Lys₈ and D-Asn₁₁), based on the presence of epimerase domains in the respective NRPS modules.⁹ The predicted locations of the D-amino acids coincide with the locations of the three D-amino acids in daptomycin,¹⁶ suggesting that A54145 and daptomycin may have similar three-dimensional structures. The prominent natural variants have either *n*-decanoic (*n*C₁₀), *iso*-decanoic (*i*C₁₀) or *anteiso*-undecanoic (*a*C₁₁) acids attached to the terminal amino group of Trp₁.^{12,14} By acyl tail modification,¹³ a few A54145 analogs were semisynthesized for evaluation of *in vitro* antibacterial activity.¹⁵ In addition, several A54145 analogs were prepared by chemoenzymatic synthesis with the excised A54145 NRPS thiolation-thioesterase di-domain to cyclize peptide thioester substrates.¹⁷ Nevertheless, none of these derivatives have advanced to clinical studies.

While testing of daptomycin derivatives for antibacterial activity in the presence of bovine surfactant, we also tested some of the natural A54145 factors. A54145E (3mGlu₁₂/Ile₁₃) was inhibited 32-fold by 1% surfactant, whereas A54145D (Glu₁₂/Ile₁₃) was inhibited only by twofold.⁸ This prompted us to develop a gene cloning and ectopic transcomplementation system in *S. fradiae* that could be used to generate combinatorial libraries of A54145 analogs, containing different combinations of amino acid substitutions in the tridecapeptide.^{8,18}

We have extended the use of this genetic system to generate mutants blocked in different combinations of amino acid modifications to

Cubist Pharmaceuticals, Lexington, MA, USA

¹Current address: Actelion Pharmaceuticals, Hegenheimermtweg 91, Allschwil 4123, Switzerland.

²Current address: CognoGen Biotechnology Consulting, Indianapolis, IN 46220, USA.

Correspondence: Dr RH Baltz, CognoGen Biotechnology Consulting, 6438 North Olney Street, Indianapolis, IN 46220, USA.

E-mail: rbaltz923@aol.com

or Dr J-Q Gu, Cubist Pharmaceuticals, 65 Hayden Avenue, Lexington, MA 02421, USA.

E-mail: jian-qiao.gu@cubist.com

Dedicated to the late Dr C Richard 'Dick'/'Hutch' Hutchinson for his exceptional contributions to natural product biosynthesis, engineering, and drug discovery.

Received 9 September 2010; revised 19 October 2010; accepted 20 October 2010; published online 24 November 2010

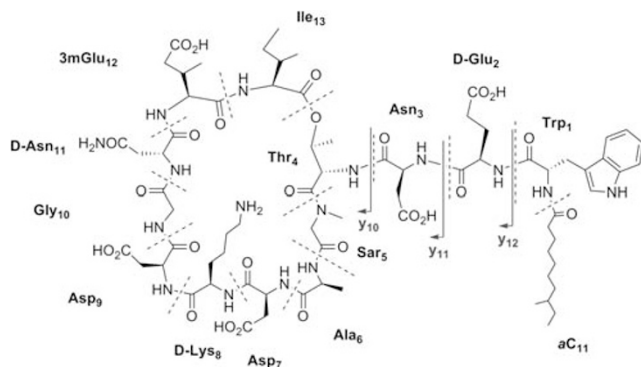


Figure 1 Chemical structure of **1** and three MS-MS product ions y_{10} – y_{12} .

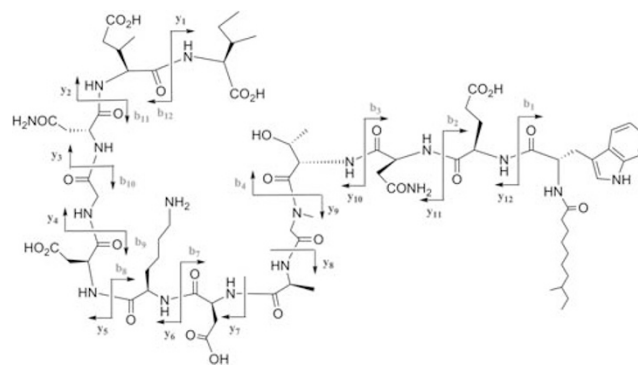


Figure 2 Chemical structure of **1a** and the MS-MS fragmentation patterns.

establish the functions of the *lptJ*, *lptK* and *lptL* genes, and to generate additional novel derivatives of A54145.¹⁸ One of the novel compounds, CB-182,390 (**1**, Figure 1) had very good antibacterial activity against *S. aureus*, and was not inhibited by bovine surfactant.¹⁸ In this study, we present the purification and structural characterization of **1**, and demonstrate that it has two amino acid substitutions: Asn₃ for hAsn₃ and Asp₉ for moAsp₉. Furthermore, the presence of D-amino acids in A54145 at positions 2, 8 and 11 was established.

RESULTS AND DISCUSSION

S. fradiae DA1489 was engineered to delete *lptJ*, *lptK* and *lptL* genes responsible for hydroxylation of Asn₃ and methoxylation of Asp₉. The strain produced a novel lipopeptide with the predicted mass ion at m/z 1626.8 $[M+H]^+$ consistent with the presence of Asn₃ and Asp₉ substituted for hAsn₃ and moAsp₉.¹⁸ DA1489 was fermented in DSF medium supplemented with 0.79% L-Ile in 125-ml shake flasks, and fermentation broth was collected after 6 days. It has been shown previously that the addition of L-Ile to the medium causes preferential incorporation of the *anteiso*-undecanoyl side chain and of L-Ile at position 13.^{10,14} An agar-based bioassay showed that the crude broth inhibited growth of *S. aureus* (data not shown). LC-MS analysis identified the target compound with a A54145-like UV spectra: λ_{max} 221, 279 and 288 nm¹² and ESI-MS data at m/z 1626.8 for the $[M+H]^+$ ion.

To generate sufficient quantities of the lipopeptide for characterization, structure elucidation and assessment of biological activities, 5 l of a fermentation broth of DA1489 was sequentially fractionated by Diaion HP-20 resin column chromatography, Sephadex LH-20 gel filtration size exclusion chromatography and reverse-phase C₈ HPLC to give compound **1** (16 mg). The purity was determined as greater than 95% by HPLC analysis as described in the Materials and methods section.

Compound **1**, $[\alpha]_D^{25}$ -5 (c 0.1, H₂O), was obtained as a white amorphous powder. The molecular formula was deduced as C₇₃H₁₁₁O₂₅N₁₇ from the HR-ESI-MS data at m/z 1626.7976 $[M+H]^+$ (C₇₃H₁₁₂O₂₅N₁₇, calcd for 1626.8015) consistent with the predicted biosynthetic pathway product.¹⁸ Analysis of the IR spectrum suggested that it contained one or more amine groups (3290 cm⁻¹), carboxylic acid groups (1714 cm⁻¹) and amide functionalities (1643 cm⁻¹). The UV absorption at λ_{max} 221, 279 and 288 nm was consistent with the presence of conjugation systems contributed by the Trp moiety.¹² Standard amino acid analysis of **1** indicated the presence of one Ala, four Asx (Asp and/or Asn), one Gly, one Glx (Glu or Gln), one Ile, one Lys, one Thr and three unidentified amino acid residues because of the lack of amino acid standards. The results confirmed that compound **1** contains two additional Asx residues in comparison

with A54145E.¹² The amino acid sequence of **1** was determined by the MS-MS data of the parent molecule and its linear hydrolysate **1a**. The nomenclature for the amino acid sequence-determining fragmental ions was described elsewhere.^{19,20} The MS-MS spectrum of **1** showed distinct fragment peaks at m/z 1608.6 ($[M+H-H_2O]^+$), 1272.6 (y_{12}), 1143.4 (y_{11}), 1029.3 (y_{10}) (Figure 1) along with their corresponding water-loss peaks. The three fragments y_{12} , y_{11} and y_{10} confirmed the side-chain amino acid sequence of **1** as *N*-undecanoyl-Trp-Glu-Asn. Furthermore, the amino acid sequence of **1** was determined by the analysis of the MS-MS data of the linear lipopeptide **1a** (m/z 1644.8 $[M+H]^+$), which was produced by hydrolysis of **1** with lithium hydroxide. As demonstrated in many studies,^{21,22} sequential cleavage of amide bonds of a linear peptide allowed observations of two sets of fragmental ions y_n and b_n , resulting from both C- and N-termini, respectively, to identify the amino acid sequence. The observed experimental MS-MS product ions of **1a** included y_n ($n=1$ –12) at m/z 132.0 (y_1), 275.0 (y_2), 389.2 (y_3), 446.4 (y_4), 561.2 (y_5), 689.4 (y_6), 804.4 (y_7), 875.4 (y_8), 946.4 (y_9), 1047.8 (y_{10}), 1161.5 (y_{11}), 1290.9 (y_{12}) and b_n at m/z 355.3 (b_1), 483.9 (b_2), 598.4 (b_3), 699.4 (b_4), 956.4 (b_7), 1084.8 (b_8), 1199.5 (b_9), 1256.4 (b_{10}), 1370.6 (b_{11}) and 1513.5 (b_{12}), which agreed with their respective theoretical fragment ions (Figure 2). Therefore, the amino acid sequence of **1** was strongly supported by its linear hydrolysate **1a** as *N*-undecanoyl-Trp-Glu-Asn-Thr-Sar-Ala-Asp-Lys-Asp-Gly-Asn-3mGlu-Ile.

Miao *et al.*⁹ found epimerase domains in modules 2, 8 and 11 of the A54145 NRPS, suggesting the presence of D-Glu₂, D-Lys₈ and D-Asn₁₁. In this study, to confirm the presence of the three D-stereochemistry amino acids, Glu, Lys and Asp, were specifically selected for determination of their optical purity using an enantiomer-labeling method.²³ In this method, the measured optical purity of Asp was derived from both Asp and Asn residues, as Asn was converted to Asp during the DCl/D₂O hydrolysis of **1**. The GC-MS results showed the optical purity of D-Glu, D-Lys and D-Asp as 96, 99 and 28%, respectively. The obtained optical data of D-Glu and D-Lys obviously suggested the presence of D-Glu₂ and D-Lys₈ in **1**. The measured 28% optical purity of D-Asp for **1** could be reasonably interpreted from the contribution of one D-Asn₁₁ and three L-Asp/Asn residues, including L-Asp₇, L-Asp₉ and L-Asn₃. The above-mentioned data are consistent with the results predicted by the analysis of the biosynthetic gene clusters,^{8,9} and with the precedent of determining D-Ser₁₁ in daptomycin and its analog,^{16,21} but provide no additional evidence to confirm D-Asn₁₁ in A54145. It is noteworthy that these three D-amino acids are present at the same positions as D-Asn₂, D-Ala₈ and D-Ser₁₁ in daptomycin, suggesting that A54145 and daptomycin may have similar three-dimensional structures.

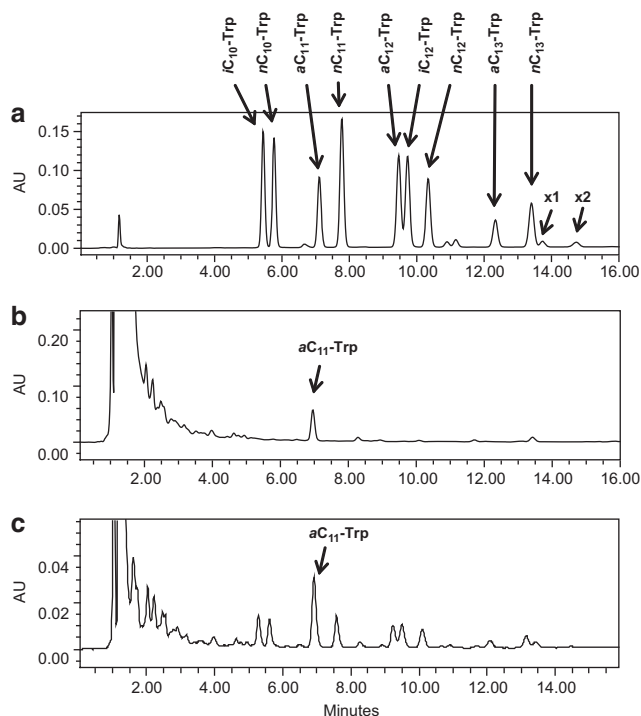


Figure 3 Determination of *N*-acyl substituent of **1** by HPLC analyses of fatty acid (FA)-Trp amides. (a) HPLC profile of a panel of authentic FA-Trp amides. (b) HPLC profile of Trp coupling resultant for compound **1**. (c) HPLC profile after spiking the tested lipopeptide reaction mixture (b) into a panel of authentic FA-Trp amides (a). x1 and x2 stands for two unidentified side products in (a).

Naturally occurring A54145 factors have acyl groups, which consist of *n*-decanoyl (*n*C₁₀), *iso*-decanoyl (*i*C₁₀) and *anteiso*-undecanoyl (*a*C₁₁).¹² The acyl substituent of **1** was determined by acidic hydrolysis of the lipopeptide followed by coupling the resultant FA with Trp to produce a corresponding FA-Trp amide.²¹ The HPLC retention time (*t*_R) and UV spectra of the FA-Trp amides from a library of authentic FAs (Figure 3a) were compared with those of amide from **1** (Figure 3b). In addition, the identification of the acyl group was further confirmed by observing a homogeneous peak at *t*_R 7.01 min after spiking the reaction resultant of **1** with the library mixture (Figure 3c). As a result, the *anteiso*-undecanoyl (*a*C₁₁) group was identified in **1**, as expected from fermentation of *S. fradiae* DA1489 in DSF medium supplemented with L-Ile.^{14,18}

Structural characterization of **1** by NMR studies was somewhat challenging because of the observation of two sets of ¹H-NMR signals in DMSO-*d*₆. Initial NMR spectra of **1** were measured at 2 mM sample concentration in a mixture of 90% H₂O and 10% D₂O (pH 5.0 adjusted by HCl), 5 °C (data not shown) as described for daptomycin NMR studies.²⁴ However, it was found that the relatively broad-resonance line widths and lack of sufficient amide proton dispersion resolution prevented the identification and full assignment of the amino acid spin systems. The broad resonance line widths may have resulted from the significant aggregation tendency of the lipopeptide molecule in aqueous solution.²⁴ DMSO-*d*₆ was, therefore, selected as an alternative NMR solvent. Interestingly, two sets of proton resonances at a ratio of approximately 3:2 were observed for **1** at room temperature in DMSO-*d*₆. This phenomenon was observed for all A54145 analogs (unpublished data). Furthermore, a series of ¹H-NMR studies were carried out in DMSO-*d*₆ at various temperatures, including 25, 30, 40, 50, 60, 70 and 80 °C. These data showed partial

coalescence of ¹H-NMR signals at elevated temperature, consistent with the presence of slowly exchanging conformers such as peptide amide rotamers.²⁵ In addition, the spectrum measured at 40 °C exhibited the optimal proton dispersion resolution. Accordingly, all subsequent NMR experiments were carried out for **1** (5 mM) in DMSO-*d*₆ at 40 °C. To identify and sequence-specifically assign the amino acid residues of **1**, NMR spectra, including ¹H, DEPT, gHSQC-DEPT, gHMBC, gDQF-COSY, TOCSY and ROESY were recorded. To identify amino acid residues of **1**, analysis of the gDQF-COSY spectrum enabled the assignment of aromatic moieties of Trp by analysis of the ³J_{H-H} scalar couplings. For example, the indole *N*-proton at δ_{H} 10.71 was correlated with δ_{H} 7.11 (H-2) in the Trp₁ residue. In turn, six amino acid residues of **1** including Glu, Thr, Lys, Ala, 3mGlu and Ile were unambiguously identified after analyses of the TOCSY and gHSQC-DEPT data. In these cases, the TOCSY spectrum reveals the intraresidue ¹H-¹H spin systems extending from each exchangeable backbone amide proton through to the nonexchangeable side-chain protons of each residue. gHSQC-DEPT is a phase-sensitive 2D NMR experiment similar to HSQC (CH and CH₃ in one phase, and CH₂ in the opposite phase). For example, the amino acid residue Thr of **1** was identified as described below. The TOCSY spectrum revealed spin systems for the Thr residue in both the major and minor conformers, with assignments for the minor conformer listed in parentheses [] as follows; an exchangeable NH at δ 7.85 [7.62], H- α at δ 5.05 [4.84], H- β at δ 5.14 [5.00] m and H- γ at δ 1.14 [1.03] d (*J*=6.2 Hz). Assignment of the Thr residue was further confirmed by gHSQC-DEPT data showing C- α at δ 53.41 [53.33] (CH), the oxygenated C- β at δ 72.76 [72.97] (CH) and C- γ at δ 18.30 [17.70] (CH₃). In addition, one Gly residue was also identified from the gHSQC-DEPT spectrum by virtue of H- α signals at δ_{H} 3.78 [3.77] and 3.66 [3.65] both correlated to δ_{C} 45.37 [45.11] t. These resonances were characteristically upfield relative to α -position chemical shifts of the other nonglycine amino acid residues in the molecule. Furthermore, the Sar residue was identified by the characteristic Me signals at δ_{H} 3.06 [2.82] s and two coupled methylene signals at δ_{H} 4.20 [4.51] d (*J*=17 Hz) and 3.86 [4.00] d (*J*=17 Hz). Nevertheless, assignments of the remaining Asp and Asn residues were challenging because of their heavily overlapping H- α and H₂- β signals attributable to the repeating units containing the same spin system pattern. To resolve this issue, the ROESY spectral data were collected to trace the sequence-specific resonance assignment of **1**, especially useful for two Asp and two Asn residues. Summarized as a chart in Figure 4, most backbone connectivities were confirmed by the inter-residue dipolar correlations (NOE) of amide protons with the side chain and H- α protons of the neighboring residues in the sequence. More importantly, spatial correlations between the Thr₄ NH at δ 7.85 and the Asn₃ H- α at δ 4.63 and NH at δ 8.08 led to the assignment of NMR signals of the Asn₃ residue after carefully deducing the spin system containing the NH at δ 8.08 from the TOCSY spectrum. Using a similar approach, NMR signals of Asp₇, Asp₉ and D-Asn₁₁ residues were individually assigned based on their backbone NOE correlations between H- α and neighboring amide protons. Furthermore, a series of NOEs (Figure 4) observed for **1** should facilitate the calculation of the three-dimensional structure of **1** in the future. Thus, the structure of **1** was determined as *anteiso*-undecanoyl-L-tryptophanyl-D-glutamyl-L-asparaginyll-L-threonyl-sarcosinyll-L-alanyl-L-aspartyl-D-lysinyll-L-aspartyl-glycyl-D-asparaginyll-*threo*-3-methyl-L-glutamyl-L-isoleucine- ϵ 1-lactone, to which the trivial name A54145E(Asn₃Asp₉) has been assigned.

Compound **1** obtained from *S. fradiae* DA1489 was tested for *in vitro* antibacterial activity against *S. aureus* in the presence of pulmonary surfactant using a microdilution technique, following

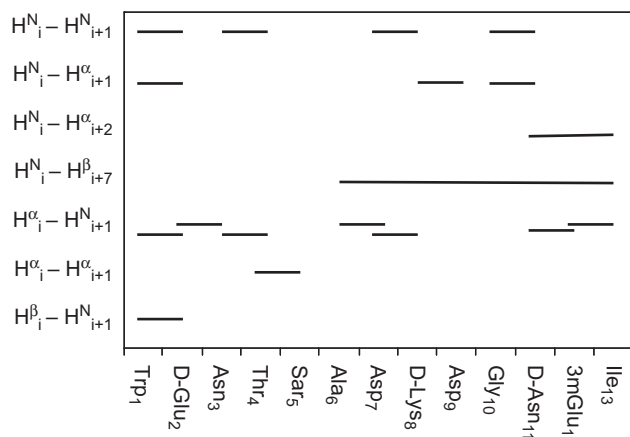


Figure 4 Inter-residue NOEs observed in the ROESY data set of **1** at a mixing time of 100 ms. H^N_i stands for the amide bond proton of a certain amino acid residue within the molecule. H^α_i and H^β_i stand for protons at positions α and β in an amino acid residue.

CLSI standards as described,⁴ and it exhibited a MIC of $2 \mu\text{g ml}^{-1}$ in the presence of 1% of pulmonary surfactant,¹⁸ whereas the daptomycin control showed an MIC of $64 \mu\text{g ml}^{-1}$. The superior antibacterial activity in the presence of bovine surfactant suggests that further investigations of compound **1** in animal models would be warranted.

To our knowledge, this is first example characterizing the novel lipopeptide analog by the complete assignment of NMR proton and carbon signals and determination of the D-stereochemistry of three amino acids using an enantiomer-labeling method. The current study supports the predictions made on the basis of the gene deletions in the engineered *S. fradiae* strain, and points out the predictive power of DNA sequence analysis of NRPS genes to assign the stereochemistry of the individual amino acids by the presence or absence of epimerase domains in individual modules.^{9,16}

EXPERIMENTAL PROCEDURE

General experimental procedures

Melting points were determined using a Thomas-Hoover capillary melting point apparatus (Thomas Scientific, Swedesboro, NJ, USA) and are uncorrected. Optical rotations were recorded by a Perkin-Elmer polarimeter 341 (Perkin-Elmer, Waltham, MA, USA). UV spectra were measured on an Evolution 300 UV spectrophotometer (Thermo Corporation, Waltham, MA, USA) and IR spectra on a Thermo Nicolet 380 FT-IR spectrometer. NMR spectra were recorded on a Varian 600 MHz NMR spectrometer (Agilent Technologies, Santa Clara, CA, USA) with standard pulse programs and 1D- and 2D NMR data were processed by MestReNova software (Mestrelab Research, Santiago de Compostela, Spain). HR-ESI-MS data were obtained from a Finnigan LTQ Orbitrap mass spectrometer (Thermo Corporation). Column chromatography was conducted on Diaion HP20 resins (Mitsubishi Chemical, Tokyo, Japan) and Sephadex LH-20 (GE Healthcare Biosciences, Piscataway, NJ, USA). Analytical HPLC was performed at ambient temperature using a Waters Alliance 2690 HPLC system and a 996-photodiode array detector (Waters, Milford, MA, USA). Semi-preparative HPLC was performed on a Varian system equipped with two-model SD-1 PrepStar solvent delivery modules, a PDA ProStar detector, a ProStar injector and a model 701 fraction collector.

Strains, media and fermentation conditions

Media and growth conditions were as described previously.^{10,14,18} Engineered strain *S. fradiae* DA1489 was constructed and fermented in DSF production

medium supplemented with 0.79% (w/v) of L-Ile as described.¹⁸ *S. aureus* ATCC 29213 was used for bioassay of lipopeptide antibiotic activity.⁸

HPLC analyses

Target compound **1** was analyzed by a Waters HPLC system with a Waters symmetry C₈ column (4.6×250 mm, $5 \mu\text{m}$) and a Waters guard symmetry C₈ cartridge (4.6×20 mm, $5 \mu\text{m}$). Mobile solvent systems included CH₃CN plus 0.01% TFA (v/v) modifier (solvent A) and water plus 0.01% TFA (v/v) modifier (solvent B). The mobile phase, flowing at 1.5 ml min^{-1} , was linearly changed from 30 to 45% A over 14 min, washed at 90% A (2 min) and equilibrated with 30% A (4 min). In HPLC chromatograms, a prominent peak with retention time (t_R) at 8.82 min was identified as the predicted lipopeptide **1** by the A54145-like characteristic UV spectrum (λ_{max} 221, 279 and 288 nm) and ESI-MS data at m/z 1626.8 [M+H]⁺.

Extraction and isolation

Production culture broth of DA1489 (5 l) was centrifuged to remove biomass, and the supernatant was loaded onto an open glass column packed with preconditioned 500 ml Diaion HP20 resin (60×500 mm) in water. The column was sequentially eluted with 1.5 l each of water, 10% MeOH, 30% MeOH and MeOH. The MeOH eluate was concentrated by rotary evaporation and lyophilized to give 2 g crude material, which was further subjected to a 500 ml Sephadex LH-20 column chromatography eluting with a mixture of MeOH-H₂O (1:1). Fractions containing target component were collected, giving 720 mg of lyophilized powder. Further purification was achieved by semipreparative HPLC using a Waters SymmetryPrep C₈ column (19×300 mm, $7 \mu\text{m}$) at a flow rate of 20 ml min^{-1} with a linear gradient from 33 to 40% A over 20 min, washing the column with 90% A for 4 min and equilibrating with 33% A for 6 min. HPLC traces were recorded at UV 220 nm. Compound **1** (16 mg) was obtained at t_R 17.6 min.

A54145E(Asn₃Asp₉) (1). White amorphous powder; m.p. 205°C (dec.); $[\alpha]_D^{25}$ -5 (c 0.1, H₂O); UV H₂O λ_{max} (ϵ) 221 (38905), 279 (5888) and 288 (5248) nm; IR (dried film) ν_{max} 3290, 3054, 2930, 1714, 1643, 1519, 1455, 1403, 1336, 1267 and 1189 cm^{-1} ; ¹H and ¹³C NMR data, see Table 1; HR-ESI-MS m/z 1626.7976 [M+H]⁺ (C₇₃H₁₁₂O₂₅N₁₇, calcd for 1626.8015).

Alkali hydrolysis of 1. A 1 ml aliquot of **1** ($10 \mu\text{g ml}^{-1}$) was prepared in a 1.5 ml vial, to which 1 mg lithium hydroxide monohydrate was added. After 20 min at room temperature, the hydrolysis reaction was quenched by adding 10 μl of formic acid, and the hydrolysate **1a** was subjected to LC-MS-MS analysis.

LC-MS-MS analysis of 1a. LC-MS-MS analysis was carried out using a Finnigan Surveyor HPLC system interfaced to a Finnigan LTQ Orbitrap mass spectrometer. A Waters Sunfire C₈ column (2.1×100 mm, $3.5 \mu\text{m}$) was used for HPLC separations. The sample injection volume was 10 μl . The solvent systems were water containing 0.1% formic acid (solvent C) and CH₃CN containing 0.1% formic acid (solvent D). The linear gradient for separation was set up as from 10 to 90% D over 10 min, and then the column was cleaned with 90% D for 2 min and equilibrated with 10% D for 3 min. The flow rate was $350 \mu\text{l min}^{-1}$. Positive ESI source conditions were sheath gas flow rate at 30 (in arbitrary units), auxiliary gas flow rate at 2 (in arbitrary units), ion spray voltage at 2 kV, capillary temperature at 350°C , capillary voltage at 40 V and tube lens voltage at 250 V. Normalized collision energy was 60%.

MS-MS data used for amino acid sequence analysis of **1** were acquired from its hydrolysate **1a**, which was confirmed by the [M+H]⁺ m/z at 1644.8 along with its characteristic UV spectra at λ_{max} 221, 279 and 288 nm.

Analyses of amino acids optical purity

A modified enantiomer-labeling method was employed to quantitatively determine the enantiomeric purity of amino acids according to previously established protocols.²³ In brief, compound **1** (100 nmol) was hydrolyzed in 0.5 ml of 6 N DCl in D₂O at 110°C for 24 h and dried under a stream of nitrogen. The residue was esterified with 0.5 ml of 4 N DCl in MeOH at 110°C for 15 min. After drying by a gentle nitrogen steam, the ester was acylated with 0.25 ml of trifluoroacetic anhydride in ethyl trifluoroacetate (1:1, v/v) at 130°C

Table 1 ¹H and ¹³C NMR data for **1** in DMSO-*d*₆ (δ in p.p.m., mult., *J* in Hz)^a

Residue	Position	δ _C [minor conformer]	δ _H [minor conformer]
Trp ₁	NH		7.86 d (7.5)
	α	56.96 [56.88] d	4.50 m
	β	30.98 [31.10] t	2.92 dd (8.8, 15), 3.07 dd (5.6, 15)
	1		10.71 br s
	2	126.72 [126.64] d	7.11 d (2.1)
	3	113.2 s	
	4	121.57 d	7.56 d (7.8)
	4a	131.68 s	
	5	121.24 d	6.94 dd (7.8)
	6	123.89 d	7.02 dd (7.8)
	7	114.37 d	7.30 d (7.8)
	7a	140.38 s	
	C=O	176.27 s	
Glu ₂	NH		8.10 ^c
	α	55.19 [55.25] d	4.19 ^c
	β	30.38 [30.30] t	1.64 ^c , 1.86 ^c
	γ	32.99 t	2.05 ^c
	C=O	175.78 s	
Asn ₃	COOH	178.16 s	12.10, br s
	NH		8.08 [8.24] ^c
	α	52.44 d	4.63 [4.58] ^c
	β	39.59 t	2.41 [2.39] m, 2.54 [2.58] ^c
C=O	174.33 s		
	CONH ₂	175.83 s	6.79 [6.85] br s, 7.22 [7.31] br s
Thr ₄	NH		7.85 [7.62] d (8.1)
	α	53.41 [53.33] d	5.05 [4.84] dd (4.6, 8.1)
	β	72.76 [72.97] d	5.14 [5.00] m (6.2)
	γ	18.30 [17.70] q	1.14 [1.03] d (6.2)
C=O	172.87 [172.60] s		
	Sar ₅	N-CH ₃	39.49 [38.14] q
α		53.08 [54.98] t	4.20 [4.51] d (17), 3.86 [4.00] d (17)
Ala ₆	C=O	172.76 [172.32] s	
	NH		8.09 [8.15] ^c
	α	51.96 [51.87] d	4.24 [4.11] m
	β	20.6 [21.0] q	1.22 [1.24] d (6.3)
C=O	176.80 s		
	Asp ₇	NH	
α		52.97 d	4.56 ^c
β		38.67 t	2.49 ^c , 2.67 m
C=O	174.94 s		
	COOH	176.09 s	12.10, br s
Lys ₈	NH		7.84 ^c
	α	56.13 [55.49] d	4.12 ^c
	β	34.03 [33.51] t	1.52 m, 1.66 m
	γ	25.18 [25.04] t	1.25 ^c
	δ	29.71 t	1.49 ^c
	ε	41.81 t	2.73 ^c
NH ₂			7.71
	C=O	176.15 s	
Asp ₉	NH		7.89 ^c
	α	53.05 d	4.65 ^c
	β	40.50 t	2.42 ^c , 2.52 ^c
	C=O	175.09 s	
COOH	173.21 s	12.10, br s	

Table 1 Continued

Residue	Position	δ _C [minor conformer]	δ _H [minor conformer]
Gly ₁₀	NH		7.95 [7.76] m
	α	45.37 [45.11] t	3.78 [3.77] m, 3.66 [3.65] m
C=O	173.18 s		
	Asn ₁₁	NH	
α		52.44 d	4.56 ^c
β		39.02 t	2.52 ^c , 2.68 ^c
C=O	174.27 s		
	CONH ₂	175.76 s	6.84 br s, 7.25 br s
3mGlu ₁₂	NH		7.89 [7.69] ^c
	α	59.00 [58.76] d	4.41 m
	β	35.64 [35.80] d	2.33 ^c
	γ1	40.87 [40.95] t	1.94 [2.01] m, 2.31 [2.37] ^c
γ2	14.37 q		0.79 d (6.3)
	C=O	174.74 [175.17] s	
COOH	177.80 s		12.10, br s
	Ile ₁₃	NH	
α		60.15 [60.02] d	4.12 m
β		39.03 d	1.74 m
γ1	28.09 [28.37] t		1.12 ^c , 1.37 ^c
	γ2	17.78 ^b q	0.80 ^c
δ	18.30 ^b q	0.80 ^c	
C=O	174.31 s		
	aC11	C=O	176.74 s
2		38.33 t	2.04 ^c (7.0)
3		28.29 t	1.35 m
4		32.06 t	1.08 ^c
5		29.48 t	1.13 ^c
6		29.48 t	1.13 ^c
7		39.14 t	1.19 ^c , 1.01 ^c
8		36.85 d	1.24 ^c
9		31.77 t	1.26 ^c , 1.07 ^c
10		14.20 q	0.81 t (6.8)
CH ₃ -8	22.22 q	0.80 d (6.8)	

^aAssignments were made on the basis of gDQF-COSY, TOCSY, gHSQC-DEPT, gHMBC and ROESY data.^bInterchangeable assignment due to lack of observed HMBC correlations.^cOverlapped signals.

for 10 min. After the excess of reagents were removed by a stream of nitrogen, the residue was dissolved in 0.15 ml of toluene and injected into the GC. All of the GC-MS analyses were performed using a Varian Saturn 2000/HP5973 GC-MS system (Agilent Technologies, Santa Clara, CA, USA), including a Saturn 2000 software/HP Chem integrator. The GC was fitted with a deactivated glass capillary column (20 000 × 0.31 mm) coated with Chirasil-Val (film thickness 0.2 μm) (Applied Sciences Laboratories, State College, PA, USA). Hydrogen was used as the carrier gas at a flow rate of 1.5 ml min⁻¹. Sample introduction (0.5 μl) was made in split mode with a 25:1 split ratio and an injector temperature of 190 °C. The temperature of the column oven was programmed at 65 °C for 3 min, followed by a linear gradient of 65–190 °C over 31 min.

Determination of *N*-acyl substituent

A panel of nine authentic fatty acid-Trp (FA-Trp) amides was prepared as described previously.²¹ All chemicals used for coupling reactions were obtained in analytical grade. Decanoic-(*n*C₁₀), lauric-(*n*C₁₂), tridecanoic-(*n*C₁₃) acids, L-tryptophan (Trp), 1-hydroxy-benzotriazole hydrate (HOBt) and *O*-(7-azabenzotriazol-1-yl)-*N,N,N',N'*-tetramethyl-uronium PF₆ (HATU) were purchased from Sigma-Aldrich (St Louis, MO, USA). The compounds 8-methylnonanoic

acid (*i*C₁₀) and undecanoic acid (*n*C₁₁) were bought from Fisher Scientific (Pittsburgh, PA, USA). Compounds 8-methyldecanoic-(*a*C₁₁), 10-methyldecanoic-(*a*C₁₃) and 10-methylundecanoic-(*i*C₁₂) acids were purchased from Indofine Chemical Company (Hillsborough, NJ, USA). The compound 9-methylundecanoic acid (*a*C₁₂) was provided by Dr Wai-Kit Pang (Medicinal Chemistry Department, Cubist Pharmaceuticals, Lexington, MA, USA). Analytical HPLC was performed on a YMC-Pack ODS-A column (4.6×150 mm, 5 μm) (YMC America, Allentown, PA, USA) and monitored at 280 nm. Mobile solvent systems included CH₃CN plus 0.01% TFA (v/v) modifier (solvent A) and water plus 0.01% TFA (v/v) modifier (solvent B). The mobile phase was linearly delivered from 55 to 70% A over 15 min. At a flow rate of 1.5 ml min⁻¹, typical retention times were observed for nine authentic FA-Trp amides, including *i*C₁₀-Trp at 5.37 min, *n*C₁₀-Trp at 5.69 min, *a*C₁₁-Trp at 7.01 min, *n*C₁₁-Trp at 7.67 min, *a*C₁₂-Trp at 9.33 min, *i*C₁₂-Trp at 9.59 min, *n*C₁₂-Trp at 10.20 min, *a*C₁₃-Trp at 12.18 min and *n*C₁₃-Trp at 13.25 min.

The *N*-acyl substituent of **1** was determined by acidic hydrolysis of **1** to produce the corresponding FA tail, followed by coupling with Trp to further produce the resultant tryptophanyl amide. The HPLC retention times (*t*_R) and UV spectra were compared with those of authentic FA-Trp amides. Compound **1** (3 mg) was dissolved with 10 N HCl (1 ml) in a 1 ml vacuum hydrolysis tube. The tube was evacuated and then placed in a 110 °C heating block for 16 h. After cooling, CH₂Cl₂ (1 ml × three times) was added to extract the resultant FA. After the CH₂Cl₂ extract was dried under a stream of N₂, 10 mM Trp, 10 mM HOBt, 10 mM HATU, 0.75 ml DMF and 10 μl DIPEA were added sequentially. After 2 h incubation at room temperature with shaking at 200 r.p.m., a 10 μl aliquot of the reaction mixture was analyzed by analytical HPLC as described above to acquire the *t*_R and UV spectrum. To confirm the presence of the *N*-acyl group, a panel of authentic FA-Trp amides was further spiked into the lipopeptide reaction mixture for HPLC analysis.

ACKNOWLEDGEMENTS

We are grateful to C Mascio and J Silverman for MIC testing and H Cheng for initial LC-MS analyses. We also thank V Rajgarhia and C Gandhi for providing fermentation broth. We are also grateful to I Parsons and M Varoglu for their critical review of this manuscript. We acknowledge GJ Heffron of Harvard Medical School NMR facility for recording NMR spectra, WS Lane of Harvard Microchemistry and Proteomics analysis facility for standard amino acid analysis, and J Gerhardt of CAT GmbH, Tübingen, Germany for determination of optical purity of selected amino acids. This work was supported by Cubist Pharmaceuticals, Inc.

- 1 Arbeit, R. D., Maki, D., Tally, F. P., Campanaro, E. & Eisenstein, B. I. The safety and efficacy of daptomycin for the treatment of complicated skin and skin-structure infections. *Clin. Infect. Dis.* **38**, 1673–1681 (2004).
- 2 Fowler, V. G. *et al.* Daptomycin versus standard therapy for bacteremia and endocarditis caused by *Staphylococcus aureus*. *N. Engl. J. Med.* **355**, 653–655 (2006).

- 3 Pertel, P. E. *et al.* Effects of prior effective therapy on the efficacy of daptomycin and ceftriaxone for the treatment of community-acquired pneumonia. *Clin. Infect. Dis.* **46**, 1142–1151 (2008).
- 4 Silverman, J. A., Morton, L. I., Vanpraagh, A. D., Li, T. & Alder, J. Inhibition of daptomycin by pulmonary surfactant: *in vitro* modeling and clinical impact. *J. Infect. Dis.* **191**, 2149–2152 (2005).
- 5 Miao, V. *et al.* Genetic engineering in *Streptomyces roseosporus* to produce hybrid lipopeptide antibiotics. *Chem. Biol.* **13**, 269–276 (2006).
- 6 Nguyen, K. T. *et al.* Identification of a glutamic acid 3-methyltransferase gene by functional analysis of an accessory gene locus important for daptomycin biosynthesis in *Streptomyces roseosporus*. *Mol. Microbiol.* **61**, 1294–1307 (2006).
- 7 Nguyen, K. *et al.* Combinatorial biosynthesis of lipopeptide antibiotics related to daptomycin. *Proc. Natl Acad. Sci. USA* **103**, 17462–17467 (2006).
- 8 Nguyen, K. T. *et al.* Genetically engineered lipopeptide antibiotics related to A54145 and daptomycin with improved properties. *Antimicrob. Agents Chemother.* **54**, 1404–1413 (2010).
- 9 Miao, V., Brost, R., Chapple, J., Coëffet-LeGal, M.-F. & Baltz, R. H. The lipopeptide antibiotic A54145 biosynthetic gene cluster from *Streptomyces fradiae*. *J. Ind. Microbiol. Biotechnol.* **33**, 129–140 (2006).
- 10 Alexander, D. C. *et al.* Development of a genetic system for lipopeptide combinatorial biosynthesis in *Streptomyces fradiae* and heterologous expression of the A54145 biosynthetic gene cluster. *Appl. Environ. Microbiol.* **76**, 6877–6887 (2010).
- 11 Boeck, L. D. *et al.* A54145, a new lipopeptide antibiotic complex: discovery, taxonomy, fermentation and HPLC. *J. Antibiot.* **43**, 587–593 (1990).
- 12 Fukuda, D. S., Du Bus, R. H., Baker, P. J., Berry, D. M. & Mynderse, J. S. A54145, a new lipopeptide antibiotic complex: isolation and characterization. *J. Antibiot.* **43**, 594–600 (1990).
- 13 Fukuda, D. S., Debono, M., Molloy, R. M. & Mynderse, J. S. A54145, a new lipopeptide antibiotic complex: microbial and chemical modification. *J. Antibiot.* **43**, 601–606 (1990).
- 14 Boeck, L. D. & Wetzel, R. W. A54145, a new lipopeptide antibiotic complex: factor control through precursor directed biosynthesis. *J. Antibiot.* **43**, 607–615 (1990).
- 15 Counter, F. T. *et al.* A54145, a new lipopeptide antibiotic complex: microbiological evaluation. *J. Antibiot.* **43**, 616–622 (1990).
- 16 Miao, V. *et al.* Daptomycin biosynthesis in *Streptomyces roseosporus*: cloning and analysis of the gene cluster and revision of peptide stereochemistry. *Microbiology* **151**, 1507–1523 (2005).
- 17 Kopp, F., Grünwald, J., Mahlert, C. & Marahiel, M. A. Chemoenzymatic design of acidic lipopeptide hybrids: new insights into the structure-activity relationship of daptomycin and A54145. *Biochemistry* **45**, 10474–10481 (2006).
- 18 Alexander, D. C. *et al.* Production of novel lipopeptide antibiotics related to A54145 by *Streptomyces fradiae* mutants blocked in biosynthesis of modified amino acids and assignment of *lptJ*, *lptK* and *lptL* gene functions. *J. Antibiot.* (e-pub ahead of print 24 November 2010; doi:10.1038/ja.2010.138).
- 19 Biemann, K. & Martin, S. Mass spectrometric determination of the amino acid sequence of peptides and proteins. *Mass Spectrom. Rev.* **6**, 1–76 (1987).
- 20 Roepstorff, P. & Fohlman, R. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.* **11**, 601 (1984).
- 21 Gu, J.-Q. *et al.* Structural characterization of daptomycin analogues A21978C₁₋₃ (D-Asn₁₁) produced by a recombinant *Streptomyces roseosporus* strain. *J. Nat. Prod.* **70**, 233–240 (2007).
- 22 Tseng, J. L., Yan, L., Fridland, G. H. & Desiderio, D. M. Tandem mass spectrometry analysis of synthetic opioid peptide analogs. *Rapid Commun. Mass Spectrom.* **9**, 264–275 (1995).
- 23 Frank, H., Nicholson, G. J. & Bayer, E. Enantiomer labelling, a method for the quantitative analysis of amino acids. *J. Chromatogr.* **167**, 187–196 (1978).
- 24 Ball, L. J., Goult, C. M., Donarski, J. A., Micklefield, J. & Ramesh, V. NMR structure determination and calcium binding effects of lipopeptide antibiotic daptomycin. *Org. Biomol. Chem.* **2**, 1872–1878 (2004).
- 25 Smith, M. B. & March, J. in *March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure* 5th edn, 73–77 (John Wiley & Sons, Inc., New York, 2001).