Synthesis and Characterization of the Arylomycin Lipoglycopeptide Antibiotics and the Crystallographic Analysis of Their Complex with Signal Peptidase

Jian Liu,[†] Chuanyun Luo,[‡] Peter A. Smith,[†] Jodie K. Chin,[†] Malcolm G. P. Page,[§] Mark Paetzel,[*,‡] and Floyd E. Romesberg[*,†]

[†]Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, United States
[‡]Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada
[§]Basilea Pharmaceutica International Ltd., Grenzacherstrasse 487, CH-4058, Basel, Switzerland

Ⓢ *Supporting Information*

ABSTRACT: Glycosylation of natural products, including antibiotics, often plays an important role in determining their physical properties and their biological activity, and thus their potential as drug candidates. The arylomycin class of antibiotics inhibits bacterial type I signal peptidase and is comprised of three related series of natural products with a lipopeptide tail attached to a core macrocycle. Previously, we reported the total synthesis of several A series derivatives, which have unmodified core macrocycles, as well as B series derivatives, which have a nitrated macrocycle. We now report the synthesis and biological evaluation of lipoglycopeptide arylomycin variants whose macrocycles are glycosylated with a deoxy-α-mannose substituent, and also in some cases hydroxylated. The synthesis of the derivatives bearing each possible deoxy-α-mannose enantiomer allowed us to assign the absolute stereochemistry of the sugar in the natural product and also to show that while glycosylation does not alter antibacterial activity, it does appear to improve solubility. Crystallographic structural studies of a lipoglycopeptide arylomycin bound to its signal peptidase target reveal the molecular interactions that underlie inhibition and also that the mannose is directed away from the binding site into solvent which suggests that other modifications may be made at the same position to further increase solubility and thus reduce protein binding and possibly optimize the pharmacokinetics of the scaffold.
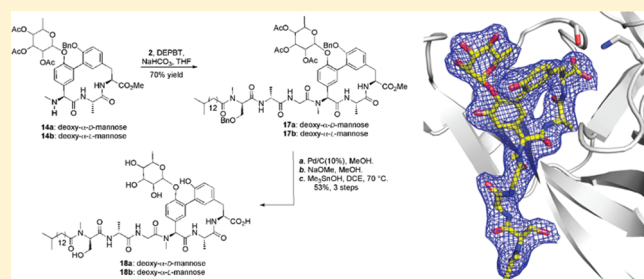
■ INTRODUCTION

The clinical and agricultural use of antibiotics imposes a relentless selection pressure on bacteria that has driven the evolution of multidrug resistance in many pathogens, and novel classes of antibiotics are needed.[1,2] Bacteria produce a large assortment of antibiotics, possibly to gain advantage over competing microorganisms for limited resources,[3-9] and these compounds have proven to be the richest source of antimicrobials for development into therapeutics. While most, if not all, these natural products are produced as families of related compounds, the significance of this diversity is debated.[10-14] The arylomycins, first isolated in 2002 from a strain of Streptomyces, consist of three related series of compounds, each of which has a conserved C-terminal tripeptide macrocycle attached to an N-terminal lipopeptide (Figure 1).[15-17] The macrocycle of the A series compounds is unmodified, while those of the B series and lipoglycopeptides are nitrated and glycosylated (and in some cases hydroxylated), respectively.[15,16]

The arylomycins inhibit type I signal peptidase (SPase, EC 3.4.21.89),[16,18,19] which is an essential membrane-bound serine endopeptidase with a highly conserved active site that is required to remove the amino-terminal leader (signal) sequence during, or shortly after, protein translocation across the cytoplasmic membrane. SPase acts via a unique Ser-Lys catalytic dyad with an unusual nucleophilic attack on the *si*-face of the substrate, as opposed to the *re*-face attack characteristic of the more common Ser-His-Asp catalytic triad serine proteases.[20] Moreover, its position on the outer surface of the cytoplasmic membrane should make it relatively accessible to inhibitors, although penetration of the outer membrane could limit accessibility in Gram-negative bacteria. While their novel mechanism of action originally generated much enthusiasm, excitement for developing the arylomycins waned when it was concluded that their activity was limited to only a few Gram-positive bacteria.[16,17] However, after reporting the first synthesis of an arylomycin, the A series member arylomycin A₂ (Figure 1), as well as several derivatives,[21] we demonstrated that they actually have a remarkably broad spectrum of activity,[19] including potent activity against both Gram-positive and Gram-negative bacteria, but that activity is limited in some cases by one of at least two resistance mechanisms: target mutation, specifically, the presence of a proline residue in SPase;[19] or a second as yet undefined mechanism that

Arylomycin A-C$_{16}$:   R$_1$ = H, R$_2$ = H, R$_3$ = H
                     R$_4$ = n-C$_{15}$

Arylomycin B-C$_{16}$:   R$_1$ = NO$_2$, R$_2$ = H, R$_3$ = H
                     R$_4$ = n-C$_{15}$

Lipoglycopeptides: R$_1$ = H, R$_2$ = deoxy-α-mannose, R$_3$ = H or OH
                     R$_4$ = iso- or n-C$_{13-15}$

Arylomycin C-C$_{16}$:   R$_1$ = H, R$_2$ = deoxy-α-mannose, R$_3$ = H
                     R$_4$ = n-C$_{15}$

BAL4850C: R$_1$ = H, R$_2$ = deoxy-α-mannose, R$_3$ = OH
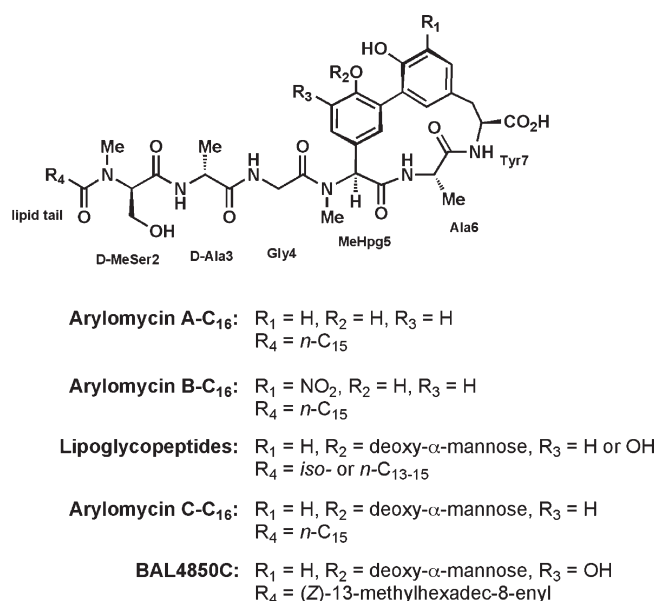   R$_4$ = (Z)-13-methylhexadec-8-enyl

Figure 1. Structure of the arylomycin derivatives characterized in this study. Arylomycin A-C$_{16}$ and arylomycin B-C$_{16}$ correspond to A and B series arylomycins, respectively, while arylomycin C-C$_{16}$ and BAL4850C correspond to the lipoglycopeptide series of arylomycins (see text for details).

confers *Streptococcus agalactiae* (*S. agalactiae*) with resistance to the A series derivatives.[22] Moreover, following the synthesis of a B series arylomycin, arylomycin B-C$_{16}$, we demonstrated that the nitro group does not negatively impact activity against bacteria that are sensitive to arylomycin C$_{16}$, and importantly, that it overcomes the resistance of *S. agalactiae* and imparts the scaffold with a reasonably potent minimum inhibitory concentration (MIC) of 8 μg/mL.[22]

Like the lipoglycopeptide arylomycins, many antibiotics are glycosylated, and in some cases the sugar substituents are required for activity.[23,24] From a medicinal chemistry perspective, glycosylation can also impact an antibiotic's potential for development as a therapeutic by affecting its pharmacokinetic properties, including absorption, distribution, metabolism, and excretion, at least in part due to changes in solubility and serum binding.[25] The most common sugar substituents are 6-deoxysugars, of which more than a hundred have been identified among different secondary metabolites,[26] and indeed the sugar substituent of the lipoglycopeptide arylomycins was identified as deoxy-α-mannose,[16] although its absolute stereochemistry was not determined.

The lipoglycopeptide arylomycins have been shown to have moderate activity against several bacteria, inhibiting a strain of *Streptococcus pneumoniae* (*S. pneumoniae*) with MICs ranging from 8 to >64 μM; a strain of *Staphylococcus aureus* (*S. aureus*), with MICs ranging from 32 to >64 μM; and a strain of *Haemophilus influenzae* (*H. influenzae*), with an MIC of 64 μM.[16] While they are not active against intact *Escherichia coli* (*E. coli*), they were shown to have activity against permeabilized mutant strains, leading to the suggestion that their development as therapeutics would require the optimization of outer membrane penetration.[16] However, the potential role of the resistance conferring proline has not been examined.

E. coli SPase is 324 amino acids in length (molecular weight, 35 960 Da; pI 6.9)[27] and contains two amino-terminal transmembrane segments (residues 4–28 and 59–77), one small cytoplasmic

region (residues 29–58), and a large carboxyl-terminal periplasmic catalytic domain (residues 78–324).[28,29] Proteinase K digestion,[29,30] gene-fusion,[31] and disulfide cross-linking studies[32,33] are all consistent with both the N- and C-termini of *E. coli* SPase facing the periplasmic space. The catalytically active periplasmic domain of *E. coli* SPase (SPase Δ2-76) has a molecular weight of 27 952 Da[34] and pI of 5.6.[35] It has been subcloned, purified,[34] characterized,[35] and crystallized.[36] To date, four crystal structures of *E. coli* SPase have been reported (all with the Δ2-76 enzyme), including the unbound protein,[37] a binary complex with a β-lactam inhibitor,[20] a binary complex with an A family arylomycin (arylomycin A$_2$),[18] and a ternary complex with arylomycin A$_2$ and β-sultam.[38] While these structures have helped elucidate the mechanisms of the molecular recognition underlying the inhibition of SPase by the arylomycins, the effects of macrocycle glycosylation remained unclear.

We now report the first synthesis of an arylomycin lipoglycopeptide and its biological characterization, as well as the structural analysis of the binary complex with a related glycosylated and hydroxylated derivative. Total synthesis allowed us to assign the absolute stereochemistry of the deoxy-α-mannose substituent and to determine that the spectrum of activity of the glycosylated derivative is limited by the same mechanisms of resistance as are the A series compounds. The structural analysis revealed that the inhibitor binds in a fashion similar to that previously reported for the A series derivative and that the sugar is oriented away from the active site and into the aqueous environment. Consistent with these structural studies, in addition to finding that glycosylation does not interfere with SPase binding or activity, we find that it appears to increase aqueous solubility and reduce protein binding on the basis of MIC values in the presence of serum proteins. In all, the data reveal that contrary to previous conclusions,[16] glycosylation does not interfere with the antibacterial activity of the arylomycins, including activity against Gram-negative bacteria, and that similar types of modifications might be used to optimize the pharmacokinetics of this promising scaffold.

■ RESULTS AND DISCUSSION

Lipoglycopeptide Synthesis. To synthesize the lipoglycopeptide arylomycins, we modified our previously reported syntheses of the arylomycin A[21] and B[22] series compounds to increase flexibility and material throughput. Because the absolute stereochemistry of the sugar was not known, we targeted the synthesis of both the deoxy-α-L-mannose and the deoxy-α-D-mannose variants. The required lipopeptide **2** was prepared using the procedure of Zhu and co-workers (Scheme 1).[39,40] Briefly, the previously reported tripeptide **1** was acylated with isopalmitic acyl chloride, generated in situ from isopalmitic acid and oxalyl chloride in dichloromethane (DCM), yielding the fully protected fatty tail, which was hydrolyzed under Nicolaou's conditions (Me$_3$SnOH/ 1,2-dichloroethane (DCE)) to afford the lipopeptide **2** in 51% yield.

Macrocycle synthesis commenced with the preparation of the hydroxyphenylglycine-alanine iododipeptide **8** from commercially available 4-hydroxyphenylglycine **3** (Scheme 2). Briefly, *t*-butoxycarbonyl (Boc) protection of **3**, followed by monoiodization,[41] afforded **4** in 80% yield, which was subsequently converted into oxazolidinone **5**. After reduction by triethylsilane in trifluoroacetic acid (TFA), followed by reinstallation of the Boc group, the desired *N*-methylamino acid **6** was obtained without racemization. Acid **6** was then coupled to L-Ala-OMe to afford dipeptide ester **7**. Finally, hydrolysis of the methyl ester using Me$_3$SnOH

Scheme 1
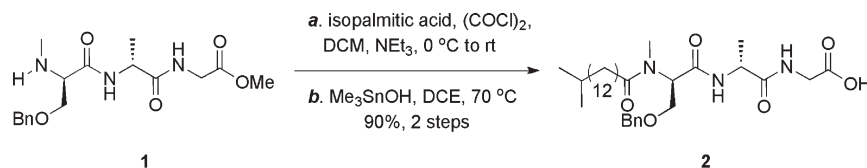
Scheme 2
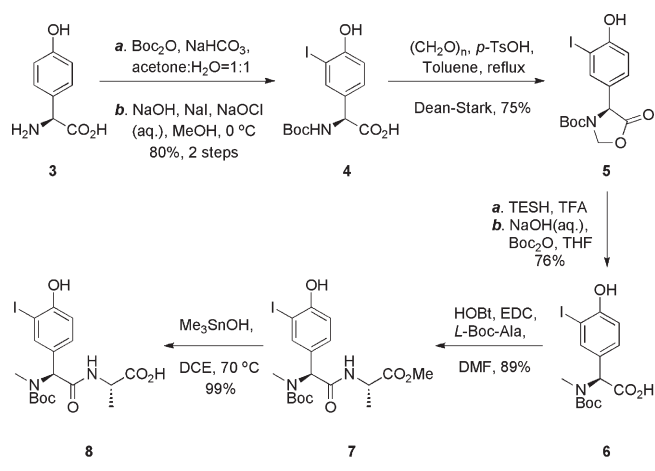
provided iododipeptide **8** in quantitative yield. This sequence produced **8** in seven steps with 40% overall yield, which required only one column purification for product **7**.

The protected tyrosine pinacol boronic ester **11** was prepared from iodotyrosine in four steps (Scheme 3). Briefly, triply protected iodotyrosine **10** was prepared from commercially available iodotyrosine **9** in greater than 90% yield with only one column purification. The boronic ester of compound **11** was installed via a Miyama reaction,[42] and then deprotection, followed by HOBt/EDC-mediated coupling to dipeptide **8** provided tripeptide **12**. To optimize macrocyclization of **12**, we first employed conditions that we found to be optimal for the cyclization of the bis-methyl phenol protected A or B series macrocycle cores ($PdCl_2(dppf)$/$NaHCO_3$, dimethylformamide (DMF)). However, in this case the desired product was obtained in less than 25% yield, suggesting that while the free phenol may minimize the epimerization of **12**,[39] it also appears to adversely affect the reaction. Thus, we reoptimized the cyclization conditions (Table 1) and found that $Pd(tBu_3P)_2$ and K_2CO_3 provided **13** in a satisfactory 48% yield.

The only remaining challenge was to glycosylate the free phenol of the macrocycle without O- to C-glycoside rearrangement under the required Lewis acidic reaction conditions (Table 2).[43] Due to the low nucleophilicity of the phenol and the steric hindrance of the neighboring O-benzyl group, we first attempted to glycosylate **13** using glycosyl bromides **15** in the presence of AgOTf and 4 Å molecular sieves in DCM, but no glycosylated product was detected. We then investigated trichloroacetimidate **16a** as a glycosyl donor. TMSOTf-promoted glycosylation, either in catalytic or superstoichiometric quantities, yielded the desired macrocycle in less than 40% yield. However, 10 equiv of BF_3−Et_2O[44] resulted in simultaneous glycosylation and Boc deprotection and yielded the desired glycosylated macrocycle

14a in 76% yield. Similarly, **14b** was obtained in 70% yield using the same conditions.

With lipopeptide tail **2** and both glycosylated macrocycles **14a** and **14b** in hand, it only remained to couple and deprotect the two halves of the molecule. Compound **2** was coupled to **14a** or **14b** to provide the fully protected natural products **17a** and **17b** (Scheme 4). Global deprotection was carried out in three mild reactions to avoid the possible elimination of the D-Ser hydroxyl group. Pd/C catalyzed hydrogenation, NaOMe induced deacetylation, and hydrolysis with Me_3SnOH provided the candidate natural products **18a** and **18b** in greater than 50% yield. Comparison of the ^{1}H NMR spectra of the two candidate natural products with that of the authentic natural product (kindly provided by Dr. Sheng-Bing Peng, Eli Lilly) revealed that while almost all of the proton resonances of the **18a** spectrum are nearly identical to the natural product, those of the sugar are significantly different (Figure S1 of the Supporting Information). In contrast, the ^{1}H spectra of **18b** and the natural product are virtually identical, as are the ^{13}C spectra. Thus, we conclude that the lipopeptide arylomycins are glycosylated with deoxy-α-L-mannose. Furthermore, similarities in the NMR spectra in the region corresponding to the sugar of the different members of the lipoglycopeptides,[16] suggest that they are all glycosylated with deoxy-α-L-mannose.

While the arylomycins, including the lipoglycopeptides, are naturally lipidated with different fatty acids ranging in length from 12 to 16 carbons, our analysis of the A and B series compounds was performed with a straight chain C_{16} tail.[21] Thus, for systematic comparison of biological activity, we used the above protocol but with the straight chain C_{16} lipid, to synthesize the corresponding lipoglycopeptide derivative. As expected, this synthesis proceeded with indistinguishable yields. For simplicity, we refer to these derivatives as arylomycin A-C_{16}, arylomycin B-C_{16}, and arylomycin C-C_{16}, corresponding to the A, B, and lipoglycopeptide compounds, respectively (Figure 1).

Structural Analysis of a Lipoglycopeptide Arylomycin Bound to SPase. Structural studies focused on analysis of the glycosylated and hydroxylated lipoglycopeptide antibiotic BAL4850C and SPaseΔ2−76. BAL4850C contains the same sugar that is characteristic of the lipoglycopeptide class of arylomycins, but is differentiated from arylomycin C-C_{16} by macrocyclic hydroxylation and lipidation by an unsaturated C_{16} fatty acid (Figure 1). All attempts to soak BAL4850C into preformed crystals of SPaseΔ2-76 were unsuccessful, so we instead developed a cocrystallization method which yielded well-ordered crystals that diffracted to beyond 2.4 Å resolution (Table 3). The initial $F_o − F_c$ difference map revealed a well-defined rod-shape density corresponding to the N-terminal tripeptide, as well as a ring-shape electron density corresponding to the three residue macrocycle core of the inhibitor with clear electron density for the mannose substituent (Figure 2A). The modeled L-stereochemistry for the mannose substituent is consistent with the electron density, but the resolution
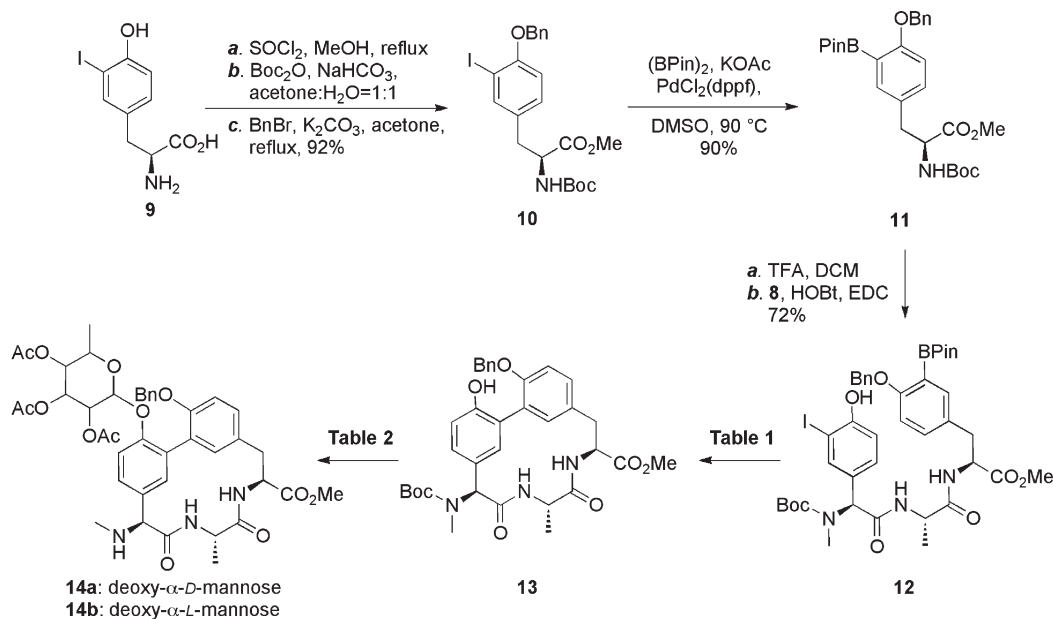
**Scheme 3**

Table 1. Cyclization Conditions

entry[a]	catalyst	base	yield, %
1	PdCl$_2$(dppf)	K$_2$CO$_3$	25[c]
2	Pd(0)(Pt-Bu$_3$)$_2$	K$_2$CO$_3$	48[c]
3	PEPPSI^{TM}-IPr[b]	K$_2$CO$_3$	<10[d]
4	PdCl$_2$(PPh$_3$)$_2$	K$_2$CO$_3$	<10[d]
5	Pd(OAc)$_2$(SPhos)$_2$	K$_2$CO$_3$	<10[d]
6	Pd$_2$(dba)$_3$+SPhos	K$_2$CO$_3$	<10[d]
7	Pd(PPh$_3$)$_4$	K$_2$CO$_3$	<10[d]
8	Pd(0)(Pt-Bu$_3$)$_2$	Cs$_2$CO$_3$	<10[d]
9	Pd(0)(Pt-Bu$_3$)$_2$	CsF	<10[d]
10	Pd(0)(Pt-Bu$_3$)$_2$	NaOH	<10[d]
11	Pd(0)(Pt-Bu$_3$)$_2$	K$_3$PO$_4$	<10[d]

[a] All the reactions were carried out in dimethyl sulfoxide (DMSO) at 90 °C for 24 h. [b] PEPPSI-IPr: (1,3-diisopropylimidazol-2-ylidene)-(3-chloropyridyl)palladium(II) dichloride.[58] [c] Isolated yield. [d] Based on LC-MS.

Table 2. Glycosylation Conditions

entry[a]	glycosyl donor	lewis acid	product[b]	yield[c]
1	(15)	AgOTf (4 eq)	NR[d]	0
2	(16a)	TMSOTf (0.5 eq)	Mixture	ND[e]
3	16a	TMSOTf (2.2 eq)	Mixture	40%
4	16a	BF$_3$-Et$_2$O (10 eq)	14a	76%
5	(16b)	BF$_3$-Et$_2$O (10 eq)	14b	70%

[a] All reactions were carried out in anhydrous DCM with 4 Å molecular sieves; details are provided in the Supporting Information. [b] Product ratios determined by LC-MS. [c] Yields were isolated yields. [d] No product detected. [e] Not determined.

is not high enough to independently confirm the assigned stereo-chemistry. The electron density of the C$_{16}$ fatty acid tail of the lipoglycopeptide is weak at 1.0 σ, suggesting the fatty acid tail is disordered. Similarly, no density was observed for the fatty acid tail in the reported complex with arylomycin A$_2$.[18]

The refined structure revealed that the arylomycin is positioned within the SPase binding site with its C-terminal macrocycle oriented toward the catalytic residues and both the peptide backbone and side chains tightly packed within the substrate binding groove (Figures 2 and 3). Similar to the complex of SPaseΔ2-76 and arylomycin A$_2$,[18] one C-terminal carboxylate oxygen (O45) interacts with three catalytically essential residues of the enzyme, including the nucleophile Ser91 Oγ, the general base Lys146 Nζ, and a component of the oxyanion hole, Ser89 Oγ. (Note that the numbering system used with previous SPase structures[18,20,37,38] is different by one residue due to an error in the originally reported sequence of the *E. coli* protein.[30] The

sequencing error occurred in the cytoplasmic region, between the two N-terminal transmembrane segments, which is not present in SPaseΔ2-76, and therefore does not affect the register of the residues within the crystal structures, only the residue numbering system. The numbering system used in the currently reported structure matches that in the Swiss-Prot sequence database (accession number, P00803).) The other C-terminal carboxylate oxygen (O44) is hydrogen-bonded to Ile145 N and the general base Lys146 Nζ. The peptide backbone of the arylomycin forms 11 direct hydrogen bonds with SPase, forming parallel β-strand interactions with the β-sheets that make up the substrate binding groove (SPase residues 143−146 and 82−86), seven of which are mediated by the macrocycle tripeptide and four by the N-terminal tripeptide. Macrocycle residues N33 and N28 form hydrogen bonds with Asp143 O and Gln86 O,
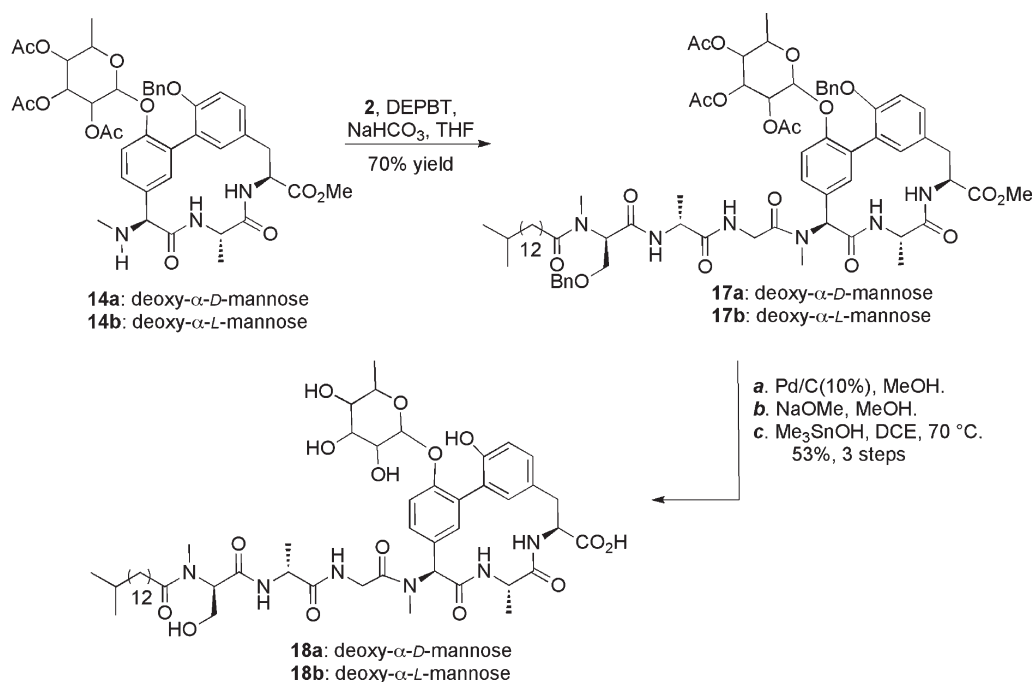
Scheme 4

14a: deoxy-α-D-mannose
14b: deoxy-α-L-mannose

2, DEPBT,
NaHCO₃, THF
70% yield

17a: deoxy-α-D-mannose
17b: deoxy-α-L-mannose

a. Pd/C(10%), MeOH.
b. NaOMe, MeOH.
c. Me₃SnOH, DCE, 70 °C.
53%, 3 steps

18a: deoxy-α-D-mannose
18b: deoxy-α-L-mannose

Table 3. Data Collection and Refinement Statistics

crystal parameters	
space group	$P4_3 2_1 2$
a, b, c (Å)	72.0, 72.0, 262.6
data collection statistics	
wavelength (Å)	1.5418
resolution (Å)	$32.2 - 2.4 (2.5 - 2.4)^a$
total reflections	247952 (28360)
unique reflections	26680 (2605)
$R_{merge}{}^b$	0.105 (0.44)
mean $I/\sigma(I)$	10.4 (5.1)
completeness (%)	99.6 (100.0)
redundancy	9.3 (10.9)
refinement statistics	
protein molecules (chains) in A.U.	2
residues	432
inhibitors	2
water molecules	57
total number of atoms	3617
$R_{cryst}{}^c/R_{free}{}^d$ (%)	24.5/26.5
average B-factor (Å²) (all atoms)	56.5
rmsd on angles (deg)	1.098
rmsd on bonds (Å)	0.007

^{a} The data collection statistics in brackets are the values for the highest resolution shell. $^b R_{merge} = \Sigma_{hkl}\Sigma_i |I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an individual reflection and $\langle I(hkl)\rangle$ is the mean intensity of that reflection. $^c R_{cryst} = \Sigma_{hkl} ||F_{obs}| - |F_{calc}||/\Sigma_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively. $^d R_{free}$ is calculated using 5% of the reflections randomly excluded from refinement.

respectively, and there is a conserved water molecule (water 14 in chain A and water 15 in chain B) within the hydrogen-bonding
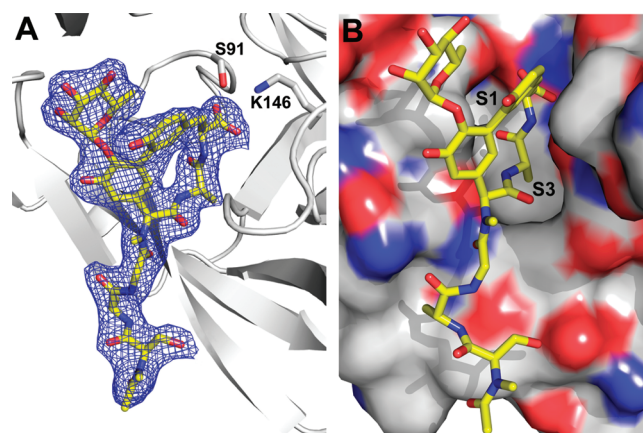

Figure 2. View of lipoglycopeptide BAL4850C within the active site and substrate-binding groove of SPase. Due to weak electron density, the fatty acid tail is not included. (A) Electron density for lipoglycopeptide BAL4850C bound within the active site and substrate-binding groove of SPase. A cross-validated 2F₀−Fᶜ electron density map contoured at 1σ (blue) is shown with the lipoglycopeptide shown in stick representation and colored by element (carbon, yellow; oxygen, red; nitrogen, blue). The SPase protein is shown in cartoon representation and colored in gray. The catalytic residues (Ser91 and Lys146) are labeled and colored by element (carbon, gray; oxygen, red; nitrogen, blue). (B) Molecular surface of SPase with basic residues in blue, acidic residues in red, and all others in gray. The bound lipoglycopeptide is shown and colored as in panel A. The S1 and S3 binding sites of SPase are labeled. Chain B within the asymmetric unit was used to make the figure.

distance of O45. Finally, the C30 methyl group of Ala6 is directed toward the S3 substrate-specificity pocket and is in van der Waals contact with the side chain of SPase residue Ile145. Clearly, the macrocycle provides the majority of the interactions involved in SPase recognition. Interestingly, although the deoxy-α-L-mannose
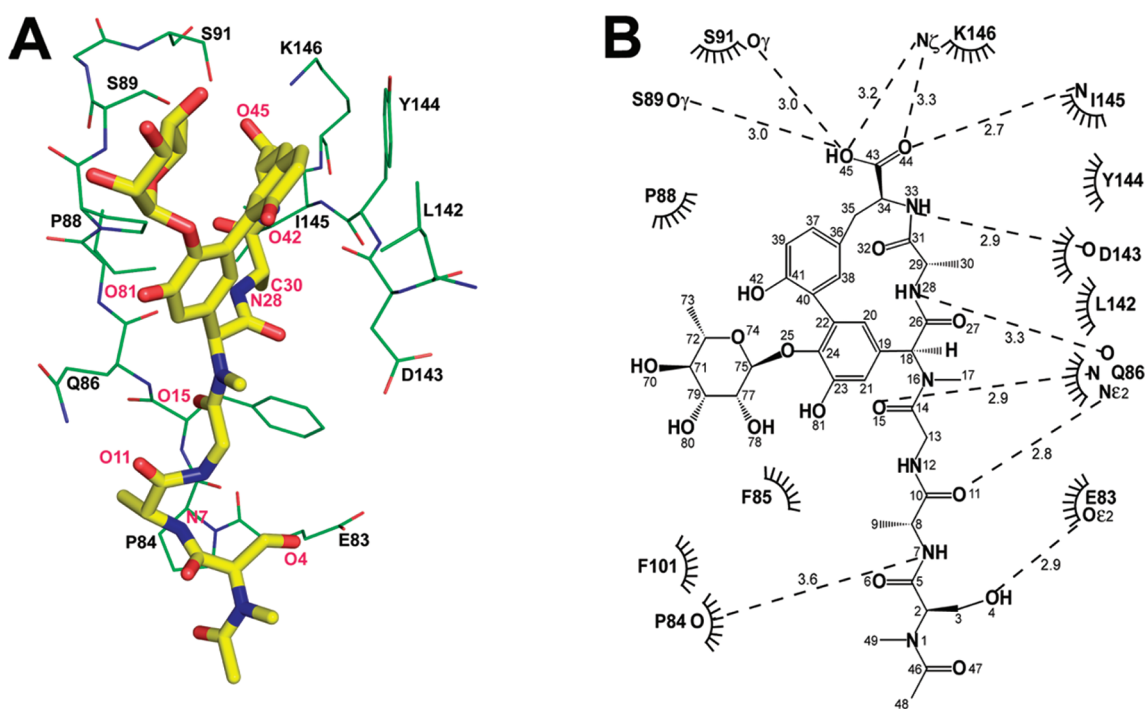
Figure 3. Interactions between lipoglycopeptide BAL4850C and SPase. Due to weak electron density, the fatty acid tail is not included. (A) SPase is shown in line representation and colored by element (carbon, green; oxygen, red; nitrogen, blue; residues labeled in black). The lipoglycopeptide is shown in stick and colored by element (carbon, yellow; nitrogen, blue; oxygen, red; atoms labeled in red). Chain B within the asymmetric unit was used to make the figure. (B) Schematic representation of lipoglycopeptide BAL4850C interactions with the active site and substrate binding groove of SPase. Dashed lines indicate hydrogen bonds and distances are in angstroms (average value for the two molecules in the asymmetric unit). The residues involved in van der Waals interactions are symbolized by arches.

is in van der Waals contact with SPase residue Pro88, it is predominantly solvent exposed (Figures 2 and 3).

A superposition of the structures of the three SPase—arylomycin complexes solved to date reveals that very little adjustment within the protein is needed to accommodate the sugar (Figure 4). Moreover, the presence of the sugar does not appear to significantly alter the structure of the macrocycle, which the superposition reveals is similar, and engages the protein in a similar way, in each complex. However, the superposition also reveals that the binding mode for the three N-terminal residues (D-MeSer2-D-Ala3-Gly4) of the inhibitor is more variable, particularly at D-Ala3. These observations are consistent with the macrocycle mediating the majority of interactions between the inhibitor and the enzyme and suggest that the N-terminal peptidic tail may be more flexible.

Biological Activity. The antibacterial activity of the lipoglycopeptide arylomycin C-C_{16} was characterized by determining the MIC required to inhibit the growth of several Gram-positive bacteria (Table 4). Against *Staphylococcus epidermidis* (*S. epidermidis*), *Streptococcus pyogenes* (*S. pyogenes*), *S. pneumoniae*, *Corynebacterium glutamicum* (*C. glutamicum*), and *Rhodococcus opacus* (*R. opacus*) arylomycin C-C_{16} has activity that is indistinguishable from the analogous A series compound. Moreover, like arylomycin A-C_{16}, but unlike arylomycin B-C_{16}, arylomycin C-C_{16} has no activity against *Streptococcus agalactiae* (*S. agalactiae*) strain COH1, demonstrating that the ability of the nitro substituent to impart the B series scaffold with activity against this pathogen is unique.

We next determined the activity of arylomycin C-C_{16} against *Staphylococcus aureus* (*S. aureus*), *E. coli*, and *Pseudomonas aeruginosa* (*P. aeruginosa*). As with both the A and B series compounds, arylomycin C-C_{16} has no activity against these pathogens.

Also as with the A and B series compounds, arylomycin C-C_{16} does have significant activity against the mutant pathogens where the arylomycin resistance-conferring proline residue is mutated to a residue that does not confer resistance (P29S in the *S. aureus* protein, and P84L in the *E. coli* and *P. aeruginosa* proteins)[19] (Table 5). Moreover, the addition of polymyxin B nonapeptide, which permeabilizes the outer membrane of Gram-negative bacteria, did not significantly affect the MICs. Thus, while the spectrum of arylomycin C-C_{16} is limited by the same resistance mechanisms as is the A series compound, in contrast to previous conclusions,[16] the presence of the sugar does not impair the inhibitor's ability to penetrate the outer membrane of Gram-negative bacteria.

Although not unprecedented among therapeutics,[45−48] from a drug development perspective, the fatty acid tails of the arylomycins might prove to be a liability, for example due to decreased solubility and increased serum binding. Thus, to begin to examine the effects of glycosylation on aqueous solubility and serum binding, we re-determined the MICs of arylomycin A-C_{16} and arylomycin C-C_{16} against wild type *S. epidermidis* and sensitized *E. coli* in the presence of pooled human serum (25−100%) or bovine serum albumin (4−10%) in MHIIB. While the value of MICs varied under the different conditions tested, the MIC observed for arylomycin C-C_{16} was consistently 2- to 4-fold lower than that of arylomycin A-C_{16} under these conditions. These small but reproducible effects are consistent with glycosylation increasing the concentration of the free inhibitor available for SPase binding.

■ CONCLUSIONS

Our previous demonstration that the arylomycin class of antibiotics has a broader spectrum of antibacterial activity than
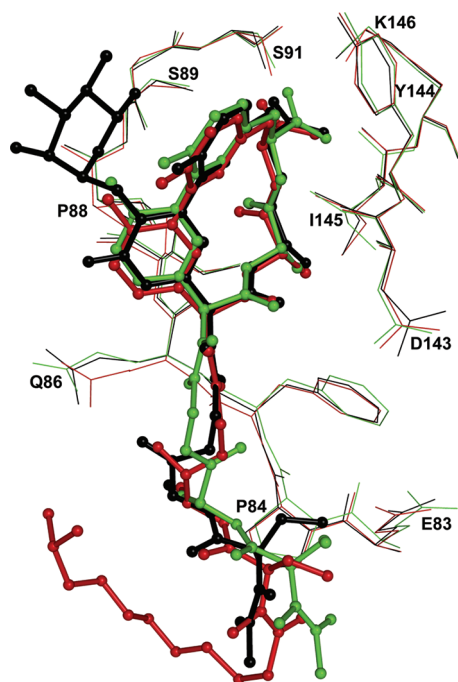
Figure 4. Superposition of the active sites of SPase−arylomycin complexes. The SPase active site residues are shown as thin lines and labeled. The arylomycin molecules are shown in ball and stick representation. The lipoglycopeptide BAL4850C complex (PDB ID, 3S04) is black, the ternary complex with arylomycin A$_2$ and a β-sultam (PDB ID, 3IIQ)[38] is red, and the arylomycin A$_2$ complex (PDB ID, 1T7D)[18] is green. Due to weak electron density, the fatty acid tail is not included in the BAL4850C and arylomycin A$_2$ complexes.

previously appreciated, and that where resistance does exist, it results from the presence of a specific proline mutation in the target SPase protein,[19] makes the arylomycin scaffold promising for development as a therapeutic. Toward the further exploration of this class of natural product antibiotics, we found that the lipoglycopeptide variants can be synthesized in reasonable yield, with the key steps being a Pd(tBu$_3$P)$_2$/K$_2$CO$_3$-mediated macrocycle cyclization and a BF$_3$−Et$_2$O-mediated trichloroacetimidate glycosylation. Characterization of the synthetic product allowed us to unambiguously determine that the natural products are glycosylated with deoxy-α-L-mannose, and contrary to previous reports,[16] we found that glycosylation does not appear to interfere significantly with activity. The structural analysis revealed that the lipoglycopeptides bind SPase in a manner analogous to the arylomycin A series compounds with the sugar moiety oriented away from the enzyme active site and largely solvent exposed. Nonetheless, the structural analysis also revealed that the hydrophobic portion of the sugar interacts with active site residues and that glycosylation does affect the interactions between the peptidic position of the inhibitors tail and SPase. Thus, it remains possible that glycosylation affects activity against bacteria that were not examined in the current study. While the selection pressure, if any, that favors glycosylation of the arylomycin scaffold in nature remains unclear, it does appear that glycosylation increases the solubility of the scaffold, an important pharmacokinetic attribute for any candidate therapeutic. Thus, derivatization at the same position with other substituents, e.g., other sugars, phosphates, and sulfates, etc., might further improve the pharmacokinetic properties of the arylomycin scaffold and aid in its potential development as a therapeutic. Experiments to test this hypothesis are currently in progress.

Table 4. MICs of Arylomycin A- and C-C$_{16}$ (μg/mL) [a]

strain	arylomycin A-C$_{16}$	arylomycin C−C$_{16}$
S. epidermidis RP62A	0.25	0.25
S. epidermidis PAS9001 [b]	8	8
S. agalactiae COH-1	>128	>128
S. pyogenes MGAS-1	8	16
S. pneumoniae R800	8	8
C. glutamicum DSM 44475	4	4
R. opacus DSM 1069	2	1

[a] Arylomycin A-C$_{16}$ included for comparison.[19] [b] Evolved from *S. epidermidis* RP62A to be resistant to arylomycin A-C$_{16}$.[19]

Table 5. MICs of Arylomycin C-C$_{16}$ (μg/mL) against Mutants Known To Be Sensitive to Arylomycin A- and B-C$_{16}$

strain	arylomycin A-C$_{16}$	arylomycin C−C$_{16}$
S. aureus NCTC 8325	>128	>128
S. aureus P29S PAS8001 [b]	2	4
E. coli MG1655	>128	>128
E. coli P84L PAS0260 [b]	2	2
P. aeruginosa PAO1	>128	>128
P. aeruginosa P84L PAS2008 [b]	8	16

[a] Arylomycin A-C$_{16}$ included for comparison.[19] [b] Constructed from their respective wild type strains by mutation of the resistance-conferring proline to either serine (*S. aureus*) or leucine (*E. coli* and *P. aeruginosa*).[18]

■ EXPERIMENTAL SECTION

General Experimental Procedures. Dry solvents were purchased from Acros. Commercially available amino acids were purchased from Bachem (Torrence, CA, USA), Chem-Impex (Wood Dale, IL, USA), or Novabiochem (EMD Chemicals, Gibbstown, NJ, USA). Celite 545 filter aid (not acid-washed) was purchased from Fisher. Anhydrous 1-hydroxybenzotriazole (HOBt) was purchased from Chem-Impex. All other chemicals were purchased from Fisher/Acros or Aldrich. Reactions were magnetically stirred and monitored by thin layer chromatography (TLC) with 0.25 mm Whatman precoated silica gel (with fluorescence indicator) plates. Flash chromatography was performed with silica gel (particle size, 40−63 μm; EMD Chemicals). ^{1}H and ^{13}C NMR spectra were recorded on Bruker DRX 500 or Bruker DRX 600 spectrometers. Chemical shifts are reported relative to either chloroform (δ 7.26) or methanol (δ 3.31) for ^{1}H NMR, and either chloroform (δ 77.16) or methanol (δ 49.00) for ^{13}C NMR. High-resolution time-of-flight mass spectra were measured at the Scripps Center for Mass Spectrometry. Electrospray ionization (ESI) mass spectra were measured on either an HP Series 1100 MSD or a PESCIEX API/Plus. For all compounds exhibiting atropisomerism or isolated as semipure mixtures, NMR spectra are provided in the Supporting Information. Yields refer to chromatographically and spectroscopically pure compounds unless otherwise stated.

All preparative reverse-phase chromatography was performed using Dynamax SD-200 pumps connected to a Dynamax UV-D II detector (monitoring at 220 nm) on a Phenomenex Jupiter C$_{18}$ column (10 μm, 2.12 × 25 cm^2, 300 Å pore size). All solvents contained 0.1% TFA; solvent A, H$_2$O; and solvent B, 90/10 acetonitrile/H$_2$O. All samples were loaded onto the column at 0% B, and the column was allowed to equilibrate ∼10 min before a linear gradient was started. Retention times are reported according to the linear gradient used and the percent B at the time the sample eluted.

Production and Cocrystallization of SPase Δ2-76 and Lipoglycopeptide BAL4850C and Data Collection. SPase Δ2-76 was expressed and purified as described previously.[36] Prior to

cocrystallization, SPase Δ2-76 (18.0 mg/mL in 20 mM Tris-HCl, pH 7.4, and 0.5% Triton-100) was combined with the lipoglycopeptide BAL4850C (10 mM in dimethyl sulfoxide (DMSO)) in a 1:1 molar ratio and incubated on ice for 1 h.

Cocrystallization trials for SPase Δ2-76 in complex with BAL4850C (provided by Basilea Pharmaceutica International Ltd., Basel, Switzerland) were carried out by the sitting-drop vapor diffusion method. The final optimized reservoir condition that produced high-quality crystals for data collection was 22% (w/v) PEG 4000 and 0.2 M KCl. The drop consisted of 2 μL of protein/inhibitor complex described above, 2 μL of reservoir solution, and 2 μL of 0.025 M *n*-dodecyl-β-D-maltoside (DDM). This 6 μL drop was equilibrated over 1 mL of reservoir solution at 20 °C. The crystals formed from a light precipitate after approximately 2 weeks and had an average size of ~0.2 × 0.1 × 0.5 mm³.

Before data collection, the crystal was transferred by a pipet from the growth drop to a cryoprotectant composed of 24% (w/v) PEG 4000, 0.2 M KCl, 0.008 M DDM, 20% glycerol) for 30 s. The crystal was mounted on a Hampton Research loop and flash-cryo-cooled by directly placing it into a gaseous nitrogen stream at 100 K. The X-rays (wavelength, 1.5418 Å) were generated from Cu K$_\alpha$ radiation via a Rigaku MicroMax-007 Microfocus X-ray rotating-anode generator running at 40 kV and 20 mA and equipped with Osmic Confocal VariMax High Flux optics. The crystal-to-detector distance was set to 200 mm. All frames (280) were recorded on a R-AXIS IV++ imaging-plate detector with a 0.5° oscillation angle and an exposure time of 240 s/frame. The data revealed diffraction to beyond a resolution of 2.4 Å. Data were collected, indexed, and scaled using the program CrystalClear.[49] The crystals belong to the tetragonal space group $P4_32_12$. The unit cell dimensions were determined to be a = 72.0 Å, b = 72.0 Å, and c = 262.6 Å. The Matthews coefficient (V_m) is 3.03 Å³/Da for two molecules in the asymmetric unit. The fraction of the crystal volume occupied by solvent was 59.3%, calculated by the program *Matthews* in the CCP4i suite of programs.[50,51] For crystal and data collection statistics see Table 3.

Phasing, Model Building, and Refinement. A molecular replacement solution was found using the program Phaser in the CCP4i suite of programs.[51] The atomic coordinates used for the search model were taken from a 2.5 Å crystal structure of SPase Δ2-76 (PDB code, 1T7D; molecule A).[18] The topology and parameter files for the inhibitor were generated using the program PRODRG.[52] Coordinates for the inhibitor were manually docked into clear electron difference density $(F_o - F_c)$ near the active site. In addition, the main chain trace and the side chain assignments for the dynamic regions corresponding to residues Phe197-Asn201 and Asp305-Leu315 in chain B were built in manually. Water molecules were added to well-defined peaks (2.0σ and greater into the $F_o - F_c$ maps). Model building and analysis was performed with the program Coot.[53,54] Refinement of the structure was carried out using the program Refmac 5 in the CCP4i suite, as well as CNS.[51,55] The cycles of refinement were carried out for both protein model and inhibitor model using rigid body and restrained NCS refinement in the program Refmac 5, and simulated annealing, energy minimization, and B-factor refinement were performed in CNS. In addition, a cycle of TLS refinement was carried out using the TLS Motion Determination Sever and the restrained TLS refinement protocol of Refmac 5 within the CCP4i suite.[51] In all cycles of refinement, 5% of the reflections were set aside for cross-validation. Final refinement and analysis statistics of the complex are provided in Table 3. The stereochemistry of the structure model was analyzed with the program PROCHECK.[56] No stereochemical outliers were observed in the Ramachandran plot, with 96.3% of the residues in the preferred regions. An all atom superposition of the arylomycin complexes was performed with the program Pymol[57] using molecule B of the lipoglycopeptide arylomycin complex (PDB, 3S04), molecule A of the ternary complex with arylomycin A$_2$ and a β-sultam (PDB, 3IIQ),[38] and molecule A of the arylomycin A$_2$ complex structure (PDB, 1T7D).[18] Figures were prepared using the programs

ISIS Draw version 2.5 (MDL Information Systems, Inc.), and PyMol.[57] The atomic coordinates (accession code: 3S04) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ, USA (http://www.rcsb.org).

Determination of Antimicrobial Activity. Antimicrobial activity was examined using 13 bacterial strains, *S. epidermidis* RP62A, *S. aureus* NCTC 8325, *E. coli* MG1655, *P. aeruginosa* PAO1, *S. epidermidis* RP62A SpsIB(S29P) (PAS9001), *S. aureus* NCTC 8325 SpsB(P29S) (PAS8001),[19] *E. coli* MG1655 LepB(P84L) (PAS0260),[19] *P. aeruginosa* PAO1 LepB(P84L) (PAS2008),[19] *R. opacus* DSM 1069, *S. agalactiae* COH-1, *S. pyogenes* M1-5448, *S. pneumoniae* R800, and *C. glutamicum* ATCC 44475. Minimum inhibitory concentrations were determined from at least three independent experiments using the CLSI broth microdilution method. Briefly, inocula were prepared by suspending bacteria growing on solid media into the same type of broth used in the MIC experiment and diluting to a final concentration of 1×10^7 colony forming units/mL. A 5 μL aliquot of this suspension were added to the wells of a 96-well plate containing 100 μL of media with the appropriate concentrations of compound. The MICs of *E. coli*, *P. aeruginosa*, *S. aureus*, *S. epidermidis*, *Rhodococcus equi* (*R. equi*), *R. opacus*, and *C. glutamicum* were determined in cation-adjusted Mueller Hinton II broth. MICs of *S. pyogenes* and *S. pneumoniae* were determined in Todd Hewitt broth. The MICs of *S. agalactiae* were determined in cation-adjusted Mueller Hinton II broth and in Todd Hewitt broth (MIC values differed by at most 2-fold between these two media). In all cases MICs were defined as the lowest concentration of compound to inhibit visible growth.

ASSOCIATED CONTENT

Ⓢ **Supporting Information.** Full characterization of synthetic intermediates and compound **18a**, supporting data and NMR spectra, and complete ref 16 This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
mpaetzel@sfu.ca, floyd@scripps.edu

ACKNOWLEDGMENT

REFERENCES

(1) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. *Nat. Rev. Drug Discovery* **2007**, *6*, 29–40.

(2) Hancock, R. E. W. *Nat. Rev. Drug Discovery* **2007**, *6*, 28.

(3) Allen, H. K.; Donato, J.; Wang, H. H.; Cloud-Hansen, K. A.; Davies, J.; Handelsman, J. *Nat. Rev. Microbiol.* **2010**, *8*, 251–259.

(4) Baltz, R. H. *J. Ind. Microbiol. Biotechnol.* **2006**, *33*, 507–513.

(5) Czaran, T. L.; Hoekstra, R. F.; Pagie, L. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 786–790.

(6) D'Costa, V. M.; Griffiths, E.; Wright, G. D. *Curr. Opin. Microbiol.* **2007**, *10*, 481–489.

(7) Laskaris, P.; Tolba, S.; Calvo-Bado, L.; Wellington, L. *Environ. Microbiol.* **2010**, *12*, 783–796.

(8) Lynch, M. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (Suppl 1), 8597–8604.

(9) Martinez, J. L. *Proc. Biol. Sci.* **2009**, *276*, 2521–2530.

(10) Fischbach, M. A.; Clardy, J. *Nat. Chem. Biol.* **2007**, *3*, 353–355.

(11) Firn, R. D.; Jones, C. G. *Mol. Microbiol.* **2000**, *37*, 989–994.

(12) Firn, R. D.; Jones, C. G. *Nat. Prod. Rep.* **2003**, *20*, 382–391.

(13) Stone, M. J.; Williams, D. H. *Mol. Microbiol.* **1992**, *6*, 29–34.

(14) Williams, D. H.; Stone, M. J.; Hauck, P. R.; Rahman, S. K. *J. Nat. Prod.* **1989**, *52*, 1189–1208.

(15) Holtzel, A.; Schmid, D. G.; Nicholson, G. J.; Stevanovic, S.; Schimana, J.; Gebhardt, K.; Fiedler, H. P.; Jung, G. *J. Antibiot.* **2002**, *55*, 571–577.

(16) Kulanthaivel, P.; et al. *J. Biol. Chem.* **2004**, *279*, 36250–36258.

(17) Schimana, J.; Gebhardt, K.; Holtzel, A.; Schmid, D. G.; Sussmuth, R.; Muller, J.; Pukall, R.; Fiedler, H. P. *J. Antibiot.* **2002**, *55*, 565–570.

(18) Paetzel, M.; Goodall, J. J.; Kania, M.; Dalbey, R. E.; Page, M. G. P. *J. Biol. Chem.* **2004**, *279*, 30781–30790.

(19) Smith, P. A.; Roberts, T. C.; Romesberg, F. E. *Chem. Biol.* **2010**, *17*, 1223–1231.

(20) Paetzel, M.; Dalbey, R. E.; Strynadka, N. C. *Nature* **1998**, *396*, 186–190.

(21) Roberts, T. C.; Smith, P. A.; Cirz, R. T.; Romesberg, F. E. *J. Am. Chem. Soc.* **2007**, *129*, 15830–15838.

(22) Roberts, T. C.; Smith, P. A.; Romesberg, F. E. *J. Nat. Prod.* **2011**, *74*, 956–961.

(23) Weymouth-Wilson, A. C. *Nat. Prod. Rep.* **1997**, *14*, 99–110.

(24) Baltz, R. H. *Chem. Biol.* **2002**, *9*, 1268–1270.

(25) Cao, H.; Hwang, J.; Chen, X. In *Opportunity, Challenge and Scope of Natural Products in Medicinal Chemistry*; Tiwari, V. K., Mishra, B. B., Eds.; Research Signpost: Kerala, India, 2011; pp 411–431.

(26) Piepersberg, W.; Distler, J. In *Biotechnology: Products of Secondary Metabolism*; Rehm, H.-J., Reed, G., Eds.; Wiley-VCH Verlag GmbH: Weinheim, Germany, 1997; Vol. 7, pp 397–488.

(27) Barrett, A. J.; Rawlings, N. D. *Arch. Biochem. Biophys.* **1995**, *318*, 247–250.

(28) Bilgin, N.; Lee, J. I.; Zhu, H. Y.; Dalbey, R.; von Heijne, G. *EMBO J.* **1990**, *9*, 2717–2722.

(29) Moore, K. E.; Miura, S. *J. Biol. Chem.* **1987**, *262*, 8806–8813.

(30) Wolfe, P. B.; Wickner, W.; Goodman, J. M. *J. Biol. Chem.* **1983**, *258*, 12073–12080.

(31) San Millan, J. L.; Boyd, D.; Dalbey, R.; Wickner, W.; Beckwith, J. *J. Bacteriol.* **1989**, *171*, 5536–5541.

(32) Dalbey, R. E.; Wickner, W. *Science* **1987**, *235*, 783–787.

(33) Whitley, P.; Nilsson, L.; von Heijne, G. *Biochemistry* **1993**, *32*, 8534–8539.

(34) Kuo, D. W.; Chan, H. K.; Wilson, C. J.; Griffin, P. R.; Williams, H.; Knight, W. B. *Arch. Biochem. Biophys.* **1993**, *303*, 274–280.

(35) Tschantz, W. R.; Paetzel, M.; Cao, G.; Suciu, D.; Inouye, M.; Dalbey, R. E. *Biochemistry* **1995**, *34*, 3935–3941.

(36) Paetzel, M.; Chernaia, M.; Strynadka, N.; Tschantz, W.; Cao, G.; Dalbey, R. E.; James, M. N. *Proteins* **1995**, *23*, 122–125.

(37) Paetzel, M.; Dalbey, R. E.; Strynadka, N. C. *J. Biol. Chem.* **2002**, *277*, 9512–9519.

(38) Luo, C.; Roussel, P.; Dreier, J.; Page, M. G.; Paetzel, M. *Biochemistry* **2009**, *48*, 8976–8984.

(39) Dufour, J.; Neuville, L.; Zhu, J. *Chem.—Eur. J.* **2010**, *16*, 10523–10534.

(40) Dufour, J.; Neuville, L.; Zhu, J. P. *Synlett* **2008**, *15*, 2355–2359.

(41) Edgar, K. J.; Falling, S. N. *J. Org. Chem.* **1990**, *55*, 5287–5291.

(42) Ishiyama, T.; Murata, M.; Miyaura, N. *J. Org. Chem.* **1995**, *60*, 7508–7510.

(43) Li, Y.; Wei, G.; Yu, B. *Carbohydr. Res.* **2006**, *341*, 2717–2722, and references therein.

(44) Nicolaou, K. C.; Mitchell, H. J.; Jain, N. F.; Winssinger, W.; Hughes, R.; Bando, T. *Angew. Chem., Int. Ed.* **1999**, *38*, 240–244.

(45) Beauregard, D. A.; Williams, D. H.; Gwynn, M. N.; Knowles, D. J. *Antimicrob. Agents Chemother.* **1995**, *39*, 781–785.

(46) Breukink, E.; de Kruijff, B. *Nat. Rev. Drug Discovery* **2006**, *5*, 321–332.

(47) Kim, S. J.; Schaefer, J. *Biochemistry* **2008**, *47*, 10155–10161.

(48) Nagarajan, R. *Jpn. J. Antibiot.* **1993**, *46*, 1181–1195.

(49) Pflugrath, J. W. *Acta Crystallogr., D: Biol. Crystallogr.* **1999**, *55*, 1718–1725.

(50) Matthews, B. W. *J. Mol. Biol.* **1968**, *33*, 491–497.

(51) *Acta Crystallogr., D: Biol. Crystallogr.* **1994**, *50*, 760–763.

(52) van Aalten, D. M.; Bywater, R.; Findlay, J. B.; Hendlich, M.; Hooft, R. W.; Vriend, G. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 255–262.

(53) Emsley, P.; Cowtan, K. *Acta Crystallogr., D: Biol. Crystallogr.* **2004**, *60*, 2126–2132.

(54) Kleywegt, G. J. *Acta Crystallogr., D: Biol. Crystallogr.* **2007**, *63*, 94–100.

(55) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr., D: Biol. Crystallogr.* **1998**, *54*, 905–921.

(56) Laskowski, R. A.; McArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283–291.

(57) DeLano, W. L. *The PyMol Molecular Graphics System*; DeLano Scientific: San Carlos, CA, USA, 2002.

(58) *ChemFiles*; Sigma-Aldrich: St. Louis, MO, USA, 2006; Vol. 3, No. 6.