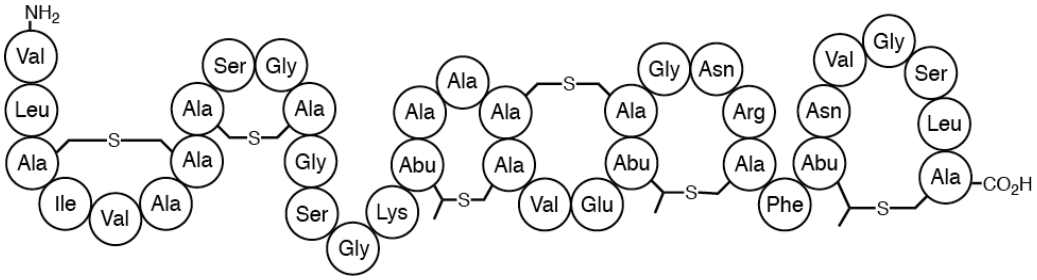
Supporting Information

Structural Characterization of the Highly Cyclized Lantibiotic Paenicidin A via a Partial Desulfurization/Reduction Strategy

Christopher T. Lohans,[†] Zedu Huang,[†] Marco J. van Belkum,[†] Maude Giroud,[†] Clarissa S. Sit,[†] Erika M. Steels,[†] Jing Zheng,[†] Randy M. Whittal,[†] Lynn M. McMullen,[‡] and John C. Vederas[*†]*

[†]Department of Chemistry, [‡]Department of Agricultural, Food & Nutritional Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G2

*To whom correspondence should be directed. Email: john.vederas@ualberta.ca



Supporting Information Contents:

Bacterial Strains and Culture Conditions

Paenibacillus polymyxa NRRL B-30509, obtained from the U.S. Department of Agriculture Agricultural Research Service Patent Culture Collection, was grown aerobically in BD BBL Brucella broth at 34 °C with shaking (200 rpm). The indicator organisms used are listed in the Supporting Information section entitled "Spectrum of Activity of Paenicidin A". All Gram-positive indicator organisms were grown in APT (All Purpose Tween; BD) broth, except for *Staphylococcus aureus*, which was grown in LB broth, Miller (Novagen). All Gram-positive indicator organisms were grown at room temperature without shaking, except for *S. aureus*, *Pediococcus acidilactici*, and *Bacillus subtilis*, which were grown at 30 °C without shaking. All Gram-negative indicators except *Campylobacter jejuni* were grown in LB broth at 37 °C with shaking (225 rpm). *C. jejuni* was streaked onto Mueller Hinton (MH) agar plates from glycerol stocks, and grown for two days at 37 °C under an atmosphere of 10 % CO_2, 5 % O_2, and the remainder N_2.

Sequencing of Genomic DNA

Genomic DNA from *P. polymyxa* NRRL B-30509 was isolated using the DNeasy Blood & Tissue Kit (QIAGEN) according to the manufacturer's instructions. 454 sequencing was performed at GenoSeq (UCLA Genotyping and Sequencing Core, Los Angeles, CA, USA) on a Roche GS FLX Titanium system. The sequencing reads were then assembled into contigs using the GS De Novo Assembler software (Roche). Artemis[1] was used to browse the sequencing data and to predict open reading frames.

Genetic Probing for SRCAM 602

Genomic DNA was probed for the presence of an SRCAM 602 structural gene by PCR using degenerate primers MVB212 (5'-GCNACNTAYTAYGGNAAYGG) and MVB213 (5'-TGYTGNACCCANCCRTTNAC). The amplification consisted of 30 cycles of denaturing at 94 °C for 1 min, annealing at 45 °C for 30 s, and extension at 68 °C for 15 s. HiFi Platinum Taq Polymerase (Invitrogen) was used according to the manufacturer's instructions. A known sequence from the genome (the paenicidin A structural gene, *paeA*) was amplified using degenerate primers MVB224 (5'-TTYGAYCTNGAYATHCARGT) and MVB225 (5'-CCNACRTTNGTRAARCANCG) as a positive control. The results are shown below in Figure S1.

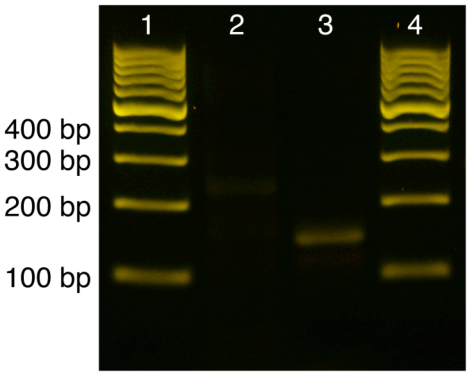


Figure S1. Agarose gel depicting the PCR-based genetic probing of *P. polymyxa* NRRL B-30509. Lanes 1 & 4: 100 base pair (bp) ladder. Lane 2: Probing for a SRCAM 602 structural gene (expected size: 116 bp). Lane 3: Probing for *paeA*, the paenicidin A structural gene (expected size: 149 bp). Despite using degenerate primers, no amplicon is observed of the expected size for a putative SRCAM 602 structural gene. As a control for this experiment, the genome of *P. polymyxa* NRRL B-30509 was probed (also using degenerate primers) for a sequence it was known to contain (the paenicidin A structural gene, *paeA*). This control yielded an amplicon of the expected size.

Antimicrobial Activity Assays

Antimicrobial activity was tested using spot-on-lawn assays. Molten soft agar (10 mL; 0.75 % agar supplemented with the appropriate media) was inoculated with 100 μL of a fully grown culture of the indicator organism, and poured onto an agar plate supplemented with the same medium. For the testing of activity against *C. jejuni*, the cells from a Mueller Hinton (MH) agar plate confluent with *C. jejuni* were suspended in MH broth, which was then added to molten soft agar and poured. Once the soft agar solidified, 10 μL of each sample was pipetted onto the plate and allowed to dry. The plate was incubated overnight at a temperature suitable for the indicator organism.

For the deferred inhibition studies, a BD BBL Brucella agar plate was stabbed with an overnight culture of *P. polymyxa* NRRL B-30509 and incubated overnight at 34 °C. Proteinase K (10 μL; 1.5 mg/mL) was spotted ~1 cm away from the stab locations, and allowed to dry. This plate was overlayered with 5 mL of molten soft agar inoculated with 50 μL of the indicator organism culture, allowed to set, and incubated overnight at a temperature appropriate for the indicator organism. A suspension of *C. jejuni* was used as described above for the inoculation of molten soft agar.

Isolation of Polymyxins E1 and E2

P. polymyxa NRRL B-30509 was cultured in 1 L BD BBL Brucella broth (0.1 % inoculum) at 34 °C with shaking (200 rpm) for 40 h. The culture was centrifuged (10,000 x g, 10 min, 4 °C) to remove the cells. Ammonium sulfate was added to the supernatant to 80 % saturation, and stirred at 4 °C for 16 h. The precipitate was collected by centrifugation (10,000 x g, 10 min, 4 °C), and redissolved in 100 mL of deionized water. This solution was loaded onto a 20 mL CM-Sephadex C-25 (GE Healthcare) column pre-equilibrated with 50 mM sodium phosphate, pH 7.8. The column was washed with 100 mL of 50 mM sodium phosphate, pH 7.8. The polymyxins were eluted with 100 mL of 50 mM sodium phosphate, pH 7.8, 500 mM NaCl. This eluent was loaded onto a 4 g column of Amberlite XAD16 (Sigma-Aldrich), washed with 100 mL of deionized water, and eluted with 100 mL of 80 % isopropyl alcohol (IPA) with 0.1 % trifluoroacetic acid (TFA). Following the removal of IPA and TFA *in vacuo*, the resulting solution was further fractionated via reversed phase HPLC. A C_{18} column (Vydac, 300 Å, 5 μm, 4.6 mm i.d. x 250 mm) was used with mobile phases A (0.1 % TFA) and B (acetonitrile with 0.1 % TFA), monitored at 220 nm, with a flow rate of 1 mL/min. Following 5 minutes at 20 % B, the level of B was increased to 45 % over the course of 40 minutes. Polymyxins E2 and E1 were collected at 18.0 min and 20.6 min respectively and purity was confirmed by MALDI-TOF MS. These peptides were identified via comparison of MS/MS data with commercially available polymyxins E1 and E2 (see Figure S2 below). Further, isolated polymyxin E1 co-eluted with commercial polymyxin E1 through reversed phase HPLC (data not shown).

MS/MS Analysis of Polymyxins E1 and E2

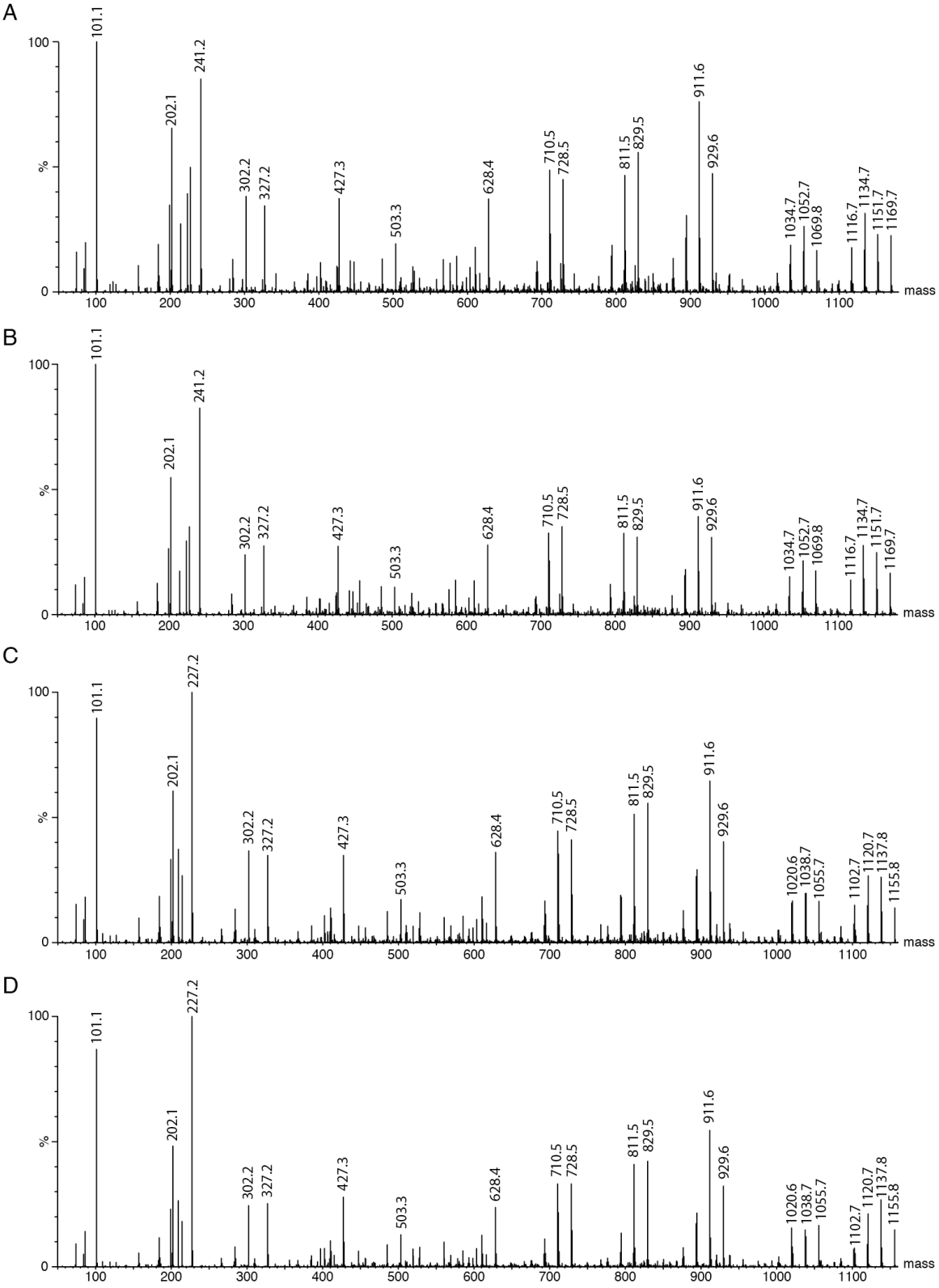


Figure S2. MS/MS analysis of polymyxin isolates from *P. polymyxa* NRRL B-30509 and comparison to commercial polymyxins E1 and E2. MS/MS data for (A) commercial polymyxin E1, (B) isolated polymyxin E1, (C) commercial polymyxin E2, and (D) isolated polymyxin E2.

Isolation of Tridecaptin A Variants

The culture supernatant was prepared as described for the polymyxin isolation. This supernatant was loaded onto a column packed with 40 g of Amberlite XAD16 resin (Sigma-Aldrich) at a flow rate of 1 mL/min at 4 °C. The column was washed (with a flow rate of 10 mL/min) with 250 mL of each of the following: water, 20 % IPA, 40 % IPA, 60 % IPA, and 80 % IPA with 0.1 % TFA. The 60 % and 80 % IPA with 0.1 % TFA fractions were concentrated *in vacuo* to 50 mL, and loaded onto a Phenomenex Strata SPE C18-E cartridge preconditioned with 50 mL of methanol (MeOH) followed with 100 mL of milli-Q water. The cartridge was washed with 50 mL of the following: 30 % ethanol (EtOH), 30 % acetonitrile (ACN), 40 % IPA, and 80 % IPA with 0.1 % TFA. The 80 % IPA with 0.1 % TFA fraction was concentrated *in vacuo* to approximately 10 mL. The tridecaptin variants were isolated by HPLC, using conditions identical to the polymyxin isolation, with the different components eluting from 36 to 41 minutes. Two tridecaptin variants were produced in relatively greater amounts (tridecaptin A_3 and A_4). These peptides were identified using MS/MS (see Figure S3). The lipid chains of these peptides were identified as described in the next section.

Derivatization and Characterization of the Tridecaptin Lipid Chains

The lipid chains of the tridecaptin analogues were characterized following an approach similar to that used for tridecaptin A (Kato et al., 1978). The combined tridecaptin analogues were dissolved in 4 mL of 6 M HCl and hydrolyzed at 110 °C for 2 hours. The hydrolysate was extracted with 5 mL diethyl ether, and the layers were separated. The organic layer was dried with anhydrous sodium sulfate, and then filtered. To the filtrate, 2 mL of methanol and 50 μL of 2 M trimethylsilyldiazomethane in hexanes was added. The mixture was stirred at room temperature for 30 minutes, then quenched with 50 μL of 2 M acetic acid and dried under a stream of argon. The residue was dissolved in 100 μL of dichloromethane, and analyzed using GC-MS as described in the mass spectrometry section. This yielded two major peaks, as shown in Figure S4 below.

MS/MS Analysis of Tridecaptins A_3 and A_4

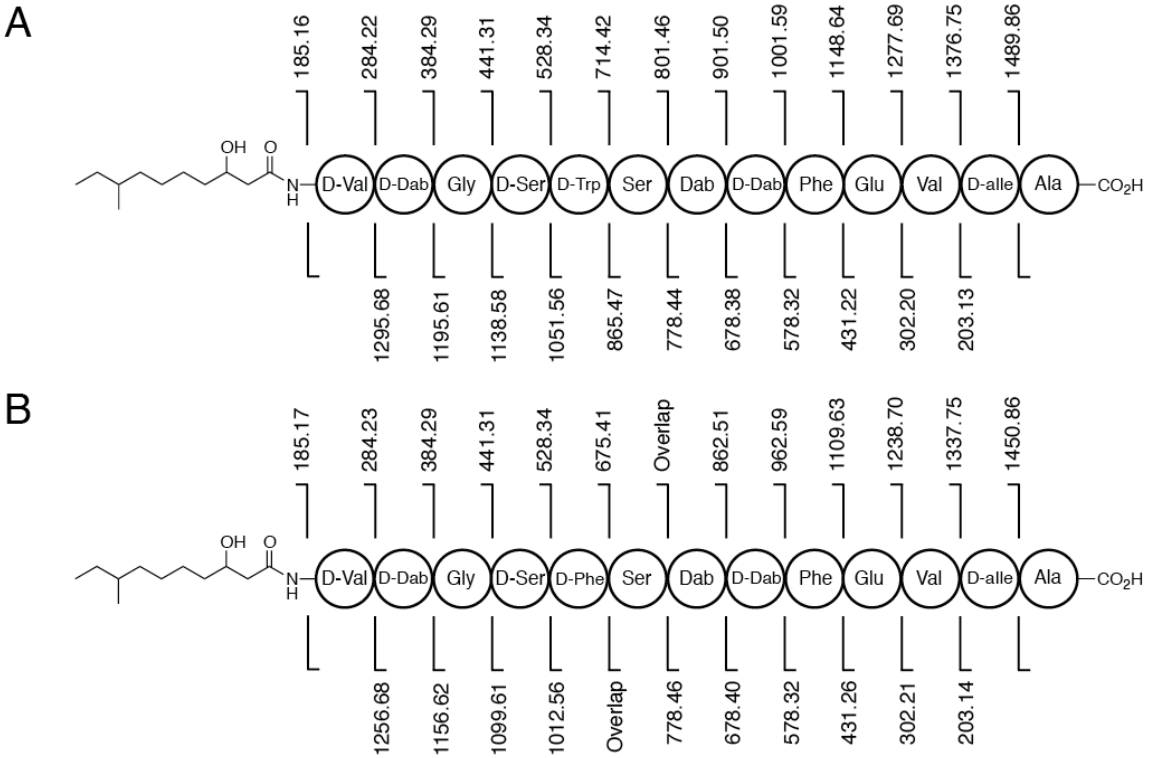


Figure S3. MS/MS characterization of tridecaptins from *P. polymyxa* NRRL B-30509. Summary of MS/MS data obtained for (A) tridecaptin A_3 and (B) tridecaptin A_4, with b-ions indicated above the structure, and y-ions indicated below.

GC-MS and EI-MS Data for the Derivatized Tridecaptin A Lipid Chains

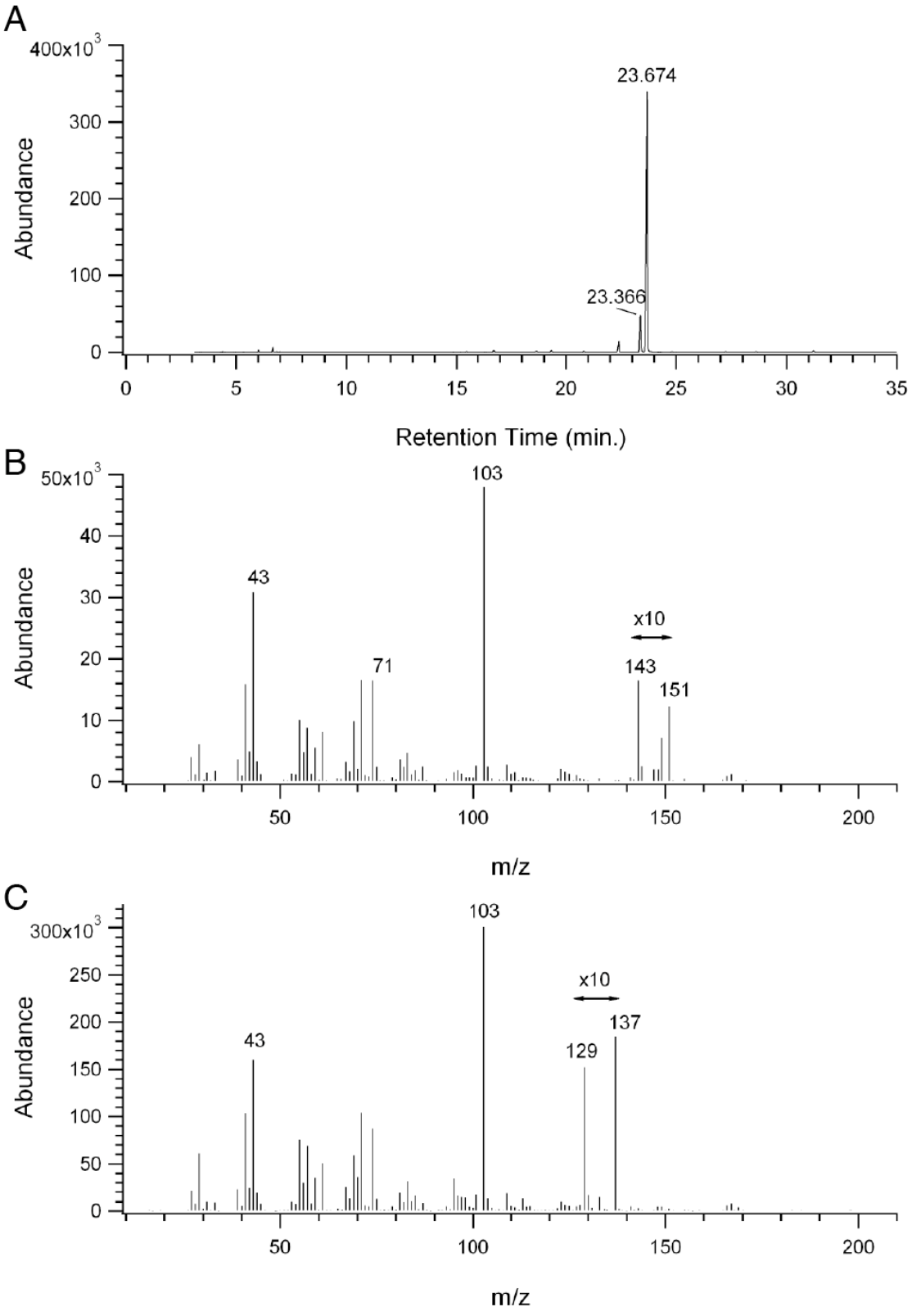


Figure S4. GC-MS and EI-MS data obtained for the derivatized tridecaptin lipid chains. (A) Extracted ion chromatogram from an injection of the derivatized tridecaptin lipid chains via GC-MS (monitoring at 103.0). EI-MS data for (B) the first peak at 23.366 min, methyl 3-hydroxy-9-methyldecanoate, and (C) the second peak at 23.674 min, methyl 3-hydroxy-8-methyldecanoate. A region of the EI-MS spectra has been magnified 10X to show key fragmentations allowing for structural identification (see also Fig. S5). Furthermore, the relative elution order of the iso- (methyl 3-hydroxy-9-methyldecanoate) and the anteiso- (methyl 3-hydroxy-8-methyldecanoate) structural isomers of the tridecaptin lipid chain are as expected based on literature precedent.[2]

Tridecaptin A Lipid Chain Fragmentation

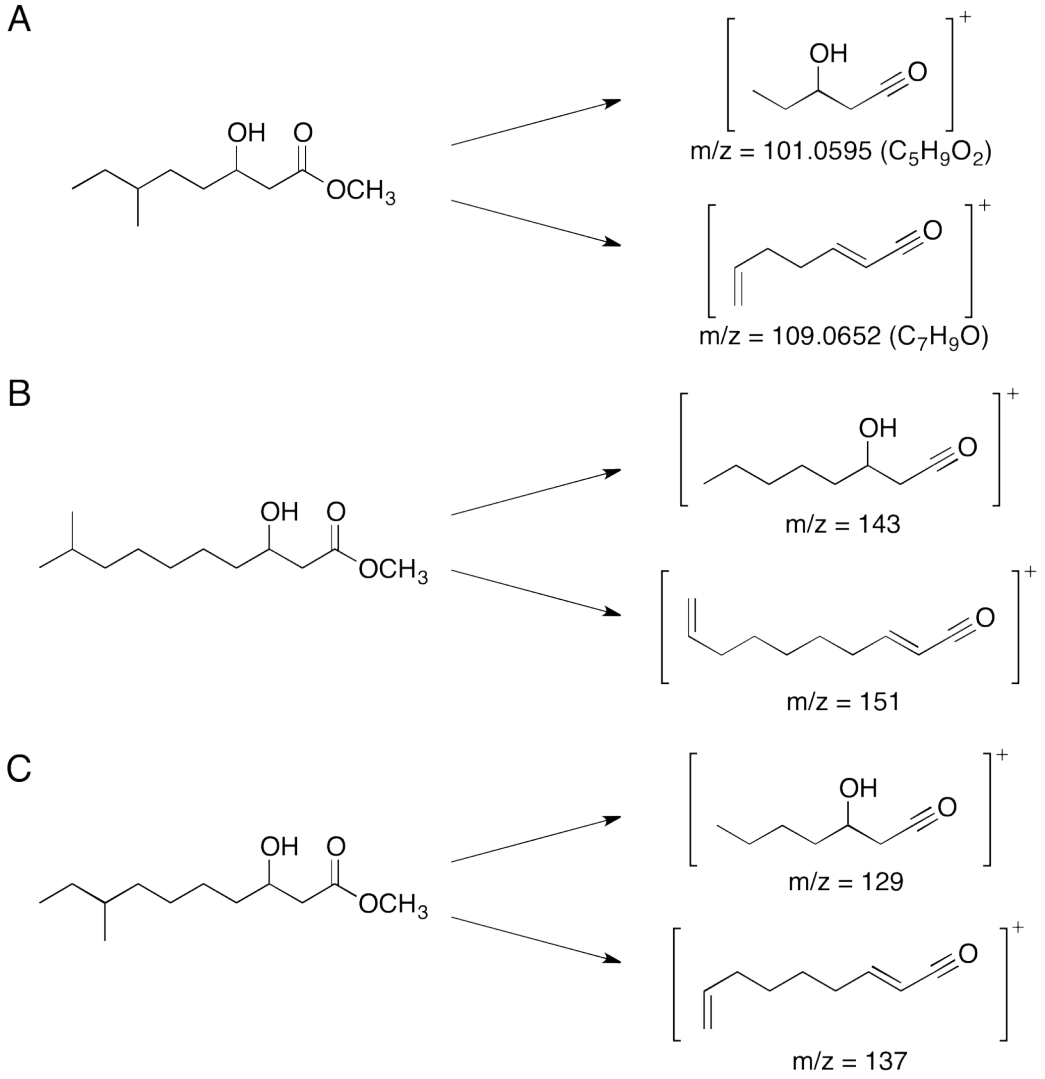


Figure S5. Predicted fragmented structures from EI-MS analysis of tridecaptin A lipid chain derivatives. (A) Fragments (with elemental composition) observed by high-resolution EI-MS data obtained for a synthetic sample of methyl 3-hydroxy-6-methyloctanoate, obtained on a Kratos Analytical MS-50G double focusing sector instrument (unpublished data). Key fragments are indicated for (B) methyl 3-hydroxy-9-methyldecanoate and (C) methyl 3-hydroxy-8-methyldecanoate that allowed for the differentiation of their structures.

Isolation of Paenicidin A

The 40 % IPA fraction from the Amberlite XAD16 step described in the tridecaptin isolation section was concentrated to 50 mL and loaded onto a Phenomenex Strata SPE C18-E cartridge, preconditioned with 50 mL of MeOH and 100 mL of water. The cartridge was washed with 50 mL of 30 % EtOH, 30 % ACN, 40 % IPA, and 80 % IPA with 0.1 % TFA. The 80 % IPA with 0.1 % TFA fraction was concentrated to 10 mL. Paenicidin A was isolated by HPLC, using conditions identical to the polymyxin isolation, eluting at 26 min. The sequence and structural gene of paenicidin A were determined as described in the main text.

Spectrum of Activity of Paenicidin A

Indicator strain	**Activity Result**[a]
Gram-positive	
Bacillus subtilis JH642	++
Carnobacterium maltaromaticum UAL26	-
Enterococcus faecium ATCC 19434	+
Lactococcus lactis subsp. *cremoris* HP	++
Listeria monocytogenes ATCC 15313	+
Pediococcus acidilactici PAC 1.0	-
Staphylococcus aureus ATCC 25923	-
Gram-negative	
Campylobacter jejuni 11168	-
Escherichia coli DH5α	-
Salmonella enterica serovar Typhimurium ATCC 13311	-
Pseudomonas aeruginosa ATCC 15442	-

[a] Inhibition of the indicator strain at 100 μM paenicidin A in a spot-on-lawn assay indicated by: ++, inhibition; +, faint inhibition; -, no inhibition.

Description of the Proteins Encoded by the Paenicidin A Gene Cluster

Protein Name	Size (a.a.)	Size (kDa)	Proposed function
PaeA	58	5.9	Lantibiotic structural gene
PaeF	242	27.1	ABC transporter ATP binding protein
PaeE	266	30.9	ABC transporter membrane-bound subunit
PaeG	255	28.9	ABC transporter membrane-bound subunit
PaeB	1058	122.3	Lantibiotic dehydratase
PaeT	625	71.4	Lantibiotic ABC transporter
PaeC	460	52.0	Lantibiotic cyclase

Reductive Alkylation of Paenicidin A

Approximately 0.1 mg of paenicidin A was dissolved in 50 μL of 100 mM NH_4HCO_3. 5 μL of 100 mM DTT was added to this solution, which was heated at 50 °C for 20 minutes. Iodoacetamide (5 μL; 200 mM) was added, and the resulting mixture was left to react in the dark for 20 minutes. The extent of alkylation was monitored by MALDI-TOF MS. No alkylation of paenicidin A was observed after treatment with these conditions (data not shown).

Reductive Desulfurization of Paenicidin A

Paenicidin A (0.5 mg) was suspended in 2.0 mL of CH_3OH/H_2O (1:1), to which 10 mg of $NiCl_2$ and 10 mg of $NaBH_4$ were added. This mixture was stirred under 1 atm of H_2 at room temperature. A 400 μL aliquot was removed after 30 min. After 8.5 h, the remaining mixture was taken off the hydrogenation apparatus. Each fraction was then centrifuged, followed by the removal of the supernatant. CH_3OH/H_2O (2.0 mL; 1:1) was added to the black nickel boride pellets, and these suspensions were sonicated for 15 minutes. These suspensions were centrifuged, and the supernatants were collected. The organic solvent was removed *in vacuo*, and the desulfurization progress was judged by MALDI-TOF MS.

Chiral Analysis of Paenicidin A

The hydrolysis of 0.5 mg of paenicidin A and the derivatization of the component amino acids to pentafluoropropionamide methyl esters was performed as has been previously reported, as was the preparation of Lan and MeLan standards.[3,4] The retention times of the Lan and MeLan residues from paenicidin A were compared with DL-Lan, LL-Lan, DL-MeLan, LL-MeLan, D-*allo*-L-MeLan and L-*allo*-L-MeLan standards, using GC-MS. The instrumentation used is described in the mass spectrometry section. The residues in the natural peptide were determined to be DL-Lan and DL-MeLan exclusively (data not shown).

NMR Spectroscopy

A 0.6 mM solution of paenicidin A was prepared in 500 μL of a 4:1 mixture of H_2O and CD_3OH with 100 μM 2,2-dimethyl-2-silapentanesulfonic acid. NMR experiments were performed at 25 °C on a Varian VNMRS 700 MHz spectrometer with a triple-resonance HCN cold probe and Z-axis pulsed-field gradients. Solvent signal was suppressed with transmitter presaturation. Chemical shifts were assigned based on homonuclear TOCSY (mixing time 100 ms) and NOESY (mixing time 200 ms) experiments. The partial proton chemical shift assignments for paenicidin A are shown below. Only the chemical shifts that were unambiguously assigned from the TOCSY and NOESY datasets are listed.

Proton Chemical Shifts of Paenicidin A

	HN	Hα	Hβ	others
Val 1		3.78	2.18	γCH_3 1.03
Leu 2	8.68	4.51		
Ala(S) 3	8.58	4.54	2.98, 3.12	
Ile 4	8.11	4.04	2.02	γCH 1.16
Val 5	7.88	4.19	2.08	γCH_3 0.93, 0.96
Ala 6	8.67	4.11	1.40	
Ala(S) 7	8.21	4.53	2.95, 3.00	
Ala(S) 8	9.01	5.00	3.09	
Ser 9	8.50	4.34	3.91, 3.96	
Gly 10	9.02	3.88, 4.21		
Ala(S) 11	7.58	4.10	2.93, 3.54	
Gly 12	8.62	3.97		
Ser 13	8.19	4.45	3.86, 3.88	
Gly 14	8.45	3.97		
Lys 15	8.22	4.34		
Abu(S) 16	8.86	4.76	3.56	γCH_3 1.32
Ala 17	8.07	4.27	1.41	
Ala 18	7.82	4.55	1.35	
Ala(S) 19	8.47	4.58	2.85, 3.26	
Ala(S) 20	8.57		3.03, 3.34	
Val 21	7.55	4.43	2.17	γCH_3 0.86, 0.93
Glu 22	8.50	4.26		
Abu(S) 23	8.59	4.68	3.67	γCH_3 1.29
Ala(S) 24	8.01	4.62	3.08, 3.18	
Gly 25	8.65	4.04, 4.14		
Asn 26	8.58	4.38	2.81, 2.85	
Arg 27	8.36	4.20		
Ala(S) 28	7.52	4.40	2.88, 2.94	
Phe 29	8.28	4.44	3.07, 3.17	
Abu(S) 30	8.20	4.57	4.12	γCH_3 1.24
Asn 31	8.18	4.55	3.42	
Val 32	8.41	3.99	2.05	γCH_3 0.95, 1.01
Gly 33	8.72	3.81, 4.20		
Ser 34	8.17	4.51	3.84, 4.03	
Leu 35	8.29			
Ala(S) 36	7.58			

NOESY Crosspeaks Supporting Structural Characterization

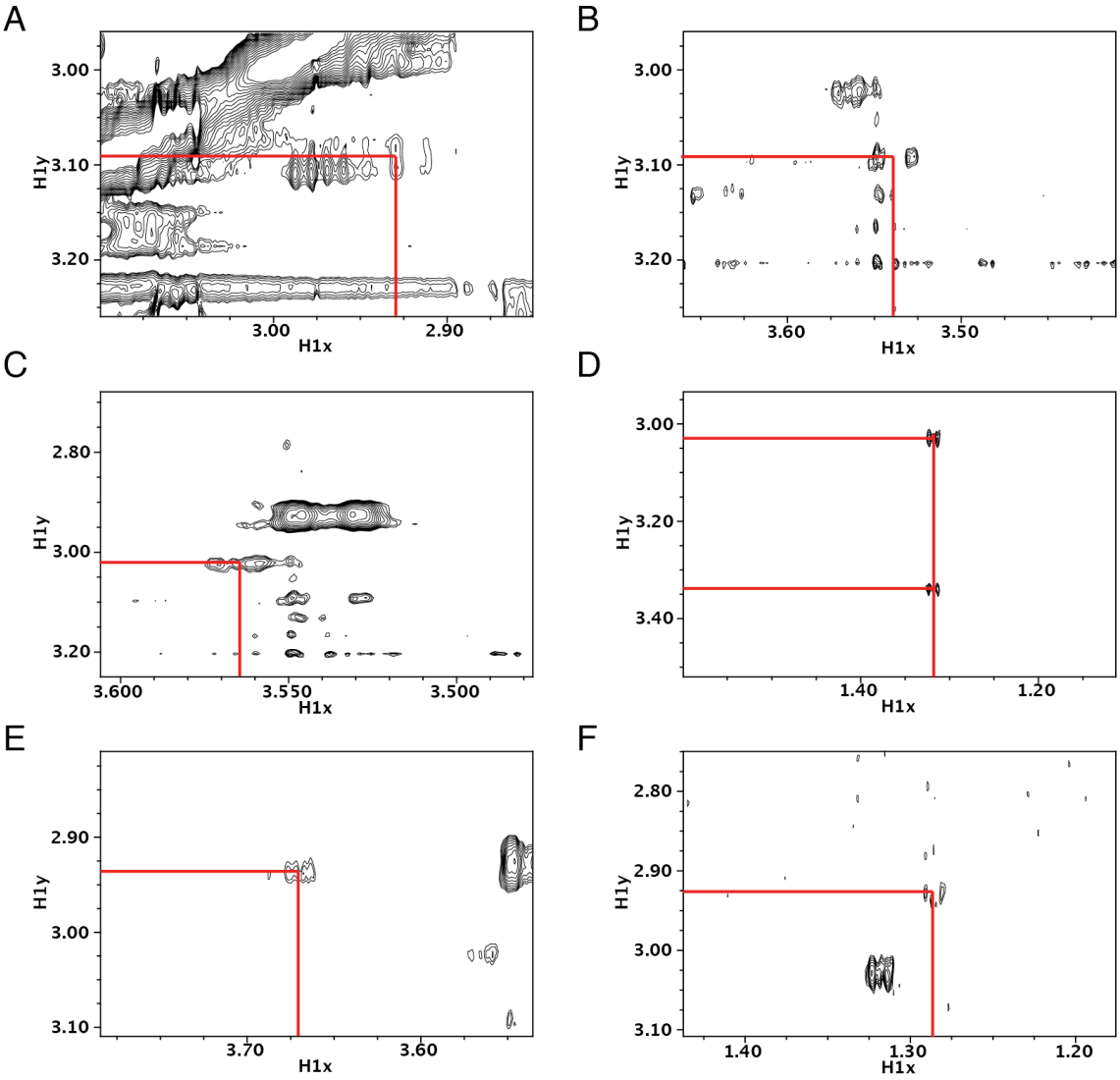


Figure S6. NOE data supportive of the proposed structure of paenicidin A. NOE crosspeaks between (A) 11.Hβ1 and 8.Hβ, (B) 11.Hβ2 and 8.Hβ, (C) 16.Hβ and 20.Hβ1, (D) 16.Hγ and 20.Hβ1, 20.Hβ2, (E) 23.Hβ and 28.Hβ1, and (F) 23.Hγ and 28.Hβ1.

Mass Spectrometry

MALDI-TOF MS analysis was conducted using the two-layer method with 3,5-dimethoxy-4-hydroxycinnamic acid as the matrix.[5] An AB Sciex Voyager Elite MALDI-TOF MS (Foster City, CA, USA) was operated in positive ion reflectron mode for spectral acquisition, using delayed extraction.

LC-MS/MS analysis was conducted on a Q-Tof Premier mass spectrometer (Waters, Milford, MA, USA) coupled with a nanoAcquity UPLC system (Waters, Milford, MA, USA). The sample (5 μL) was loaded onto a nano trap column (180 μm x 20 mm, Symmetry® C18 nanoAcquity™ column, Waters, Milford, MA, USA) followed by a nano analytical column (75 μm × 150 mm, Atlantis™ dC18 nanoAcquity™ column, Waters, Milford, MA, USA). Desalting on the peptide trap was achieved by flushing the trap with 1 % ACN, 0.1 % formic acid in water at a flow rate of 10 μL/min for 2-5 minutes. Peptides were separated with a gradient of 1-60 % solvent B (ACN with 0.1 % formic acid) over 45 to 55 min at a flow rate of 350 nL/min.

GC-MS experiments were performed on a 7890A GC system with a 5975C VL MSD (Agilent Technologies, Santa Clara, CA, USA). For the analysis of the tridecaptin lipid chain, 1 μL of sample was loaded onto a Zebron ZB-5MS capillary column (30 m length, 0.25 mm i.d., 0.25 μm film thickness; Phenomenex Inc., Torrance, CA, USA) by pulsed splitless injection, and detected in dual scan/single-ion monitoring mode (SIM ions at m/z 103.0, 43.0, and 29.0). The temperature programme started at 70 °C and was immediately ramped to 250 °C at a rate of 3 °C/min, and held at 250 °C for 5 min. MeLan stereochemistry was examined by loading 1 μL of sample onto the column described above via a 10:1 split injection, and detected in dual scan/single-ion monitoring

mode (SIM ions at m/z 379.0 and 248.0). The temperature programme started at 70 °C and was immediately raised to 190 °C at a rate of 10 °C/min, followed by a ramp to 205 °C at a rate of 1 °C/min. To analyze Lan stereochemistry, 1 μL of sample was loaded onto a Zebron ZB-1701 capillary column (30 m length, 0.25 mm i.d., 0.25 μm film thickness; Phenomenex Inc., Torrance, CA, USA) by a 10:1 split injection, and detected in dual scan/single-ion monitoring mode (SIM ions at m/z 365.0 and 248.0). The temperature programme started at 70 °C, was raised to 170 °C at a rate of 10 °C/min, followed by a ramp to 200 °C at a rate of 3 °C/min, and finished with a ramp to 300 °C at a rate of 10 °C/min.

References

(1) Rutherford, K.; Parkhill, J.; Crook, J.; Horsnell, T.; Rice, P.; Rajandream, M. A.; Barrell, B. *Bioinformatics* **2000,** *16,* 944.

(2) Massart-Leën, A. M.; De Pooter, H.; Decloedt, M.; Schamp, N. *Lipids* **1981,** *16,* 286.

(3) Ross, A. C.; Liu, H.; Pattabiraman, V. R.; Vederas, J. C. *J. Am. Chem. Soc.* **2010,** *132,* 462.

(4) Liu, W.; Chan, A. S.; Liu, H.; Cochrane, S. A.; Vederas, J. C. *J. Am. Chem. Soc.* **2011,** *133,* 14216.

(5) Dai, Y.; Whittal, R. M.; Li, L. *Anal. Chem.* **1999,** *71,* 1087.