Lucentamycins A–D, Cytotoxic Peptides from the Marine-Derived Actinomycete Nocardiopsis lucentensis

ARTICLE *in* JOURNAL OF NATURAL PRODUCTS · SEPTEMBER 2007

Impact Factor: 3.8 · DOI: 10.1021/np070101b · Source: PubMed

5 AUTHORS, INCLUDING:

Paul R Jensen
University of California, San Diego
202 PUBLICATIONS **8,933** CITATIONS
SEE PROFILE

William Fenical
University of California, San Diego
488 PUBLICATIONS **20,963** CITATIONS
SEE PROFILE

Lucentamycins A−D, Cytotoxic Peptides from the Marine-Derived Actinomycete *Nocardiopsis lucentensis*

Ji Young Cho, Philip G. Williams,† Hak Chol Kwon, Paul R. Jensen, and William Fenical*

Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093-0204

Received March 7, 2007

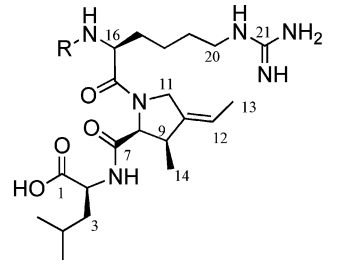
Four new 3-methyl-4-ethylideneproline-containing peptides, lucentamycins A−D (**1**−**4**), have been isolated from the fermentation broth of a marine-derived actinomycete identified by phylogenetic methods as *Nocardiopsis lucentensis* (strain CNR-712). The planar structures of the new compounds were assigned on the basis of 1D and 2D NMR spectroscopic techniques, while the absolute configurations of the amino acid residues were determined by application of the advanced Marfey method. Lucentamycins A (**1**) and B (**2**) showed significant *in vitro* cytotoxicity against HCT-116 human colon carcinoma.

Actinomycete bacteria collected in the marine environment are known to include obligate marine taxa[1] that by 16s rDNA sequence analysis are closely related to terrestrial bacteria. Actinomycetes have traditionally represented one of the most important resources for the discovery of new biologically active metabolites.[2] The genus *Nocardiopsis* has been shown to be phylogenetically coherent and to represent a distinct lineage within the order Actinomycetales.[3] *Nocardiopsis* strains are ubiquitous in the environment and are frequently isolated from habitats with moderate to high salt concentrations such as saline soils, marine sediments, and salterns.[4] Strains of the actinobacterial genus *Nocardiopsis* have also been reported to produce a variety of antimicrobial and other biologically active agents including the cytotoxic antifungal antibiotic kalafungin,[5] the antibiotic 3-trehalosamine,[6] the protein kinase C inhibitor methylpendolmycin,[7] and a staurosporine-like inhibitor of a cyclic AMP-dependent protein kinase.[7]

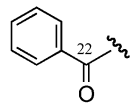
In 2003, we isolated an actinomycete identified as *Nocardiopsis lucentensis* (strain CNR-712) from sediment collected from a shallow saline pond on the island of Little San Salvador, in the Bahamas. A 50 mL culture of this organism provided a whole culture broth extract (EtOAc) that was cytotoxic toward the human colon tumor cell line HCT-116. In this paper, we report the results of the bioassay-guided fractionation of culture extracts of this strain, which led to the isolation and structure determination of four structurally unique, cytotoxic peptides, lucentamycins A−D (**1**−**4**). The majority of the cytotoxicity of this extract was found to be derived from lucentamycin A.

Results and Discussion

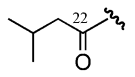
Bacterial strain CNR-712 was cultured at 27 °C by rotary shaking in 20 replicate 2.8 L Fernbach flasks each containing 1 L of culture medium. The cultures were extracted by adding 20 g/L Amberlite XAD-7 resin on the seventh day of the fermentation, and after low-speed shaking for 3 h, the resin was collected by filtration and washed with 1 L of DI water before it was eluted with acetone. The acetone eluent was concentrated under reduced pressure to give a crude extract (1.85 g from 20 L), which was subjected to repeated reversed-phase chromatography to yield lucentamycins A−D (**1**−**4**) in yields of 4.5, 3.5, 6.0, and 6.5 mg/20 L, respectively.



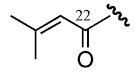
Lucentamycin A (1) R =

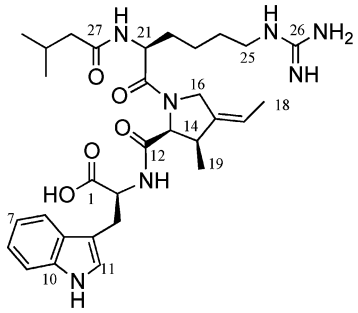


Lucentamycin C (3) R =



Lucentamycin D (4) R =





Lucentamycin B (2)

Lucentamycin A (**1**) was isolated as an optically active yellow oil ($[\alpha]_D$ −6.3, *c* 0.175, MeOH). The molecular weight of **1** was obtained from the HRESITOF mass spectrum, which showed pseudomolecular ions at *m*/*z* 543.3298 $[M + H]^+$ and 565.3109 $[M + Na]^+$. On this basis, the molecular formula was defined as $C_{28}H_{42}N_6O_5$, which indicated that **1** contained 11 double-bond equivalents. A peptide structure was evident from the ^1H and ^{13}C NMR data, recorded in DMSO-d_6 (Table 1). Diagnostic resonances for four carbonyl carbons (δ_C 165.9, 171.3, 167.2, and 175.8) and

* Corresponding author. Tel: 858-534-2133. Fax: 858-558-3702. E-mail: wfenical@ucsd.edu.

† Current address: Department of Chemistry, University of Hawaii at Manoa, Honolulu, Hawaii 96822.

10.1021/np070101b CCC: $37.00
Published on Web 07/14/2007

Table 1. NMR Spectroscopic Data (500 MHz, DMSO-d_6) for Lucentamycin A (**1**)[a]

C/H no.	δH mult (*J* in Hz)	δC		COSY	HMBC	ROESY
1		175.8	qC			
2	3.86, t (6.5)	52.4	CH	3, 2-NH	C1, C3, C4	
2-NH	6.95, d (6.5)			2	C7	
3	1.39, m	43.1	CH_2	2, 4	C1, C2, C4	
4	1.72, m	24.5	CH	3, 5	C2, C3, C6	
5	0.85, d (6.5)	22.7	CH_3	4		
6	0.85, d (6.5)	23.1	CH_3	4	C3, C4, C5	
7		167.2	qC			
8	4.22, d (9.0)	65.6	CH	9	C7, C9, C10, C14, C15	9
9	3.21, dq (9.0, 7.5)	35.01	CH	8, 14	C10, C11, C12, C14	8, 14
10		139.4	qC			
11a	4.50, d (13.2)	50.9	CH_2	11b, 12, 13	C9, C10, C12, C15	
11b	4.40, d (13.2)			11a, 12, 13		
12	5.38, q (7.5)	116.2	CH	11, 13	C9, C10, C11, C13	13, 10
13	1.61, d (7.5)	13.2	CH_3	11, 12	C10, C12	12
14	1.02, d (7.5)	14.9	CH_3	9	C8, C9, C10	9
15		171.3	qC			
16	4.78, dd (7.5, 3.0)	51.3	CH	17, 16-NH	C15, C17	
16-NH	8.43, d (7.5)			16	C15, C22	
17a	1.62, m	31.3	CH_2	17b	C16, C15	
17b	1.97, m			17a		
18a	1.40, m	28.5	CH_2	18b		
18b	1.58, m			18a		
19a	1.70, m	35.00	CH_2	19b		
19b	1.70, m			19a		
20a	3.12, m	41.1	CH_2	19, 20-NH	20-NH	
20b	3.01, m			19, 20-NH		
20-NH	10.39, br s			20	C20, C21	
21		157.4	qC			
21-NH_2[b]	4.12, br s					
21-NH[b]	7.78, br s					
22		165.9	qC			
23		133.7	qC			
24/28	7.85, d (7.5)	127.5	CH	25	C22, C26, C28	
25/27	7.40, dd (8.0, 7.5)	128.2	CH	24, 26	C23, C27	
26	7.51, td (8.0, 7.5)	131.4	CH	25, 27	C24, C28	

[a] Assignments made by a combination of proton and carbon 1D and 2D experiments (COSY, TOCSY, DEPT, HMBC, HMQC, etc.). [b]These resonances were assigned on the basis of their chemical shifts and integration.

three α-amino acid methine resonances ($\delta_{C/H}$ 51.3/4.78, 65.6/4.22, and 52.4/3.86) all indicated the peptide structure of **1**. Absorptions at 3374 and 1641 cm^{-1} in the IR spectrum of **1** were also characteristic of hydroxy, amide, and carbonyl functional groups. On the basis of the number of sp^2 carbons and their chemical shifts, the remaining degrees of unsaturation could be ascribed to a monosubstituted benzene ring, one carbon−carbon double bond (C-10, δ_C 139.4; C-12, δ_C 116.2), one carbon−nitrogen double bond (C-21, δ_C 157.4), and, by elimination, one additional ring system.

HMBC and COSY experiments (Table 1) established a series of partial structures. The methyl proton resonances of an isopropyl group (H-5 and H-6) showed HMBC correlations to one another, to C-4, and to C-3. The protons attached to C-3 showed a COSY correlation to H-2 and an HMBC correlation to a carboxyl carbon at δ_C 175.8 (C-1), while the α-proton (H-2) showed a COSY correlation to a secondary amide proton resonance (δ_H 6.95). These data established this amino acid unit as leucine. A homoarginine (Har) unit was also assigned by interpretation of COSY, TOCSY, and HMBC spectroscopic data. Specifically, four methylene protons were observed between δ 1.40 and 3.12 that, on the basis of a TOCSY experiment, comprised a contiguous alkyl chain containing four methylene groups connected to a nitrogen atom bearing an exchangeable proton. This exchangeable proton (δ_H 10.39) showed a COSY correlation to H-20 and an HMBC correlation to C-21 (δ_C 157.4). The chemical shift of C-21 was characteristic of a guanidine unit,[8] the remaining protons of which were assigned to the broad resonances at δ_H 4.12 and 7.78 in the proton spectrum. The methylene protons (H_2-17) most distant from the guanidine group showed a COSY correlation to a deshielded methine proton (H-16), which in turn showed a COSY correlation to a secondary amide proton (δ_H 8.43) and an HMBC correlation to the carbonyl (C-15) of this amino acid unit.

A third fragment was constructed by analysis of the NMR spectroscopic data starting with the C-14 methyl proton resonance. The methyl group was clearly attached to the β-position of a modified amino acid residue, since this resonance showed a COSY correlation to H-9 (δ_H 3.21) and HMBC correlations to the α-carbon (C-8, δ_C 65.6) and a quaternary sp^2 carbon signal (C-10, δ_C 139.4) in the γ-position. This γ-quaternary carbon (C-10) in turn showed HMBC correlations from two downfield methylene proton resonances (H-11a and H-11b, δ_H 4.50 and 4.40) that were α to an amino functionality. Together, these data indicated this amino acid was a β,γ-disubstituted proline unit. The substituent in the γ-position was an exocyclic ethylidene group, based on HMBC correlations from an allylic methyl (H_3-13, δ_H 1.61) and a vinyl proton (H-12, δ_H 5.38) to the γ-quaternary sp^2 carbon (C-10) of the proline ring. These comprehensive NMR correlations allowed the assignment of this proline derivative as 3-methyl-4-ethylideneproline. The remaining protons in **1** were assigned from the COSY, HMBC, and HSQC spectra, which showed a two-proton aromatic doublet at δ_H 7.85 (H-24, 28) in the HSQC spectrum with a $^1J_{CH}$ correlation to a resonance at δ_C 127.5 (C-24, 28). COSY correlations from H-25 to H-24 and to H-26 as well as HMBC correlations from H-25 to C-23 and from H-28 to C22 established the benzoic acid fragment.

Sequencing the amino acid residues in **1** was accomplished by routine HMBC NMR analyses using correlations between the secondary amide proton and the carbonyl carbon resonances. Specifically, cross-peaks between the secondary amide proton of Leu (δ_H 6.95) and the carbonyl carbon (C-7) of 3-methyl-4-

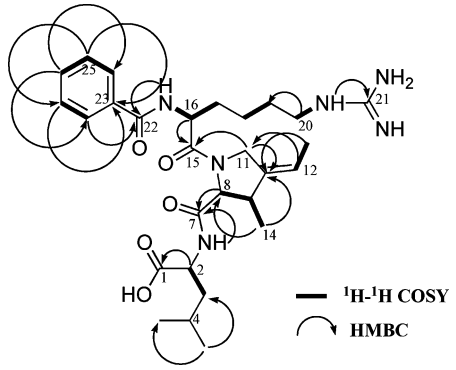


Figure 1. Key ^{1}H−^{1}H COSY and HMBC correlations used to establish the structure of **1**.

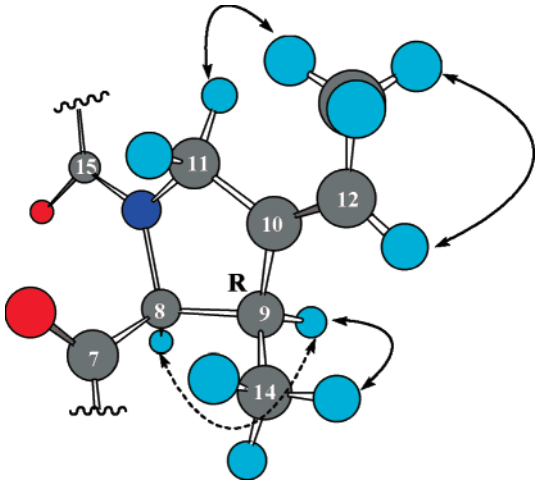


Figure 2. ROESY correlations used to establish the relative configuration of **1**.

ethylideneproline linked these two amino acids. This dipeptide fragment was in turn linked to the Har unit by the observation of an HMBC correlation between the γ-proton (H-11) of 3-methyl-4-ethylideneproline and the carbonyl carbon (C-15) of the Har unit. Finally, the secondary amide proton of the Har unit (δ_H 8.43) correlated to the carbonyl carbon (δ_C 165.9) of the benzoic acid unit to reveal the complete sequence starting from the C-terminus as Leu/3-methyl-4-ethylideneproline/benzoyl Har.

The relative configuration of the modified proline unit in **1** was established by analysis of ROESY spectroscopic data. The *trans* configuration of the (2*S*,3*R*)-3-methyl-4-ethylideneproline residue was confirmed by ROESY correlations from the methyl proton (H_3-14) to H-9 and from this proton to the α-proton (H-8) (Figure 2). A $\Delta_{10} = Z$ double-bond geometry was also confirmed by ROESY correlations.

To determine the absolute configurations of the amino acid constituents in **1**, the advanced Marfey method was applied.[9] The acid hydrolysate (1 mg, 6 N HCl, 110 °C, 16 h) of **1** was derivatized with 1-fluoro-2,4-dinitrophenyl-5-L-leucinamide (L-FDLA) and D-FDLA[10] and analyzed by positive mode electrospray ionization (ESI) LC/MS using a C8 Luna column. All derivatives were identified by their retention times, molecular weights, and UV spectra and by co-injection of the appropriate standards, if available. The configuration of the Leu unit was determined by comparing the retention times of L-derivatized (31.94 min) and D-derivatized L-Leu (44.10 min) with the L-derivatized hydrolysate (L-Leu unit, 31.88 min). The configuration of the Har unit was determined by comparing the retention times of L- (14.09 min) and D-derivatized L-Har (11.09 min) with the Har unit in the L-derivatized hydrolysate (14.18 min). To determine the configuration of the 3-methyl-4-ethylideneproline unit, the L- and D-FDLA derivatives of this unit were eluted from the Luna C8 reversed-phase column and identified on the basis of their molecular weights. On the basis of the proposed separation mechanism for Marfey's technique,[9] the absolute configuration of the amino-bearing center can be determined based on the elution order of the L- versus D-FDLA derivatives. The elution order of the enantiomers of any amino acid (L versus D) can be determined by comparison of the hydrophobicity between the α-carboxyl group and the side chain of the amino acid. In this case, the 2*R*-3-methyl-4-ethylideneproline diastereomers should show a longer retention time than the 2*S* diastereomers due to the *cis* orientation of the nonpolar groups in the former derivative. This elution order is consistent with the reported elution order for methylproline.[11] Since the 3-methyl-4-ethylideneproline peak in the L-FDLA-derivatized hydrolysate eluted before the 3-methyl-4-ethylideneproline peak in the D-FDLA-derivatized hydrolysate, the absolute configuration of the amino carbon in this residue was assigned as *S*. On this basis, the absolute configuration of lucentamycin A (**1**) could be assigned as 2*S*,8*S*,9*R*,16*S*.

Lucentamycin B (**2**) was isolated as a yellow oil ($[\alpha]_D$ −15, *c* 0.100, MeOH). The molecular weight of **2** was obtained from the HRESITOF mass spectrum, which showed pseudomolecular ions $[M + H]^+$ at *m*/*z* 596.3532 and $[M + Na]^+$ 618.3343, respectively. On the basis of the HRESITOFMS data, the molecular formula was defined as $C_{31}H_{45}N_7O_5$, which indicated 13 degrees of unsaturation. The structure of compound **2** was assigned by analysis of the 1D and 2D NMR spectra recorded in DMSO-d_6 (Table 2). These data indicated that **2** was an analogue of **1** but possessed an additional 53 mass units. For example, HMBC and COSY experiments indicated the presence of Har and 3-methyl-4-ethylideneproline units as in **1**. Key resonances that were diagnostic for the Har unit were the four methylene protons of the alkyl chain, which resonated between δ_H 1.32 and 3.11, an exchangeable proton (δ_H 10.62), a secondary amide proton (δ_H 8.05), and an α-proton at δ 4.59 that were observed in the ^{1}H NMR spectrum, plus guanidine quaternary (C-26) and carbonyl carbon (C-20) resonances that were observed in the ^{13}C NMR spectrum. The 3-methyl-4-ethylideneproline unit was evident on the basis of the two methyl doublets (H_3-18 and H_3-19) and the vinyl proton H-17 that were observed in the ^{1}H NMR spectrum.

Differences between **1** and **2** were clearly visible in the spectroscopic data. For example, compound **2** contained four aromatic protons from H-6 to H-9 (δ_H 7.60, 7.33, 7.15, and 6.91), which showed a series of correlations to each other to establish a four-carbon unit (C-6 through C-9). This unit was expanded into a disubstituted benzene ring by the observation of HMBC correlations from H-6 and H-9 to C-5 and C-10. These carbons also showed HMBC correlations from the downfield aromatic proton singlet H-11, which verified that this was an indole ring system. The α-(H-2) and β-resonances (H_2-3) of this amino acid unit were rapidly identified through a combination of COSY and HMBC correlations. These data allowed this amino acid unit to be assigned as tryptophan. From the ^{1}H NMR spectroscopic data, the last fragment contained two methyl doublets at δ_H 0.82 (H_3-30) and 0.80 (H_3-31), which were assigned to an isopropyl group. As expected, these methyl doublets (H_3-30 and H_3-31) showed HMBC correlations to one another and to C-28 and COSY correlations to H-29. The H-28 methylene protons showed a COSY correlation to H-29 and an HMBC correlation to a carbonyl carbon (C-27) at δ 171.1 to establish this unit as 3-methylbutyric acid.

As in **1**, the sequence of these units was assigned by analysis of HMBC data. Cross-peaks between the secondary amide proton of Trp (δ_H 7.02) and the carbonyl carbon of 3-methyl-4-ethylideneproline (C-12) and between the γ-proton (H-16) of 3-methyl-4-ethylideneproline and the carbonyl carbon of Har (C-20) were observed, which allowed a tetrapeptide fragment to be defined. Finally, 3-methylbutyric acid was attached to the N-terminus of

Table 2. NMR Spectroscopic Data (500 MHz, DMSO-d_6) for Lucentamycin B (**2**)[a]

C/H no.	δ mult (*J* in Hz)	δC		COSY	HMBC	ROESY
1		175.4	qC			
2	4.15, dd (11.5, 5.5)	54.1	CH	3, 2-NH	C1, C3, C4	
2-NH	7.02, d (11.5)			2	C1, C12	
3	3.02, d (5.5)	27.7	CH_2	2	C1, C2, C4, C5, C11	
4		112.0	qC			
5		127.9	qC			
6	7.33, d (8.5)	111.2	CH	7	C5, C8, C10	
7	7.15, dd (8.5, 5.5)	120.4	CH	6	C5, C6	
8	6.91, dd (7.5, 5.5)	117.7	CH	7, 9	C7	
9	7.60, d (7.5)	118.6	CH	8	C5, C7, C10	
10		139.3	qC			
11	7.11, s	123.4	CH		C4, C10, C5	
12		167.4				
13	4.15, d (8.5)	65.6	CH	14	C12, C14, C19	
14	3.13, dq (8.5, 6.5)	34.9	CH	19	C15, C17, C16, C19	14
15		139.3	qC			13, 19
16a	4.36, d (13.5)	51.4	CH_2	16b, 17, 18	C13, C14, C15, C17, C20	
16b	4.24, d (13.5)			16a, 17, 18		17
17	5.32, q (6.5)	116.1	CH	16, 18	C14, C16, C18	
18	1.60, d (6.5)	13.1	CH_3	16, 17	C15, C17	16, 18
19	0.81, d (6.5)	14.7	CH_3	14	C13, C14, C15	17
20		171.1	qC			14
21	4.59, br m	50.7	CH	22, 21-NH	C20	
21-NH	8.05, d (7.5)			21	C27	
22a	1.79, m	31.8	CH_2	22b, 23, 24		
22b	1.44, m			22a, 23, 24		
23a	1.35, m	22.9	CH_2	22		
23b	1.48, m					
24a	1.32, m	28.5	CH_2	25		
24b	1.51, m					
25a	2.99, m	41.1	CH_2	24, 25-NH	C23, C24, C26	
25b	3.11, m			25-NH		
25-NH	10.33, br s					
26		157.1	qC			
26-NH	10.62, br s					
27		171.1	qC			
28	1.95, d (6.5)	44.0	CH_2	29	C27, C29, C30	
29	1.92, m	25.6	CH	30, 31	C28	
30	0.82, d (6.5)	22.1	CH_3	29	C28, C29, C31	
31	0.80, d (6.5)	22.1	CH_3	29	C28, C29, C30	

[a] Assignments made by a combination of proton and carbon 1D and 2D experiments (COSY, TOCSY, DEPT, HMBC, HMQC, etc.).

the peptide on the basis of an HMBC correlation between the secondary amide proton of Har and the carbonyl carbon of the butyric acid unit. These data revealed the following sequence starting from the C-terminus: Trp/3-methyl-4-ethylideneproline/3-methylbutanoyl Har.

To assign the configuration of **2**, the peptide (1 mg) was hydrolyzed and divided into two portions. Each portion was derivatized with 1-fluoro-2,4-dinitrophenyl-5-L-leucinamide (L-FDLA) and D-FDLA. The L-FDLA derivatives were analyzed using electrospray ionization LC/MS in the positive mode. In the hydrolysate, the retention times of L- and D-derivatized Har units were 14.18 and 11.90 min, respectively, as determined by comparison with authentic standards. The 3-methyl-4-ethylideneproline residue in **2** had an identical retention time to the same unit derived from **1**, indicating that these stereocenters were conserved in **2**. The stereochemistry of the Trp unit was determined by comparing the retention time of L- (32.60 min) and D-FDLA-derivatized L-Trp (37.63 min) with the L-derivatized hydrolysate (32.57 min). While this derivatized peak was small when compared to the other amino acid derivatives, due presumably to Trp degradation during hydrolysis,[12] it could still be clearly identified in the hydrolysate, at the correct retention time, by single-ion monitoring of the LC-MS trace.

Examination of other fractions from the C18 reversed-phase flash column led to the isolation of lucentamycin C (**3**). High-resolution HRESTOFMS established the molecular formula of this analogue as $C_{26}H_{46}N_6O_5$ based on a pseudomolecular ion peak at 523.3591 ($[M + H]^+$), which revealed seven degrees of unsaturation. The structure of **3** was assigned on the basis of its molecular formula as well as by analysis of ^1H, ^{13}C, and 2D NMR spectroscopic data in DMSO-d_6 (Table 3). All of these data indicated **3** was an analogue of **1**, since analysis of the HMBC and COSY experiments revealed the presence of Har, 3-methyl-4-ethylideneproline, and Leu units in **3**. Further inspection of the NMR data revealed the differences could be attributed to a terminal isopropyl group in **3**. The ^1H NMR spectrum of **3** lacked the resonances corresponding to the benzoic acid unit and instead contained two methyl doublets at δ_H 0.83 (H_3-25/H_3-26), a methylene at δ_H 2.02, and a methine proton at δ 1.97. These methyl doublets (H_3-25 and H_3-26) showed COSY correlations to H-24 and an HMBC correlation to C-23, which established an isobutanoic acid moiety. Given the similarity of the NMR spectroscopic data for **3** with the data from **1**, lucentamycin C was assumed to possess the same absolute configuration as in **1**.

The structure of the final peptide isolated, lucentamycin D (**4**), was also assigned by analysis of the 1D and 2D NMR spectra recorded in DMSO-d_6 (Table 4). The elemental composition of this compound, as determined by HRESITOFMS, was $C_{26}H_{44}N_6O_5$ on the basis of a pseudomolecular ion peak $[M + H]^+$ at 521.3437 (eight degrees of unsaturation). The NMR spectroscopic data for compound **4** were almost identical to those of **3**. The only difference that could be discerned was the presence of two olefinic methyl groups (C-25 and C-26), indicating that **4** contained an isobutenoic rather than an isobutanoic acid moiety.

Lucentamycins A and B (**1**, **2**) showed significant *in vitro* cytotoxicity against HCT-116 human colon carcinoma with IC_{50}

Table 3. NMR Spectroscopic Data (500 MHz, DMSO-d_6) for Lucentamycin C (**3**)[a]

C/H no.	δH mult (*J* in Hz)	δC		COSY	HMBC
1		175.9	qC		
2	3.87, t (6.5)	52.4	CH	3, 2-NH	C1, C3, C4
2-NH	6.89, d (6.5)			2	
3	1.38, m	43.1	CH_2	2, 4	C1, C2, C4
4	1.65, m	24.5	CH	3, 5	C2, C3, C6, C5
5	0.83, d (6.5)	22.7	CH_3	4	C3, C4, C6
6	0.83, d (6.5)	23.1	CH_3	4	C3, C4, C5
7		167.2	qC		
8	4.16, d (9.0)	65.6	CH	9	C7, C9, C14, C15
9	3.20, m	34.9	CH	8, 14	C10, C11, C12, C14
10		139.4	qC		
11a	4.37, d (13.5)	50.7	CH_2	11b, 12, 13	C9, C10, C12, C15
11b	4.31, d (13.5)			11a, 12, 13	
12	5.33, q (7.5)	116.2	CH	11, 13	C9, C11, C13
13	1.60, d (7.0)	13.2	CH_3	11, 12	C10, C12
14	0.99, d (7.5)	14.9	CH_3	9	C8, C9, C10
15		171.2	qC		
16	4.60, dd (7.5, 3.0)	50.2	CH	17, 16-NH	C15, C17
16-NH	8.02, d (7.5)			16	
17a	1.51, m	31.7	CH_2	18, 16	C16, C17
17b	1.86, m			18, 16	
18a	1.41, m	22.9	CH_2	17	
18b	1.52, m			17	
19a	1.39, m	28.6	CH_2	20	
19b	1.50, m			20	
20a	3.19, m	41.2	CH_2	19, 20-NH	NH
20b	3.03, m			19, 20-NH	
20-NH	10.48, br s			20	
21		157.2	qC		
21-OH	6.99, br s				
22		171.1	qC		
23	2.02, m	44.0	CH_2		C22, C24, C25
24	1.97, m	25.6	CH_2	25, 26	C23, C25, C26
25	0.83, d (6.5)	22.2	CH_3	24	
26	0.83, d (6.5)	22.1	CH_3	24	

[a] Assignments made by a combination of proton and carbon 1D and 2D experiments (COSY, TOCSY, DEPT, HMBC, HMQC, etc.).

values of 0.20 and 11 μM, respectively. Interestingly, **3** and **4** were not cytotoxic against this cell line at concentrations up to 150 μM. A comparison of the structural differences between these compounds suggests that the aromatic ring is essential for the biological activity of this class of compounds since **1** and **2** contain phenyl and indole rings, respectively, while **3** and **4** lack such a moiety.

The structurally most intriguing part of these compounds is clearly the 3-methyl-4-ethylideneproline moiety. These compounds represent the first report of this amino acid unit. The constitution of this unit was unambiguously established by a wealth of 2D NMR spectroscopic data, confirming that the ethylidene fragment is attached at the γ-carbon. Compounds containing 3-methylproline units are relatively common in nature,[13] but ethylideneproline units are rare. The closest structurally related modified proline unit is a constituent of the benzodiazepine antibiotic tomamycin,[14] which contains a γ-ethylideneproline unit. It has been previously established that β-methylproline is biosynthetically derived from isoleucine[15] via oxidation of the δ-methyl group to an aldehyde, cyclization to the imine, and subsequent reduction (Figure 3a). The γ-ethylideneproline unit in tomamycin has been shown to be derived from tyrosine on the basis of extensive feeding studies with radioactive precursors.[14] Biosynthetic conversion of tyrosine to γ-ethylideneproline is a complex process involving initial conversion to L-DOPA, cyclization to the 2,3-dihydroindole ring system, oxidative cleavage of the aromatic ring, and several more steps (Figure 3b). In the case of the 3-methyl-4-ethylideneproline units found in **1** and **2**, both general routes are possible. One can envision this moiety being derived from SAM methylation of the β-position of hydroxyphenylpyruvate followed by transamination to give β-methyltyrosine.[16] This is subsequently oxidized to the analogous tyrosine derivative and after several more steps converted into the 3-methyl-4-ethylideneproline unit (Figure 3d). An alternative pathway for the formation of this amino acid would be via oxidation of isoleucine and cyclization to the enamine. Nucleophilic attack of the enamine on acetyl-CoA provides the necessary remaining carbons, which after reduction and dehydration would give the desired amino acid (Figure 3c). Of the two routes, the former seems more likely given the precedents.

Experimental Section

General Experimental Procedures. The optical rotations were measured on an Autopol automatic polarimeter (Rudolph Research, Flanders, NJ). UV spectra were measured on a Varian Cary UV visible spectrophotometer with a 1 cm cell. IR spectra were obtained with a Perkin-Elmer Spectrum BX spectrometer. ^1H and ^{13}C NMR spectra were obtained in DMSO-d_6 on a Varian Inova spectrometer operating at 500 and 125 MHz, respectively. HRESITOFMS data were obtained at The Scripps Research Institute, La Jolla, CA. Reversed-phase HPLC separations were performed using a semipreparative C8 Luna column (250 × 10 mm) at a flow rate of 2 mL/min using HP Series 1050 pump and UV detector.

Collection, Identification, and Cultivation of Strain CNR712. Strain CNR712 was isolated from a sediment sample collected in 2003 from a shallow saline pond on the island of Little San Salvador, The Bahamas. Nearly complete 16S rRNA gene sequencing revealed 99.5% sequence identity with *Nocardiopsis lucentensis* (sequence data has been deposited in GenBank under accession number Bankit 877589). Strain CNR712 was cultured in 20 replicate 2.8 L Fernbach flasks each containing 1 L of fermentation medium CKA (5 g of starch, 4 mL of 50% condensed fish solubles (CropMaster, Hudson, FL), 2 g of menhaden meal, 2 g of kelp powder, 2 g of chitosan, 1 L of seawater) for 7 days, after which 20 g/L Amberlite XAD-7 resin was added to the culture and the slurry shaken for 3 h. The resin was then collected by filtration through cheesecloth, washed with 1 L of deionized H_2O to remove salts, and eluted with acetone to generate a crude extract.

Table 4. NMR Spectroscopic Data (500 MHz, DMSO-d_6) for Lucentamycin D (**4**)[a]

C/H no.	δH mult (*J* in Hz)	δC		COSY	HMBC
1		175.9	qC		
2	3.87, t (6.0)	52.4	CH	3, 2-NH	C1, C3, C4
2-NH	6.89, d (6.0)			2	
3	1.38, m	43.1	CH_2	2, 4	C1, C2, C4
4	1.72, m	24.5	CH	3, 5, 6	C2, C3, C5, C6
5	0.84, d (6.5)	22.7	CH_3	4	C3, C4, C6
6	0.84, d (6.5)	23.1	CH_3	4	C3, C4, C5
OH	6.98, br s				
7		167.2	qC		
8	4.17, d (8.0)	65.6	CH	9	C7, C9, C14, C15
9	3.21, dq (8.0, 7.5)	35.0	CH	8, 14	C10, C11, C12, C14
10		139.4	qC		
11a	4.39, d (13.5)	50.9	CH_2	11b, 12, 13	C9, C10, C12, C15
11b	4.32, d (13.5)			11a, 12, 13	
12	5.35, q (7.5)	116.2	CH	11, 13	C9, C11, C13
13	1.61, d (6.5)	13.2	CH_3	11, 12	C10, C12
14	1.01, d (6.5)	14.9	CH_3	9	C8, C9, C10
15		171.4	qC		
16	4.64, dd (7.5, 3.0)	49.8	CH	17, 16-NH	C15, C17
16-NH	7.92, d (8.0)			16	C16, C22
17a	1.47, m	31.7	CH_2	16, 18	C16, C17
17b	1.78, m			18	
18a	1.38, m	22.8	CH_2	17	
18b	1.54, m			17	
19a	1.33, m	28.5	CH_2	20	
19b	1.54, m			20	
20a	2.98, m	41.2	CH_2	19, 20-NH	NH, C18, C19
20b	3.12, m			19, 20-NH	
20-NH	10.47, br s			20	
21		157.2	qC		
21-NH_2	4.09, br s				
22		165.3	qC		
23	5.74, s	118.5	CH		C22, C25, C26
24		149.3	qC		
25	2.04, s	19.3	CH_3		C23, C24, C26
26	1.74, s	26.8	CH_3		C23, C24, C25

[a] Assignments made by a combination of proton and carbon 1D and 2D experiments (COSY, TOCSY, DEPT, HMBC, HMQC, etc.).

Table 5. Retention Times (t_R, min) of L- and DL-FDLA-Derivatized Amino Acids in **1** and **2**

	m/z $[M + H]^+$	std FDLA	lucentamycin A L-FDLA	lucentamycin B L-FDLA
L-Leu	426.2	31.94 (L)	31.88	
		44.10 (D)		
L-Trp	499.1	32.60 (L)		32.57
		37.63 (D)		
L-HomoArg	483.2	14.09 (L)	14.18	14.47
		11.90 (D)		
3-methyl-4-ethylidene-proline	450.1		33.45	33.34
			36.94 (D)	

Isolation of Compounds 1−4. The crude extract (1.85 g from 20 L), obtained by resin elution, was adsorbed onto diatomaceous earth (Celite) and subjected to C-18 reversed-phased flash chromatography eluting with a step gradient from 20 to 100% MeOH in H_2O. Compounds **1** and **2** were isolated from the 60% MeOH/H_2O fraction by reversed-phase HPLC chromatography using a C-8 column (250 × 10 mm) eluting with 35% acetonitrile at a flow rate of 2 mL/min with detection at 254 nm by a UV detector. Lucentamycin A eluted at 25 min (**1**, 4.5 mg, 0.24% yield) and lucentamycin B at 32 min (**2**, 3.5 mg, 0.19% yield). Compounds **3** and **4** were isolated from the 40% MeOH/H_2O fraction using a C-8 column eluting with 25% CH_3CN in water at a flow rate of 2 mL/min. Lucentamycin C eluted at a retention time of 42 min (**3**, 6.0 mg, 0.32% yield) and lucentamycin D at 44 min (**4**, 6.5 mg, 0.35% yield).

Lucentamycin A (**1**, 4.5 mg, 0.24% yield): yellow oil; $[\alpha]^{25}_D$ −6.3 (*c* 0.175, MeOH); IR (neat) ν_{vmax} 3374, 1641, 1204, 624 cm^{-1}; UV (MeOH) λ_{max} (log ϵ) 206 (4.1), 217 sh (3.9), 300 (2.7) nm; ^1H (500 MHz, DMSO-d_6) and ^{13}C NMR (125 MHz, DMSO-d_6), see Table 1; HRESITOF $[M + H]^+$ *m/z* 543.3298 and $[M + Na]^+$ 565.3109 ($C_{28}H_{43}N_6O_5$, calcd 543.3295).

Lucentamycin B (**2**, 3.5 mg, 0.19% yield): yellow oil; $[\alpha]^{25}_D$ −15 (*c* 0.100, MeOH); IR (neat) ν_{max} 3374, 1640, 1204, 624 cm^{-1}; UV (MeOH) λ_{max} (log ϵ) 206 (4.1), 209 (3.9), 282 (3.0) nm; ^1H (500 MHz, DMSO-d_6) and ^{13}C NMR (125 MHz, DMSO-d_6), see Table 2; HRESITOF $[M + H]^+$ *m/z* 596.3532 and $[M + Na]^+$ 618.3343 ($C_{31}H_{46}N_7O_5$, calcd 596.3560).

Lucentamycin C (**3**, 6 mg, 0.32% yield): yellow oil; $[\alpha]^{25}_D$ −32 (*c* 0.250, MeOH); IR (neat) ν_{max} 3362, 1640, 1180, 624 cm^{-1}; UV (MeOH) λ_{max} (log ϵ) 206 (4.2), 217 sh (3.9) nm; ^1H (500 MHz, DMSO-d_6) and ^{13}C NMR (125 MHz, DMSO-d_6), see Table 3; HRESITOF $[M + H]^+$ *m/z* 523.3591 and $[M + Na]^+$ 545.3405 ($C_{26}H_{47}N_6O_5$, calcd 523.3607).

Lucentamycin D (**4**, 6.5 mg, 0.35% yield): yellow oil; $[\alpha]_D$ −36 (*c* 0.250, MeOH); IR (neat) ν_{max} 3338, 1638, 1198, 624 cm^{-1}; UV (MeOH) λ_{max} (log ϵ) 206 (4.2), 217 sh (4.0) nm; ^1H (500 MHz, DMSO-d_6) and ^{13}C NMR (125 MHz, DMSO-d_6), see Table 4; HRESITOF $[M + H]^+$ *m/z* 521.3437 and $[M + Na]^+$ 543.3251 ($C_{26}H_{45}N_6O_5$, calcd 521.3451).

Acid Hydrolysis and Advanced Marfey Analysis. A 0.5 mg sample of each compound was subjected to acid hydrolysis at 110 °C for 16

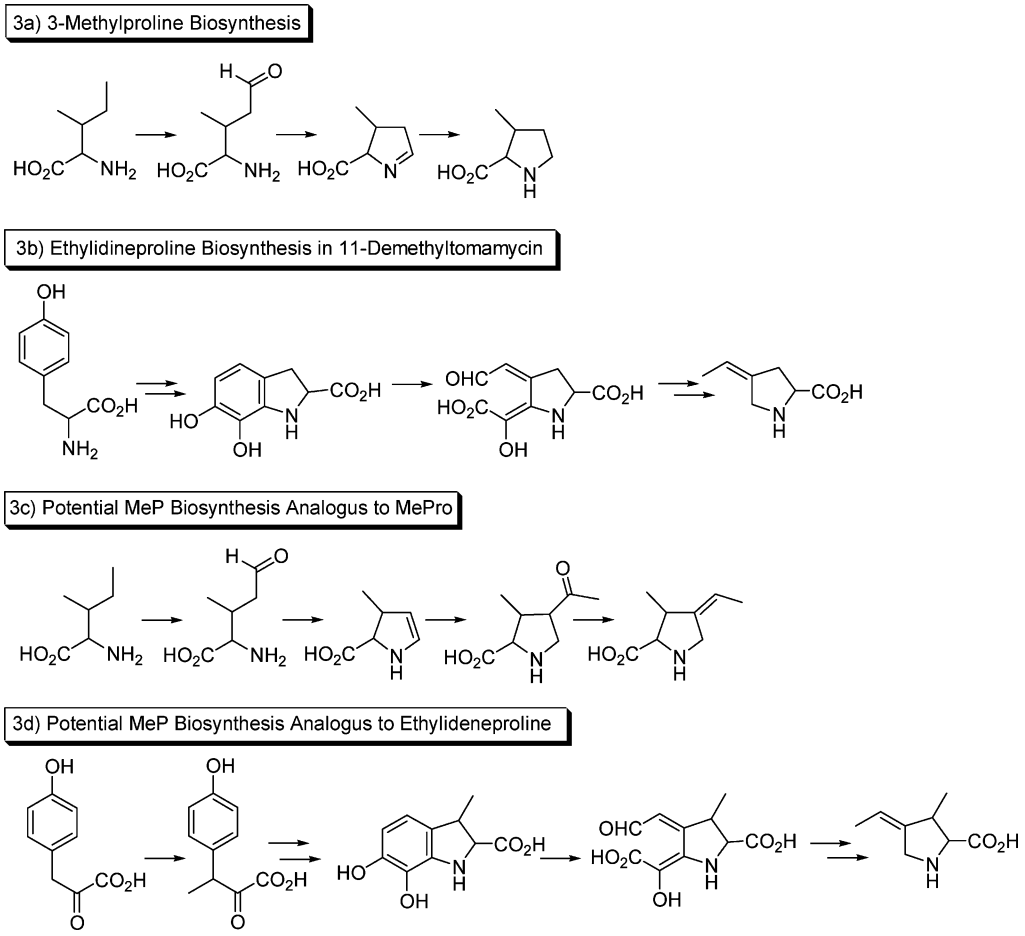


Figure 3. Biosynthetic schemes for modified prolines.

h with 6 N HCl (0.5 mL), and then the hydrolysates were evaporated to dryness and resuspended in H_2O (200 μL). To one portion (100 μL) was added 50 μL of a 1% (v/v) solution of 1-fluoro-2,4-dinitrophenyl-5-L-leucinamide (L-FDLA or D-FDLA) solution in acetone and 1 M $NaHCO_3$ (100 μL), after which the mixtures were heated to 80 °C for 3 min. The reaction mixtures were cooled, neutralized with 2 N HCl (50 μL), and diluted with MeCN (300 μL). About 10 μL of each solution of FDLA derivatives was analyzed on a C8 column (XBD-C8, 3.5 μm, 4.6 × 50 mm) by LC/MS. Aqueous MeCN containing 0.01 M TFA was used as a mobile phase with a linear gradient elution mode (30−60% for 50 min) at a flow rate of 0.7 mL/min. A Hewlett-Packard Series 1100MSD mass spectrometer was used for detection in API-ES (positive) mode. The retention times (min) of the L-FDLA-derivatized standards were L-Leu (31.94 min), L-Har (14.09 min), and L-Trp (32.60 min). The retention times of the D-FDLA-derivatized standards were L-Leu (44.10 min), L-Har (11.90 min), (2*S*,3*S*)-3-Me-4-ethylideneproline (36.94 min), and L-Trp (37.63 min). The retention times of the amino acid constituents of the L-FDLA-derivatized hydrolysate of **1** were L-Leu (31.88 min), L-Har (14.18 min), and 3-Me-4-ethylideneproline (33.45 min). The retention times of the amino acid constituents of the D-FDLA-derivatized hydrolysate of **1** were L-Leu (44.10 min), L-Har (11.90 min), and 3-Me-4-ethylideneproline (36.94 min). The retention times of the amino acid constituents of the L-FDLA-derivatized hydrolysate of **2** were L-Har (14.47 min), 3-methyl-4-ethylideneproline (33.34 min), and L-Trp (32.57 min).

Bioassay. The cytotoxic activity of test samples toward HCT-116 colon adenocarcinoma cells was determined using an *in vitro* assay that quantified cell viability via bioreduction of 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2*H*-tetrazolium (MTS) to formazan in the presence of the electron coupling reagent phenazine methyl sulfate (PMS). In the MTS assay system, the quantity of formazan product in the cell supernatant as measured by the absorbance at 490 nm is directly proportional to the number of living cells in culture, and by serial dilution, the GI_{50} of a test compounds can be determined.[17]

Acknowledgment. This work is the result of generous support from the NIH, National Cancer Institute, under grant CA44848. J.Y.C. was supported, in part, by a postdoctoral fellowship from the Korean Science and Engineering Foundation (KOSEF). The NMR spectrometer used in this research was provided by a grant from the NIH Shared Resources Program (Grant S10RR017768). We thank C. A. Kauffman (SIO), L. Zeigler (SIO), and P. Baran, The Scripps Research Institute, for assistance with the fermentations, bioassays, and IR spectra measurements, respectively.

Supporting Information Available: The ^1H, ^{13}C, and 2D NMR spectra of **1** and ^1H and ^{13}C NMR spectra of **2**−**4** are available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Berdy, J. *J. Antibiot.* **2005**, *58*, 1−26.
(2) Ayuso, A.; Clark, D.; Gonzalez, I.; Salazar, O.; Anderson, A.; Genilloud, O. *Appl. Microbiol. Biotechnol.* **2005**, *67*, 795−806.
(3) Rainey, F. A.; Ward-Rainey, N.; Kroppenstedt, R. M.; Stackebrandt, E. *Int. J. Syst. Bacteriol.* **1996**, *46*, 1088−1092.
(4) Sabry, S. A.; Ghanem, N. B.; Abu-Ella, G. A.; Schumann, P.; Stackebrandt, E.; Kroppenstedt, R. M. *Int. J. Syst. Evol. Microbiol.* **2004**, *54*, 453−456.
(5) Bergy, M. E. *J. Antibiot.* **1968**, *21*, 454−457.
(6) Dolak, L. A.; Castle, T. M.; Laborde, A. L. *J. Antibiot.* **1980**, *33*, 690−694.
(7) Peltola, J. S.; Andersson, M. A.; Kampfer, P.; Augling, G.; Kroppenstedt, R. M.; Busse, H. U.; Salkinoja-Salonen, M. S.; Rainey. F. A. *Appl. Mocrobiol. Biotechnol.* **2001**, *67*, 4293−4304.
(8) Silverstein, R. M.; Bassler, C. G.; Morrill, T. C. *Spectrometric Identification of Organic Compounds*; Wiley & Sons: New York, 1981.
(9) (a) Fujii, K.; Ikai, Y.; Oka, H.; Suzuki, M.; Harada, K. *Anal. Chem.* **1997**, *69*, 5146−5151. (b) Fujii, K.; Ikai, Y.; Mayumi, T.; Oka, H.; Suzuki, M.; Harada, K. *Anal. Chem.* **1997**, *69*, 3346−3352.
(10) FDLA was prepared by the procedure given in ref 10.

(11) Tan, L. T.; Cheng, X. C.; Jensen, P. R.; Fenical, W. *J. Org. Chem.* **2003**, *68*, 8767−8773.
(12) Cuq, J. C.; Vie, M.; Cheftel, J. C. *Food Chem.* **1983**, *12*, 73−88.
(13) For representative examples see: (a) Gray, C. R; Sanz-Cervera, J. F.; Silks, L. A.; Williams, R. M. *J. Am. Chem. Soc.* **2003**, *125*, 14692−12693. (b) Stocking, E. M.; Sanz-Cervera, J. F.; Silks, L. A.; Williams, R. M. *J. Am. Chem. Soc.* **1996**, *118*, 7008−7009.
(14) Hurley, L. H.; Gairola, C.; Das, N. V. *Biochemistry* **1976**, *15*, 3760−3769.
(15) Stocking, E. M.; Sanz-Cervera, J. F.; Unkefer, C. J.; Williams, R. M. *Tetrahedron* **2001**, *57*, 5303−5320.
(16) (a) For examples of β-methylphenylalanine-containing compounds see: Andres, N.; Wolf, H.; Zähner, H.; Rössner, E.; Zeeck, A.; König, W. A.; Sinnwell, V. *Helv. Chem. Acta* **1989**, *72*, 426−437. (b) Houge-Frydych, C. S. V.; Gilpin, M. L.; Skett, P. W.; Tyler, J. W. *J. Antibiot.* **2000**, *53*, 364−372.
(17) Mosmann, T. *J. Immunol. Methods* **1983**, *65*, 55−63.

NP070101B