STRUCTURES OF CEPAFUNGINS I, II AND III

YOSHIHIRO TERUI, JUNKO NISHIKAWA, HIROSHI HINOO,
TOSHIYUKI KATO and JUN'ICHI SHOJI

Shionogi Research Laboratories, Shionogi & Co., Ltd.,
Fukushima-ku, Osaka 553, Japan

(Received for publication December 16, 1989)

Structures of interesting acylpeptide antibiotics cepafungins I, II and III were elucidated by NMR spectroscopic studies and some degradation experiments. The antibiotics contain a common peptide part that consists of threonine and two unusual amino acid residues, γ-hydroxylysine and 4-amino-2-pentenoic acid. The unusual amino acid residues compose an interesting 12-membered ring with an exocyclic *N*-terminus to which the threonine is connected. Different fatty acyl groups connected to the *N*-terminus of the threonine distinguish the three cepafungins. The major component I and minor component III are new substance, but the minor component II has a structure identical with that of the recently reported antibiotic glidobactin A.

The present report describes the structure elucidation of cepafungins I, II and III produced by *Pseudomonas* sp. CB-3[1].

On the Major Component Cepafungin I

Elemental analysis and SI-MS showed the molecular formula of cepafungin I(**1**) to be $C_{28}H_{46}N_4O_6$.

The ^1H NMR analysis in DMSO-d_6, including temperature variation, D_2O addition, homonuclear chemical shift correlation (HOMCOR)[2], and decoupling experiments confirmed the 46 protons of **1** (40 nonexchangeable, 4 amide and 2 hydroxy protons). The ^{13}C NMR analysis, including single frequency

Fig. 1. One-bond C-H HETCOR spectrum of **1** in DMSO-d_6 at 25°C.

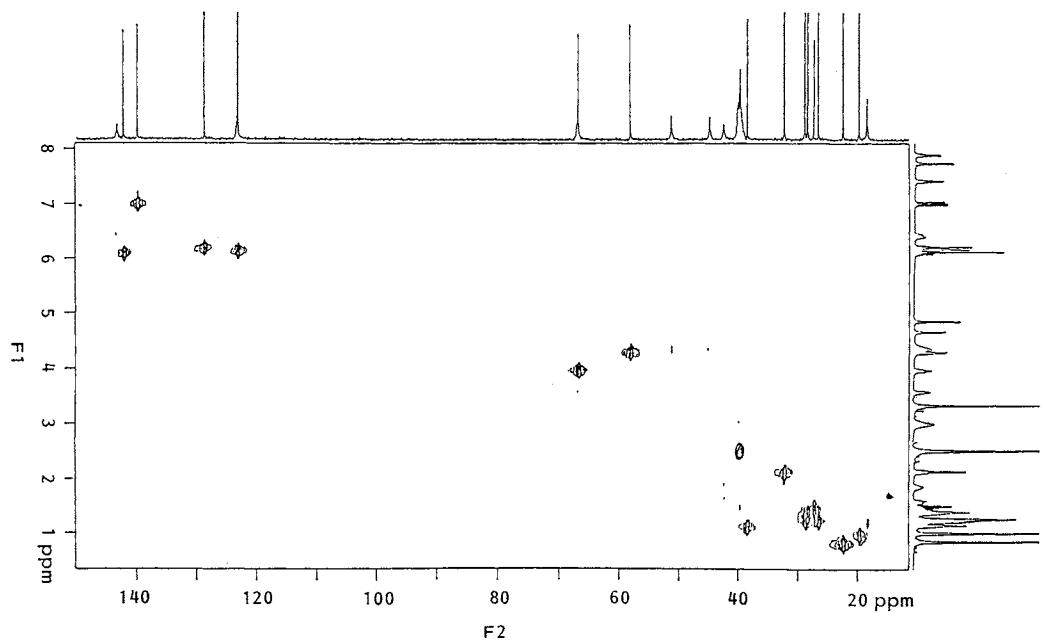
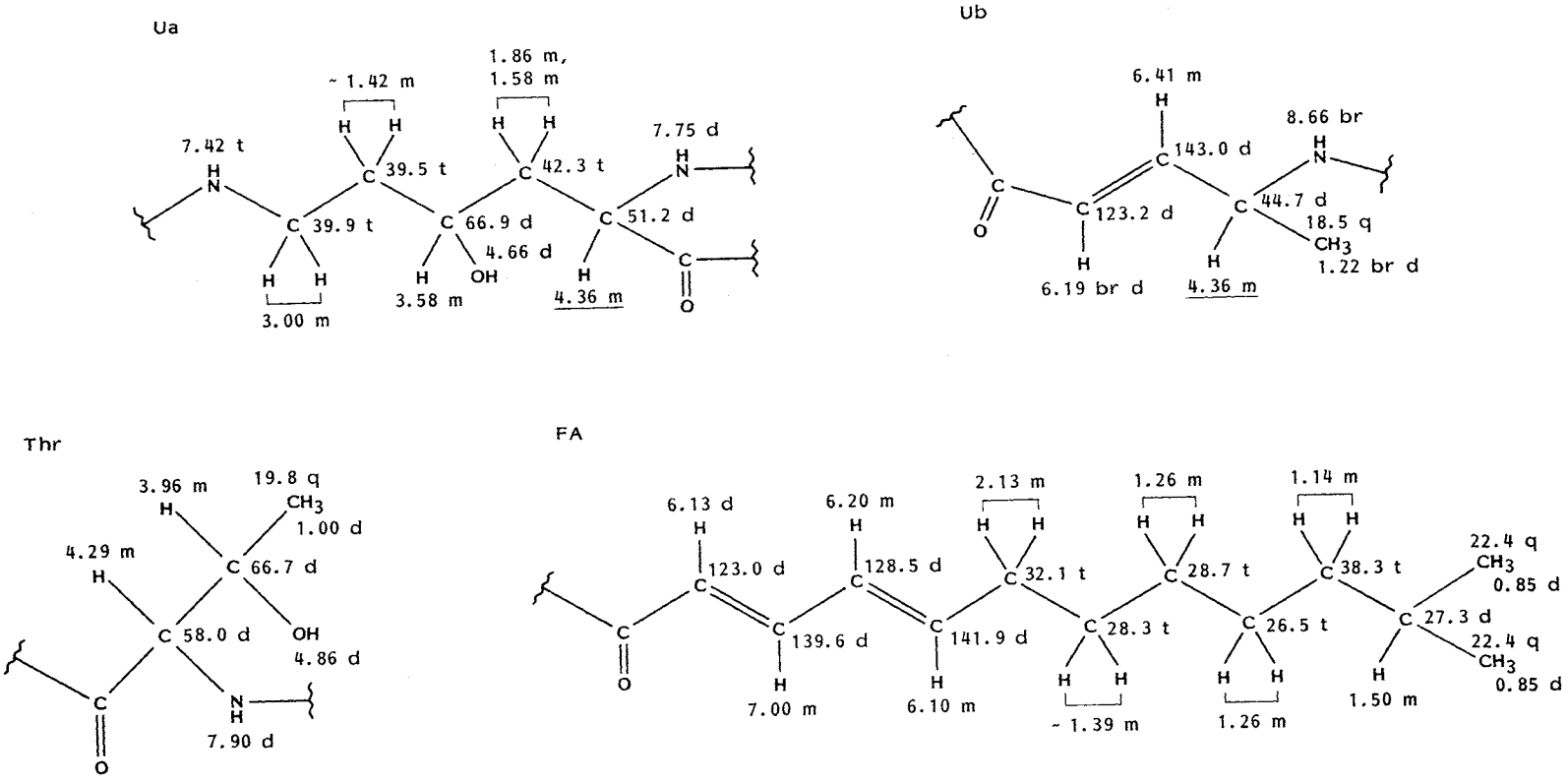


Fig. 2. Partial structures shown by NMR data*.

* With the carbonyls, see text.



off-resonance decoupling and HETCOR[3] experiments, confirmed the 28 carbons, which consisted of 4 methyl, 8 methylene, 12 methine and 4 quaternary carbons. The HETCOR spectrum focusing on one-bond C-H correlation is presented in Fig. 1. Conventional ^{13}C and ^{1}H NMR spectra which cover whole range were presented in the preceding paper[1]. Note that part of the ^{13}C NMR signals exhibited the broadening.

The ^{1}H-^{1}H and ^{13}C-^{1}H connectivities given by the NMR analyses led to substructures Ua, Ub, Thr and FA shown in Fig. 2. Structures Ua and Ub, however, needed confirmation for the following reason. One of the methine protons of Ua and that of Ub exhibited identical chemical shifts, δ 4.36, which caused an uncertainty about the connectivity. The structures were confirmed by the detailed analyses on a diacetyl derivative described later.

Based on the molecular formula, the 4 quaternary carbons of **1**, appearing at δ 165.3, 167.5, 169.3 and 171.0 in DMSO-d_6 at 25°C, were amide carbonyls. Accordingly, the question to be answered was how to connect the parts, *i.e.*, the determination of the sequence. To find the answer, a method examining the long-range C-H correlation should be useful. However, there were problems in using the method for **1** in DMSO-d_6 due to the broadening and signal overlap. In addition, **1** exhibited low solubility in various solvents. Since **1** possessed two secondary hydroxy groups, acetylation with acetic anhydride in pyridine was carried out to obtain diacetylcepafungin I (**1'**), which exhibited improved solubilities for some solvents and enabled us to obtain useful data for the structure analysis. The NMR data obtained in some solvents are listed in Tables 1, 2 and 3. Since **1'** exhibited well separated NMR signals in $CDCl_3$ or $CDCl_3$-CD_3OH (5:1), we could easily confirm the structures of the deduced parts using the connectivity among the signals. Fig. 3 is the HETCOR spectrum optimal for the long-range C-H correlation in $CDCl_3$-CD_3OH (5:1) at 25°C. Using the spectrum, the NMR signals of **1'** were assigned. To confirm the assignments of the carbonyls, we conducted selective decoupling experiments at 60°C with weak irradiation power. The key long-range C-H correlations for answering our question are illustrated in Fig. 4. Thus, the structure having an interesting 12-membered ring composed of two unusual amino acid residues, γ-hydroxylysine and

Table 1. ^{13}C and ^{1}H NMR data for **1'** in $CDCl_3$-CD_3OH (5:1) at 25°C.

Assignment	^{13}C[a]	^{1}H[b]	Assignment	^{13}C[a]	^{1}H[b]
Ua-α-CO	171.5 s		β	70.7 d	5.27 (m)
α-NH		8.05 (br)	γ	16.9 q	1.24 (d, 6.8)
α	51.5 d	4.37 (m)	Ac-CO	170.8 s	
β	39.2 t	2.03 (m), 1.85 (m)	Ac-Me	20.9 q	2.03 (s)
γ	70.1 d	4.88 (m)	FA-1	167.7 s	
δ	37.1 t	1.83 (m), 1.71 (m)	2	120.9 d	5.96 (d, 15.8)
ε	39.6 t	3.21 (m)	3	143.0 d	7.21 (dd, 15.8, 11.5)
ε-NH		6.95 (br)	4	128.5 d	6.18 (dd, 11.5, 15.1)
Ac-CO	172.1 s		5	144.7 d	6.13 (dt, 15.1, 7.0)
Ac-Me	21.1 q	2.07 (s)	6	33.2 t	2.18 (m)
Ub-CO (**1**)	170.2 s		7	29.7 t	1.44 (m)
2	123.2 d	6.42 (d, 16.0)	8	29.0 t	1.30 (m)
3	145.8 d	6.66 (dd, 16.0, 6.8)	9	27.4 t	1.30 (m)
4	45.8 d	4.73 (m)	10	39.2 t	1.17 (m)
4-NH		8.42 (d, 7.2)	11	28.2 d	1.53 (m)
5	18.7 q	1.36 (d, 7.2)	12	22.7 q	0.87 (d, 7.0)
Thr-α-CO	169.4 s		12'	22.7 q	0.87 (d, 7.0)
α-NH		7.45 (d, 8.6)			
α	56.7 d	4.70 (dd, 8.6, 5.8)			

[a] δ and multiplicity.
[b] δ (multiplicity and J/Hz).

4-amino-2-pentenoic acid, was elucidated. Since the broad NMR signals were assigned to the components of the 12-membered ring, a considerably slow conformational exchange of the ring in solution was indicated. The *trans* orientation of all the double bonds was evident from the $J=14.8\sim16.2$ Hz, vicinal ^{1}H-^{1}H coupling constants through the double bond.

On the Minor Components Cepafungins II and III

Structures of cepafungins II and III (**2** and **3**) were also analyzed using the diacetyl derivatives, **2'** and **3'**. Comparison of the NMR data for **2'** and **3'** with that of **1'** enabled us to determine the structures (see Tables 2 and 3). Clearly identifiable signals were found for all parts, especially ^{13}C NMR signals,

Table 2. ^{13}C NMR data for **1'**, **2'** and **3'** in CD_3OD at 25°C.

Carbon	1'	2'	3'	Carbon	1'	2'	3'
	δ and multiplicity				δ and multiplicity		
Ua-CO	172.5 s	172.5 s	172.6 s	Ac-CO	171.7 s	171.7 s	171.8 s
α	52.7 d	52.8 d	52.8 d	Ac-Me	21.1 q	21.0 q	21.1 q
β	39.4 t	39.6 t	39.5 t	FA-1	169.2 s	169.2 s	169.3 s
γ	71.5 d	71.5 d	71.6 d	2	122.2 d	122.2 d	122.2 d
δ	37.7 t	37.8 t	37.8 t	3	143.5 d	143.5 d	143.5 d
ε	40.6 t	40.5 t	40.5 t	4	129.7 d	129.8 d	129.8 d
Ac-CO	172.9 s	172.8 s	172.8 s	5	144.9 d	145.0 d	145.0 d
Ac-Me	21.1 q	21.0 q	21.1 q	6	34.0 t	34.0 t	34.2 t
Ub-1	171.6 s	171.5 s	171.6 s	7	30.5 t	30.3 t	27.9 t
2	124.0 d	124.1 d	124.1 d	8	29.9 t	30.0 t	39.9 t
3	146.3 d	146.2 d	146.3 d	9	28.3 t	30.3 t	29.1 d
4	47.0 d	47.0 d	47.0 d	10	40.0 t	33.0 t	23.0 q
5	18.8 q	18.7 q	18.7 q	10'			23.0 q
Thr-CO	170.8 s	170.7 s	170.8 s	11	29.1 d	23.7 t	
α	57.9 d	57.9 d	57.9 d	12	23.0 q	14.5 q	
β	71.6 d	71.6 d	71.6 d	12'	23.0 q		
γ	17.2 q	17.2 q	17.2 q				

Table 3. ^{1}H NMR data for **1'**, **2'** and **3'** in $CDCl_3$ at 25°C.

Proton	1'	2'	3'	Proton	1'	2'	3'
	δ and multiplicity				δ and multiplicity		
Ua-NH	7.20 d	7.09 d	7.03 d	β	5.30 m	5.31 m	5.31 m
α	4.37 m	4.35 m	4.34 m	γ	1.22 d	1.23 d	1.23 d
β	1.99 m	1.98 m	1.96 m	Ac-Me	2.06 s	2.07 s	2.07 s
	1.93 m	1.93 m	1.94 m	FA-2	5.83 d	5.82 d	5.81 d
γ	5.05 m	5.06 m	5.06 m	3	7.23 m	7.23 m	7.23 m
δ	1.78 m	1.78 m	1.78 m	4	6.14 m	6.14 m	6.14 m
	1.70 m	1.70 m	1.69 m	5	6.12 m	6.12 m	6.12 m
ε	3.22 m	3.22 m	3.22 m	6	2.16 m	2.16 m	2.15 m
ε-NH	5.86 t	5.74 t	5.63 t	7	1.42 m	1.42 m	1.49 m
Ac-Me	2.12 s	2.13 s	2.13 s	8	1.28 m	1.28 m	1.18 m
Ub-2	6.44 d	6.43 d	6.43 d	9	1.28 m	1.28 m	1.54 m
3	6.71 m	6.72 m	6.72 m	10	1.16 m	1.28 m	0.88 d
4	4.97 m	5.00 m	5.00 m	10'			0.88 d
4-NH	7.47 d	7.41 d	7.38 d	11	1.52 m	1.28 m	
5	1.38 d	1.39 d	1.39 d	12	0.86 d	0.89 t	
Thr-NH	6.45 d	6.34 d	6.30 d	12'	0.86 d		
α	4.69 m	4.64 m	4.61 m				

except for some differences in the FA part. No signals assigned to the isopropyl type terminal were found in the hydrocarbon chain of **2'**. Instead, there were signals indicating an ethyl type terminal. With **3'**, the signals indicating the isopropyl type terminal exsited. Compared with **1'**, however, **3'** lacked two methylene signals in the hydrocarbon chain. Thus, the structures of **2'** and **3'** (*i.e.*, those of **2** and **3**) were determined. Hydrolysis of the diacetyl derivatives with weak alkali regenerated **1**, **2** and **3**. The structures are shown in Fig. 5. The SI-MS data are consistent with the structures: **1** ($(M+H)^+$, 535), **2** ($(M+H)^+$, 521), and **3** ($(M+H)^+$, 507).

On the Absolute Configuration of Asymmetric Carbons

The absolute configurations of asymmetric carbons were determined by the chemical evidence shown in Scheme 1. L-Threonine was isolated and characterized, establishing the *S*-configuration at the α-carbon and the *R*-configuration at the β-carbon of substructure Thr. The γ-hydroxylysine obtained by the acid hydrolysis gave a retention time and CD spectrum similar to those of the reference compound, L-*threo*-γ-hydroxylysine[4], but exhibited an NMR spectrum clearly different from that of the *threo* isomer. This substance was therefore L-*erythro*-γ-hydroxylysine, which indicated the *S*-configuration at the α- and γ-carbons of the Ua fragment. To determine the configuration of the Ub-4 carbon, oxidative cleavage of the double bond with sodium periodate and potassium permanganate was carried out, followed by acid hydrolysis. Amino acid analysis confirmed L-alanine indicating the *S*-configuration of the Ub-4 carbon.

Alkaline hydrolysis of **1** gave 11-methyl-2(*E*),4(*E*)-dodecadienoic acid, which confirmed the structure of the FA fragment of **1**.

Fig. 3. Long-range C-H HETCOR spectra of 1'*.

* Carbonyl region.

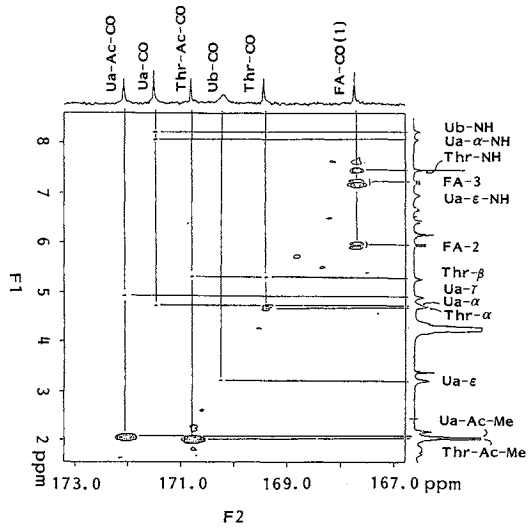


Fig. 4. The key long-range C-H correlations for determining the assignment of carbonyl signals and sequence*.

* The correlations detected by HETCOR and those confirmed by selective decoupling are shown by solid arrows and broken arrows, respectively

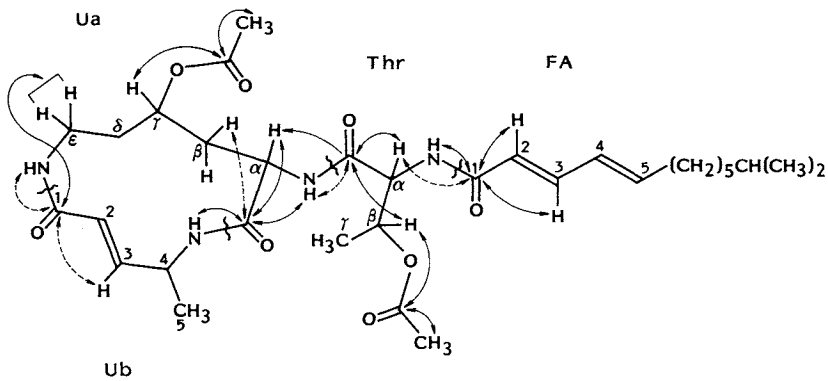
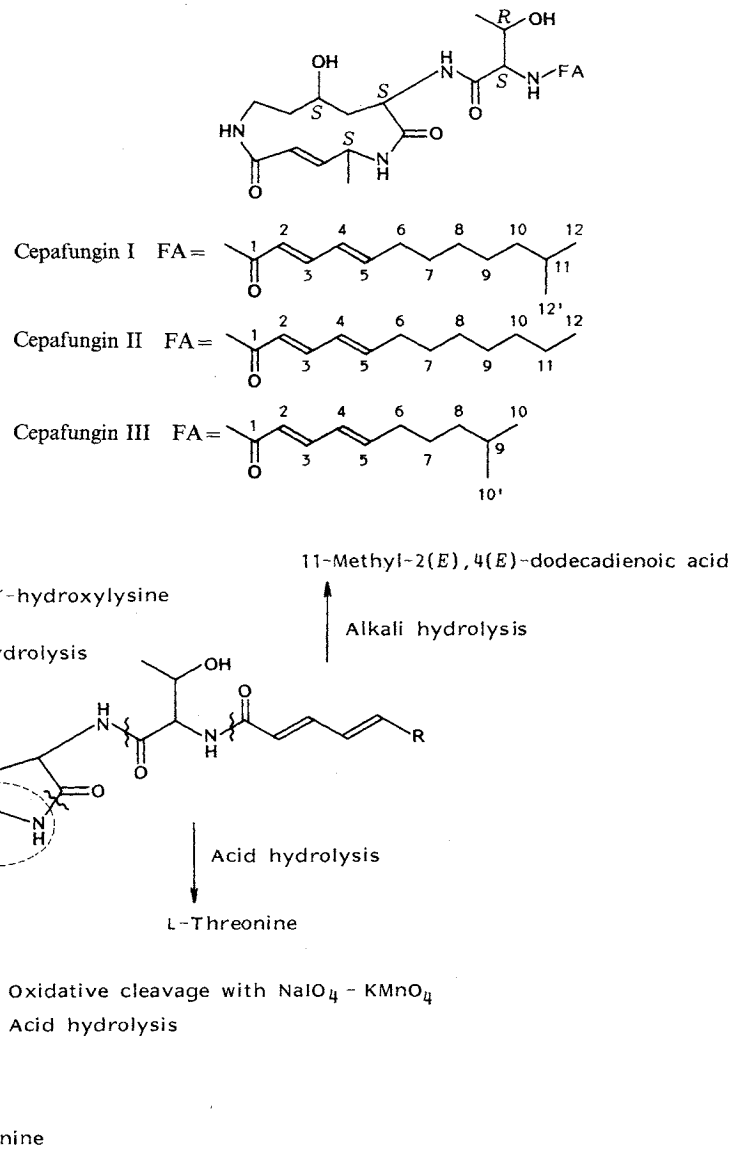


Fig. 5. Structures of cepafungins I, II and III.



The identical ^{13}C NMR data of **1'**, **2'** and **3'**, differences in the hydrocarbon chain excepted, support identical configurations.

Thus, the structures of cepafungins I, II and III were elucidated.

Inspection of the antibiotics literature indicates that the major component I and minor component III are new substance, and that the minor component II has a structure identical with that of the recently reported antibiotic glidobactin A[5,6].

Experimental

Spectral Measurements

The ^1H and ^{13}C NMR spectra were recorded with a Varian XL-400 spectrometer operating at 399.948

MHz for ^1H and at 100.579 MHz for ^{13}C. In HETCOR experiments, the J values, 140 and 7 Hz, were used to detect the one-bond and long-range C-H correlation. In all the measurements, TMS was used as an internal reference.

The CD and MS spectra were recorded with a Jasco J-40C automatic recording spectropolarimeter and a Hitachi M-68 mass spectrometer, respectively.

Materials

Preparation of Diacetylcepafungins I, II and III (**1'**, **2'** and **3'**): To a suspension of 90 mg of cepafungin complex consisting of **1**, **2** and **3** (*ca.* 3 : 1 : 1) in 3 ml of pyridine, acetic anhydride (0.2 ml) was added. The mixture was stirred at room temperature for 1 hour and became homogeneous. The reaction mixture was kept at 4°C for 16 hours, poured into ice water, and then extracted with EtOAc at pH 3.0. The EtOAc layer was washed with dilute HCl, water, dilute $NaHCO_3$, and then water. After drying with Na_2SO_4, the EtOAc solution was evaporated and a colorless powder (94 mg) was obtained. A portion of the product (70 mg) was separated by HPLC using a Nucleosil 10 C_{18} column (10 × 300 mm) and 60% acetonitrile as the mobile phase. The peaks of **1'**, **2'** and **3'** appeared in that order. The fractions corresponding to the peaks were collected. Each fraction was concentrated under reduced pressure to a nearly aqueous solution that was extracted with EtOAc at pH 3.0. The extract was washed successively with water, dilute $NaHCO_3$ and water and then dried with Na_2SO_4. Evaporation of the EtOAc gave colorless powders of **1'** (32 mg), **2'** (11 mg) and **3'** (10 mg).

Isolation of Fatty Acids: Cepafungin complex (200 mg) was hydrolyzed with 1 N NaOH at 110°C for 17 hours. The hydrolysate was acidified with 6 N HCl to pH 3.0 and extracted with EtOAc. The extract was water-washed, dried with Na_2SO_4 and concentrated to give a brown oily substance that was subjected to preparative TLC on a Merck precoated Silica gel F254 plate, $CHCl_3$ - MeOH (9 : 1). A hydrophobic zone, Rf 0.45, detected by spraying with water, was extracted with $CHCl_3$ - MeOH (1 : 1). The extract was, evaporated and then dissolved in EtOAc. The EtOAc solution was washed with 0.1 N HCl and then water dried with Na_2SO_4 and evaporated to give an oily substance (*ca.* 30 mg). When it was subjected to HPLC on a Nucleosil 10 C_{18} column (10 × 300 mm), eluting with acetonitrile - 20 mM K_2HPO_4 (4 : 6), three peaks were obtained. The substances of the peaks were named FA-1, FA-2 and FA-3 according to their retention times. Each peak fraction was collected and evaporated at pH 7.0 to a nearly aqueous solution that was extracted with EtOAc at pH 3.0. Each extract was water-washed, dried and evaporated to give pale yellow substances. FA-1 (6 mg), FA-2 (2 mg) and FA-3 (3 mg) were obtained. The major fraction FA-1 was determined to be 11-methyl-2,4-dodecadienoic acid from the ^1H and ^{13}C NMR data in $CDCl_3$ at 25°C: ^1H, δ 7.34 (m, 3-H), *ca.* 6.20 (m, 4-H, 5-H), 5.78 (d, J=15.5 Hz, 2-H), 2.19 (m, 6-H_2), 1.49 (m, 11-H), 1.45 (m, 7-H_2), *ca.* 1.28 (m, 8-H_2, 9-H_2), 1.17 (m, 10-H_2) and 0.86 (d, J=6.7 Hz, 12-H_3, 12'-H_3); ^{13}C, δ 172.2 (s, C-1), 147.6 (d, C-3), 146.3 (d, C-5), 128.2 (d, C-4), 118.1 (d, C-2), 38.9 (t, C-10), 33.1 (t, C-6), 29.4 (t, C-7), 28.7 (t, C-8), 27.9 (d, C-11), 27.2 (t, C-9) and 22.6 (q, C-12, C-12'). Alkaline hydrolysis followed by HPLC experiments showed that **1'**, **2'** and **3'** produced FA-1, FA-2 and FA-3, respectively.

Detection of Amino Acids: 1. Threonine and γ-Hydroxylysine: Cepafungin complex (60 mg) was hydrolyzed with 6 N HCl at 110°C for 7 hours. The hydrolysate, after elimination of oily substances by EtOAc extraction, was subjected to preparative chromatography on paper (Toyo Roshi No. 51A) with BuOH - AcOH - H_2O (4 : 1 : 2). Two ninhydrin positive zones (Rf 0.15 and 0.31) were cut out and extracted with 50% aq MeOH. Each extract was adsorbed on a Dowex 50X8 (NH_4) column and eluted with 0.5 N NH_4OH. Concentration of the eluates to dryness gave an Rf 0.15 substance (17 mg) and an Rf 0.31 substance (8 mg). CD and NMR data for the former indicating L-*erythro*-γ-hydroxylysine were: CD (*c* 0.1579, 0.5 N HCl); $[\theta]_{196}$ +3,080, $[\theta]_{206}$ +4,770, $[\theta]_{250}$ 0; ^1H NMR (HCl salt, in D_2O at 25°C); 4.03 (m, γ-H), 3.88 (*t*-like, α-H), 3.12 (m, ε-H_2), 2.13 (m, β-H(a)), 1.89 (m, β-H(b), δ-H(a)), 1.82 (m, δ-H(b)). The later coincided with threonine in amino acid analysis and in paper chromatography with Toyo Roshi No. 51A and BuOH - acetone - 28% NH_4OH - H_2O (8 : 1 : 1 : 6) to distinguish threonine from allo-threonine[7]. The CD data (*c* 0.1451, 0.5 N HCl) indicating the L series were: $[\theta]_{197}$ +2,600, $[\theta]_{208}$ +3,770, $[\theta]_{250}$ 0.

2. Alanine: Compound **1'** (6 mg) was dissolved in 1.2 ml of *tert*-BuOH. To the solution 0.2 ml of 1 M Na_2CO_3, 0.2 ml of 0.36 M aq $NaIO_4$, and 0.5 ml of 0.05 M aq $KMnO_4$ were added, and the mixture was stirred at 27°C for 16 hours. To the reaction mixture was added 3 ml of BuOH. Precipitates were removed by centrifugation, and the supernatant liquid was concentrated under reduced pressure. The residue was

hydrolyzed with 6N HCl at 110°C for 20 hours. The hydrolysate was subjected to preparative chromatography on Toyo-roshi No. 51A paper with PrOH - pyridine - AcOH - H_2O (15 : 10 : 6 : 12). The ninhydrin positive zone (Rf 0.55) was cut out and extracted with 50% aq MeOH. The extracted substance was subjected to chromatography again on Toyo-roshi No. 51A paper with phenol - H_2O (7 : 3), and the ninhydrin positive zone (Rf 0.46) was cut out and extracted with 50% aq MeOH. The extracted substance was adsorbed on Dowex 50X8 (NH_4) column at pH 2.0. After washing with water, the substance was eluted with 0.3 N NH_4OH. Amino acid analysis confirmed two components, alanine and threonine (ratio, 1.0 : 0.3). HPLC examination using the column MCI-gel CRS 10W (4.6 × 50 mm), mobile phase 0.1 mM aq $CuSO_4$ and flow rate 0.45 ml/minute confirmed that retention time of the major component was identical with that of L-alanine[8].

References

1) SHOJI, J.; H. HINOO, T. KATO, T. HATTORI, K. HIROOKA, K. TAWARA, O. SHIRATORI & Y. TERUI: Isolation of cepafungins I, II and III from *Pseudomonas* species. J. Antibiotics 43: 783~787, 1990
2) BAX, A.; R. FREEMAN & G. MORRIS: Correlation of proton chemical shifts by two-dimensional fourier transform NMR. J. Magn. Reson. 42: 164~168, 1981
3) BAX, A. & G. A. MORRIS: An improved method for heteronuclear chemical shift correlation by two-dimensional NMR. J. Magn. Reson. 42: 501~505, 1981
4) SHOJI, J. & H. HINOO: Chemical characterization of new antibiotics, cerexins A and B. (Studies on antibiotics from the genus *Bacillus*. II.). J. Antibiotics 28: 60~63, 1975
5) OKA, M.; K. YAGINUMA, K. NUMATA, M. KONISHI, T. OKI & H. KAWAGUCHI: Glidobactins A, B and C, new antitumor antibiotics. II. Structure elucidation. J. Antibiotics 41: 1338~1350, 1988
6) OKA, M.; H. OHKUMA, H. KAMEI, M. KONISHI, T. OKI & H. KAWAGUCHI: Glidobactins D, E, F, G and H; minor components of the antitumor antibiotic glidobactin. J. Antibiotics 41: 1906~1909, 1988
7) SHAW, K. N. F. & S. W. FOX: Stereochemistry of the β-phenylserines: Improved preparation of allophenylserine. J. Am. Chem. Soc. 75: 3421~3424, 1953
8) KATOH, H.; T. ISHIDA, Y. BABA & H. KINIWA: Optical resolution of amino acids, peptides and hydroxy-carboxylic acids using a new chiral column for ligand-exchange chromatography. J. Chromatogr. 473: 241~250, 1989