Antibiotics A21459 A and B, New Inhibitors of Bacterial Protein Synthesis

II. Structure Elucidation

P. FERRARI, K. VÉKEY†, M. GALIMBERTI, G. G. GALLO,
E. SELVA and L. F. ZERILLI*

Marion Merrell Dow Research Institute, Lepetit Research Center,
Gerenzano (VA), Italy

(Received for publication June 1, 1995)

The structures of the antibiotics, active against a few Gram-negative bacteria and *Clostridium difficile*, were determined on the basis of physicochemical analyses on the intact molecules and on the acid hydrolysate of A21459 A. FAB-MS and ^{1}H and ^{13}C NMR investigations identified the amino acid units and determined their sequence.

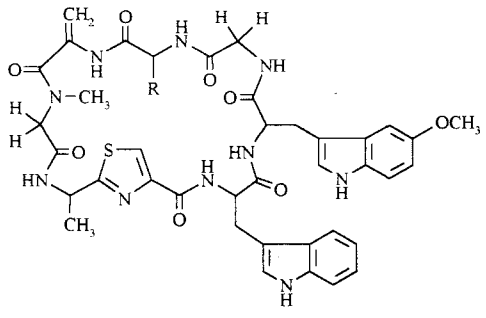
Antibiotics A21459 A and B are homodetic cyclic peptides constituted by eight amino acid units. They are glycine, methoxytryptophan, tryptophan, cysteine, alanine, sarcosine, dehydroalanine, and α-aminobutyric acid for A21459 A (alanine for A21459 B). Cysteine and alanine condensed to form a thiazole moiety, according to the biosynthesis of thiazole containing antibiotics.

Antibiotic A21459 A and B are new protein synthesis inhibitors, active against a few Gram-negative bacteria and *Clostridium difficile*. They have structures shown below. Their structure elucidation, which was based on physico-chemical analyses on the intact antibiotics and on their acid hydrolysate, is reported here.

Materials and Methods

Isolation and Purification

A21459 was extracted with methanol from the mycelium as a complex of two antibiotics, A and B, present in about equal amounts[1]. The two antibiotics, after separation by flash chromatography and purification[1], were obtained as white powders.



(1) R = -CH_2-CH_3 A21459 A
(2) R = -CH_3 A21459 B

UV and IR Spectra

UV absorption spectra were recorded with a Perkin-Elmer spectrophotometer mod. 320 in 0.1 N HCl or 0.1 N KOH. IR absorption spectra were obtained from a mineral oil suspension and in chloroform solution with a Perkin Elmer spectrophotometer mod. 580.

Potentiometric Titration and Elemental Analysis

Sample solutions in methylcellosolve - water 4 : 1 were titrated with 0.1 N KOH or with 0.1 N HCl. Titration in non-aqueous medium was carried out in glacial acetic acid and in pyridine with 0.1 N $HClO_4$ and tetrabutylammonium hydroxide, respectively. Elemental analyses were carried out by the combustion method using a C. Erba elemental analyzer mod. 1106.

Mass Spectrometry

Fast atom bombardment mass spectra (FAB-MS) were obtained in positive ion mode on a Kratos MS-50 double focusing mass spectrometer of 3000 u mass range, using 8 kV accelerating voltage, a saddle field atom gun (with Xe gas) at 6 kV voltage, and 1 mA current. The instrument was operating under computer control. DS-90 data system was used in "raw-data acquisition" mode. Using a mixture of CsI and NaI as a calibration compound[2], the accuracy of mass determination in the scanning mode was increased to about 0.1 u. Samples were dissolved in dimethylformamide containing 0.1 N acetic acid. 1 μl of this solution was mixed with 1 μl of thioglycerol matrix on the target.

† Present address: Central Research Institute for Chemistry, Hungarian Academy of Sciences, P.O. Box 17, H.1525 Budapest (Hungary).

GC-MS of Acid Hydrolysate

The antibiotic (2 mg) was hydrolysed at 105°C for 20 hours with 6 N HCl containing 1% phenol. The acid hydrolysate was derivatized in two steps: n-propyl esters were prepared by incubation at 90°C for 1 hour with 2 M HCl in anhydrous propanol; amino groups were derivatized subsequently in pentafluoropropionic anhydride - anhydrous dichloromethane 10 : 90 (v/v). GC-MS analysis was performed with a Hewlett Packard HP5985B GC-MS system equipped with a fused silica capillary column n-propionyl-L-valine t-butylamide polysiloxane-coated (25 m × 0.2 mm i.d.; OS6411, C.G.C. Analytic). Temperature program was: 80°C for 4 minutes, then increasing at 4°C per minute. Standard amino acids were used as external standards.

^{1}H and ^{13}C NMR Studies

^{1}H and ^{13}C NMR spectra were recorded in $CDCl_3$ solution at 22°C on a Bruker AM-500 or on a Bruker AM-250 spectrometer, equipped with an Aspect 3000 computer[3]. Phase-sensitive double quantum filter ^{1}H-^{1}H COSY spectra were run using time-proportional-phase-incrementation in f_1[4]. In the relayed-coherence-transfer (RCT) ^{1}H-^{1}H COSY[5a] delay was 50 ms, while in the long range COSY[5b] the fixed delay was 200 ms. HETCOR[5c] was obtained using composite 180° ^{13}C pulse (90-240-90°) and ^{1}H decoupling in f_1 domain. COLOC experiment[6] was obtained using the delays D1 = 40 ms and D2 = 30 ms.

Results and Discussion

Basic Physico-chemical Data

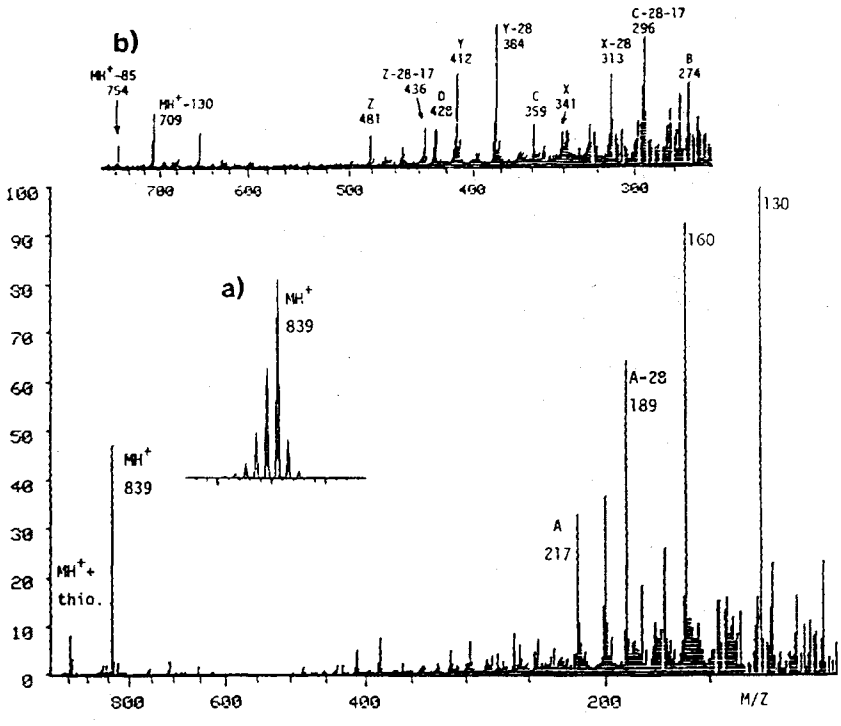
Preliminary physico-chemical analyses showed that A21459 A (**1**) and B (**2**) were structurally similar.

Investigation was initially focused on **1**. Its FAB mass spectrum (Fig. 1) showed a characteristic peak at m/z 839.3 ± 0.1. After addition of KI and CsI to the sample, peak shifts of 38 and 132 u were observed corresponding to MK^+ and MCs^+ ions, respectively. This indicated that the 839 peak referred to the protonated molecular ion, MH^+. The isotope pattern of this ion indicated the possible presence of some other element besides C, H and N. High resolution (13.000) peak matching[2] was applied to the MCs^+ peak and the precise mass of the molecular ion was calculated. The lowest molecular isotope resulted m/z 838.323 ± 0.003, which corresponds to the elemental formula $C_{41}H_{46}N_{10}O_8S$. This was consistent with the elemental analysis (58.92% C, 5.72% H, 16.21% N, and 3.77% S) and with the presence of 41 C and 46 H atoms, as determined by ^{1}H and ^{13}C NMR spectra.

IR spectra (Table 1) were typical of a peptide structure. This was confirmed by deuteration shift of the amide bands (νNH, Amide I, Amide II) upon addition of heavy water to the deuterated chloroform solution. No free

Fig. 1. FAB mass spectrum of A21459 A.

The molecular ion part (a) and the central part (b) of the spectrum are shown enlarged. Letters indicate fragmentation processes (see Fig. 4).



carboxyl and amino groups were detectable.

Absence of any ionizable function was confirmed by ionization studies in water and in non-aqueous solvents.

Identification of Amino Acid Moieties

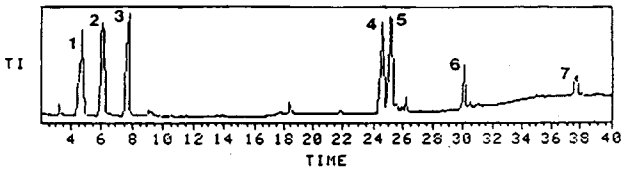
Compound **1** was hydrolysed with acid under standard conditions for peptides. GC-MS analysis of the hydrolysate (Fig. 2) revealed the presence of glycine, tryptophan, α-aminobutyric acid, an apparent alanine isomer, and two peaks consistent with racemized thiazole-alanine amino acid moieties. Thiazole-alanine units are resistant to classical acid hydrolysis with 6 N HCl[7]. Tryptophan was detected both as mono- and bis-pentafluoropropionic derivatives (Table 2). ^{1}H and ^{13}C NMR showed the presence of 5-methoxytryptophan moiety, which was not detected into the hydrolysate, probably because of its instability under the hydrolysis conditions. The presence of two tryptophan moieties appeared consistent with the values of molar absorption at the two absorption frequencies expected[8] for two tryptophan chromophores: ε_{mol} 10040 at 280 nm and 8380 at 290 nm. Accordingly, ions at *m/z* 130 and 160 observed by FAB-MS were attributed to methylene-indole and methylene-methoxyindole fragments. Furthermore, IR absorptions at 3470 and 3390 cm^{-1} were consistent with the heterocyclic NH's of the two tryptophan moieties.

Complete identification of amino acid residues by NMR analyses was based on ^{1}H-^{1}H correlation (COSYPHDQ[4], COSYRCT[5a], COSYLR[5b]), ^{1}H-^{13}C correlation through one bond (XHCORRDC[5c]), and

Table 1. Attributions of the main functional IR absorption bands of A21459 A in mineral oil suspension and in chloroform solution.

ν, (cm^{-1})		Assignment
Nujol	Chloroform	
3380	3470 & 3390	Tryptophan heterocyclic νNH
3300	3300	Peptide νNH
1650	1695~1630	Amide I (νC=O)
1530	1530	Amide II (δNH)
1505~1490	1505~1490	Aromatic νC=C
1255	1250	Aromatic δCH
740	—	Aromatic γCH

Fig. 2. GC-MS tracings of A21459 A hydrolysate showing derivatized amino acids: total ion current (TI) *vs.* time (minutes).



Relevant MS data are reported in Table 2.

Table 2. Relevant GC-MS data of derivatized amino acids obtained by hydrolysis from A21459 A.

Peak	Amino acid	MW	Characteristic ions: *m/z* (attribution) rel. int. %
1	Alanine isomer	277	236 (M−C_3H_5) 6; 218 (M−OC_3H_7) 3; 190 (M−$CO_2C_3H_7$) 100
2	α-Aminobutyric acid	291	250 (M−C_3H_5)$_5$; 232 (M−OC_3H_7) 2; 204 (M−$CO_2C_3H_7$) 100
3	Glycine	263	222 (M−C_3H_5) 9; 204 (M−OC_3H_7) 7; 176 (M−$CO_2C_3H_7$) 100
4	Thiazole-alanine enantiomer	360	360 (M) 15; 301 (M−OC_3H_7) 17; 272 (M−$CO_2C_3H_8$) 100
5	Thiazole-alanine enantiomer	360	360 (M) 14; 301 (M−OC_3H_7) 18; 272 (M−$CO_2C_3H_8$) 100
6	Tryptophan monoderivative	392	392 (M) 20; 305 (M−$CO_2C_3H_7$) 100; 130 (C_9H_8N) 7
7	Tryptophan bis-derivative	538	451 (M−$CO_2C_3H_7$) 1; 375 (M−$NH_2COC_2F_5$) 5; 276 ($C_{12}H_7NOF_5$) 2; 130 (C_9H_8N) 100

Table 3. Identification of the moieties of A21459 A by attribution of ^{1}H NMR signals*.

Tryptophan	10.64 (N^1-H); 8.60 (αNH); 7.05 (C^7-H); 6.96 (C^2-H); 6.88 (C^6-H); 6.35 (C^5-H); 5.41 (C^4-H); 5.06 (αCH); 3.55 & 2.85 (CH_2)
Thiazole	8.05 (C^5-H)
(thiazole)-Alanine	8.83 (NH); 5.48 (αCH); 1.72 (CH_3)
Sarcosine	4.97 & 3.40 (CH_2); 3.09 (CH_3)
Dehydroalanine	9.35 (NH); 5.00 & 4.71 (CH_2)
α-Aminobutyric acid	6.82 (NH); 3.98 (αCH); 2.00 & 1.88 (CH_2); 0.88 (CH_3)
Glycine	4.54 (NH); 3.48 & 1.05 (CH_2)
5-Methoxytryptophan	8.80 (αNH); 8.57 (N^1-H); 7.34 (C^4-H); 7.33 (C^7-H); 6.90 (C^6-H); 6.83 (C^2-H); 4.33 (OCH_3); 3.48 & 3.32 (CH_2)

* Chemical shift δ, ppm and (attribution).

Table 4. Attribution of the ^{13}C NMR signals of antibiotics A21459 A and B.

Attribution	Carbon	δ (ppm) A	δ (ppm) B
Tryptophan	C=O	172.93	172.87
	αCH	52.12	52.15
	CH_2	26.66	26.60
	C-2	125.51	125.47
	C-3	105.71	105.78
	C-4	115.93	116.03
	C-5	119.23	119.24
	C-6	121.22	121.21
	C-7	111.37	111.33
	C-8	134.78	134.76
	C-9	126.58	126.58
Thiazole-alanine	C=O	159.80	159.83
	C-2	170.81	170.92
	C-4	150.40	150.45
	C-5	122.77	122.84
	αCH	45.23	45.32
	CH_3	20.48	20.37
Sarcosine	C=O	166.84	166.83
	CH_2	51.01	51.07
	CH_3	37.24	37.40
Dehydroalanine	C=O	168.25	168.30
	C=(CH_2)	136.92	136.89
	CH_2	99.40	99.64
α-Aminobutyric acid or alanine	C=O	169.29	169.83
	αCH	54.37	48.37
	CH_2	21.11	—
	CH_3	10.47	13.95
Glycine	C=O	171.38	171.03
	CH_2	40.51	40.50
5-Methoxytryptophan	C=O	170.03	170.09
	αCH	57.80	57.78
	CH_2	26.95	26.96
	C-2	123.69	123.73
	C-3	108.33	108.13
	C-4	123.64	123.54
	C-5	152.35	152.35
	C-6	101.31	101.16
	C-7	106.61	106.69
	C-8	138.41	138.51
	C-9	117.39	117.42
	OCH_3	56.13	56.10

^1H-^{13}C correlation through two and three bonds (COLOC[6]). Thus, the presence of glycine and α-aminobutyric acid was confirmed. Furthermore, tryptophan, 5-methoxytryptophan, sarcosine and dehydroalanine were identified. As far as the thiazole-alanine moiety is concerned, the proton of thiazole showed connectivities with all carbons in the heterocycle and the carbon C-2 with the methyl group of alanine. The chemical shift of thiazole C-4[9] ($\delta = 150.40$ ppm) indicated that the carbonyl group at 159.8 ppm is in α position.

Complete ^1H and ^{13}C NMR data are reported in Tables 3 and 4, respectively, ordered for single amino acid units.

Fig. 3. Structure of A21459 A showing the sequencing process by COLOC (arrows indicate coupling) and the amino acidic moieties constituting the cyclic peptide (dotted lines).

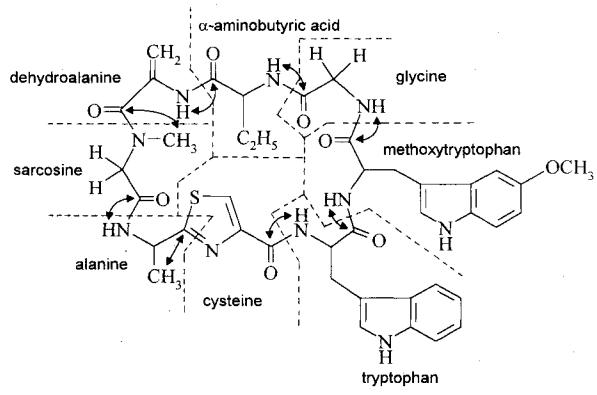
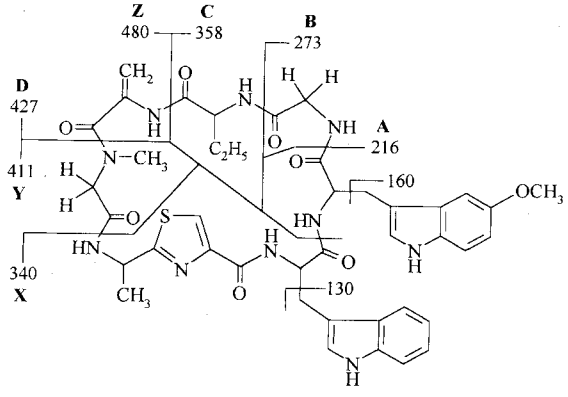


Fig. 4. Suggested FAB-MS fragmentations of A21459 A.



The fragment ions appear in the spectrum in Fig. 1 at m/z values one unit higher than those calculated from the present scheme.

Sequence Determination

Linkage of amino acid moieties of **1** was obtained from long range coupling between the peptidic carbonyl of a given amino acid and the peptidic NH (or N-CH_3) of the adjacent amino acid. Connectivities were observed between NH of thiazole-alanine and CO of sarcosine, between NCH_3 of sarcosine and CO of dehydroalanine, and so on. The entire sequence was defined as shown in Fig. 3.

While FAB-MS of cyclic peptides usually do not give significant fragmentation, FAB-MS of **1** gave information on the sequence, as indicated by the scheme in Fig. 4. The main fragmentation processes were consistent with literature data[10,11] and were often accompanied by further losses of CO and NH_3 (see Fig. 1).

Structure Determination of Antibiotic B

Considering B (**2**), the ^1H NMR signals at δ 1.42 ppm (CH_3), 4.22 ppm (αCH) and 6.88 ppm (NH) indicated that alanine in **2** substitutes for α-aminobutyric acid in **1** (Table 3). ^{13}C NMR signals at δ 169.83 ppm (C=O), 48.37 ppm (αCH), and 13.95 ppm (CH_3) (Table 4) were in agreement with this assignment. The remaining parts of the ^1H and ^{13}C NMR spectra of **1** and **2** were practically identical (Tables 3 and 4). FAB-MS indicated a molecular weight 14 u less than that of **1**. These data demonstrated that the structures of **1** and **2** are distinct in a single amino acid.

Conclusion

The complete identification of all the amino acid residues and their sequence was determined both by chemical hydrolysis and spectroscopic methods on the intact antibiotics.

Antibiotics A21459 A and B are homodetic cyclic peptides. Their structures are composed of eight amino acids. The thiazole-alanine amino acid is likely derived by amide formation between the NH_2 of cysteine and the carboxyl group of alanine, followed by condensation of this amide with the thiol group and dehydrogenation of thiazoline, *i.e.*, in analogy with the biosynthesis of thiazole containing antibiotics[12~14].

References

1) SELVA, E.; L. GASTALDO, G. S. SADDLER, G. TOPPO, P. FERRARI, G. CARNITI & B. P. GOLDSTEIN: Antibiotics A21459 A and B, new inhibitors of bacterial protein synthesis. I. Taxonomy, isolation and characterization. J. Antibiotics 49: 145~149, 1996
2) VÉKEY, K.: Calibration in positive and negative ion fast atom bombardment using salt mixtures. Org. Mass Spectrometry 24: 183~185, 1989
3) Bruker NMR software manual for Aspect 3000 computer and Bruker user's manual for AM console
4) MARION, D. & K. WÜTHRICH: Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^1H-^1H spin-spin constants in proteins. Biochem. Biophys. Res. Commun. 113: 967~974, 1983
5) WILLIAM, E. H.: Experimental Aspects of Two-dimensional NMR in "Two-dimensional NMR Spectroscopy" a) p. 203, b) p. 148, c) p. 178, 1987, W. R. CROASMUN & R. M. K. CARLSON *Ed.*, VCH Publs., New York
6) KESSLER, H.; C. GRIESINGER & J. LAUTS: Determination of Connectivities via Small Proton-Carbon Couplings with a New Two-Dimensional NMR Technique. Angew. Chem. Int. Ed. Engl. 23: 444~445, 1984
7) BREITER, J.; H. METZ & J. GRIGO: Staphylococcal micrococcins. II. Isolation, purification and identification. Arzneim. Forsch. (Drug Res.) 25: 1244~1248, 1975
8) Organic Electronic Spectral Data, Vol. I (1946~1952). MORTIMER J. KAMLET *Ed.*, Interscience Publs., New York
9) SHOJI, J.; T. KATO, Y. YOSHIMURA & K. TORI: Structural studies on thiocillin I, II, and III. J. Antibiotics 34: 1126~1136, 1981
10) ECKART, K.: Mass spectrometry of cyclic peptides. Mass Spec. Rev. 13: 23~55, 1994
11) ECKART, K.; U. SCHMIDT & H. SCHWARZ: Sequenzierung von Thiazolaminosäuren enthaltenden Cyclopentapeptiden ("Dolastatin 3") durch FAB/MSMS. Liebigs Ann. Chem. 1986: 1940~1949, 1986
12) HOLZAPFEL, C. W.; W. J. VAN ZYL & M. ROOS: The synthesis and conformation in solution of cyclo[L-Pro-L-Leu-L-(Gln)Thz-(Gly)Thz-L-Val] (dolastatin 3). Tetrahedron 46: 649~660, 1990
13) MOCEK, U.; A. R. KNAGGS, R. TSUCHIYA, T. NGUYEN, J. M. BAELE & H. G. FLOSS: Biosynthesis of the modified peptide antibiotic Nosiheptide in *Streptomyces actuosus*. J. Am. Chem. Soc. 115: 7557~7568, 1993
14) MOCEK, U.; Z. ZENG, D. O'HAGAN, P. ZHOU, L-D. G. FAN, J. M. BEALE & H. G. Floss: Biosynthesis of the modified peptide antibiotic thiostrepton in *Streptomyces azureus* and *Streptomyces laurentii*. J. Am. Chem. Soc. 115: 7992~8001, 1993