ANTIBIOTIC GE2270 A: A NOVEL INHIBITOR OF BACTERIAL PROTEIN SYNTHESIS

II. STRUCTURE ELUCIDATION

JÜRGEN KETTENRING, LUIGI COLOMBO, PIETRO FERRARI*, PAOLO TAVECCHIA, MARINO NEBULONI, KÁROLY VÉKEY†, GIAN GUALBERTO GALLO and ENRICO SELVA

MMDRI, Lepetit Research Center,
21040 Gerenzano (Varese), Italy

(Received for publication February 5, 1991)

GE2270 A, produced by *Planobispora rosea* ATCC 53773, inhibits Gram-positive bacteria and anaerobes by acting on the bacterial protein synthesis. The structure has been determined by physico-chemical methods applied to the intact molecule and to the main hydrolysis products.

Characterization by UV, IR, NMR (double quantum filter COSY), acid-base ionization, elemental analysis and FAB-MS indicated that GE2270 A is a highly modified peptide having MW 1,289 and formula $C_{56}H_{55}N_{15}O_{10}S_6$, and a weak basic function, and that it belongs to the thiazolyl peptide group of antibiotics. Acid hydrolysis yielded a main product (MW 634), responsible for the chromophoric absorption, and a number of hydrolyzed products of lower MW.

^{13}C NMR inverse techniques and MS studies (EI, positive ion chemical ionization, and collision induced dissociation FAB-MS-MS experiments) on GE2270 A, the chromophoric compound, and the other hydrolysis products led to the complete identification of the various amino acid residues and their sequence. Two out of the six chiral centers have been determined.

The structure is thought to originate from modification of a chain of 14 amino acids in a process which creates 6 thiazole rings and one pyridine. The modification process also closes the linear polypeptide to form a cyclic part with an attached side-chain. GE2270 A plausibly has a similar biosynthetic origin to that of other thiazolyl peptide antibiotics such as nosiheptide and micrococcin.

GE2270 A is a new antibiotic isolated from the *Planobispora rosea* strain ATCC 53773[1]. The mechanism of action of GE2270 A is the specific inhibition of bacterial protein biosynthesis by acting on the elongation factor Tu (EF-Tu)[1].

GE2270 A was extracted with methanol from the mycelium and was purified by column chromatography on silica gel. After purification it is obtained as a white powder.

GE2270 A is active *in vitro* against Gram-positive and a few Gram-negative bacteria. It is particularly active against anaerobes[1].

Experimental

The UV absorption spectra were recorded with a Perkin-Elmer spectrophotometer model 320 in methanol solutions as such and at the extreme pH values obtained by adding a trace of HCl or KOH.

The IR absorption spectra were obtained from a mineral oil suspension and in chloroform solution with a Perkin-Elmer spectrophotometer model 580.

The acid-base titrations in aqueous medium were obtained with 0.1 N KOH or with 0.1 N HCl on a

† Central Research Institute for Chemistry, Hungarian Academy of Sciences, H-1025 Budapest, Pusztaszeri No. 59-67, Hungary.

solution of the sample in Methyl Cellosolve - water (4 : 1), while that in non-aqueous medium was carried out in glacial acetic acid with 0.1 N $HClO_4$.

The elemental analysis was carried out by combustion method, using a C. Erba elemental analyzer model 1106.

The positive ion FAB-MS spectra were obtained on a Kratos MS-50 double focusing mass spectrometer of 3,000 dalton mass range, using 8 kV accelerating voltage, a saddle field atom (with Xe gas) at 6 kV voltage and 1 mA current. To obtain high quality data, the instrument was operated under computer control and the MSS data system was used. Using a mixture of CsI and NaI as calibration compound[2] the accuracy of mass determination was increased to about 0.1 dalton in scanning mode. The samples were dissolved in a 1 : 1 mixture of DMSO - nitro benzyl alcohol.

The ^1H and ^{13}C NMR experiments were performed, using a solution of 10 mg of lyophilized GE2270 A in 0.5 ml of DMSO-d_6, on a Bruker AM-500 spectrometer equipped with an Aspect 3000 computer, in the temperature range 297 ~ 318°K. For the homonuclear experiments double quantum filter (DQF)-COSY[3] and NOESY[4] the sweep width in t_1 was 5,500 Hz and 512 increments with 32 or 48 transitions were collected per experiment. The 90° (H) pulse was 11.5 μseconds. For the ^1H-detected ^1H-^{13}C correlations standard pulse sequences were used. The sweep width in t_2 was 5,500 Hz and in t_1 11,300 Hz. 512 increments with 80 transitions each were collected for the ^1H-detected heteronuclear relay experiment[5]. For the ^1H-detected correlation *via* long range coupling (COLOC) experiment 196 transitions per increment were collected[6,7]. The 90° (H) pulse was 10.5 μseconds.

The acid hydrolysis was carried out at 100 ~ 120°C overnight with HCl 18.5%. After cooling, the reaction mixture was diluted with water and the precipitate filtered. The cake was washed with water, acetone and finally ethyl ether, and submitted to NMR (see above) and to the collisionally induced dissociation FAB-MS-MS experiment. This measurement was performed on a triple stage quadrupole mass spectrometer Finnigan TSQ 70, fitted with an Ion-Tech neutral atom gun operated with Xe gas at 8 kV. The sample was dissolved in thioglycerol immediately before the analysis. Positive daughter ions of the protonated molecular ion were recorded using the second quadrupole as collision cell (collision gas Ar). The aqueous solution from the acid hydrolysis was lyophilized, suspended in acetonitrile, and treated with pentafluoropropionic anhydride. After 15 minutes, the solvent was evaporated off and the residue redissolved in ethyl ether and an excess of diazomethane added. After evaporation, the crude material was analyzed by a HP5985B quadrupole GC-MS system. GC: capillary column (Supelco SP1, 15 m × 0.32 mm i.d., 0.25 μm thin film); helium as carrier gas (column pressure 0.56 kg/cm^2, split vent 80 ml/minute, split valve time 30 seconds, splittless mode injection). The injector and interface temperatures were set at 280 and 250°C, respectively. The oven temperature was set at 50°C for 1 minute and increased at 10°C/minute to 240°C. EI-MS: source temperature 200°C, electron energy 70 eV. Positive ion chemical ionization (PICI)-MS: source temperature 140°C, electron energy 180 eV, ionization gas CH_4 (as make-up gas), ion source pressure 1×10^{-4} torr. For determining the configurations of serine and proline, the GC-MS analysis was repeated on a Hewlett-Packard HP5985B GC-MS system equipped with a *n*-propionyl-L-valine *tert*-butylamide polysiloxane-coated fused silica capillary column (25 m × 0.2 mm i.d.; OS6411, C.G.C. Analytic); temperature program: 80°C for 4 minutes, then increasing at 4°C/minute to 200°C.

Physico-chemical Characteristics for Basic Structure Information

The preliminary characterization of the fermentation product was aimed at the novelty recognition of the discovered antibiotic substance and was used as the starting point for the structure determination.

The UV spectra (see Table 1) are indicative of a chromophore including a weak heterocyclic base. In

Table 1. UV absorption bands in MeOH solution at different pH values.

	λ_{max} (nm)	$E^{1\%}_{1cm}$	λ_{max} (nm)	$E^{1\%}_{1cm}$	λ_{max} (nm)	$E^{1\%}_{1cm}$
MeOH	240 (sh)	456	310	265	340 (sh)	120
MeOH + KOH	240 (sh)	456	310	265	340 (sh)	120
MeOH + HCl	240 (sh)	420	309	267	340 (sh)	107

fact, the variation passing from the acidic solution to the basic one is not bathochromic and only small differences can be observed. In addition, the main absorption band at about 310 nm suggests a chromophore formed by a number of conjugated heterocycles.

The IR spectra (see Table 2) indicate that GE2270 A is a peptide without free carboxyl or amino groups. Ester or ketone groups are absent. The absorption at about 1545 cm^{-1} is indicative of heterocycles. The peptide nature is confirmed by the deuteration shifts in chloroform solution of the amide bands (νNH, Amide I, Amide II).

The ionization studies in water solution did not reveal any ionizable function of strong or medium strength, while in acetic acid a weak basic function was revealed.

The elemental analysis gave the following results: C 52.02, H 4.39, N 16.26, S 15.03, O (diff) 12.3, from which the following possible minimal formulae can be calculated:

C 9.235, H 9.288, N 2.475, O 1.640, S 1.000.
C 46.17, H 46.44, N 12.37, O 8.20, S 5.00.
C 55.41, H 55.73, N 14.85, O 9.84, S 6.00.
C 64.64, H 65.02, N 17.32, O 11.48, S 7.00.

The discussion of these formulae is postponed to the following sections on MS and NMR.

The FAB spectrum of GE2270 A is shown in Fig. 1. The lowest isotope of the main peak has an atomic mass of 1,290.3±0.1 dalton. By adding KI salt to the sample this peak shifted 38 daltons ($(M+K)^+$ ion) indicating that the observed peak at 1,290.3 is the protonated (and not *e.g.* sodiumated) molecular ion.

The isotope pattern of the protonated molecular ion indicates that besides C, H, N, O some

Table 2. Attributions of the main functional IR absorption bands in mineral oil suspension.

ν (cm^{-1})	Assignment
3600~3100	νNH, νOH
1650	Amide I, νC=O
1545	Heterocyclic, νC=C and νC=N
1525, 1495	Amide II, δNH
1250, 1205	Aromatic, δCH
870	Heterocyclic, γCH
745, 700	Aromatic, γCH

Fig. 1. FAB-MS of GE2270 A.

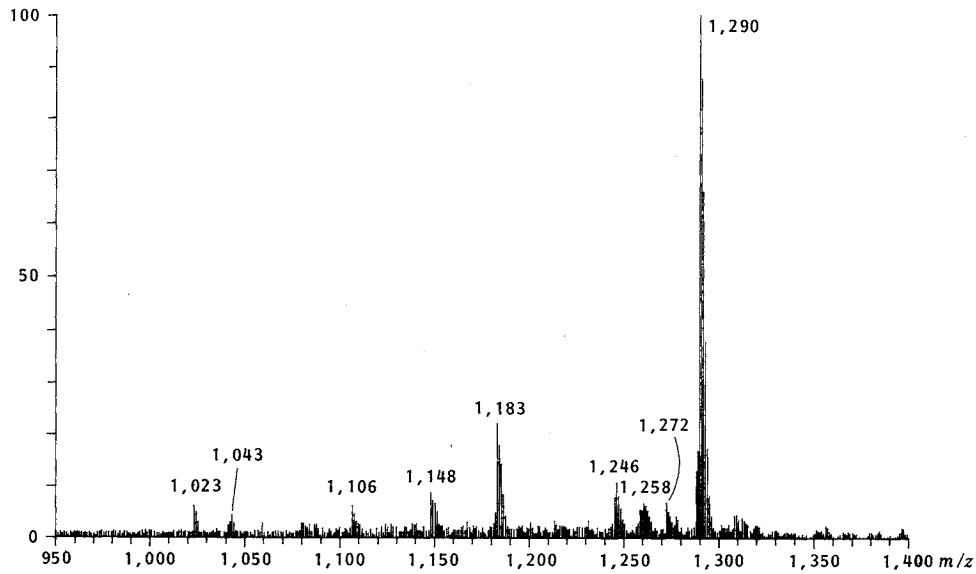
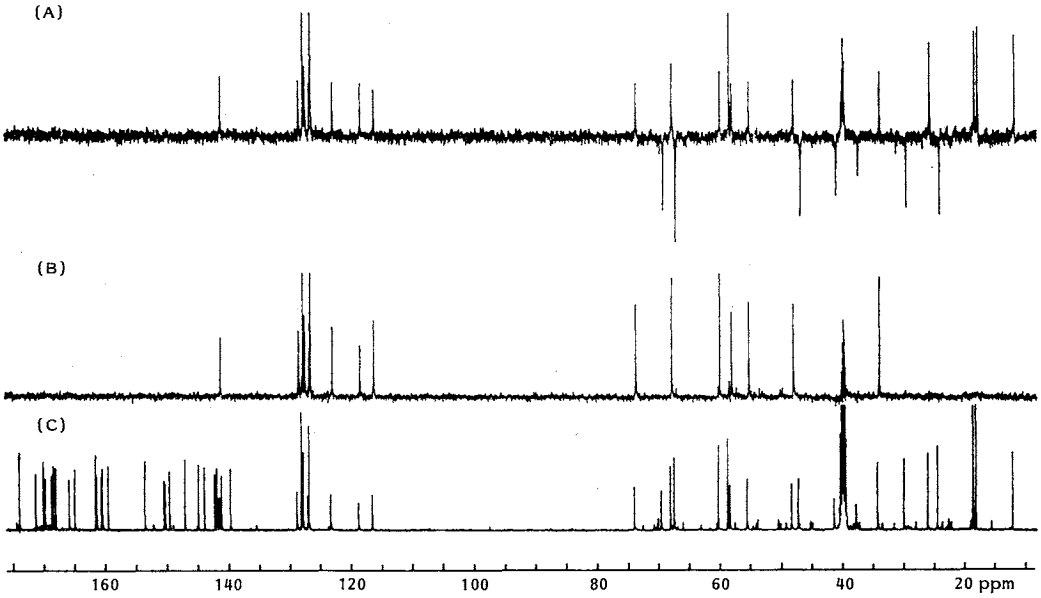


Fig. 2. DEPT ^{13}C NMR spectra of GE2270 A in DMSO-d_6 solution.

(A) DEPT spectrum showing CH and CH_3 positive, CH_2 negative, (B) DEPT spectrum showing CH only, (C) normal ^{13}C spectrum.



element containing higher mass isotopes should be present, *e.g.*, 1 chlorine atom or 5~8 sulfur atoms. Elemental analysis clearly established the presence of sulfur atoms, while excluding chlorine. Based on the isotope pattern found, the mean mass of the $(M+H)^+$ isotope cluster is $1,291.6\pm0.3$. From this value the "chemical" molecular mass can be estimated as $1,290.6\pm0.3$ dalton.

Using HR (13,000) peak matching the precise mass of the lowest isotope of the protonated molecular ion cluster was determined as 1,290.25 with about ±0.02 dalton precision. This high precision, combined with information from elemental analysis, makes it possible to determine an elemental formula for the sample. The molecular mass, combined with the knowledge that molecules are even electron species and that GE2270 A contains C, H, N, O, and S atoms only, implies that the number of N atoms in this molecule needs to be an odd number. This fact, together with the precise mass of the molecule indicates that the elemental formula is $C_{56}H_{55}N_{15}O_{10}S_6$. This is in agreement with the experimentally found isotope pattern.

The ^1H and ^{13}C spectra confirmed some of the data described above and added new important evidences. The calculation of the total numbers of H and C atoms indicated C 56 and H 55, thus making a definite and final choice for the elemental composition from those possibilities discussed under elemental analysis and FAB-MS. The assignment of the various groups based on the DEPT ^{13}C NMR spectra (Fig. 2) gave the following results: 5 CH_3, 7 CH_2, 7 CH, 11 aromatic CH, 26 quaternary C, 5 amidic NH, 1 amidic NH_2, 1 alcoholic OH.

All the above data strongly suggested that GE2270 A is a highly modified peptide and that likely it belongs to the group of thiazolyl peptide antibiotics[8]. In fact, the presence of 6 sulfur atoms led to the plausible hypothesis that a number of the original amino acids are masked in the molecule in the form of the so-called "thiazole-amino acid" unit, as examplified by micrococcin[9].

Identification of the Amino Acid Residues by NMR

The examination of the ^{1}H and ^{13}C NMR spectra for the assignment of the signals and their attribution to the single amino acid units was implemented by an NMR investigation based on ^{1}H-^{1}H homonuclear and ^{13}C-^{1}H heteronuclear short and long range coupling experiments. As method of choice for ^{13}C NMR spectroscopy the so-called inverse NMR techniques were used. The heteronuclei are detected *via* the proton response, which leads to a drastic increase in sensitivity. This allows to carry out this NMR experiment in dilute NMR solutions[10]. The proton-bearing-carbon atoms were assigned with a ^{1}H-detected relay experiment[5], establishing also the connectivities between protonated carbon atoms bound to each other by a common relay peak in this 2D experiment.

Due to the highly modified amino acids in GE2270 A only two of them could be easily identified in a DQF-COSY spectra[3] as a valine and glycine moiety (see Fig. 3). Two more amino acids show crosspeaks to amide protons, an ABMX and an AMX spin system. The ^{1}H-detected inverse experiment optimized for long range coupling[6,7] demonstrated that the ABMX spin system is connected to a $CONHCH_3$ group and this moiety was identified as *N*-methyl asparagine. The AMX spin system shows long range crosspeaks to the phenyl unit and to a OH group and was therefore identified as phenylserine. Furthermore, the DQF-COSY and the ^{1}H-detected relay experiment indicated the presence of a proline unit and an ABM

Fig. 3. DQF ^{1}H-^{1}H COSY spectrum of GE2270 A in DMSO-d_6 solution.
Identification patterns of glycine and valine shown.

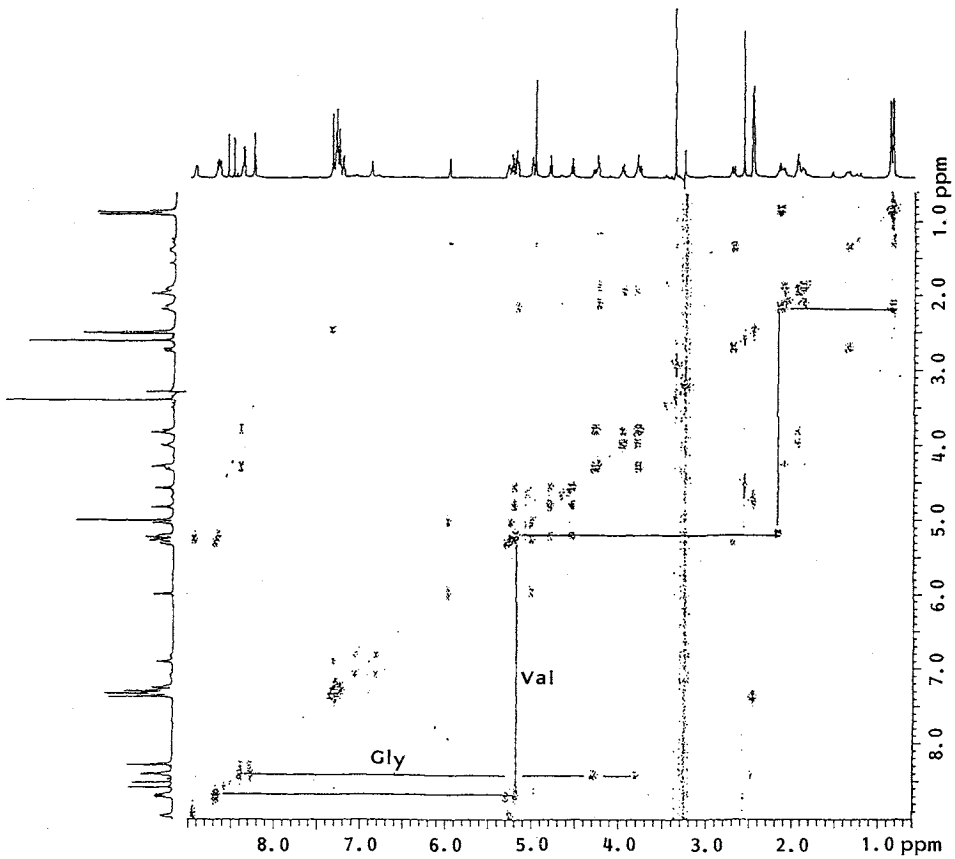
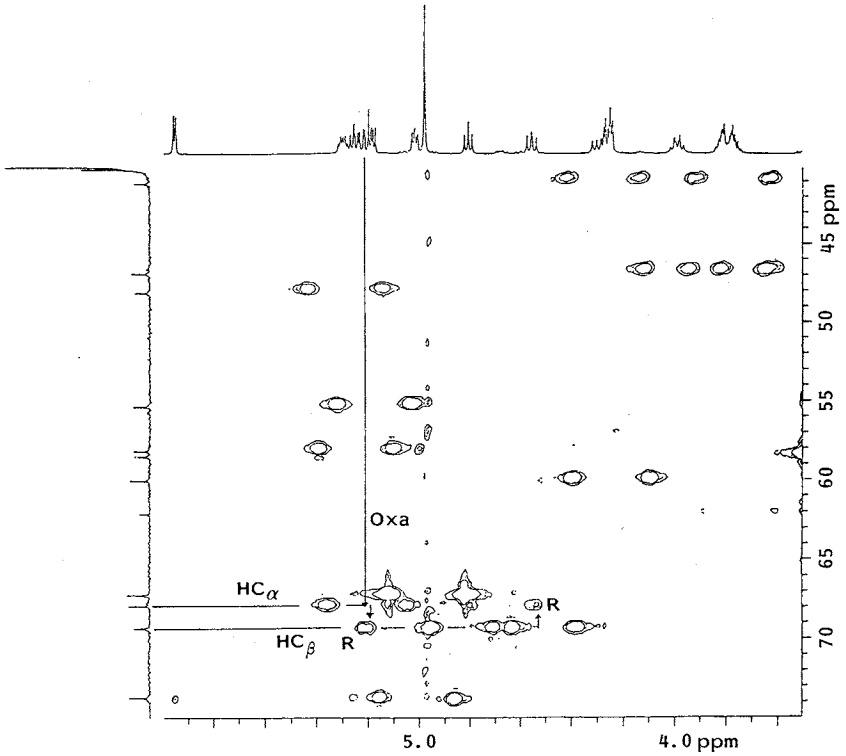


Fig. 4. Part of a ^{1}H-detected heteronuclear relay experiment.

The identification of the ABM spin system and their ^{13}C signals are shown.



spin system not correlated with an amide proton (see Fig. 4). The proline unit is attached to an amide group and the ABM spin system is a serine masked as oxazoline, which shows long range correlations between the protons and two low field quarternary carbons at 167.7 ppm and 159.2 ppm, assigned to the serine CO and to the C=N of this moiety, respectively.

Present in the spectra are also four singlets in the aromatic region, an aromatic AB system, a CH_3, CH_2-OCH_3 group, a phenyl ring and a CH_3 group bound to an NH group, as was demonstrated by exchange with D_2O. The assignment and identification of these moieties was done by ^{1}H-detected inverse experiment optimized for long range coupling[6,7]. The aromatic AB system shows the typical long range coupling pattern of a 2,3,6-substituted pyridine. The four aromatic singlets are correlated *via* ^{1}J and two long range couplings to three aromatic carbon atoms, in agreement with the presence of four thiazole units. The remaining CH_3 and CH_2-OCH_3 groups are connected *via* long range couplings to two different aromatic carbon atoms, indicating the presence of two additional substituted thiazole units.

By applying this approach, all the protons and almost all of the ^{13}C signals could be identified and assigned, except for five signals at 168.3, 165.7, 161.1 and 160.2 ppm. Their assignment, however, could be done in an indirect way with the knowledge of the amino acid sequence. The results are collected in Table 3, where, for the sake of clarity, the labeling system adopted for the final structure is used.

Acid Hydrolysis of GE2270 A

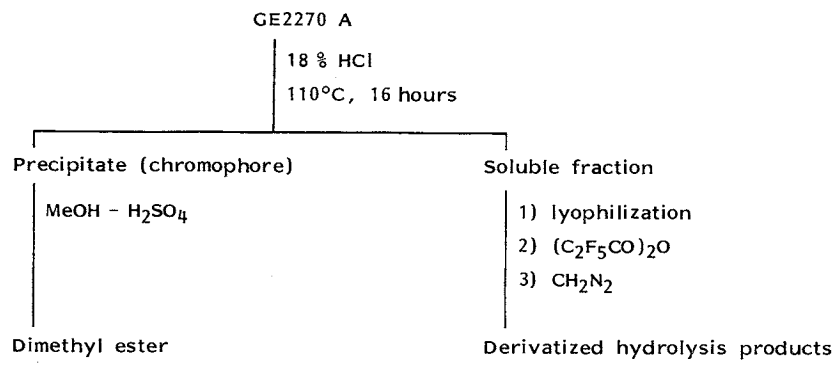
GE2270 A was hydrolyzed under acidic conditions (see Experimental) and the precipitated product was submitted to analysis as such, while the soluble fraction was derivatized and analyzed by GC-MS, as

Table 3. NMR assignments of GE2270 A (in DMSO, TMS internal standard, δ ppm, T=33°C).

^{13}C	^{1}H	Assignment	^{13}C	^{1}H	Assignment
173.5		$CONH_2$Pro	127.8	7.26	Phe 3, 5
170.9		C2	127.6		Pyr 3
169.7		$CONHCH_3$Asn	127.5	7.23	Phe 4
169.4		CO Gly	126.6	8.56, 7.31	F5, Phe 2, 6
168.3		D2*	122.9	8.26	B5
168.1		A2	118.5	8.27	Pyr 5
167.7		CO Oxa	116.3	7.35	C5
165.7		E2*	73.8	5.02	Cβ-PheSer
164.7		F2	69.3	4.81, 4.56	β Oxa
161.3		CO D*	67.9	5.22	α Oxa
161.1		CO F*	67.2	4.98	E-OCH_2
160.4		B2	60.0	4.26	α Pro
160.2		CO E*	58.5	3.38	E-CH_3
159.2		C=N Oxa	58.1	5.24	α PheSer
153.3		B4	55.3	5.19	α Val
150.3		Pyr 2,6	48.1	5.31	α Asn
150.1			46.9	3.96, 3.82	δ Pro
149.4		F4	41.1	4.29, 3.80	α Gly
146.7		C4	37.6	2.71, 1.39	β Asn
144.7		A4	33.9	2.17	β Val
143.5		E4	29.6	2.14, 1.91	β Pro
142.0		D4	25.7	2.47	CH_3Asn
141.6		Phe 1	24.1	1.97	Pro
141.1	8.38	Pyr 4	18.4	0.85	Val
141.0		E5	17.9	0.88	
139.4		D5	11.8	2.59	D-CH_3
128.5	8.50	A5			

* Assignment based on sequence information obtained by NOE and GC-MS. Others: 8.92, NH-PheSer; 8.69, NH Asn; 8.65, NH Val; 8.38, NH Gly; 7.34 $NHCH_3$ Asn; 7.34, 6.86, NH_2CO Pro; 5.95, OH PheSer. A, B, C, D, E, and F refer to the thiazole rings.

Scheme 1.



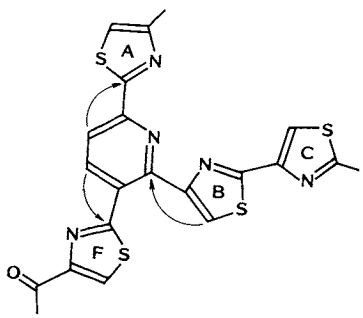
shown in Scheme 1.

The UV spectrum in ethanol, obtained from the fully methylated ester of the precipitate, shows the maxima: 304 nm, $E^{1\%}_{1cm}$ 420; 340 nm (sh), $E^{1\%}_{1cm}$ 200, indicating, by comparison with the data of GE2270 A, that it represents the central core of the natural antibiotic. This is called the chromophore.

The structure of this compound was determined on the basis of the ^{1}H and ^{13}C NMR and FAB-MS studies. The ^{13}C signals of the chromophore of GE2270 A was assigned using the ^{1}H-detected heteronuclear correlation experiments. The results are collected in Table 4. The ^{1}H-detected heteronuclear experiment,

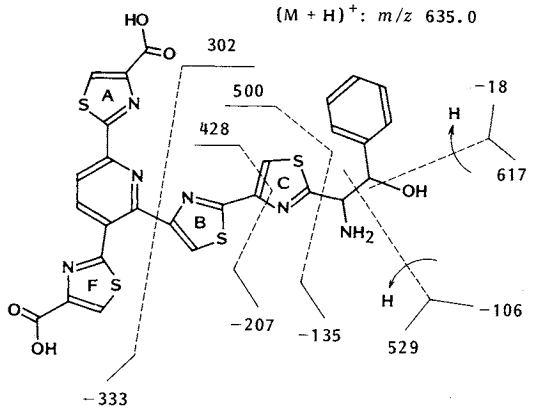
Table 4. NMR assignments of the chromophore of GE2270 A in DMSO (TMS internal standard, δ in ppm, T=33°C).

Assignment	^{13}C	^{1}H	Assignment	^{13}C	^{1}H
A2	167.7		α PheSer	57.6	5.04
A4	149.1		β PheSer	71.7	5.28
A5	130.0	8.63	Phe 1	140.2	
Pyr 2, 6	150.8, 150.2		Phe 2, 6, 4	127.6	7.28~7.38
Pyr 3	128.7		Phe 3, 5	126.3	
Pyr 4	141.8	8.56	F2	163.8	
Pyr 5	118.3	8.34	F4	146.8	
B2	161.6		F5	131.2	8.52
B4	152.7				
B5	122.8	8.27	two COOH	162.0	
C2	164.6				
C4	146.8				
C5	119.2	7.89			



optimized for long range couplings, allowed to establish the connectivities between the 2,3,6-substituted pyridine ring and the thiazole rings A, B, and F, as outlined as above. The arrows indicate the long range correlations between these moieties.

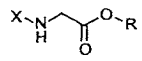
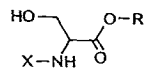
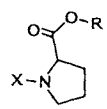
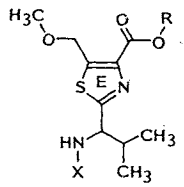
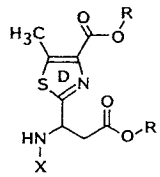
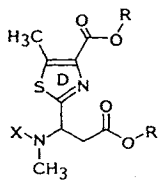
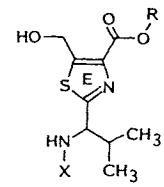
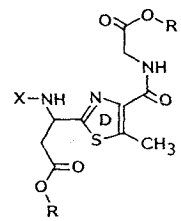
Fig. 5. Fragmentation pattern from the FAB-MS experiment on the precipitate.

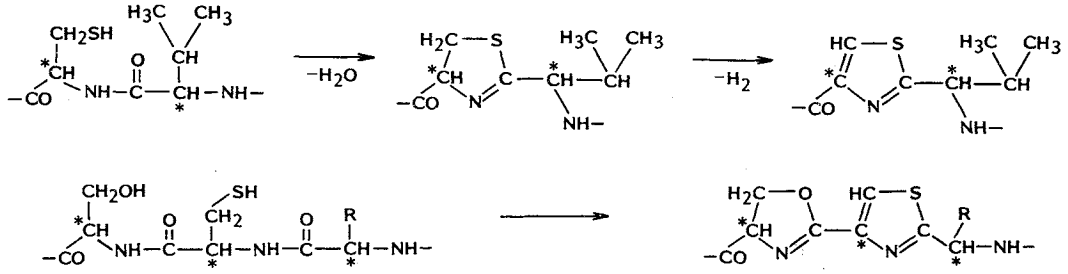


The FAB-MS analysis of the chromophore showed little fragment ions relative to the protonated molecular ion $(M+H)^+$. To increase the percentage of dissociating ions and also the number of fragmentation routes accessible to a given ion, the most common method is to admit a gas to a reaction region of the mass spectrometer to induce dissociation by collision of the parent ion (in this case the protonated molecular ion, $(M+H)^+$, m/z 635) with a neutral target gas[11]. This process known as collision induced dissociation (CID) was applied with argon, as collision gas, at 120 eV. The significant daughters found could be attributed to the breakage of the longer pyridine substituent. A suggested interpretation is presented in Fig. 5. The losses of 18, 106, and 135 confirm the presence of a phenylserine residue. The losses of 207 and 333 correspond to the fission of ring C and to the breakage between pyridine and ring B, respectively. Thus, the structure of the chromophore proposed by NMR was confirmed.

The GC-MS analysis of the soluble products from hydrolysis, as pentafluoropropionyl derivatives and methyl esters[12], confirmed and completed the previous assignments by identifying proline, serine, glycine, and a number of thiazole-amino acids. The amino acids were identified by comparison with standards, while the thiazole-amino acids by the mass spectral (EI and PICI) fragmentations. The structure

Table 5. Structure assignments to the relevant GC-MS peaks from the soluble products of hydrolysis.

Peak No.	R=CH_3, X=COC_2F_5	Rt (minutes)	Ions *m/z* (relative abundance)	
			EI	PICI
1		2.13	235 (18): $M^{+\cdot}$ 204 (30): $M-RO^{\cdot}$ 176 (100): $M-ROCO^{\cdot}$	236 (100): MH^+ 216 (15): $M-F^{\cdot}$
2		4.98	235 (42): $M-CH_2O^{\cdot}$ 206 (51): $M-ROCO^{\cdot}$ 203 (100): $235-RO^{\cdot}$	266 (100): MH^+ 234 (39): $M-RO^{\cdot}$ 206 (41): $M-ROCO^{\cdot}$
3		5.45	275 (38): $M^{+\cdot}$ 256 (2): $M-F^{\cdot}$ 216 (100): $M-ROCO^{\cdot}$	276 (100): MH^+ 244 (32): $M-RO^{\cdot}$ 216 (44): $M-ROCO^{\cdot}$
4		12.38	404 (20): $M^{+\cdot}$ 361 (33): $M-C_3H_7^{\cdot}$ 372 (49): M−ROH 329 (100): $372-C_3H_7^{\cdot}$	405 (100): MH^+ 385 (8): $M-F^{\cdot}$ 373 (22): $M-RO^{\cdot}$
5		13.15	404 (48): $M^{+\cdot}$ 373 (27): $M-RO^{\cdot}$ 345 (49): $M-ROCO^{\cdot}$ 313 (100): 345−ROH	405 (100): MH^+ 385 (7): $M-F^{\cdot}$ 373 (10): $M-RO^{\cdot}$
6		13.30	418 (73): $M^{+\cdot}$ 403 (13): $M-CH_3^{\cdot}$ 359 (32): $M-ROCO^{\cdot}$ 327 (100): 359−ROH	419 (100): MH^+ 399 (8): $M-F^{\cdot}$ 359 (6): $M-ROCO^{\cdot}$
7		13.40	390 (35): $M^{+\cdot}$ 372 (11): $M-H_2O$ 347 (62): $M-C_3H_7^{\cdot}$ 329 (100): $347-H_2O$	391 (100): MH^+ 373 (22): MH^+-H_2O 359 (11): $M-RO^{\cdot}$
8		16.60	461 (42): $M^{+\cdot}$ 446 (100): $M-CH_3^{\cdot}$ 402 (44): $M-ROCO^{\cdot}$	462 (100): MH^+ 430 (46): $M-RO^{\cdot}$



assignments to the main ions found for the eight relevant chromatographic peaks are presented in Table 5.

The stereochemistry of serine and proline was assigned as S from the Rt's on the chiral column[13] described under Experimental.

The biogenesis of thiazole-amino acids would occur after peptide formation, *e.g.*, from cysteine and valine by condensation of the thiol with the adjacent amide group under formation of thiazoline and successively of thiazole after dehydrogenation.

Thus, these precursor amino acids are irreversibly changed during such reaction and will not be identified after hydrolysis.

Similarly, the oxazoline ring would derive from a dipeptide containing serine and, *e.g.*, cystein. In this case the condensation is not followed by dehydrogenation and serine will be identified after hydrolysis.

Sequence Determination of the Amino Acids

As expected for a medium sized molecule like GE2270 A and under the measurement conditions, the NOE's are quite weak and reside on the negative side. The highly modified nature of the amino acids does not allow to establish a great part of the sequence by looking for sequential connectivities[14]. The NMR NOESY experiment[4] confirmed only some intra amino acids connectivities and that the *N*-methylasparagine-thiazole D moiety is close to the phenylserine-glycine moiety. This is demonstrated by some weak interactions between one of the H_β protons of *N*-methylasparagine and the amide proton of glycine and a weak NOE signal between the CH_3 of thiazole D and the aromatic part of the phenylserine unit.

The ^{1}H-detected COLOC experiment (see Fig. 6) allowed to establish the following parts of the sequence (for the inter and intra amino acid long range correlations see Table 6):

1) Prolineamide→oxazoline→thiazole A→pyridine[a].
2) Thiazole B[b]→thiazole C→phenylserine→glycine→CO.
3) Pyridine→thiazole F.

[a] Only a very weak crosspeak between the last two moieties was found. [b] The crosspeak correlating these two thiazole units was detected only in two experiments at the higher temperature.

The above mentioned experiments did not allow to establish the connectivities between the two remaining thiazole moieties D and E and the amino acids valine and *N*-methylasparagine. However, hydrolysis and following GC-MS studies (see above) connected valine to ring E and *N*-methylasparagine to ring D. As the latter unit (see the NOESY experiment) is close to the glycine-phenylserine subunit the sequence had to be continued as follows:

Thiazole B→thiazole C→phenylserine→glycine→CO→thiazole D→*N*-methylasparagine→CO→thiazole E→valine→CO→thiazole F→pyridine.

Fig. 6. High field part of a ^{1}H-detected heteronuclear experiment optimized for long range couplings.

The sequence thiazole B→thiazole C→PheSer→Gly→CO is outlined.

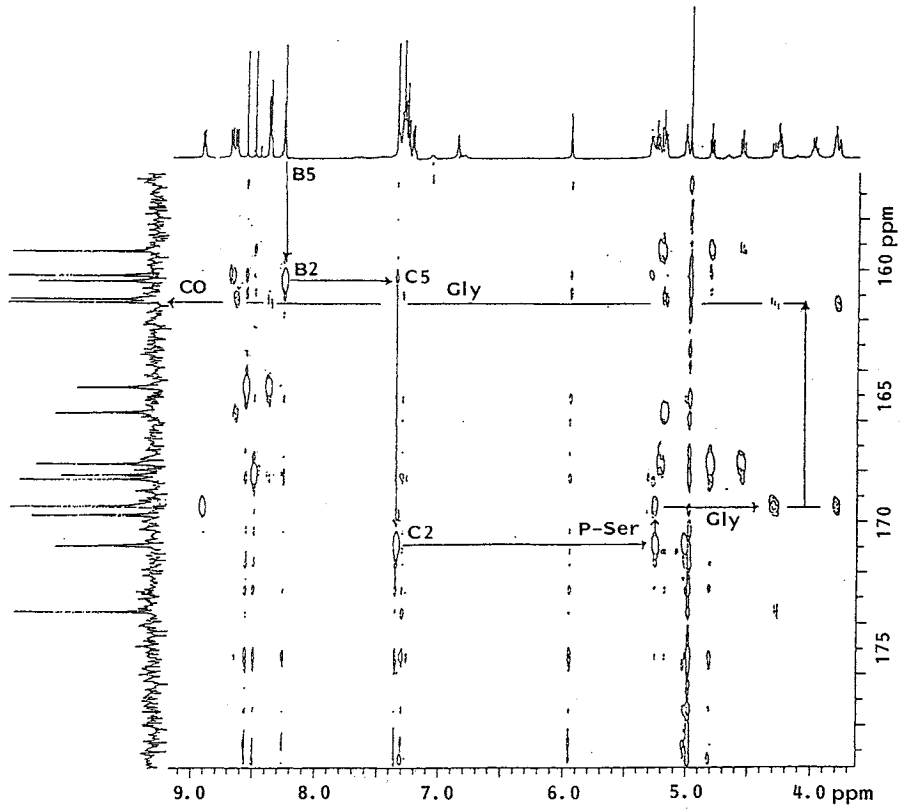


Table 6. Long range couplings observed by inverse COLOC experiment for GE2270 A (only quarternary carbon atoms listed).

Carbon atom	LR	Carbon atom	LR
$CONH_2$ Pro	Hα, β Pro	C=N Oxa	Hα, β Oxa, A5
C2	C5, Hα, β PheSer	B4	B5
$CONHCH_3$	$NHCH_3$, Hα, β Asn	Pyr 2, 6	Pyr 4
CO Gly	NH, Hα P-Ser, Hα Gly	F4	F5
D2[a]	Hα Asn	C4	C5
A2	A5, Pyr 5[b]	A4	A5
CO Oxa	Hα, β Oxa, Hα Pro[b]	E4	E-CH_2O
E2[a]	NH, Hα, β Val	D4	D-CH_3
F2	F5, Pyr 4	Phe 1	Hα, β, OH PheSer
CO D[a]	NH, Hα Gly	E5	E-CH_2O
CO F[a]	NH, Hα, β Val	D5	D-CH_3
B2	B5[c], C5	Pyr 3	Pyr 5

[a] Assignment based on sequence information obtained by NOE and GC-MS.

[b] Very weak crosspeak.

[c] The long range correlation connecting thiazoles B and C was detected only in two experiments at elevated temperatures.

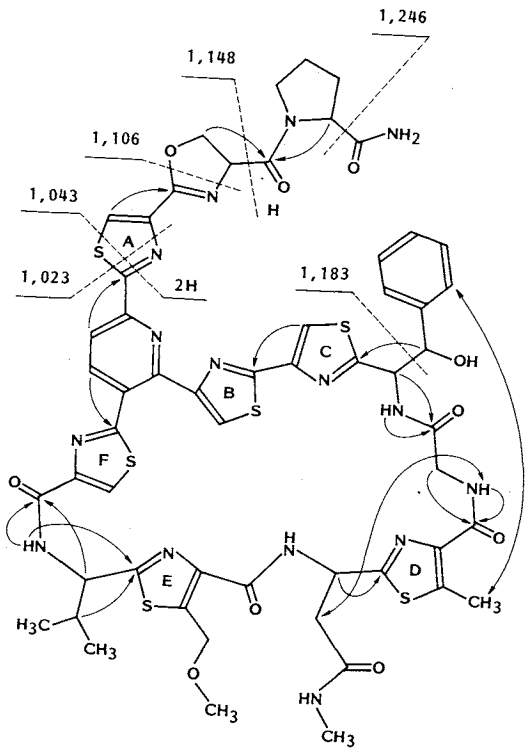
This sequence is consistent with the GC-MS results showing the presence of a fragment containing the glycine-CO-thiazole D-*N*-methylasparagine unit.

Further confirmation for the central part of GE2270 A came from the hydrolysis of this molecule where the cromophore (precipitate) was isolated. The ^{1}H and ^{13}C NMR data are in full agreement with the results obtained on the intact molecule. The central part is in fact a pyridine, substituted in positions 2, 3 and 6 with thiazole moieties. Thiazole B is connected to a further thiazole ring (C), which is substituted in position 2 by the phenylserine unit (Table 4).

In conclusion, the combined methods of NMR and GC-MS applied to the intact antibiotic and to the hydrolysis products indicate the structure shown in Fig. 7.

Fig. 7. Structure of GE2270 A.

The arrows indicate the sequencing process by ^{1}H-detected heteronuclear and NOESY NMR experiments. The lines and the mass numbers indicate the suggested FAB-MS fragmentations.



FAB-MS Fragmentation

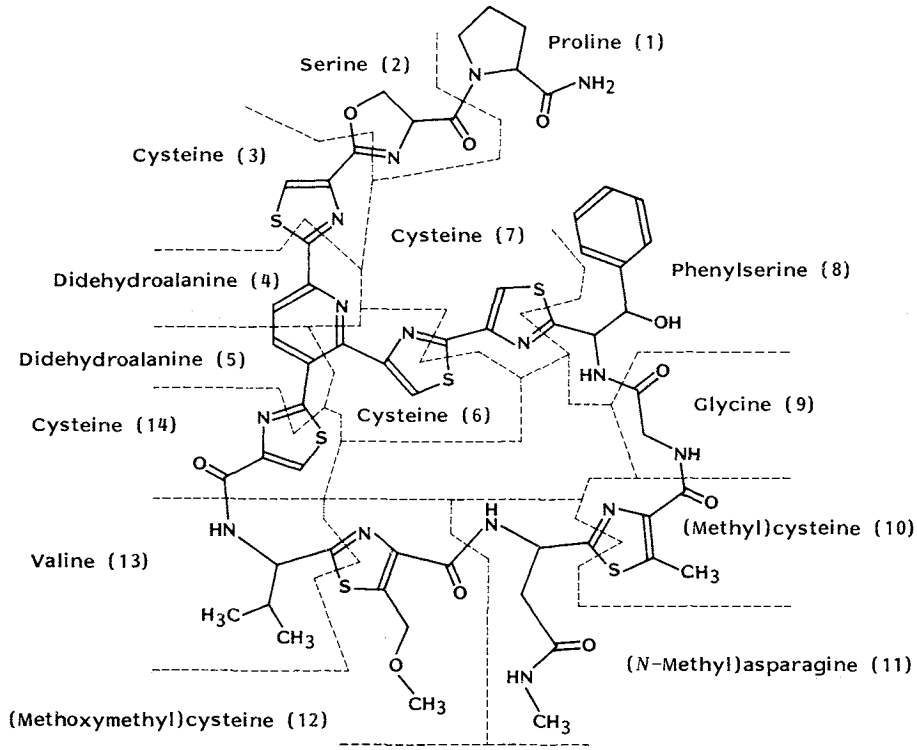
Following the structure elucidation, it appears of interest to interpret the FAB-MS fragmentation. In the FAB spectrum (Fig. 1), besides the protonated molecular ion some low abundance fragments can also be observed. FAB analysis of linear polypeptides classically gives rise to fragment ions occuring in well defined mass sequences. These ions are normally rationalized on the basis of low-energy fragmentations deriving directly from the protonated molecular ion. In the case of GE2270 A, not only are fragment ions corresponding to the common amino acids absent, but the fragments which occur are limited to higher masses. This pattern corresponds to the cyclic peptide structure of GE2270 A, which prevents lower mass fragments being seen. Observable fragments should thus be attributable to extracyclic substituents.

A suggested fragmentation scheme is presented in Fig. 7 (always the lowest isotope is indicated, with truncated mass numbers). Besides the common peaks at *m/z* 1,272 ($-H_2O$) and 1,258 ($-CH_3OH$), the peak at *m/z* 1,183 (loss of 107 from the molecular ion) is due to the loss of the hydroxybenzyl residue. The scheme invokes also the breaking of an oxazole or thiazole ring (*m/z* 1,106, 1,043, and 1,023)—a type of cleavage which, to our knowledge, has not yet been observed in modified peptides. Finally, the relatively high abundance of the fragment at *m/z* 1,246 corresponding to a loss of 44 from the protonated molecular ion, together with the loss of 142, giving *m/z* 1,148 and suggesting the loss of two nitrogen atoms, is due to the fact that the molecule does not have a free carboxy end, but an amide (prolinamide).

Conclusions and Considerations

In conclusion, the complete identification of all the amino acidic residues and their sequence follows from the NMR, FAB-MS, and GC-MS data on the natural antibiotic and its hydrolysis products. Antibiotic

Fig. 8. Structure of GE2270 A showing the original amino acids of the peptide chain.



GE2270 A shows structural similarities with the group of antibiotics classified by BÉRDY[8] as thiazolyl peptides. Notable examples are thiostrepton, nosiheptide, micrococcin, sulfomycin[15], and thioxamycin[16], which were the subject of extensive chemical degradation or X-ray studies.

In particular, GE2270 A shows identical reactivity on acid hydrolysis to micrococcin[17], leading to the isolation of the chromophoric compound. Furthermore, micrococcin[18] and thiocillin[19] molecules show remarkable similarities of GE2270 A. In fact, they contain six thiazoles and are formed by a ring spanning a pyridine moiety to which a side chain containing thiazoles is attached.

As far as biogenesis is concerned, GE2270 A and related antibiotics are of interest in that they represent highly modified peptides. The thiazole-amino acid residues formally derive from the cysteine-amino acid dipeptide. Similarly, the dipeptide containing serine gives origin to oxazoline. The unusual chromophoric part centered on pyridine is thought to take origin formally from the interaction of two didehydroalanine units in a single peptide chain[18]. Actually, pyridine was demonstrated[20] to originate from the carboxyl group of one cysteine and two serine molecules, so that didehydroalanine is to be considered a formal intermediate unit. In general, it is believed that these peptide antibiotics occur in two main groups: the thiostrepton group maintains the peptidic part referred to as "amino end", and the "chromophoric central part" is represented by a tetrahydro pyridine, while the micrococcin and GE2270 A group loses the "amino end" under aromatization to pyridine[18].

Summarizing, GE2270 A belongs to the thiazolyl peptide group of antibiotics. As hypothesis, the original linear peptide precursor of GE2270 A can be visualized from Fig. 8 where the structure is shown together with the indication of the number assigned to the amino acid units. Following these biosynthetic considerations, the peptide ring of GE2270 A would derive from ten amino acid units, whereas the corresponding rings of the micrococcin, nosiheptide, and thiostrepton group and of the thioxamycin and sulfomycin group would derive from nine and twelve amino acid units, respectively.

In the structure of GE2270 A the configuration of proline and serine is *S*, from the acid hydrolysate analyzed on a chiral column. Determination of the configurations of the other asymmetrical centers is in progress.

Acknowledgments

The authors are indebted to G. TAMBORINI for the UV and IR spectra and for acid-base titrations, to D. EDWARDS for GC-MS with chiral column of the hydrolysates, and to G. VAGO (Finnigan, Italy) for technical support on the MS-MS experiments.

References

1) SELVA, E.; G. BERETTA, N. MONTANINI, G. S. SADDLER, L. GASTALDO, P. FERRARI, R. LORENZETTI, P. LANDINI, F. RIPAMONTI, B. P. GOLDSTEIN, M. BERTI, L. MONTANARO & M. DENARO: Antibiotic GE2270 A: A novel inhibitor of bacterial protein synthesis. I. Isolation and characterization. J. Antibiotics 44: 693~701, 1991
2) VÉKEY, K.: Calibration in positive and negative ion fast atom bombardment using salt mixtures. Org. Mass Spectrosc. 24: 183~185, 1989
3) MARION, D. & K. WÜTHRICH: Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^{1}H-^{1}H spin-spin coupling constants in proteins. Biochem. Biophys. Res. Commun. 113: 967~974, 1983
4) MACURA, S.; K. WÜTHRICH & R. R. ERNST: The relevance of J crosspeaks in two-dimensional NOE experiments of macromolecules. J. Magn. Reson. 471: 351~357, 1982
5) LERNER, L. & A. BAX: Sensitivity-enhanced two-dimensional heteronuclear relayed coherence transfer NMR spectroscopy. J. Magn. Reson. 69: 375~380, 1986
6) BAX, A. & M. F. SUMMERS: ^{1}H and ^{13}C assignments from sensitivity enhanced detection of heteronuclear multiple-bond connectivity by 2D multiple quantum NMR. J. Am. Chem. Soc. 108: 2093~2094, 1986
7) BAX, A. & D. MARION: Improved resolution and sensitivity in ^{1}H-detected heteronuclear multiple-bond correlation spectroscopy. J. Magn. Reson. 78: 186~191, 1988
8) BÉRDY, J.: Thiazolyl peptides. *In* Handbook of Antibiotic Compounds. Vol. IV/I. *Ed.*, J. BÉRDI *et al.*, pp. 389~417, CRC Press Inc., 1980
9) BREITER, J.; H. METZ & J. GRIGO: Staphylococcal micrococcins. II. Isolation, purification and identification. Arzneim. Forsch. (Drug Res.) 25: 1244~1248, 1975
10) KETTENRING, J. K.; R. CIABATTI & B. CAVALLERI: Complete assignment of the ^{13}C and ^{15}N NMR signals of ramoplanin by inverse NMR spectroscopy. J. Antibiotics, in preparation
11) BUSH, K. L.; G. L. GLISH & S. A. MCLUCKEY: Mass Spectrometry/mass spectrometry. VCH Publish, 1988
12) KNAPP, D. R.: Handbook of Analytical Derivatization Reactions. John Wiley & Sons Inc., 1979
13) FRANK, H.; G. J. NICHOLSON & E. BAYER: Rapid gas chromatographic separation of amino acid enantiomers with a novel chiral stationary phase. J. Chromatogr. Sci. 15: 174~176, 1977
14) KETTENRING, J. K.; R. CIABATTI, G. WINTERS, G. TAMBORINI & B. CAVALLERI: Ramoplanin (A-16681), a new glycolipodepsipeptide antibiotic. IV. Complete sequence determination by homonuclear 2D NMR spectroscopy. J. Antibiotics 42: 268~275, 1989
15) ABE, H.; K. KUSHIDA, Y. SHIOBARA & M. KODAMA: The structures of sulfomycin I and berninamycin A. Tetrahedron Lett. 29: 1401~1404, 1988
16) MATSUMOTO, M.; Y. KAWAMURA, Y. YASUDA, T. TANIMOTO, K. MATSUMOTO, T. YOSHIDA & J. SHOJI: Isolation and characterization of thioxamycin. J. Antibiotics 42: 1464~1469, 1989
17) BROOKES, P.; A. T. FULLER & J. WALKER: Chemistry of micrococcin P. J. Chem. Soc. (Part I.) 1957: 689~699, 1957
18) BYCROFT, B. W. & M. S. GOWLAND: The structures of the highly modified peptide antibiotics micrococcin P1 and P2. J. Chem. Soc. Chem. Commun. 1978: 256~258, 1978 and references therein
19) SHOJI, J.; T. KATO, Y. YOSHIMURA & K. TORI: Structural studies on thiocillins I, II and III. (Studies on antibiotics from the genus Bacillus. XXIX). J. Antibiotics 34: 1126~1136, 1981 and references therein
20) HOUCK, D. R.; L. C. CHEN, P. J. KELLER, J. M. BEALE & H. G. FLOSS: Biosynthesis of the modified peptide antibiotic nosiheptide in *Streptomyces actuosus*. J. Am. Chem. Soc. 110: 5800~5806, 1988