THE TOTAL STRUCTURE OF CEREXIN A

(STUDIES ON ANTIBIOTICS FROM THE GENUS *BACILLUS*. XVI[1])

JUN'ICHI SHOJI, TOSHIYUKI KATO and RYUZI SAKAZAKI

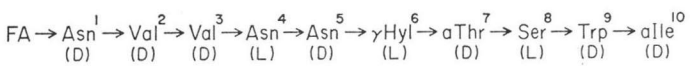
Shionogi Research Laboratory, Shionogi & Co., Ltd.,
Fukushima-ku, Osaka, 553 Japan

(Received for publication September 2, 1976)

The constituent fatty acid of cerexin A was elucidated to be β-hydroxy isoundecanoic acid by gas chromatography, mass spectrometry and nuclear magnetic resonance. The asparaginyl asparagine linkage in the amino acid sequence of the antibiotic, which has been previously described, was proved to be a normal α-carboxylpeptide bond by dehydration-reduction procedure. The three asparagine residues in the sequence were isolated separately and their chiralities were determined. The structure of cerexin A was concluded from the above results.

Cerexin A is an antibiotic active against Gram-positive bacteria produced by *Bacillus cereus* 60-6.[2] The approximate empirical formula of $C_{66}H_{109}N_{15}O_{18}$ and the amino acid sequence of the antibiotic have already been reported.[3,4] Some additional experiments described here clarified the total structure of cerexin A as in Fig. 1.

Fig. 1. Structure of cerexin A.



FA : β-Hydroxy isoundecanoyl
γHyl : L-threo-γ-Hydroxylysine

For preliminary analysis of the constituent fatty acids the ethereal extract of the hydrolyzate of the antibiotic was methylated and examined with gas chromatography. When the hydrolyzate obtained under the usual hydrolysis conditions for amino acid analysis, *e.g.* in constant boiling hydrochloric acid at 110°C for 20 hours, was used for the analysis, four majour peaks were observed (Fig. 2-b). Earlier,[3] this observation was tentatively interpreted as the presence of a variety of fatty acid components in the antibiotic. However, the sample prepared by hydrolysis for 1 hour gave primarily one peak (Fig. 2-a). Moreover, when the ethereal extract from the hydrolysis for 20 hours was hydrogenated and analyzed, one of the four peaks observed before hydrogenation shifted to a position of shorter retention time (Fig. 2-c). By comparison[5] of the retention time of the shifted peak with that of reference fatty acid esters, it was tentatively identified with methyl isoundecanoate. It has been reported that longer hydrolysis of the antibiotic EM49[6] resulted in extensive destruction of β-hydroxy fatty acids and formation of a mixture consisting largely of α,β-unsaturated fatty acids and butyrolactones. Also, similar observations have already been made in our work on the fatty acid constituent of antibiotic 333-25 which contains β-hydroxy anteisononanoic acid.[7] Therefore, the fatty acid constituent of cerexin A was assumed to be β-hydroxy isoundecanoic acid.

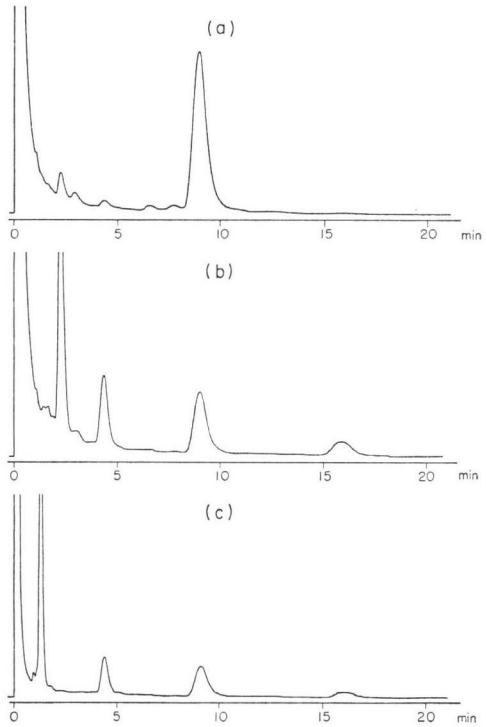
The fatty acid methyl ester was isolated by preparative gas chromatography from the methylated product of the ethereal extract obtained from a one-hour hydrolyzate. The n.m.r. spectrum (Fig. 3)

and mass spectrum (Fig. 4) of the ester are illustrated. As in the case of methyl β-hydroxy anteisononanoate,[7] the molecular ion peak was not obtained, but peaks at *m/e*, 198 ($M-H_2O$), 183 ($M-H_2O$, CH_3), 167 ($M-H_2O$, OCH_3), 155 ($M-H_2O$, C_3H_7), 141 ($M-H_2O$, C_4H_9), 103 ($CH(OH)CH_2\text{-}COOCH_3$), 74 ($CH_2=C(OH)\text{-}OCH_3$) *etc.* were observed. As already mentioned, the base peak, *m/e*, 103, attributable to the fragment ion caused by β, γ-fragmentation seems to be the common base peak of β-hydroxy fatty acid methyl esters.

Thus, all the constituents of cerexin A were determined. The molecular formula of $C_{63}H_{103}N_{15}O_{19}$ was anticipated from the constituents and the calculations for this formula nearly coincided with the elemental analyses[3] of cerexin A and its salts, as shown in Table 1.

In the previous paper,[4] we could not conclude whether the asparaginyl asparagine linkage in the amino acid sequence was a normal α-carboxyl peptide bond or β-carboxyl bond, since EDMAN degradation reaction with deacyl cerexin A hardly proceeded at the step (Asn→Asn) and thereafter. Some assumptions were made, but the problem, we thought, should be clarified by more certain evidence.

Fig. 2. Gas chromatograms of fatty acid methyl esters from cerexin A.
(a) Preparation from 1-hour hydrolysis
(b) Preparation from 20-hour hydrolysis
(c) Hydrogenated product of the preparation from 20-hour hydrolysis



RESSLER and KASHELIKAR[8] reported a simple identification of asparaginyl and glutaminyl residues

Fig. 3. N.m.r. spectrum of methyl β-hydroxy isoundecanoate from cerexin A.
The spectrum was recorded with a Varian A-60 spectrometer on solution in CCl_4 containing TMS.

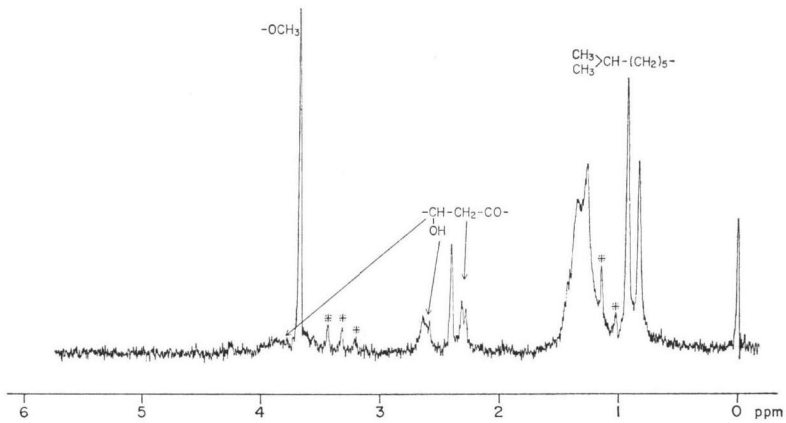


Fig. 4. Mass spectrum of methyl β-hydroxy isoundecanoate from cerexin A.

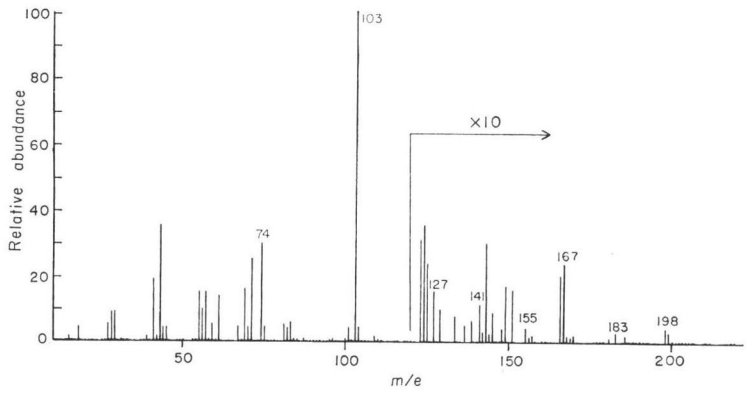


Table 1. Calculable molecular formulas and analytical data[3] of cerexin A and its salts.

Cerexin A (free form). *Anal.* Found	C, 53.31; H, 7.72; N, 14.27
$C_{63}H_{103}N_{15}O_{19}\cdot 3H_2O$	C, 52.96; H, 7.69; N, 14.71
Hydrochloride. *Anal.* Found	C, 51.77; H, 7.84; N, 14.32; Cl, 2.64
$C_{63}H_{103}N_{15}O_{19}\cdot HCl\cdot 3H_2O$	C, 51.64; H, 7.57; N, 14.34; Cl, 2.42
Sulfuric acid salt. *Anal.* Found	C, 50.25; H, 7.52; N, 13.63; S, 2.38
$C_{63}H_{103}N_{15}O_{19}\cdot H_2SO_4\cdot 3H_2O$	C, 49.56; H, 7.33; N, 13.76; S, 2.10
Sodium salt. *Anal.* Found	C, 54.35; H, 6.26; N, 13.86; Na, 2.96
$C_{63}H_{102}N_{15}O_{19}Na\cdot H_2O$	C, 53.49; H, 7.41; N, 14.85; Na, 1.63

in *endo*-position in peptides by dehydration-reduction procedure. The peptide amides are dehydrated on a microscale in triethyl phosphite with the peptide coupling agent ethylene chlorophosphite, to yield the corresponding cyanopeptide derivatives which are then subjected directly to a micro-BIRCH reduction. The treated peptides are hydrolyzed, and analyzed with an automatic amino acid analyzer. By the procedure, *endo*-asparaginyl residues give rise to 2,4-diaminobutyric acid through β-cyanoalanyl derivatives, whereas isoasparaginyl residues give rise to β-alanine.

We applied this procedure to cerexin A with a slight modification: Thionylchloride in dimethylformamide was used for the dehydration reaction. To verify our procedure, two model compounds, N-carbobenzoxyl-L-asparaginyl glycine methyl ester and N-carbobenzoxy-L-isoasparaginyl glycine methyl ester, were synthesized and subjected to the modified dehydration-reduction procedure. As expected, a reasonable amount of 2,4-diaminobutyric acid from the former and β-alanine from the latter peptide were produced. When the procedure was applied to cerexin A, the result was complicated because of extensive side reactions. However, in respect to the fate of the asparagine residues, we found a large amount of 2,4-diaminobutyric acid in addition to a small amount of aspartic acid but no β-alanine. This clearly indicated that all the three asparagine residues in cerexin A exist really as asparaginyl residues, but not as isoasparaginyl residues at all; namely, all the residues link to the next through normal α-carboxyl peptide bonds.

In deacyl cerexin A, the results were almost similar to the above, but a relatively larger amount of aspartic acid was formed in comparing with the case of cerexin A and a small amount of β-alanine was also determined. This suggests that some bond migration of an asparaginyl residue to produce largely

an isoaspartyl residue might occur, possibly at the site (Asn→Asn) during the deacylation procedure. Here, isoaspartyl residue is unaffected by the dehydration-reduction procedure. It has been reported that asparaginyl residues in the sequence of Asn→Gly of adrenocorticotropic hormones undergo preferential deamination in dilute alkaline solution to produce a mixture of α- and β-aspartyl peptides, consisting of largely β-peptides. A mechanism through a succinimide ring system formed by the asparaginyl side chain and the neighbouring C-terminal amide nitrogen of the peptide has been proposed.[9]

In the previous paper,[4] we have already mentioned the suspicion that such a bond migration as described above might occur during successive application of EDMAN degradation reaction and/or the deacylation procedure, in both of which the peptide is exposed to slightly alkaline condition. We have also encountered with a similar difficulty in successive EDMAN degradation reaction in structural studies of brevistin,[10] which contains an Asn→Asp linkage. The degradation reaction applied to deacyl brevistinic acid proceeded only in a few extent at the step (Asn→Asp), as judged by analysis of the remaining peptide. In this case, however, the next residue could be practically indicated by the PTH-amino acid produced in a small amount. Although we have only a few examples, it may be assumed that such a bond migration tends to occur particularly in the sequence, Asn→Asn or Asn→Asp.

The amino acid residues in the sequence of cerexin A are numbered from the N-terminus as shown in Fig. 1. Deacyl cerexin A was dinitrophenylated and then hydrolyzed. From the hydrolyzate, DNP-Asp^1 was isolated by TLC. The heptapeptide with Asn^4 at the N-terminus was prepared by three-step process of EDMAN degradation. Dinitrophenylation and hydrolysis afforded DNP-Asp^4 and nondinitrophenylated Asp^5, which were separated readily and then the Asp^5 was also dinitrophenylated. CD curves of these DNP-Asp specimens were measured and compared with those of synthesized DNP-D-Asp and DNP-L-Asp. From the comparison, Asn^1 was deduced to be D-form, Asn^4 to be L-form and Asn^5 to be D-form, respectively.

From these results, we concluded the total structure of cerexin A as shown in Fig. 1, except for the configuration of the fatty acyl residue.

Experimental

Gas chromatographic examination:

For analytical purpose a Perkin-Elmer Model 881 equipped with a hydrogen flame detector and a steel column (6 feet, 1/8-inch diameter) packed with 15% diethylene glycol succinate polymer on Chromosorb W 80~100 mesh was used. Conditions were: carrier gas (N_2), 5.0 kg/cm^2, approximately 30 ml/min, temperature, 160° or 130°C.

Ten mg of cerexin A was hydrolyzed with constant boiling hydrochloric acid at 110°C for 1 hour or 20 hours. The hydrolyzate was extracted with 2 ml of ether three times. The ethereal extract was dried with anhydrous sodium sulfate and filtered into a 10-ml glass-stoppered centrifuge tube. The tube was dipped in a cold water bath, and was evaporated by a stream of nitrogen. To the concentrate, a few drops of ethereal solution of diazomethane were added. After a few minutes, excess diazomethane was removed by a stream of nitrogen, and the ethereal solution was ready for GC analysis. Another preparation of the ethereal extract from 20-hour hydrolysis was evaporated to an oily film, dissolved in 2 ml of methanol and hydrogenated for 1 hour in the presence of platinum oxide. The methanol was carefully removed by a stream of nitrogen and the hydrogenated product was methylated as above.

The chromatograms of the three samples run at 160°C are illustrated in Fig. 2. When the last sample (the hydrogenated product) was analyzed by the GC at 130°C, a main peak at a retention time of 3.60 minutes was observed. Comparison of the retention time with those of suitable fatty acid esters by the usual graphical method led to tentative identification of the peak with methyl isoundecanoate.

As reference compounds, methyl isooctanoate, methyl anteisononanoate (prepared from polymyxin E), methyl isododecanoate and methyl anteisotridecanoate (prepared from amphomycin) were used in addition to commercially available normal fatty acid methyl esters.

Isolation of methyl β-hydroxy isoundecanoate:

The mass spectrum was measured with a Hitachi RMU-6 mass spectrometer.

A 500 mg sample of cerexin A dissolved in 8 ml of constant boiling hydrochloric acid in a sealed tube was heated at 110°C for 1 hour. The hydrolyzate was extracted with 20 ml of ether three times. The residual solution was concentrated to a residue, which was again hydrolyzed in the same manner. The hydrolysis and extraction with ether were repeated three times. The ethereal extracts were combined, dried (Na_2SO_4), concentrated carefully by a stream of nitrogen and methylated.

For preparative purpose of gas chromatography a Varian Aerograph Model 1520-IB equipped with a thermal conductivity detector and an aluminum column (10 feet, 3/8-inch outer diameter) packed with 5% diethylene glycol succinate polymer on Chromosorb Q 60~80 mesh was used under the following conditions: carrier gas (He), 4 kg/cm^2, approximately 200 ml/min; temperature, 180°C.

About ten runs were needed, the main peak was collected in a trapping tube as a colorless oil. The n.m.r. and mass spectra are illustrated in Figs. 3 and 4.

Dehydration-reduction procedure with cerexin A and deacyl cerexin A:

About 5 mg of cerexin A dissolved in dimethylformamide (0.3 ml) was cooled in an ice-bath. Thionylchloride (0.1 ml) was added to the solution, which was stood in the cold for 2 hours and then at room temperature for 16 hours. The reaction mixture, a dark brown solution, was concentrated under reduced pressure to a syrup, which was then dissolved in water (1 ml) and freeze-dried. To the residue in a flask, methanol (0.1 ml) was added and then a few mililiters of liquid ammonia were trapped under cooling with dry ice-acetone. Small pieces of sodium were added slowly until a blue color appeared became to persist for at least one minute. Then, the ammonia gas was allowed to evaporate spontaneously at room temperature. The reaction mixture was dissolved in a small amount of methanol and water, evaporated under reduced pressure, neutralized and slightly acidified (pH 2.0) with 6 N hydrochloric acid, and freeze-dried. A portion of the product was hydrolyzed with constant boiling hydrochloric acid at 110°C for 40 hours and the hydrolyzate was analyzed with an automatic amino acid analyzer Hitachi KLA-5. Commercially available samples of 2,4-diaminobutyric acid and β-alanine were used as references for the analysis.

Among the amino acids found in the analysis, which originated from asparagine residues were as follows (in ratio): aspartic acid (0.17), and 2,4-diaminobutyric acid (0.83). No β-alanine was detected.

Some 5 mg of deacyl cerexin A were processed with the same procedure. The amino acids originated from asparagine residues were found as follows (in ratio): aspartic acid (0.40), 2,4-diaminobutyric acid (0.54) and β-alanine (0.06).

In these analyses, in addition to other constituent amino acids, *allo*-threonine, serine, valine and *allo*-isoleucine, the following amino acids were also found: lysine, α-aminobutyric acid and alanine (probably from reduction of γ-hydroxylysine, *allo*-threonine and serine), glycine (probably from retroaldol condensation of *allo*-threonine and serine) and isoleucine (probably from α-epimerization of *allo*-isoleucine).

Dehydration-reduction procedure with N-carbobenzoxy-L-asparaginyl glycine methyl ester and N-carbobenzoxy-L-isoasparaginyl glycine methyl ester:

N-Carbobenzoxy-L-asparginyl glycine methyl ester was prepared by coupling of N-carbobenzoxy-L-asparagine *p*-nitrophenyl ester (commercially available) and glycine methyl ester. N-Carbobenzoxy-L-isoasparagine was coupled with *p*-nitrophenol by dicyclohexylcarbodiimide to yield N-carbobenzoxy-L-isoasparagine *p*-nitrophenyl ester, which was then coupled with glycine methyl ester to give N-carbobenzoxy-L-isoasparaginyl glycine methyl ester.

One mg each of the above two model peptides was dehydrated and reduced by the same manner as described in the former section. The amino acids found (in ratio) were: asparatic acid (0.05), 2,4-diaminobutyric acid (0.90), glycine (0.19) and ethanolamine (0.81) from N-carbobenzoxy-L-asparaginyl glycine methyl ester; and aspartic acid (0.11), β-alanine (0.68), glycine (0.31) and ethanolamine (0.69)

from N-carbobenzoxy-L-isoasparaginyl glycine methyl ester. Here, the molar ratio was expressed by setting sum of glycine and ethanolamine as 1.00, because the ethanolamine was derived from a part of glycine methyl ester residue.

The chiralities of Asn^1, Asn^4 and Asn^5:

The CD curve was recorded with a spectropolarimeter JASCO Model ORD/UV-6.

Deacyl cerexin A, 15 mg, prepared by enzymatic deacylation with Polymyxin Acylase,[4] was dissolved in water (1.0 ml). Sodium bicarbonate (47 mg) and then 5% 2,4-dinitrofluorobenzene solution in ethanol (1.0 ml) were added to the solution. After being stirred for 3 hours in the dark, the reaction mixture was evaporated and extracted with ether. When the aqueous solution was acidified with hydrochloric acid, the dinitrophenylated product of deacyl cerexin A precipitated (17 mg). It was hydrolyzed with hydrochloric acid and DNP-Asp^1 was extracted from the hydrolyzate with ethyl acetate. Purification of DNP-Asp was carried out by preparative TLC [on Silica gel GF with chloroform - ethanol - 14% ammoniacal water (4: 7: 2), Rf *ca.* 0.38]. DNP-Asp was extracted from the plate with a mixture of methanol and 14% ammoniacal water (1: 1) and transferred to ethyl acetate, which was washed with dil.HCl and water, dried (Na_2SO_4) and evaporated to dryness. Approximately 690 mcg of DNP-Asp^1, which was estimated from the optical density at 362 nm, was obtained, and subjected to CD measurement.

Some 15 mg of deacyl cerexin A was processed by the procedure of EDMAN degradation up to the third step. The remaining peptide, which is the heptapeptide with Asn^4 at the N-terminus,[4] was dinitrophenylated in the same manner. The reaction mixture was evaporated and extracted with ether, before and after acidification with HCl. Then, the DNP-peptide was extracted with butanol and hydrolyzed with hydrochloric acid. The DNP-Asp^4 was extracted with ethyl acetate from the hydrolyzate and purified by the TLC as described above, being obtained as a pure preparation (672 mcg). The residual aqueous solution, which contained a mixture of amino acids involving Asp^5, was dinitrophenylated and separated by the TLC. The DNP-Asp^5 was extracted from the plate and obtained as a pure preparation (1.312 mcg).

DNP-Asp^1, CD: $[\theta]_{270}0$, $[\theta]_{228}+7300$, $[\theta]_{220}+5700$ (*c* 0.0365, N NaOH);

DNP-Asp^4, CD: $[\theta]_{270}$ 0, $[\theta]_{228}-9580$, $[\theta]_{220}-6330$ (*c* 0.0319, N NaOH);

DNP-Asp^5, CD: $[\theta]_{270}$ 0, $[\theta]_{228}+5900$, $[\theta]_{220}+5000$ (*c* 0.0312, N NaOH)

The values of synthetic specimens have been reported in our previous work as below.[10]

DNP-D-Asp, CD: $[\theta]_{267}$ 0, $[\theta]_{228}+15700$, $[\theta]_{216}$ 0 (*c* 0.0785, N NaOH);

DNP-L-Asp, CD: $[\theta]_{269}$ 0, $[\theta]_{230}-12100$, $[\theta]_{216}$ 0 (*c* 0.0685, N NaOH)

Acknowledgement

The authors are indebted to Dr. YUZO NAKAGAWA for measurement of MS, and also Dr. KAORU KURIYAMA and TATSUO IWATA for measurement of CD, members of Shionogi Research Laboratory.

References

1) SHOJI, J.; H. HINOO, Y. WAKISAKA, K. KOIZUMI & M. MAYAMA: Isolation of a new peptide antibiotic complex, B-43 (Studies on antibiotics from the genus *Bacillus*. XV). J. Antibiotics 29: 813~817, 1976

2) SHOJI, J.; H. HINOO, Y. WAKISAKA, K. KOIZUMI, M. MAYAMA, S. MATSUURA & K. MATSUMOTO: Isolation of two new related peptide antibiotics cerexins A and B (Studies on antibiotics from the genus *Bacillus*. I). J. Antibiotics 28: 56~59, 1975

3) 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

4) SHOJI, J. & T. KATO: The amino acid sequence of cerexin A (Studies on antibiotics from the genus *Bacillus*. VII). J. Antibiotics 28: 764~769, 1975

5) SHOJI, J. & H. OTSUKA: Studies on tsushimycin. II. The structures of constituent fatty acids. J. Antibiotics 22: 473~479, 1969

6) PARKER, W. L. & M. L. RATHNUM: EM49, a new peptide antibiotic. II. Chemical characterization. J. Antibiotics 26: 449~456, 1973

7) SHOJI, J.; H. HINOO & R. SAKAZAKI: The constituent amino acids and fatty acid of antibiotic 333-25 (Studies on antibiotics from the genus *Bacillus*. XII). J. Antibiotics 29: 521~525, 1976
8) RESSLER, C. & D. V. KASHELIKAR: Identification of asparaginyl and glutaminyl residues in *endo*-position in peptides by dehydration-reduction. J. Am. Chem. Soc. 88: 2025~2035, 1966
9) a) GRAF, L.; S. BAJUSZ, A. PATTHY, E. BARÁT & G. CSEH: Revised amide location for porcine and human adrenocorticotropic hormone. Acta Biochim. et Biophys. Acad. Sci. Hung. 6: 415~418, 1971
b) SIEBER, P.; W. RITTEL & B. RINIKER: Die Synthese von menschlichen adrenocorticotropem Hormon (α_h-ACTH) mit revidierter Aminosäuresequenz. Helv. Chim. Acta 55: 1243~1248, 1972
10) SHOJI, J. & T. KATO: The structure of brevistin (Studies on antibiotics from the genus *Bacillus*. X). J. Antibiotics 29: 380~389, 1976