77. Complete Sequence Determination and Localisation of One Imino and Three Sulfide Bridges of the Nonadecapeptide *Ro 09-0198* by Homonuclear 2D-NMR Spectroscopy. The DQF-RELAYED-NOESY-Experiment

by **Horst Kessler*** and **Stefan Steuernagel**

Institut für Organische Chemie der Johann Wolfgang Goethe-Universität, Niederurseler Hang, D–6000 Frankfurt am Main 50

and **Dieter Gillessen**

Zentrale Forschungseinheiten, *F. Hoffmann-La Roche & Co. AG*, CH–4002 Basel

and **Tsutomu Kamiyama**

Department of Microbiology and Chemotherapy, *Nippon Roche Research Center*, Kamakura 247, Japan

Professor *Gerhard Quinkert* zum 60. Geburtstag gewidmet

(24. II. 87)

The constitution of the nonadecapeptide *Ro 09-0198* (**1**) has been established by homonuclear two-dimensional NMR spectroscopy. The peptide contains three sulfide bridges and a bridge formed from the side chain of a lysine residue (lysinoalanine moiety). Only the composition was known before, thus the complete sequencing and the elucidation of the bridges had to be performed. Of special importance was the application of the RELAYED-NOESY experiment. This technique was improved by the introduction of a double-quantum filter to reduce the tailing around the diagonal. The resulting sequence of **1** exhibits a surprizing analogy of this immunopotentiating peptide to the ACE inhibitor ancovenin.

1. Introduction. – NMR spectroscopy is increasingly important for structural elucidation of peptides [1] and proteins [2]. The assignment of NMR spectra and the determination of conformations of compounds *with known amino-acid sequence* is routinely achieved [2] with 2D correlated techniques [3] [4] in combination with 2D nuclear *Overhauser* and exchange spectroscopy (NOESY) [5]. Each amino-acid residue shows an isolated scalarly coupled spin system in the ^{1}H-NMR spectrum, which can normally be analyzed by H,H-COSY [3] [6], H-RELAYED-H,H-COSY [4], and TOCSY [7] [8] techniques. However, the assignment of all signals of side-chain protons in spectra of larger peptides is often impossible without information from heteronuclear measurements [9] [10]. An incomplete elucidation of the proton-spin systems then prevents the identification of the amino-acids residues. This problem can be overcome by the knowledge of the amino-acid sequence. In this case, NOEs between protons of adjacent residues can be used to assign sequential connectivities [1] [2]. To our knowledge, the elucidation of the constitution of a medium-sized peptide with unknown sequence has not yet been successful. Here we report our investigation of the new nonadecapeptide *Ro 09-0198* (**1**), which was found in the culture broth of *Streptoverticillum griseoverticillatum* which shows interesting immunopotentiating activities [11]. It was possible to assign all proton signals only by several homonuclear ^{1}H-NMR measurements. The analysis of sequential assign-

ments resulted in the determination of the petide sequence. In addition, the positions of 3 sulfide bridges as well as the imino bridge in the side chain of a lysinoalanine moiety could be localized [12].

The success was mainly made possible by the use of a new variant of RELAYED-NOESY technique [13]. A distinct improvement has been obtained by the introduction of a double-quantum filter between the RELAYED transfer and the acquisition. The thus obtained narrow diagonal peaks allowed the elucidation of the connectivities of H–C(3)'s across the S-atoms in the lanthionine (Ala(3-Cys(*S*))) and β-methyllanthionine (Abu(3-Cys(*S*))) moieties.

The procedure shown in this paper is especially important because the structural elucidation of polycyclic S-bridged oligopeptides is extremely difficult by conventional techniques, see *e.g.* the elucidations of duramycin [14] [15] and cinnamycin [15] or epidermin [16].

2. Materials. – *Ro 09-0198* (**1**) was obtained from *Streptoverticillum griseoverticillatum* MAR 164C-MY6 isolated from a soil sample. The strain MAR 164C-MY6 was cultured for 7 days at 27° in a medium containing 4% of sucrose, 4% of potato starch, 1% of wheat germ, 2% of toast soya, 0.25% of NaCl, and 0.32% of $CaCO_3$. After filtration, the cultured broth was applied to a column of *Diaion WK-10* and eluted with 2N AcOH/EtOH 1:1. The eluate was concentrated to a small volume and then extracted with BuOH at pH 7.6. After removal of BuOH, the residue was crystallized from MeOH to give pale brown crystalls of **1**. Recrystallization from MeCN/H_2O 3:2 (pH 4.1) yielded pure **1** · HCl (pH 4.1) as colorless needles.

3. Techniques. – 3.1. *General.* The most important homonuclear 2D NMR experiments for the analysis of molecular constitution and conformation of peptides [1] and proteins [2] in solution can be classified into two groups. The first group of experiments enables the detection and analysis of isolated amino acid spin systems by scalar couplings: from the number of available experiments, the DQF-COSY [6], the RELAYED-COSY [4], and the TOCSY [7] [8] experiments are the most frequently applied techniques. Experiments of the second group use through space effects (NOE) in order to provide a sequential analysis of these spin systems NOESY [5], RELAYED-NOESY [13], CAMELSPIN or ROESY [17].

All spectra shown in this paper were acquired and processed in the phase-sensitive mode [18]. For all spectra 512 FIDs, each of 2 K complex data points, were collected. For each t_1 increment, between 16 and 64 transients were recorded, dependent on the phase cycle. Zero-filling to yield a 2K times 1K real data matrix and multiplication with an appropriate window function (shifted squared sine bell) were done before *Fourier* transformation. All spectra were phase-corrected to yield signals with pure absorption line-shapes.

3.2. *The DQF-RELAYED-NOESY Experiment.* In spectra with overlapping signals, the assignment of the cross peaks may be difficult or impossible. For an unknown compound, it is extremely important to assign all signals unambiguously in order to avoid circular arguments [1], especially when the measurements are restricted to homonuclear experiments. At the stage of spin-system elucidation, this problem can be solved by means of the RELAYED-COSY technique [4]. To overcome this problem at the stage of *sequential* resonance assignment the RELAYED-NOESY experiment has been described recently [13]. We propose an improved variant of this technique by the introduction of a double-quantum filter between the RELAYED- step and the acquisition. The advantages of the double-quantum filter will be explained in the following by a comparison of the pulse sequences shown in *Fig. 1*. In this context, we follow the notation of the product operator formalism [19] proposed by *Sørensen et al.* We consider an *ABX* system having only one coupling J_{AB} and one NOE between *A* and *X* or between *B* and *X*.

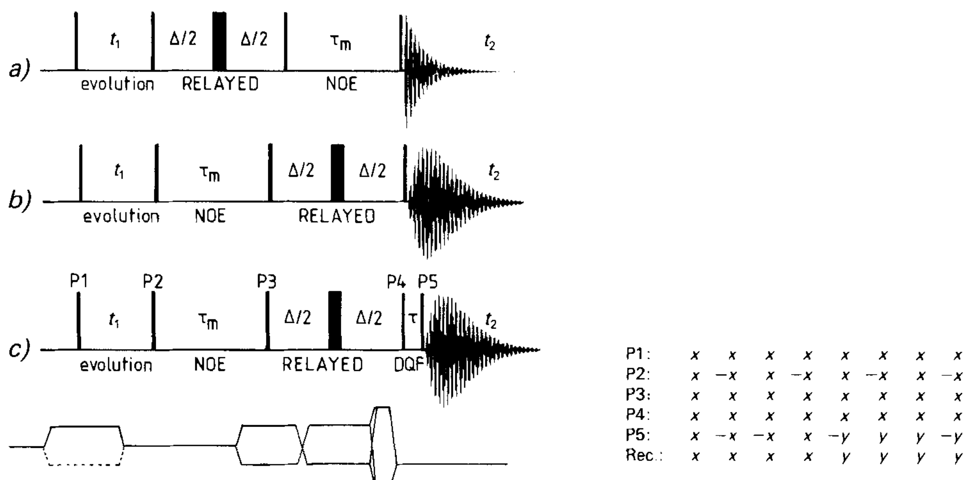


Fig. 1. *Pulse sequences for* a) *and* b) *the RELAYED-NOESY and* c) *the DQF-RELAYED-NOESY experiments.* The 90°- and 180°-pulses are symbolized by thin and thick bars, respectively. Pulse sequence *a*) describes the RELAYED-NOESY with the RELAYED step before the NOE transfer and *b*) with the RELAYED step following the NOE transfer; both taken from [13]. Pulse sequence *c*) has been used in this paper. Under this pulse sequence, the coherence transfer pathway [21] and phase cycle used for the DQF-RELAYED-NOESY experiment are shown. Phases of P1 and P2 have to be shifted in steps of 90° without incrementing the receiver phase resulting in a 32 step phase cycle. For better suppression of artifacts, phases of P3 and P4 can be cycled additionally. For phase-sensitive acquisition with quadrature detection in ω_1, TPPI [18] is executed.

In sequence *a*), the anti-phase magnetization of a nucleus A with respect to nucleus B ($2I_x^A I_z^B \sin\pi J_{AB} t_1$; chemical-shift effects are neglected) is converted by the second 90_y° pulse into antiphase magnetization of nucleus B ($-2I_z^A I_x^B \sin\pi J_{AB} t_1$), which is refocused in the Δ delay to ($-I_y^B \sin\pi J_{AB} t_1 \sin\pi J_{AB}\Delta$). This part of the magnetization is converted into polarization *via* the third 90_x° pulse. The last part of the pulse sequence is identical to the NOESY sequence.

In pulse sequence *b*), however, the first part corresponds to the NOESY sequence. Thus, the in-phase magnetization of a nucleus A is, after the third 90° pulse, modulated with the chemical shift of another nucleus, *e.g.* X, due to the transfer of NOE. This term evolves anti-phase magnetization due to *J*-coupling ($2I_x^A I_z^B \cos\Omega_X t_1 \sin\pi J_{AB}\Delta$), which is converted into antiphase magnetization of nucleus B ($-2I_z^A I_x^B \cos\Omega_X t_1 \sin\pi J_{AB}\Delta$) by the last 90° pulse. It is evident that this resembles the typical polarization transfer of the COSY experiment. This polarization transfer can be substituted by the pair of pulses of a double-quantum filter: The first pulse (P4 in *Fig. 1c*) creates multiple-quantum coherences, which are reconverted into measurable single-quantum coherences (here, the chemical-shift modulation is neglected for simplicity):

$$2I_x^A I_z^B \sin\pi J_{AB}\Delta$$

$$\downarrow (\pi/2)_x$$

$$-2I_x^A I_y^B \sin\pi J_{AB}\Delta = -\tfrac{1}{2}\cdot(2I_x^A I_y^B \sin\pi J_{AB}\Delta + 2I_y^A I_x^B \sin\pi J_{AB}\Delta) \qquad A$$

$$-\tfrac{1}{2}\cdot(2I_x^A I_y^B \sin\pi J_{AB}\Delta - 2I_y^A I_x^B \sin\pi J_{AB}\Delta) \qquad B$$

$$\downarrow (\pi/2)_x$$

$$-\tfrac{1}{2}\cdot(2I_x^A I_z^B \sin\pi J_{AB}\Delta + 2I_x^A I_z^B \sin\pi J_{AB}\Delta)$$

Terms A and B represent a linear combination of double-quantum and zero-quantum coherences. By an appropriate phase cycling, double-quantum coherences (term A) can be selected exclusively. In *Fig. 1* this is shown for pulse sequence *c*) in combination with the coherence transfer pathway [21].

The reduction of peak intensities by the use of a double-quantum filter is more than compensated by the filtering of the dispersive contributions, similar to the advantages in the DQF-COSY [6] compared to COSY. Therefore, cross peaks close to the diagonal can be much better recognized. The appearance of regular in-phase/anti-phase patterns for all peaks facilitates their recognition. An improvement for the interpretation of spectra obtained by this technique compared to a conventional NOESY spectrum results from the twofold asymmetry: *i*) Cross peaks which have symmetric positions with respect to the diagonal occur whenever scalarly coupled nuclei exhibit an NOE. However, as a consequence of the asymmetry of the pulse sequence the resulting cross-peak fine structure is also asymmetric [20], *i.e.* a reflection of a peak at the diagonal meets a mirror peak with inverted sign (see *Fig. 2*). The same will be observed for diagonal signals.

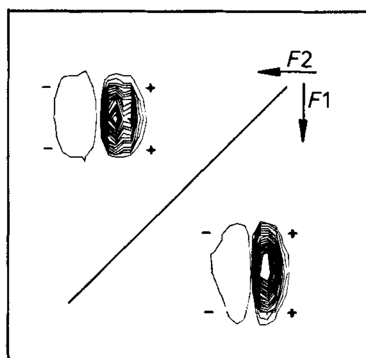


Fig. 2. *Representation of the asymmetry of cross peaks in the DQF-RELAYED-NOESY spectrum.* Two cross peaks occurring symmetrically with respect to the diagonal (*i.e.* magnetization transfer between 2 nuclei occurs *via* NOE as well as *via J* coupling) are shown indicating their characteristic mixed in-phase/anti-phase structure. Positive levels are drawn with more contour lines than negative levels. When the peaks are tried to be reflected at the diagonal, positive lines will meet with negative ones and *vice versa*.

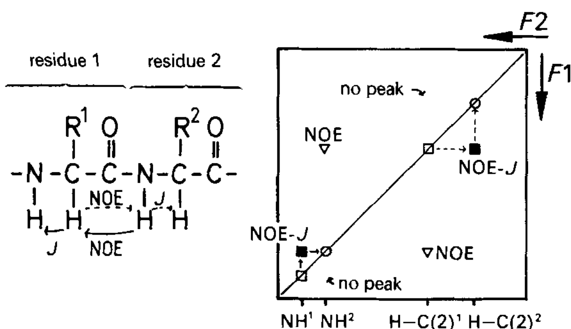


Fig. 3. *Schematic representation of the asymmetric appearance of cross peaks in the DQF-RELAYED-NOESY spectrum.* NOE-*J* cross peaks for H–C(2)^1, H–C(2)^2 and NH^1,NH^2 are obtained by the magnetization transfer shown by dashed and solid arrows, respectively. Cross peaks asymmetric with respect to the diagonal can be used to read the direction of the amino-acid sequence. This is shown in the schematic spectrum (see text).

ii) The spectrum is asymmetric with respect to the appearance of peaks on both sides of the diagonal: signals appearing on one side of the diagonal need not have a counterpart on the other side. This is shown schematically in *Fig. 3*. NOE cross peaks between H–C(2)'s and NH's preceding in the sequence are generally observed. Because these nuclei are not scalarly coupled, magnetization transfer in the DQF-RELAYED-NOESY can only occur in one direction (first step: NOE, second step: coupling, not *vice versa*).

This allows one to read the direction of the peptide sequence from the arising N*H*,N*H* and *H*–C(2),*H*–C(2) NOE-*J* cross peaks. The cross peak between the N*H* signals shown in *Fig. 3* exhibits the chemical shift of N*H*[1] in *F*2 and the chemical shift of N*H*[2] in *F*1. It originates as follows: The chemical shift of N*H*[2] of amino-acid residue 2 has been labeled in t_1. In a first step, this magnetization is transferred *via* NOE to H–C(2) of residue 1, and in a second step the magnetization is transferred to the scalarly coupled N*H* of residue 1. The direction of the peptide sequence is also given in *H*–C(2) cross peaks, however, the ordering is reversed. The 'labeled' H–C(2) of residue *i* in *F*1 polarizes the N*H* of residue *i*+1 *via* NOE, whose polarization is transferred *via* coupling to H–C(2) of residue *i*+1 (*F*2).

4. Strategy. – The NMR analysis was started without any sequential information. However, information about the amino-acid constituents of **1** as the result of acidic hydrolysis was given. This is summarized in *Table 1* together with some spectroscopical considerations. At the beginning of the investigations, the best conditions for the measurements were tested. All spectra were recorded with a 50 mmolar sample solution in (D_6) DMSO at 310 K. The spin systems could be identified only with some difficulties by interpretation of DQF-COSY spectra, because not all signals were completely resolved. Since variation in temperature did not cause significant variations in chemical shifts, 0.01, 0.02, and 0.04 ml C_6D_6 were added, and DQF-COSY spectra were recorded in order to

Table 1. *Quantity and Types of Amino Acids Detected after Acid Hydrolysis.* Some remarks on spectroscopical aspects are given. For the sake of clarity numbering of the sequence is given already here, although it is the result of the work described below.

Amino acid	Quantity	Abbreviations [a])	Remarks
Phenylalanine	3	Phe or F	expected *ABMX* spin systems
Lanthionine (=*S*-(2-amino-2-carboxyethyl)-cysteine)	1	Ala^4/Ala^{14} or A4/A14	two isolated *ABMX* spin systems
Aspartic acid[b])	1	Asn or N	an *ABMX* system, too; side chain function is amide
β-Methyllanthionine (=*S*-(2-amino-2-carboxy-1-methylethyl)-cysteine)	2	Ala^1/Abu^{18} and Ala^5/Abu^{11} or A1/X18 and A5/X11[c])	two isolated spin systems each
Glycine	2	Gly or G	two *ABX* spin systems
Glutamic acid[b])	1	Gln or Q	side chain function is amide
Proline	1	Pro or P	
Arginine	1	Arg or R	
3-Hydroxyaspartic acid	1	(3-OH) or X'[d])	
Valine	1	Val or V	
Lysinoalanine (=N^6-(2-amino-2-carboxyethyl)-lysine)	1	Lys^{19}/Ala^{10} or Lys-NH-Ala or K19/A10	will be regarded as two isolated spin systems, because of problems in the assignment of all signals of Lys; N(6)*H* is expected to be broad

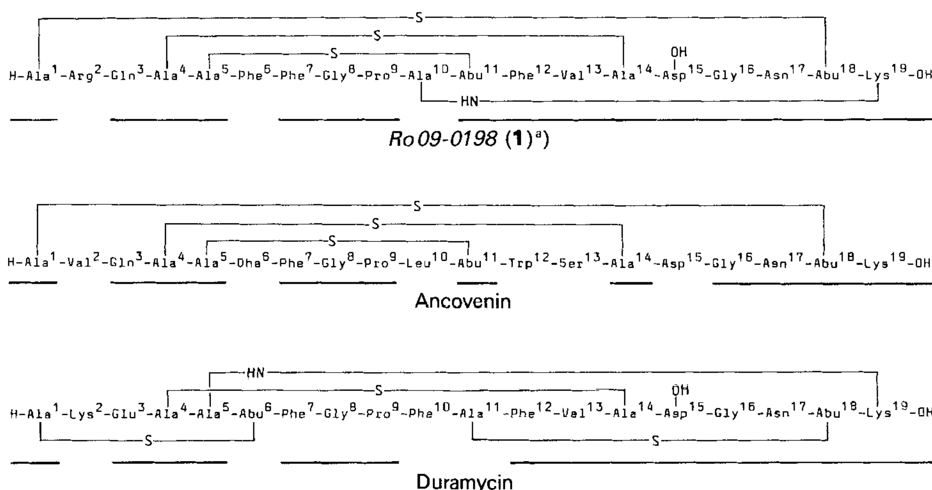
[a]) The abbreviations of the original amino-acid residues in the peptide are given, see [24]. The residue numbers are indicated only for hydrolysis products consisting of 2 amino-acid residues. The configurations at C(2) of the residues are not ascertained.

[b]) Two equivalents of NH_3 have been found that turned out to come from the side chain amide functions of the amino-acid residues asparagine and glutamine.

[c]) X = 2-aminobutyric acid.

[d]) X' = 3-hydroxyaspartic acid.

Scheme. *Comparison of the Sequences of* Ro 09-0198 **1**, *Ancovenin and Duramycin.* Homologous residues are underlined in all peptides.



[a]) The amino-acid configuration at C(2) of Ala[1], Ala[5], Ala[10], and Ala[4] (or Ala[14]) is not ascertained (see *Chapt. 6*).
[b]) Dha = 2,3-didehydroalanine.

distinguish between partially overlapping or degenerate signals. *H*–C(2), *H*–C(3) connectivities were confirmed by use of the RELAYED-COSY technique. Following this procedure, all spin systems could be recognized, however, a discrimination of identical spin systems of different types of amino acids was not possible at this stage.

Reaching this level, the difference in the strategy in our case and in peptides with known sequences becomes obvious: Identical spin systems of different amino acids in peptides *with known sequences* can be analyzed by assigning their sequential connectivities. In our case, this information can only be used for sequence determination and not for characterizing the type of amino acid. Therefore, it is not sufficient to recognize a particular spin system by its characteristic coupling pattern in the COSY spectrum. In order to find out the peptide sequence, it is absolutely necessary to start the interpretation of *connectivities* within the NOESY spectrum with all spin systems *already assigned* to the individual type of amino acid.

5. Results. – 5.1. *Preliminary Considerations.* An observation of the amino acid spin systems in *Table 1* shows that six of overall 19 isolated spin systems occur only once in the whole spectrum (Gln, Pro, Arg, Asp(3-OH), Val, Lys[19]). Two spin systems appear twice: the Abu part of β-methyllanthionine and two sets of Gly resonances. It turnes out that the presence of 9 very similar *ABMX* systems of 4 types of amino acids causes special problems.

Fig. 4 shows the 'fingerprint region' of the DQF-COSY spectrum containing 19 N*H*,*H*–C(2) cross peaks. Counting the number of expected cross peaks indicates that the peptide is not cyclic at the N-terminus: 15 spin systems of the amino acids exhibited in *Table 1* contribute 1 cross peak, proline none, and glycine and arginine each should give rise to 2 (for Arg: *H*–(2),N(2)*H* and *H*–C(5),N(5)*H*). However, as we will see later, due

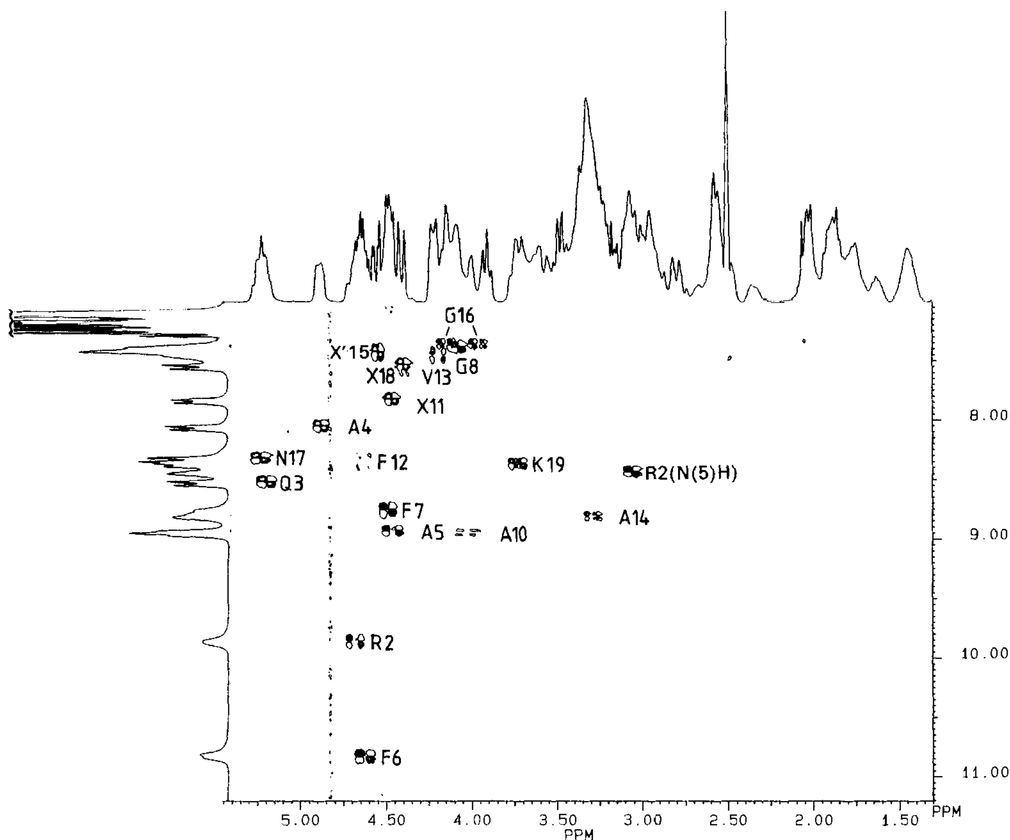


Fig. 4. *NH,H–C(2) region of the 300-MHz DQF-COSY spectrum of* **1** *at 310 K.* Negative levels are drawn with only 1 contour line, while positive levels are drawn with 8 contour lines. On the top and on the left of the two-dimensional spectrum, the corresponding regions of the one-dimensional ^{1}H-NMR spectrum are shown. All N*H*,*H*–C(2) cross peaks in this spectrum are indicated with one-letter symbols for amino-acids residues [24] according to their position in the sequence. See also abbreviations in *Table 1;* X = 2-aminobutyric acid, X' = 3-hydroxyaspartic acid.

to degeneracy of the *H*–C(2) resonances of Gly[8], the maximum number of expected cross peaks would be 20, if the peptide was cyclic at the N- and C-terminus. The presence of two terminal amide groups (–$CONH_2$) is detected by their characteristic coupling patterns in the aromatic region and, in addition, as exchange peaks in the NOESY spectrum.

5.2. *Assignments.* 5.2.1. *Identification of Scalarly Coupled Spin Systems.* The spin systems of all amino acids in this peptide could be assigned in DQF-COSY spectra. The characteristic spin systems of Val, Gln, Asp(3-OH), Lys, Arg, and Pro were assigned directly. Fortunately, all signals of Arg could be found, which allowed a clear differentiation from the spin system of Lys. The N(6)*H* signal of the lysinoalanine moiety could not be found. This signal is expected to be rather broad due to 4 coupling partners and probably due to exchange with the H_2O content of the sample. The identity of the side-chain function of Gln could be determined by NOE, observed between its 2 *H*–C(4) and the protons of one of the terminal NH_2 groups in the NOESY spectrum. There was

also no problem to assign the signals of Abu[11] and Abu[18] of the β-methyllanthionine moieties and of both Gly residues (for assignments see *Figs. 4* and *5*).

The main problem was the assignment of the *ABMX*-spin systems. In contradiction to preliminary assumptions, 8 instead of 9 *ABMX* systems were found. They were assigned to the residues Asn, Phe, the S-bridged Ala[4], Ala[5], and Ala[14], and the N-bridged Ala[10]. The missing *ABMX* spin system immediately shows that one of the residues which could be attributed to such a spin systems does not contain an amide N-atom but an N-terminal free NH_2 group (Ala[1]). The spin system of Asn was identified in analogy to the Gln residue by NOEs between its 2 H–C(3) and the protons of the other terminal NH_2 group observed in the NOESY spectrum.

For the discrimination of Phe and the S-bridged Ala, we started with the following assumptions: Due to the rigidity introduced by the S-bridges, the side chains of the lanthionine and β-methyllanthionine moieties will be frozen in one conformation. This

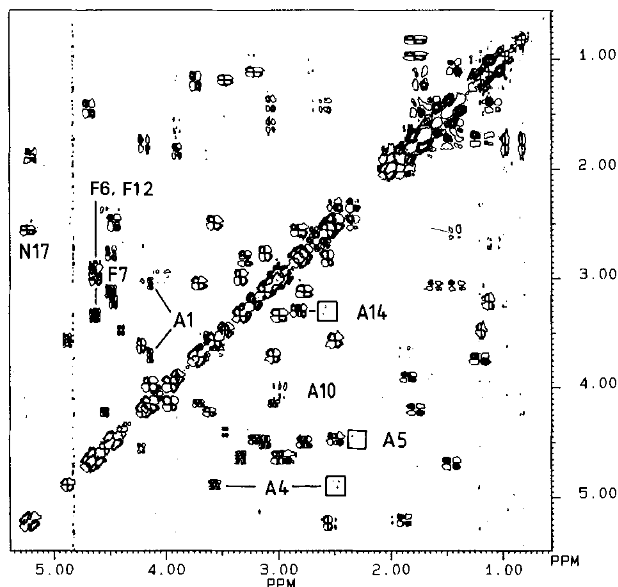


Fig. 5. *High-field part of the 300-MHz DQF-COSY spectrum of* **1** *at 310 K* (see *Fig. 4*). Negative levels are drawn with only 1 contour line, while positive levels are drawn with 8 contour lines. All H–C(2),H–C(3) cross peaks of the amino-acids residues with similar *ABMX* spin systems and of the N-terminal amino-acid moiety are indicated. Cross peaks with low or vanishing intensities because of small coupling constants are indicated by boxes. They partially appear at a lower plotting level but are all visible in the RELAYED-COSY indicating a H–C(2)→H–C(3)→H'–C(3) magnetization transfer. For abbreviations see *Fig. 4*.

will lead to dihedral angles that probably cause small coupling constants, so that one of the H–C(2),H–C(3) cross peaks may vanish. This is the explanation for the observed cross peaks of low or vanishing intensity in 3 cases as indicated in *Fig. 5* (Ala[4], Ala[5] and Ala[14]). These cross peaks appear partially at a lower plotting level but are clearly expressed in the RELAYED-COSY spectrum indicating a H–C(2)→H–C(3)→H'–C(3) magnetization transfer. In 2 cases, the chemical-shift differences between connected

Table 2. *Chemical Shifts in the 500-MHz-^{1}H-NMR Spectrum* ((D)$_6$DMSO) *of* **1** *at 310 K*

Amino acid	Chemical shifts [ppm]			
	N*H*	*H*–C(2)/*H*–C(2)	*H*–C(3)/*H'*–C(3)	Others
Ala[1]	–	4.14	3.70, 3.02	
Arg[2]	9.85	4.68	1.75, 1.45	1.60, 1.40 (2 *H*–C(4)); 3.05, 3.05 (2 *H*–C(5)); 8.41 (N(5)*H*)
Gln[3]	8.51	5.19	1.88, 1.88	2.01, 2.01 (2 *H*–C(4)); 7.05, 6.55 (N(5)*H*)
Ala[4]	8.02	4.88	3.55, 2.48	
Ala[5]	8.90	4.42	2.45, 2.30	
Phe[6]	10.8	4.6	3.31, 2.95	
Phe[7]	8.72	4.45	3.10, 2.74	
Gly[8]	7.38	4.09, 4.09		
Pro[9]	–	3.90	1.81, 1.60	1.90, 1.80 (2 *H*–C(4)); 3.75–3.70 (2 *H*–C(5))
Ala[10]	8.93	4.08 (br.)	3.04 (br.)	
Abu[11]	7.80	4.42	3.45	1.20 (2 *H*–C(4))
Phe[12]	8.31	4.6	2.95, 2.95	
Val[13]	7.45	4.20	1.75	0.95, 0.80 (3 H–C(4), 3 H–C(4'))
Ala[14]	8.80	3.29	2.80, 2.52	
(3-OH)[15]	7.40	4.55	4.20	3.60 (OH)
Gly[16]	7.35	4.15, 3.98		
Asn[17]	8.30	5.21	2.52, 2.52	7.40, 6.80 (N(4)H_2)
Abu[18]	7.51	4.39	3.20	1.10 (3 *H*–C(4))
Lys[19a])	8.32	3.70	1.70, 1.20	1.42, 1.05 (2 *H*–C(5)) 2.70–2.40 (2 *H*–C(6))

[a]) No chemical shifts given for N(6)*H* and 2 *H*–C(4) of Lys[19].

H–C(3) signals are rather large (Ala[1], Ala[4]), which is at least not incompatible with the expected rigidity of these side chains. The remaining H–C(3)'s have to be assigned to the Phe residues. Evidence for the correctness of this conclusion is provided by observed NOEs between those H–C(3)'s and aromatic protons in the NOESY spectrum.

The chemical-shift values of all assigned protons are collected in *Table 2*. In the 'fingerprint region' of the DQF-COSY the corresponding N*H*,*H*–C(2) cross peaks are indicated (*Fig. 4*, see above). In *Fig. 5*, the high field part of the DQF-COSY spectrum is shown with the assignment of some *H*–C(2),*H*–C(3) cross peaks.

At this stage, it has to be pointed out that the distinction between the *ABMX* systems obtained by the preceding argumentation has to be handled carefully to avoid mistakes when determining the positions of the amino-acid residues in the peptide sequence. Although NOEs are observed that can be used to identify the spin systems, this information should be taken into account very critically, because short through-space distances *per se* are not sufficient to identify binding within a molecule. However, in the evaluation of the information from NOESY and DQF-RELAYED-NOESY spectra, the preliminary assignment will be kept in mind, and all assignments from these spectra will be checked thoroughly.

5.2.2. *Sequential Assignments of the Amino-Acid Residues.* 5.2.2.1. *Amino-Acid Sequence.* As pointed out above, NOE or ROE are the only homonuclear information for the determination of a peptide sequence in our case. In the NOESY spectrum, most of the corresponding peaks of the COSY spectrum (*Fig. 4*) also appear because the through-

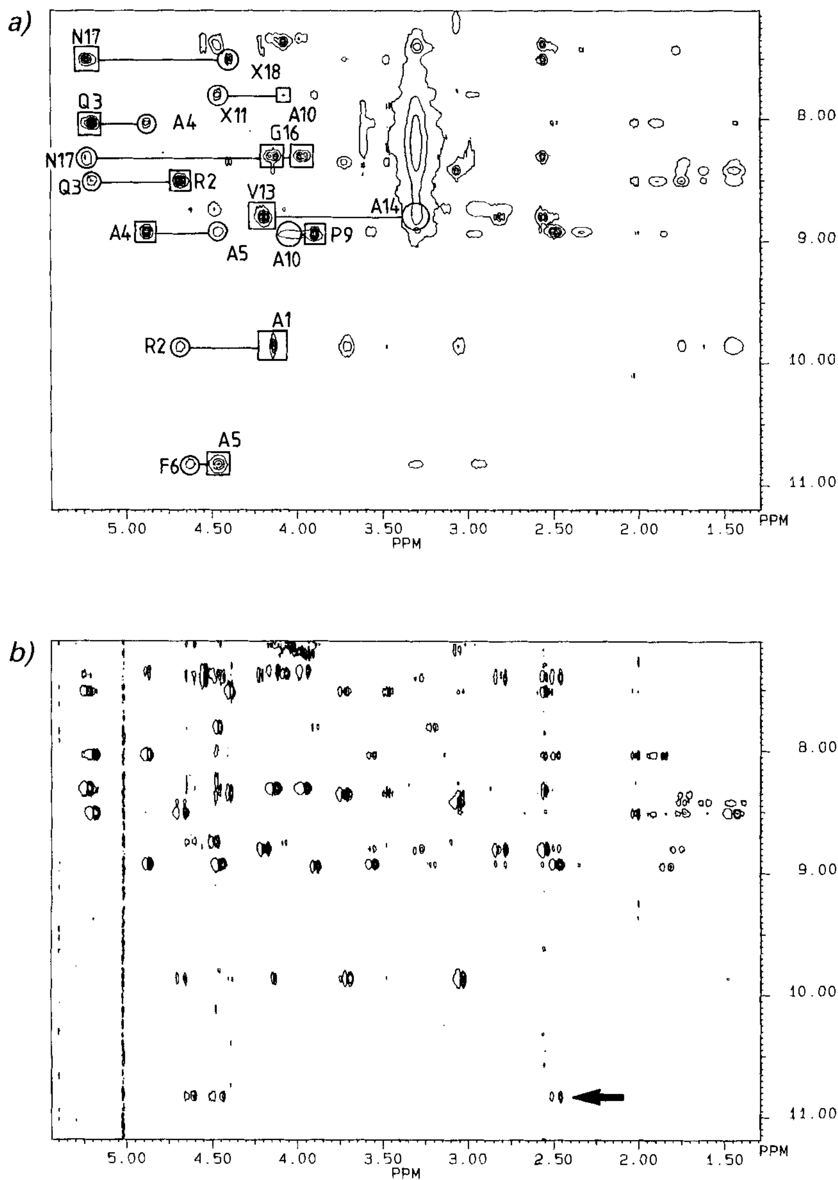


Fig. 6. *NH,H–C(2) regions of* a) *the 500-MHz NOESY and* b) *the 500-MHz DQF-RELAYED-NOESY spectrum of* **1** *at 310 K. a)* Only positive levels are shown. The mixing time was 150 ms. Those cross peaks are connected which correlate *H*–C(2) signals next to one peptide bond. Their chemical shifts are visible at the same amide signal. Cross peaks correlating N$H^{(i)}$,*H*–C(2)$^{(i)}$ are encircled, and N$H^{(i+1)}$,*H*–C(2)$^{(i)}$ are shown in boxes. *b)* Negative levels are drawn with only 1 contour line, while positive levels are drawn with 8 contour lines. The mixing-time τ_m for the NOE transfer was 300 ms, the delay Δ for the RELAYED transfer 25 ms. The arrow indicates the NOE-*J* cross peak of N*H of Phe*^{6} *and H*–C(2) of Ala^5, that identifies the otherwise ambiguous Ala^5-Phe^6 peptide bond (see text).

space distance of scalarly coupled nuclei are small enough to cause NOE cross peaks. In addition, the NOESY spectrum (*Fig. 6a*) exhibits cross peaks which correlate nuclei which do not belong to the same spin system. These signals are useful for sequential analysis [2]. In our case, 10 amide bonds can be identified by N$H^{(i+1)}$, H–C(2)$^{(i)}$ cross peaks unambiguously (*Fig. 6a*). However, several overlapping signals prevented a straightforward analysis in the NOESY spectrum. For example, the H–C(2) signals of Phe^6 and Phe^{12} (not shown in *Fig. 6a* because of the low intensity due to the broadness of the NH signal of Phe^{12}) overlap at 4.6 ppm as well as those of Ala^5 and Abu^{11} at 4.42 ppm. In the narrow region from 7.35 and 7.50 ppm, 5 amide protons resonate. The resulting ambiguities in overcrowded regions can be removed by transferring magnetization, caused by NOE into less crowded regions of the spectrum *via* J-coupling in a DQF-RELAYED-NOESY spectrum. Then, *e.g.* N$H^{(i+1)}$,N$H^{(i)}$ cross peaks appear from the following pathway: N$H^{(i+1)}$ to H–C(2)$^{(i)}$ by NOE, then to N$H^{(i)}$ by RELAYED. Ambiguities from overlapping H–C(2)'s can also be removed by the observation of N$H^{(i)}$ cross peaks in the DQF-RELAYED-NOESY (see *Fig. 6b*).

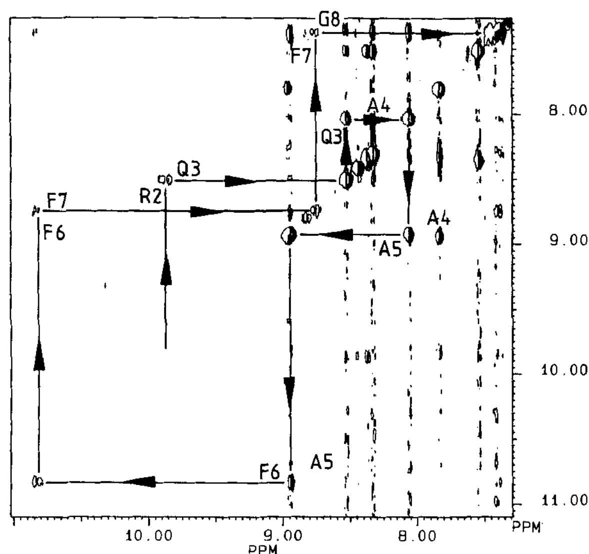


Fig. 7. *Low-field part of the 500-MHz DQF-RELAYED-NOESY spectrum of* 1 *at 310 K.* The levels are drawn in the same manner as in *Fig. 6b*. By means of the N$H^{(i+1)}$,N$H^{(i)}$ NOE-J cross peaks the N-terminal part of the peptide is indicated. The amino acid in position 9 is proline, which does not possess an amide proton.

In *Fig. 6b*, this is shown for the NH of Phe^6 at 10.85 ppm. The NOE to the H–C(2) signal of either Ala^5 or Abu^{11} is transferred to side-chain protons of the corresponding amino-acid residue (arrow in *Fig. 6b*), so that the questionable NOE (A5 in *Fig. 6a*) can be assigned as the NOE of NH of Phe^6 to H–C(2) of Ala^5. In *Fig. 7*, another region of the DQF-RELAYED-NOESY spectrum is shown. There, the sequence reading by means of NH, NH NOE-J signals is shown for the N-terminal part of the peptide sequence, from Arg^2 to Gly^8. Sequence reading is terminated at this point, because the next residue is Pro. This is proven by a NOE between H–C(2) of Gly^8 and H–C(5) of Pro^9.

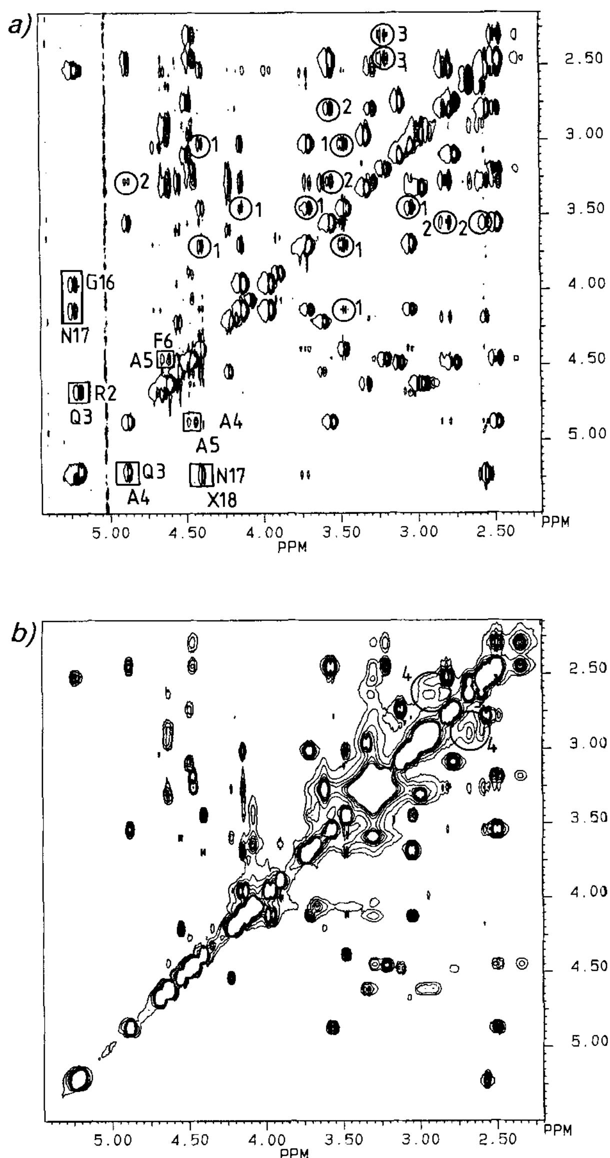


Fig. 8. *Regions of* a) *the 500-MHz DQF-RELAYED-NOESY and* b) *the 500-MHz NOESY spectrum of* **1** *at 310 K.* *a)* Region 5.4–2.25 × 5.4–2.2 ppm of the same spectrum as that shown in *Fig. 6b*, with the levels drawn in the same manner. In boxes, some $H-C(2)^{(i)}$,$H-C(2)^{(i+1)}$ NOE-*J* cross peaks are indicated demonstrating the asymmetric appearance of these peaks. Cross peaks connecting $H-C(3)$ signals of the lanthionine and β-methyllanthionine moieties adjacent *via* S-atoms are encircled. Peaks marked with '1' connect the $H-C(3)$ signals of Ala^1 and Abu^{18}, '2' of Ala^4 and Ala^{14}, and '3' of Ala^5 and Abu^{11}. Some cross peaks involving the corresponding $H-C(2)$ signals are visible (for assignments *cf. Fig. 5*). *b)* The same region as in *Fig. 8a* is shown with positive levels only. The encircled peaks of *Fig. 8a* can hardly be seen in this spectrum. The encircled cross peaks '4' indicate the NOE between $H-C(6)$ of Lys^{19} and $H-C(3)$ of Ala^{10}.

Table 3. *Correlations for the Determination of Sequential and Side-Chain Connectivities in* Ro 09-0198 (**1**)[a])

H-Ala-Arg-Gln-Ala-Ala-Phe-Phe-Gly-Pro-Ala-Abu-Phe-Val-Ala-Asp(OH)-Gly-Asn-Abu-Lys-OH (S bridges: Ala¹–Abu¹⁸, Ala⁴–Ala¹⁴, Ala⁵–Abu¹¹; HN bridge: Ala¹⁰–Lys¹⁹)

Line	1–2	2–3	3–4	4–5	5–6	6–7	7–8	8–9	9–10	10–11	11–12	12–13	13–14	14–15	15–16	16–17	17–18	18–19	
1						■	■			■	■							■	NHNH
2	■	■	■	■	■				■	■			■	■		■	■		H–C(2)NH
3		■	■	■	■	■	■						■	■		■			NHNH
4	■	■	■	■	■	■	■		■	■			■	■		■	■	■	H–C(2)H–C(2)
5	■	■	■	■	■	■	■	×	■	■	■		■	■	■	■	■	■	H–C(2)NH

Line	Ala¹	Ala⁴	Ala⁵	Ala¹⁰	Abu¹¹	Ala¹⁴	Abu¹⁸	Lys¹⁹	
6	1	2	3	4	3	2	1	4	side chain

a) In *Lines 1–5*, all NOEs from NOESY and DQF-RELAYED-NOESY spectra are given that have been used for sequence determination. *Lines 1* and *2* indicate the corresponding NOE from the NOESY spectrum, and *Lines 3* and *4* the corresponding NOE-*J* cross peaks from the DQF-RELAYED-NOESY spectrum. *Line 5* shows cross peaks that have been identified in the DQF-RELAYED-NOESY spectrum; the Gly^8-Pro^9-connection could be determined by a corresponding cross peak for *H*–C(2) of Gly^8 and H–C(5) of Pro^9 what is marked by a cross. In *Line 6*, correlations of *H*–C(3) signals across the S-atoms of the lanthionine and *β*-methyllanthionine moieties as well as the correlation of H–C(6) of Lys^{19} with H–C(3) of Ala^{10} in order to assign the side-chain connectivities are shown.

Independent prove of the sequence analysis is also provided by H–$C(2)^{(i)}$, H–$C(2)^{(i+1)}$ cross peaks (*Fig. 8a*). The combination of all these sequence information is presented in *Table 3*.

The only critical point in the sequence analysis is the lack of information about the connectivity Phe^{12}–Val^{13}. Unfortunately, the N*H* of Val^{13} does not show any NOE cross peak from which sequential connectivity could be derived[1]). We conclude the existence of a peptide bond between these amino-acids residues from the amidic nature of the N*H* of Val^{13} (chemical shift and coupling) and from the exclusion of all other possibilities.

5.2.2.2. *Assignments of Side-Chain Connections.* The last step in the determination of the constitution of **1** is the determination of the 4 side-chain connectivities. The 2 H–C(3) of Abu^{18}, Ala^{14} and Abu^{11} as well as the 2 H–C(6) of Lys^{19} are to be expected to exhibit NOE to the H–C(3)'s of the 3-substituted Ala^1, Ala^4, Ala^5, and Ala^{10}, respectively. These cross peaks are expected to lie close to the diagonal, which is rather broad in the NOESY spectrum. On comparison of the 2 identical sections in the NOESY and in the DQF-RELAYED-NOESY spectrum in *Fig. 8*, the advantage of the latter becomes evident: Due to the double-quantum filter, the intensity of the diagonal signals is reduced and the indicated signals can easily be detected (*Fig. 8a*). In addition, the magnetization is transferred in the RELAYED step to other protons which are scalarly coupled to the involved H–C(3)'s. Hence, ambiguities because of partial overlapping of signals can be overcome. Only the NOE of *H*–C(3) of Ala^{10} and *H*–C(6) of Lys^{19} can directly be observed in the NOESY spectrum as broad and small signals (encircled in *Fig. 8b*).

The correctness of the assignments described above can be confirmed by the fact that NOE between other side-chain protons of adjacent amino-acid residues are not observed

1) **Note added in proof:** A baseline corrected NOESY spectrum enabled us to detect this cross peak.

in our case. The only reasonable explanation for the observed spatial proximity of side-chain protons can be given by the S-bridges of the lanthionine and β-methyllanthionine moieties. Analogue considerations count for the lysinoalanine moiety. Therefore, the NOE indicating side-chain connectivities can, in addition, be used for the distinction of the different amino-acids residues with similar spin systems.

6. Sequence Homology with Other Peptides. – The amino-acid composition of *Ro 09-0198* (**1**) is identical with that of cinnamycin and differs from that of duramycin only by the exchange of the Arg residue by a Lys residue [15]. However, so far no structure has been established for cinnamycin, and only a tentative sequence and assignment of 3 S-bridges and the NH-bridge has been proposed for duramycin (see *Scheme* in *Chapt.4*) [15] [22]. Not considering the assignment of the amide functions of Gln^3 and Asn^{17} duramycin's sequence, in comparison to our peptide, exhibits 4 differences in the primary sequence. Three of them are represented by exchanges in positions 6, 10, and 11. The 4th distinction is the replacement of Lys for Arg in position 2 as discussed above. Furthermore, the assignment of 2 of the S-bridges and of the NH-bridge is different while only 1 S-bridge assignment, from Ala^4 to Ala^{14}, is identical in both structures.

While our work was in progress, the results of *Wakamiya et al.* [23] were published describing the chemical degradation of ancovenin (see *Scheme*, in *Chapt. 4*) a peptide that inhibits the function of ACE. It was astonishing to us that there is a large homology between both peptides. It is striking that the same residues in the same position of the peptides are connected by S-bridges. Only 6 out of 19 amino-acids residues differ, as is shown in *Table 4*. Under the residues, the synonyme codons of DNA-base triplets are given that encode for the corresponding amino acids. It is evident, that in one case (Phe→Dha), only a point-mutation is sufficient for the change of the code. For the understanding of this listing, it must be mentioned that Dha is formed in biosynthesis from serine by dehydration. The 'strange' amino-acid moieties lanthionine, β-methyllanthionine and lysinoalanine are formed by 1,2-addition of cysteine or lysine to 2,3-didehydroalanine or didehydro-amino-butyric acid, respectively. In general, this reaction is stereospecific so that the new chiral center at C(2) of the residue derived form the initial didehydro acid is D-configurated. The comparison of the peptides **1** and ancovenin gives a first hint for the configuration of the amino acids involved in the peptide **1**. However,

Table 4. *Comparison of the Nonidentical Amino Acids of* Ro 09-0198 (**1**) *and Ancovenin.* Under the amino-acid residues, the corresponding DNA base tripletts are given.

Arg^2 of **1**	→	Val^2	Phe^6 of **1**	→	Dha^{6a})	Dha^{10} of **1**	→	Leu^{10a})
CGU		GUU	UUU		UCU	UCU		CUU
CGC		GUC	UUC		UCC	UCC		CUC
CGC		GUA				UCA		CUA
CGG		GUG				UCG		CUG
Phe^{12} of **1**	→	Trp^{12}	Val^{13} of **1**	→	Ser^{13}	$Asp(3\text{-}OH)^{15}$ of **1**	→	Asp^{15}
UUU		UGG	GUU		UCU	???		GAU
UUC			GUG		UCG			GAC
			GUC		UCC			
			GUA		UCA			

[a]) Dha = 2,3-Didehydroalanine. Dha^{10} of **1** represents the 3-substituted Ala^{10}.

NMR spectroscopy can solve the configuration only simultaneously with the elucidation of the conformation, which need quantitative interpretation of NOE and coupling constants. This work is in progress in our group.

7. Conclusions. – The present paper describes the determination of the constitution of an unknown medium-sized peptide. For the first time, we succeeded in a complete assignment of all sequential connectivities of a 19 amino acids containing peptide. It must be considered that the presented results have been obtained simply from homonuclear ^{1}H-NMR measurements. This was mainly possible by the improvement of the well known RELAYED-NOESY technique. The introduction of a double-quantum filter enabled an optimized extraction of the whole qualitative information contained in a NOESY-like spectrum, that otherwise could not be analyzed. The astonishing homology to another recently investigated peptide may confirm our results.

8. Measurement Conditions. – 8.1. *General.* All spectra shown were recorded on a *Bruker AM 300* or an *AM 500* spectrometer, each equipped with an *Aspect 3000* computer. The data have been processed on a *Bruker* processing station with a 160 Megabyte hard disk. All measurements were performed at 310 K. The sample used was a 50 mM solution in $(D)_6$DMSO in a 5 mm sample tube.

8.2. *300-MHz H,H-DQF-COSY Spectra.* Sequence: $D1$-90°-t_1-90°-$D2$-90°-t_2; $D1$ = 1.5 s, $D2$ = 4 µs, 90°-pulse = 14 µs, acquisition time = 297 ms, sweep width in $F1$ and in $F2$ = 3.448.276 Hz, size = 2 K, 512 increments with 32 transients each, quad. detection in both dimensions, total measuring time 9 h; one time zerofilling and multiplication with $\pi/4$-shifted squared sine bell in both dimensions before phase-sensitive *Fourier* transformation for pure absorption line-shapes.

8.3. *500-MHz NOESY Spectra.* Sequence: $D1$-90°-t_1-90°-$D9$-90°-t_2; $D1$ = 2.5 s, $D9$ = 150 ± 20 ms, 90°-pulse = 14 µs, acquisition time = 297 ms, sweep width in $F1$ and in $F2$ = 3448.276 Hz, size = 2 K, 512 increments with 32 transients each, total measuring time 16 h; quad. detection in both dimensions; one time zerofilling and multiplication with $\pi/3$-shifted squared sine bell in both dimensions before phase-sensitive *Fourier* transformation for pure absorption line-shapes.

8.4. *500-MHz DQF-RELAYED-NOESY Spectra.* Sequence: $D1$-90°-t_1-90°-$D9$-90°-D_2-180°-$D2$-90°-$D3$-90°-t_2; $D1$ = 2.5 s, 90°-pulse = 14 µs, $D9$ = 300 ± 20 ms, $D2$ = 12.5 ms, $D3$ = 4 µs, acquisition time = 297 ms, sweep width in $F1$ and in $F2$ = 3448.276 Hz, each size = 2 K, 512 increments with 64 transients each, total measuring time 30 h; quad. detection in both dimensions; one time zerofilling in each dimension and multiplication with squared sine bell $\pi/3$-shifted in $F1$ and $\pi/4$-shifted in $F2$ before phase-sensitive *Fourier* transformation for pure absorption line-shapes.

We are indebted to Dr. *C. Griesinger* for the assistance in designing the DQF-RELAYED-NOESY experiment. We would like to thank Mr. *M. Manneberg* for determining the amino-acid composition of *Ro 09-0198*, the late Dr. *E. Gross*, Natl. Inst. Child Health, Hum. Dev., *NIH*, Bethesda, Md, USA, for samples of lanthionine and β-methyllanthionine and Dr. *J. Föhles, Deutsches Wollforschungsinstitut*, Aachen, FRG, for a sample of synthetic lysinoalanine. Financial support from the *Deutsche Forschungsgemeinschaft* and the *Fonds der Chemischen Industrie* is acknowledged.

REFERENCES

[1] H. Kessler, W. Bermel, in 'Application of NMR Spectroscopy to Problems in Stereochemical and Conformational Analysis', Eds. A. P. Marchand and Y. Takeuchi, Verlag Chemie Intern., VCH, Weinheim, 1986, pp. 179–205; H. Kessler, W. Bermel, A. Müller, K.-H. Pook, in 'The Peptides-Analysis, Synthesis, Biology', Eds. S. Udenfriend, J. Meienhofer, and V. Hruby, Academic Press, New York, 1985, Vol. 7, pp. 437–473.
[2] K. Wüthrich, in 'NMR of Proteins and Nucleic Acids', J. Wiley & Sons, New York, 1986.
[3] W. P. Aue, E. Bartholdi, R. R. Ernst, *J. Chem. Phys.* **1976**, *64*, 2229.
[4] G. Eich, G. Bodenhausen, R. R. Ernst, *J. Am. Chem. Soc.* **1982**, *104*, 3731.
[5] J. Jeener, B. H. Meier, P. Bachmann, R. R. Ernst, *J. Chem. Phys.* **1979**, *71*, 4546.
[6] M. Rance, O. W. Sørensen, G. Bodenhausen, G. Wagner, R. R. Ernst, K. Wüthrich, *Biochem. Biophys. Res. Commun.* **1983**, *113*, 967; U. Piantini, O. W. Sørensen, R. R. Ernst, *J. Am. Chem. Soc.* **1982**, *104*, 6800.

[7] L. Braunschweiler, R. R. Ernst, *J. Magn. Reson.* **1983**, *53*, 512.
[8] D. G. Davis, A. Bax, *J. Am. Chem. Soc.* **1985**, *107*, 2820.
[9] H. Kessler, H.-R. Loosli, H. Oschkinat, *Helv. Chim. Acta* **1985**, *68*, 661.
[10] H. Kessler, M. Bernd, H. Kogler, J. Zarbock, O. W. Sørensen, G. Bodenhausen, R. R. Ernst, *J. Am. Chem. Soc.* **1983**, *105*, 6944.
[11] K. Takemoto, Y. Umeda, H. Ishitsuka, Y. Yagi, *Proc. Jpn. Soc. Immunol.* **1981**, *11*, 363.
[12] S. Steuernagel, Diploma Thesis, Frankfurt, 1986.
[13] G. Wagner, *J. Magn. Reson.* **1984**, *57*, 497.
[14] E. Gross, J. H. Brown, in 'Peptides 1976, Proc. Eur. Pept. Symp. 14th.', Ed. A. Loffet, Editions de l'Université de Bruxelles, Bruxelles, Belgium, 1976, pp. 183–190.
[15] E. Gross, in 'Antibiotics, Isolation, Separation and Purification', Eds. M. J. Weinstein and G. H. Wagman, Elsevier Scientific Publishing Company, Amsterdam, 1987, pp. 415–462.
[16] H. Allgaier, G. Jung, R. G. Werner, U. Schneider, H. Zähner, *Angew. Chem.* **1985**, *97*, 1052; *Eur. J. Biochem.* **1986**, *160*, 9.
[17] A. A. Bothner-By, R. L. Stephens, J. Lee, C. D. Warren, R. W. Jeanloz, *J. Am. Chem. Soc.* **1984**, *106*, 811; A. Bax, D. G. Davis, *J. Magn. Reson.* **1985**, *63*, 207; H. Kessler, C. Griesinger, R. Kerssebaum, K. Wagner, R. R. Ernst, *J. Am. Chem. Soc.* **1987**, *109*, 607.
[18] D. Marion, K. Wüthrich, *Biochem. Biophys. Res. Commun.* **1983**, *113*, 967.
[19] O. W. Sørensen, G. W. Eich, M. H. Levitt, G. Bodenhausen, R. R. Ernst, *Prog. Nucl. Magn. Reson. Spectrosc.* **1983**, *16*, 163.
[20] C. Griesinger, Ph. D. Thesis, Frankfurt, 1986.
[21] G. Bodenhausen, H. Kogler, R. R. Ernst, *J. Magn. Reson.* **1984**, *58*, 370.
[22] E. Gross, J. H. Brown, *Fed. Proc., Fed. Am. Soc. Exp. Biol.* **1976**, *35*, 1332.
[23] T. Wakamiya, Y. Ueki, T. Shiba, Y. Kido, Z. Motoki, *Tetrahedron Lett.* **1985**, *26*, 665.
[24] IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), *Int. J. Pept. Prot. Res.* **1984**, *24*, No. 1, Appendix; or *Pure Appl. Chem.* **1984**, *56*, 595.