Sequential resonance assignments in protein 1H nuclear magnetic resonance spectra

ARTICLE *in* JOURNAL OF MOLECULAR BIOLOGY · APRIL 1982

Impact Factor: 4.33 · DOI: 10.1016/0022-2836(82)90010-9 · Source: PubMed

3 AUTHORS, INCLUDING:

Kurt Wüthrich
The Scripps Research Institute
742 PUBLICATIONS **76,162** CITATIONS

Reprinted from *J. Mol. Biol.* (1982) **155**, 367–388

Sequential Resonance Assignments in Protein ^1H Nuclear Magnetic Resonance Spectra

Glucagon Bound to Perdeuterated Dodecylphosphocholine Micelles

GERHARD WIDER, KONG HUNG LEE AND KURT WÜTHRICH

J. Mol. Biol. (1982) **155**, 367–388

Sequential Resonance Assignments in Protein ^1H Nuclear Magnetic Resonance Spectra

Glucagon Bound to Perdeuterated Dodecylphosphocholine Micelles

GERHARD WIDER, KONG HUNG LEE AND KURT WÜTHRICH

Institut für Molekularbiologie und Biophysik
Eidgenössische Technische Hochschule
ETH-Hönggerberg, CH-8093 Zürich, Switzerland

(Received 17 August 1981)

The assignment of the ^1H nuclear magnetic resonance spectrum of glucagon bound to perdeuterated dodecylphosphocholine micelles with the use of two-dimensional ^1H nuclear magnetic resonance techniques at 360 MHz is described. Sequential resonance assignments were obtained for all backbone and C^β protons except the N-terminal amino group and the amide proton of Ser2. The assignments of the non-labile amino acid side-chain protons are complete except for the γ-methylene protons of Gln20 and Gln24. These assignments provide a basis for the determination of the three-dimensional structure of lipid-bound glucagon.

1. Introduction

Glucagon is a hormone that consists of a linear polypeptide chain of 29 amino acid residues, M_r 3500. The primary target organ for glucagon is the plasma membrane of liver and other cells. Specific binding to a plasma membrane receptor site mediates activation of glycogenolysis (Pohl *et al.*, 1969). Evidence has been presented that recognition between glucagon and its receptor depends on the ordered lipid structures surrounding the receptor site in the membrane (Rodbell *et al.*, 1971; Rubalcava & Rodbell, 1973). A possible avenue to further insights into the mode of action would be *via* knowledge of the conformations adopted by glucagon in the different environments encountered on the way from the sites of its synthesis in the islets of Langerhans to the complex formation with the receptor site. Previously, an α-helical conformation of glucagon was determined in single crystals (Sasaki *et al.*, 1975). However, ^1H n.m.r.† studies showed that the α-helical form was not preserved in aqueous solutions of monomeric glucagon (Bösch *et al.*, 1978), which coincides with earlier conclusions from circular dichroism experiments that monomeric glucagon in solution adopts predominantly a flexible "random

† Abbreviations used: n.m.r., nuclear magnetic resonance; p.p.m., parts per million; NOE, nuclear Overhauser enhancement; NOESY, 2-dimensional NOE spectroscopy; COSY, 2-dimensional correlated spectroscopy.

0022-2836/82/070367-22 $02.00/0

coil'' conformation (see e.g. Panijpan & Gratzer, 1974, and references therein). In view of the functional properties of glucagon, it appears of considerable interest to complement these structural data with a determination of the molecular conformation in a lipid/water interface, for example along the surface of lipid micelles. In this paper we describe individual ^1H n.m.r. assignments for the entire amino acid sequence of glucagon bound to perdeuterated dodecylphosphocholine micelles.

In previous studies it was shown by a variety of physicochemical methods that mixed micelles of glucagon and dodecylphosphocholine in a sufficiently concentrated solution for n.m.r. studies contained one molecule of glucagon and $\sim$40 detergent molecules, which corresponds to a molecular weight of $\sim$17,000 (Bösch *et al.*, 1980). High resolution ^1H n.m.r. studies with one-dimensional techniques indicated that the micelle-bound glucagon adopted a predominantly extended conformation (Bösch *et al.*, 1980). The polypeptide backbone was found to be roughly parallel to the micelle surface, with the depth of immersion corresponding approximately to the average length of an amino acid side-chain (Brown *et al.*, 1981). From combined use of distance constraints obtained from nuclear Overhauser enhancement experiments and a distance geometry algorithm, a spatial structure for the glucagon segment residues 19 to 27 was determined (Braun *et al.*, 1981). Since at that stage the ^1H n.m.r. lines of only a small number of amino acid side-chain hydrogen atoms had been assigned individually, only low resolution could be attained for this structure. The resonance assignments in this paper are the first step in the determination of the spatial structure of micelle-bound glucagon at higher resolution (Wüthrich *et al.*, 1982).

2. Materials and Methods

Sequential resonance assignments were obtained with combined use of 2-dimensional correlated spectroscopy (Aue *et al.*, 1976) and 2-dimensional nuclear Overhauser enhancement spectroscopy (Jeener *et al.*, 1979; Anil Kumar *et al.*, 1980*a*,*b*). For a brief description of these 2 experiments, the reader is referred to the preceding paper (Wagner & Wüthrich, 1982). In contrast to the basic pancreatic trypsin inhibitor, where a sizeable proportion of the sequential assignments was obtained from studies in ^2H_2O solutions (Dubs *et al.*, 1979; Wagner *et al.*, 1981), all the amide protons in micelle-bound glucagon exchange quite rapidly, and hence the sequential assignments had to rely entirely on studies in H_2O solution.

The spin systems of the individual amino acid side-chains were identified in a COSY spectrum recorded in a ^2H_2O solution of micelle-bound glucagon. In several cases, ambiguities in the COSY connectivities were resolved by comparison with a spin echo correlated spectrum (Nagayama *et al.*, 1979,1980; Nagayama & Wüthrich, 1981) recorded under identical conditions.

Two-dimensional ^1H n.m.r. spectra at 360 MHz were recorded on a Bruker HX 360 spectrometer. The spectra were obtained from 256 measurements with t_1 values from 0 to 40 ms. Special care had to be taken for the suppression of the H_2O solvent resonance, which was much broader than, for example, in a dilute solution of a protein. Satisfactory results were obtained by selective, continuous irradiation of the H_2O line at all times except during data acquisition: 1024 data points were used to store the data for each value of t_1. In the different experiments, between 192 and 256 transients were accumulated for each t_1 value. To end up with a 512×512 point data matrix in the frequency domain, which corresponded

to a digital resolution of 6·3 Hz/point for the spectra recorded in H_2O and 5·8 Hz/point for ^2H_2O spectra, the time domain matrix was expanded to 1024 points in t_1 and 2048 points in t_2 by zero filling. Otherwise, data acquisition and data handling were done as described in the preceding paper (Wagner & Wüthrich, 1982).

Bovine glucagon was purchased from SERVA, Heidelberg. [$^2H_{38}$]dodecylphosphocholine was synthesized as described (Brown, 1979). Two samples were prepared with a mixed solvent of 90% H_2O and 10% ^2H_2O, and with ^2H_2O. The samples contained 0·015 M-glucagon, 0·7 M-[$^2H_{38}$]dodecylphosphocholine and 0·05 M-phosphate buffer (pH 6·0). In the ^2H_2O solution, the pH meter reading was used without correction for isotope effects (Kalinichenko, 1976; Bundi & Wüthrich, 1979*a*), and the concentration of residual solvent protons was minimized by repeated lyophilization from ^2H_2O. Both samples were carefully degassed and sealed under a nitrogen atmosphere. Chemical shifts are quoted relative to external 3-trimethylsilyl-[2,2,3,3-$^2H_4$]propionate, where the ϵ-methyl resonance of Met27 was taken to be at 2·04 p.p.m. at 37°C and was used as an internal reference.

Mixed micelles of glucagon and dodecylphosphocholine were previously characterized by a variety of physicochemical methods (Bösch *et al.*, 1980; Brown *et al.*, 1981). For glucagon concentrations between 0·001 and 0·006 M and a 50-fold excess of detergent molecules, it was found that an homogeneous population of micelles consisting of 1 glucagon molecule and ~40 detergent molecules prevailed, with a molecular weight of ~17,000. For the 2-dimensional n.m.r. experiments, the glucagon concentration was increased to 0·015 M. Identical ^{1}H n.m.r. spectra were obtained to those for solutions with 0·004 M-glucagon (Bösch *et al.*, 1980; Braun *et al.*, 1981). Since other methods, for example, light-scattering and the analytical ultracentrifuge, could not be used reliably at these high concentrations, we relied on the n.m.r. evidence that similar mixed micelles were formed over the entire concentration range.

3. Results and Discussions

(a) *Identification of the amino acid side-chain spin systems*

Previously, the complete spin systems of the non-labile protons of Ala19, Val23 and Trp25 in glucagon bound to perdeuterated dodecylphosphocholine micelles were identified with one-dimensional n.m.r. experiments (Bösch *et al.*, 1980). In addition, the imidazole resonances of His1, the δ-methyl resonances of the two leucine residues in positions 14 and 26, the ϵ-methyl line of Met27 and a Thr $C^\beta H$–$C^\gamma H_3$ fragment were described. In what follows, we describe how the complete spin systems of all but two of the 29 amino acid residues were identified in a COSY spectrum recorded in ^2H_2O solution (Fig. 1).

In the COSY spectrum of Figure 1, peaks corresponding to the resonance positions in the one-dimensional spectrum are located on the diagonal from the upper right to the lower left corner. Non-diagonal "cross peaks" manifest through-bond J-connectivities between distinct diagonal peaks. In Figure 1, two well-separated spectral regions from 0·5 to 5·0 p.p.m., and from 6·5 to 7·5 p.p.m. can be distinguished, which are not connected by J-couplings. The low-field region contains the resonances of the aromatic rings, the high-field region includes those of all the other protons. In the lower right triangle of Figure 1, the connectivities for seven aliphatic spin systems and the aromatic rings of tryptophan and the two tyrosine residues are indicated.

The COSY connectivity patterns for the amino acid residues in glucagon are shown in Table 1. They represent a different presentation of the previously

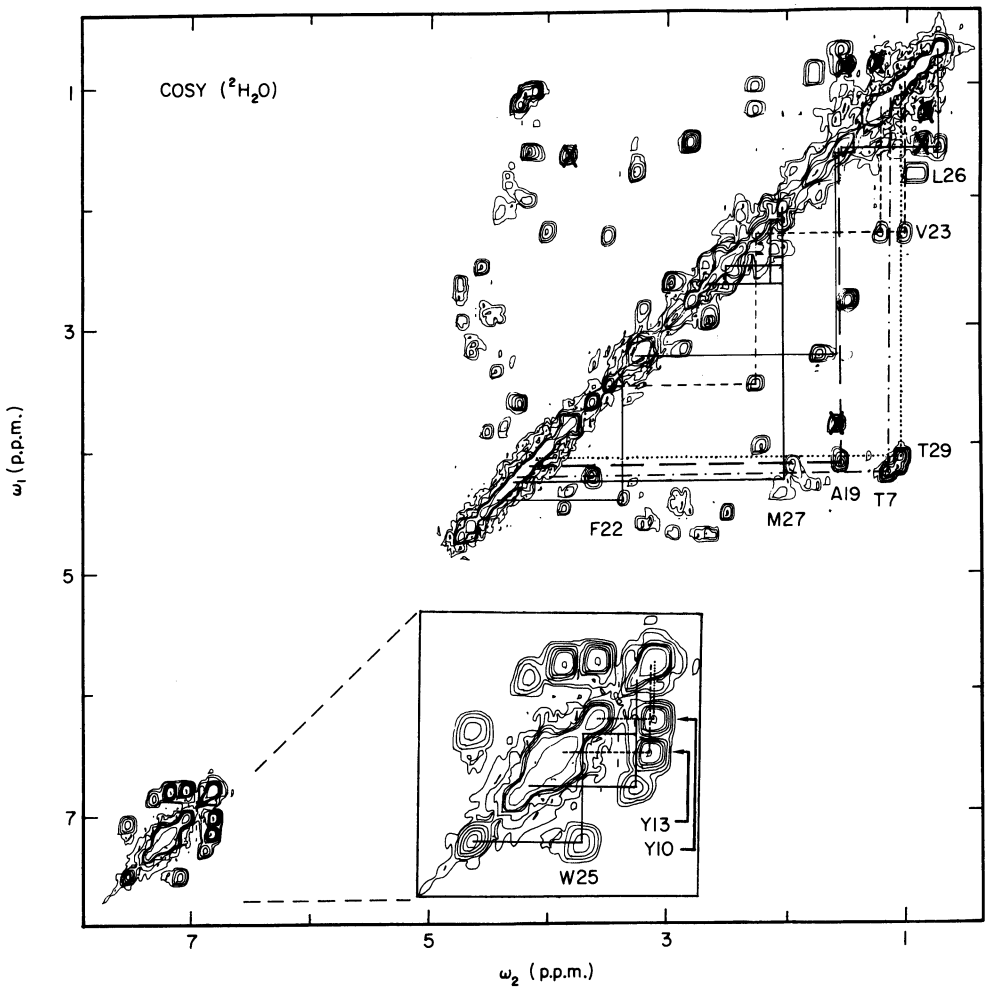
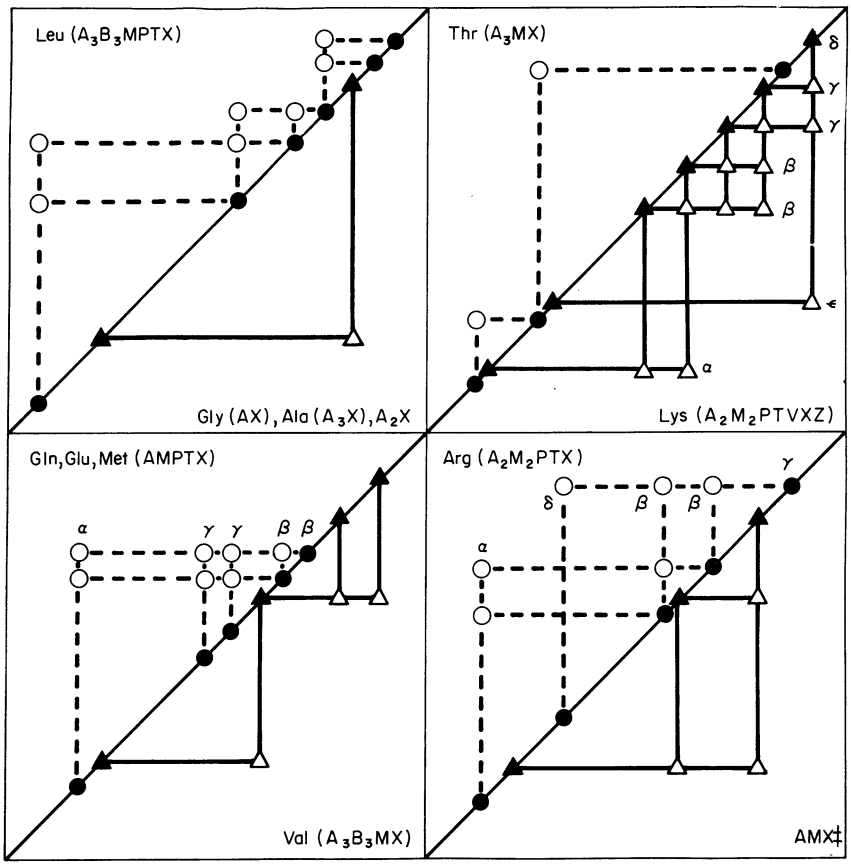


FIG. 1. Contour plot of a 360 MHz ^1H COSY spectrum of glucagon bound to perdeuterated dodecylphosphocholine micelles in ^2H_2O. The sample contained 0·015 M-glucagon, 0·7 M-[$^2H_{38}$]dodecylphosphocholine, 0·05 M-phosphate buffer, p^{2}H 6·0, t = 37°C. Under these conditions the predominant species in the solution are mixed micelles of 1 glucagon molecule and ~40 detergent molecules, with a molecular weight of about 17,000 (Bösch *et al.*, 1980). The spectrum was recorded in 24 h, the digital resolution is 5·88 Hz/point. The symmetrized (Baumann *et al.*, 1981) absolute value spectrum is shown. The aromatic region is also presented on an expanded scale. Proton–proton J-connectivities are indicated for the following residues: Thr7 (–·–·), Ala19 (— —), Phe22 (——), Val23 (– – – –), Leu26 (——), Met27 (——), Thr29 (····) and the aromatic rings of Tyr10 (····), Tyr13 (– – – –) and Trp25 (——). In order not to overcrowd the Figure, only the C$^\alpha$H connectivity with the lower field C$^\beta$H line is shown even for amino acid residues where 2 non-degenerate β-methylene resonances were observed (Table 2). Cross peaks originating from residual protons in the perdeuterated dodecylphosphocholine are marked (X).

described connectivity patterns in spin echo correlated spectroscopy (Nagayama & Wüthrich, 1981). Unique connectivities, which can provide the information needed for unambiguous identification, prevail for valine, isoleucine (not shown in Table 1), leucine, proline (not shown in Table 1), threonine, lysine and arginine.

TABLE 1

COSY connectivity diagrams for the weakly coupled spin systems of the non-labile, aliphatic protons in the common amino acid residues†



† In each of the 4 squares, 2 connectivity diagrams are shown. For the diagram in the upper left area of each square, diagonal peaks are indicated with filled circles, cross peaks with open circles and connectivities with broken lines. In the lower right area, filled and open triangles and continuous lines are used. The amino acid types and spin systems represented by the diagrams are also indicated. For better readability, the diagonal peaks of the AMPTX system, Arg and Lys, are identified. Ile and Pro, which are not present in glucagon, are not shown.

‡ A_2X or AMX spin systems may arise from Asn, Asp, Cys, His, Phe, Ser, Trp and Tyr.

For the other common amino acids, the COSY patterns allow only a classification into groups of several residues, and additional information is required to further distinguish the individual amino acid types. For instance, Gly, A_2X spin systems and Ala can be distinguished from the resonance intensities and chemical shifts, and the $C^{\alpha}H$–$C^{\beta}H_2$ fragments of the aromatic side-chains can be distinguished from the other AMX or A_2X spin systems on the basis of the NOE connectivities with the aromatic rings (see below). Alternatively, combination of the COSY

connectivities with sequential assignment of the polypeptide backbone resonances can also be used to identify the individual residue types (see below).

To illustrate the analysis, we consider the Val spin system in Figure 1. From the diagonal peak at 3·49 p.p.m., a broken, horizontal line leads to the $C^{\alpha}H$–$C^{\beta}H$ cross peak at ($\omega_1 = 3{\cdot}49$, $\omega_2 = 2{\cdot}26$ p.p.m.). From there a vertical line leads to the diagonal peak of $C^{\beta}H$ at 2·26 p.p.m. A horizontal line makes the connections to the two cross peaks with the $C^{\gamma}H_3$ resonances at ($\omega_1 = 2{\cdot}26$, $\omega_2 = 1{\cdot}21$ p.p.m.) and ($\omega_1 = 2{\cdot}26$, $\omega_2 = 1{\cdot}02$ p.p.m.). Two vertical lines then lead to the diagonal peaks of

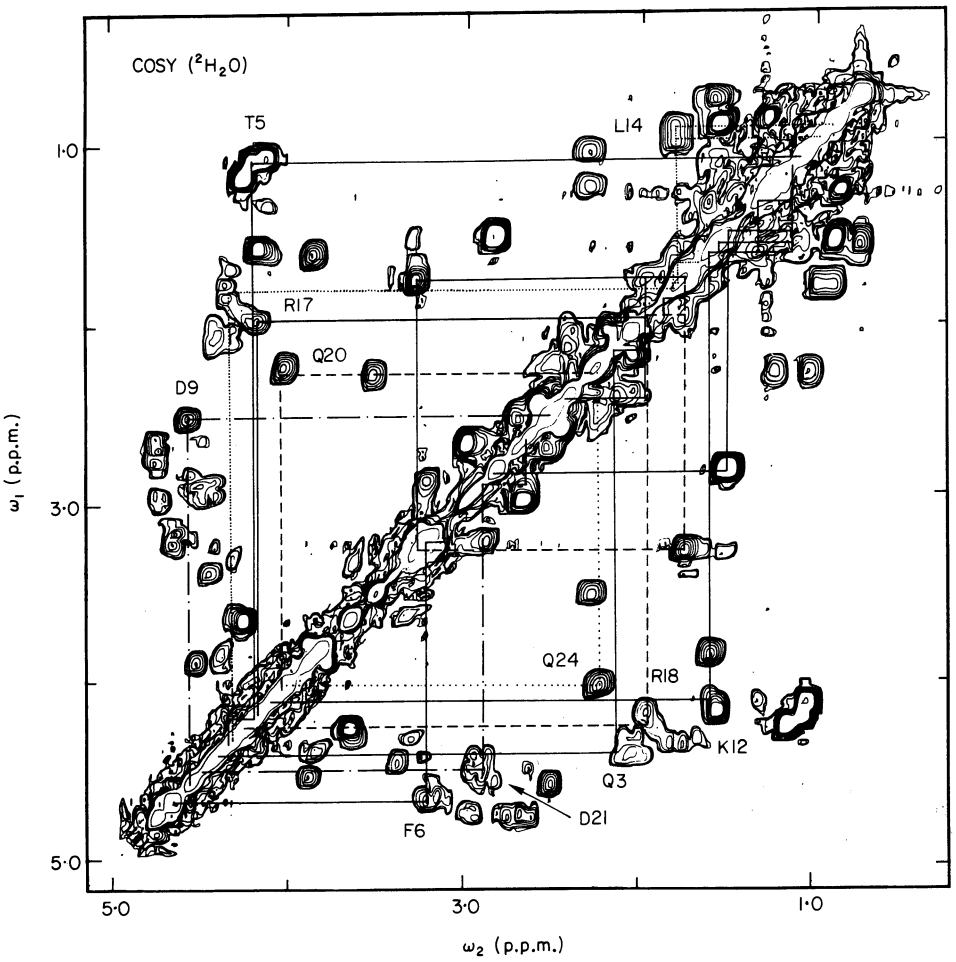


FIG. 2. Spectral region from 0·2 to 5·1 p.p.m. of the 360 MHz ^1H COSY spectrum of micelle-bound glucagon in Fig. 1. J-connectivities for the following residues are indicated in the upper left triangle of the spectrum: Thr5 (——), Asp9 (—·—·), Leu14 (····), Arg17 (——), Gln20 (— — —). The lower right triangle contains the J-connectivities for Gln3 (——), Phe6 (——), Lys12 (——), Arg18 (— — —), Asp21 (—·—), Gln24 (····). In order not to overcrowd the Figure, only the $C^{\alpha}H$ connectivity with the lower field $C^{\beta}H$ line is shown even for amino acid residues where 2 non-degenerate β-methylene resonances were observed (Table 2).

the γ-methyl groups. Similarly, the connectivities for the other 28 amino acid residues are documented in Figures 1 to 3. When comparing the connectivities in these Figures with those in Table 1, one should consider that for practical reasons only one of the connectivities between $C^{\alpha}H$ and $C^{\beta}H_2$ is outlined, even when the β-methylene protons are not degenerate.

For the aromatic side-chains, COSY provided two separate spin systems of $C^{\alpha}H$–$C^{\beta}H_2$ (Figs 1 to 3) and the ring protons (Fig. 1), respectively. Except for His1, the connectivity with the ring protons could be established *via* NOEs, as predicted by theoretical considerations (Billeter *et al.*, 1982). These results are documented in

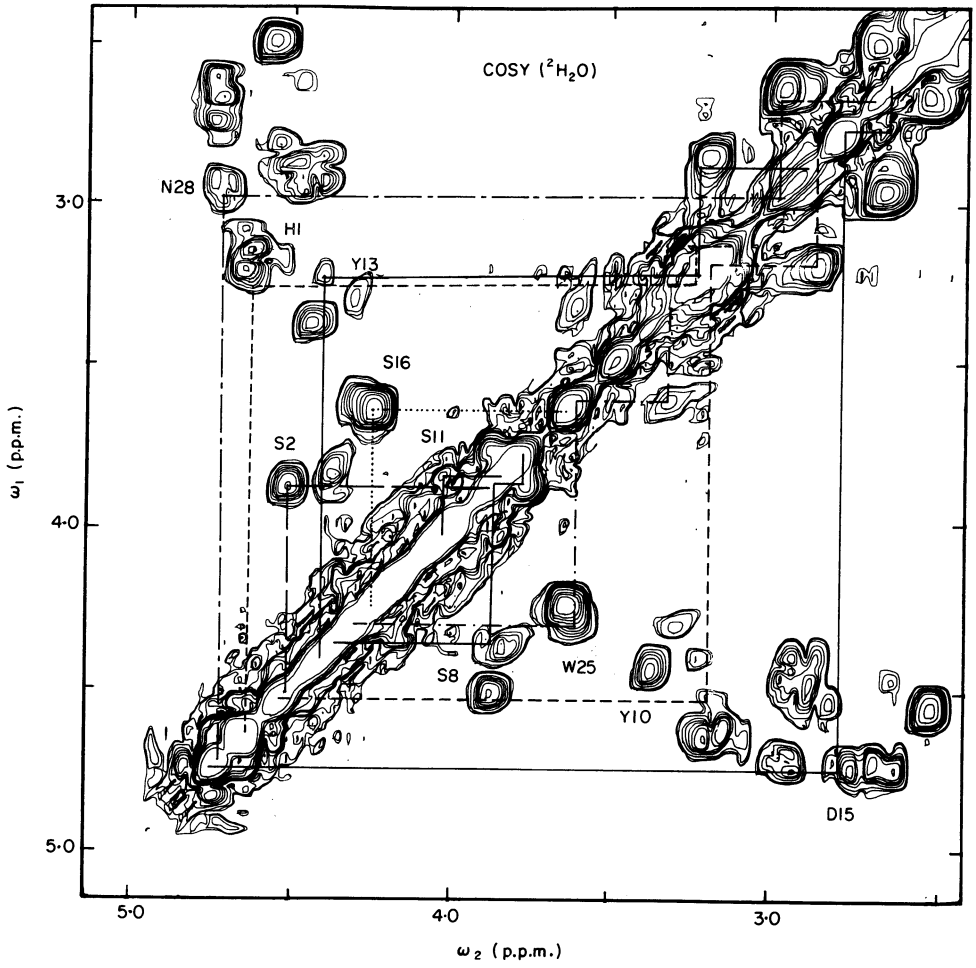


FIG. 3. Spectral region from 2·4 to 5·1 p.p.m. of the 360 MHz ^{1}H COSY spectrum of micelle-bound glucagon in Fig. 1. Connectivities for the following residues are indicated in the upper left triangle: His1 (– – –), Ser2 (— —), Ser11 (——), Tyr13 (——), Ser16 (· · ·), Asn28 (–·–). The lower right triangle contains the J-connectivities for Ser8 (——), Tyr10 (– – –), Asp15 (——), Trp25 (–·–·). In order not to overcrowd the Figure, only the $C^{\alpha}H$ connectivity with the lower field $C^{\beta}H$ line is shown even for amino acid residues where 2 non-degenerate β-methylene resonances were observed (Table 2).

Figure 4, which shows a NOSEY spectrum recorded in the same ^2H_2O solution of micelle-bound glucagon as the COSY spectrum in Figure 1. A NOESY spectrum (Anil Kumar *et al.*, 1980*a*) has a similar appearance to that of a COSY spectrum (Fig. 1), but the non-diagonal peaks manifest through-space NOE connectivities between distinct resonances on the diagonal. In Figure 4, only a small spectral region is shown, which contains all the cross peaks between $C^\beta H_2$ and ring protons within the individual aromatic side-chains.

With regard to the sequential resonance assignments, it was important that one could at the outset rely on 23 completely identified side-chain spin systems. The spin systems of Gly4, one Thr, one Leu and three of the four five-spin systems of Gln and Met could be identified only after the sequential assignments were completed. It should be added that, with the exception of the residue types that occur only once in the amino acid sequence of glucagon and give rise to unique

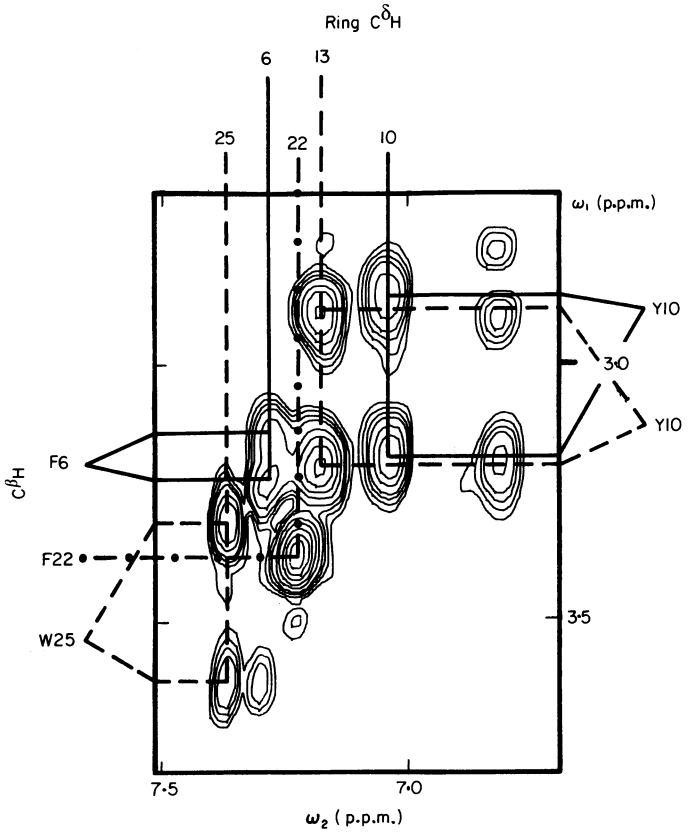


FIG. 4. Contour plot of the spectral region (2·6 to 3·8 p.p.m.) × (6·7 to 7·5 p.p.m.) of a 360 MHz ^1H NOESY spectrum of glucagon bound to perdeuterated dodecylphosphocholine micelles recorded in the sample described in the legend to Fig. 1. The temperature was 37°C. The spectrum was recorded in 26 h, using a mixing time of 200 ms. The digital resolution is 5·8 Hz/point. The symmetrized absolute value spectrum is shown. The connectivities between $C^\beta H_2$ and the C^δ ring protons of the aromatic residues are indicated as follows: Phe6 (——), Phe22 (·—·), Tyr10 (——), Tyr13 (— — —), Trp25 (— — —).

COSY patterns (Table 1; Fig. 6), i.e. Lys12, Ala19, Val23 and Trp25, the assignments to specific sequence locations indicated in Figures 1 to 5 were obtained only after completion of the sequential resonance assignments (see below).

(b) *Strategy used to assign the ^{1}H n.m.r. spectrum of micelle-bound glucagon*

Compared to the resonance assignments for basic pancreatic trypsin inhibitor described by Wagner *et al.* (1981) and in the preceding paper (Wagner & Wüthrich, 1982), a somewhat different strategy had to be selected for glucagon. For example, since the labile protons of micelle-bound glucagon exchange too rapidly with the solvent to be observed in the ^{1}H n.m.r. spectra recorded in ^2H_2O solution (Bösch *et al.*, 1980), all the sequential assignments had to be obtained with spectra recorded in H_2O. Furthermore, sequential assignments *via* NOEs between NH_{i+1} and $C^\beta H_i$ (d_3) played a much more important role than in basic pancreatic trypsin inhibitor. Otherwise the fundamental aspects of the procedures used were closely similar. Neighbouring amino acid residues were identified *via* one or two of the NOE connectivities d_1 (from NH_{i+1} to $C^\alpha H_i$), d_2 (from NH_{i+1} to either NH_{i+2} or NH_i) or d_3 (from NH_{i+1} to $C^\beta H_i$) (Billeter *et al.*, 1982), the J-connectivities with the labile protons were obtained from a COSY spectrum recorded in H_2O and, as far as possible, the sequential assignments were checked step by step against the amino acid sequence.

Figure 5 shows the region of the COSY spectrum in H_2O that contains all the NH_i–$C^\alpha H_i$ cross peaks in micelle-bound glucagon. As discussed in more detail in the preceding paper (Wagner & Wüthrich, 1982), each amino acid residue in glucagon (Fig. 6) should give rise to one such peak, except for Gly4, which might give two peaks, and the N-terminal residues His1 and Ser2, which might not be seen because of rapid exchange of the labile backbone protons (Bundi & Wüthrich, 1979*b*). From the resonance identifications in Figures 1 to 3, it was further known that the cross peaks with the C^α protons of two AMX spin systems at 4·66 and 4·64 p.p.m. would not be seen for experimental reasons, since at 37°C they coincided with the water line and were therefore bleached out by the water irradiation (Anil Kumar *et al.*, 1980*b*). Figure 5 contained all the expected cross peaks, and thus provided a complete "fingerprint" of the polypeptide chain. From the side-chain identifications, 23 peaks could be assigned to specific spin systems. Among these, Lys12, Ala19, Val23 and Trp25 could be assigned immediately to a specific location in the amino acid sequence, since glucagon contains only one residue of each of these four types (Fig. 6). These four residues provided ideal starting points for the sequential assignments, which could then be checked against the amino acid sequence from the first step.

Figure 5 affords a convincing illustration of the improved resolution achieved when going from unidimensional n.m.r. to two-dimensional n.m.r. However, because of the inherently broad lines and the small dispersion of the $C^\alpha H$ and NH chemical shifts, some overlap of different peaks persisted in the COSY and NOESY spectra of micelle-bound glucagon, which limited the analysis of the cross peaks among C^α and amide protons in some instances. Therefore, and because of the superior spectral resolution in the $C^\beta H$ region, d_3 connectivities between NH_{i+1} and

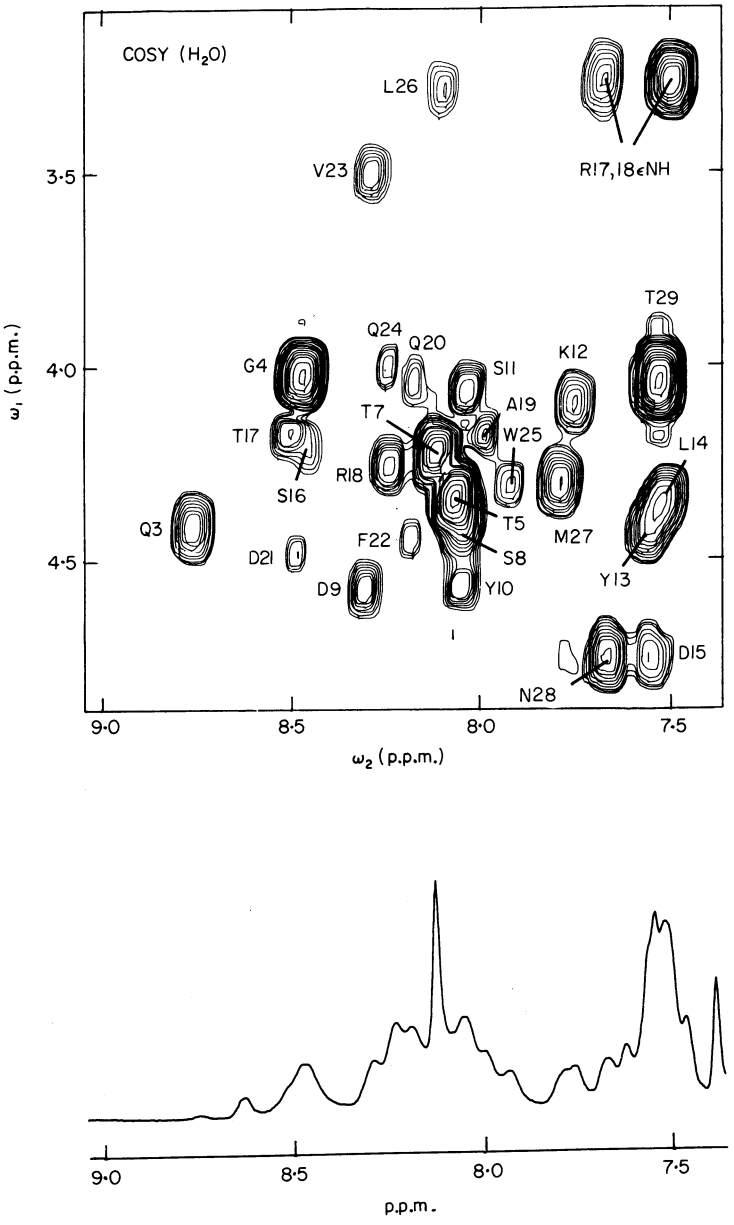


FIG. 5. Spectral region (3·1 to 4·9 p.p.m.) × (7·4 to 9·1 p.p.m.) of a 360 MHz ^{1}H COSY spectrum of glucagon bound to perdeuterated dodecylphosphocholine micelles. The composition of the sample was identical to that in Fig. 1, except that a mixed solvent of 90% H_2O and 10% ^2H_2O was used. The temperature was 37°C. The letters and numbers indicate the resonance assignments for the $C^{\alpha}H_i$–NH_i cross peaks for residues 3 to 29. In addition, the cross peaks connecting $C^{\delta}H_2$ and the guanidinium protons of the 2 arginine side-chains are seen in the upper right corner. For comparison, a conventional, unidimensional spectrum of the same sample is shown at the bottom of the Figure. It contains the well-separated resonance of the amide proton of Phe6 at 8·63 p.p.m. In the COSY spectrum, the cross peak for Ph6 was bleached out by the irradiation of the solvent line (Anil Kumar *et al.*, 1980*b*), since the chemical shift of $C^{\alpha}H$ coincides with that of the H_2O line (Table 2).

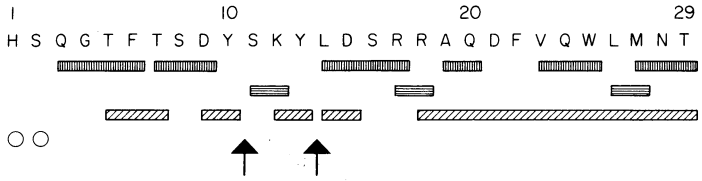


FIG. 6. Amino acid sequence of bovine glucagon and survey of the experimental data by which individual resonance assignments were obtained for the micelle-bound polypeptide. ( ) Sequential assignments *via* d_1 (NOE from NH_{i+1} to $C^{\alpha}H_i$); ( ) sequential assignments *via* d_2 (NOE from NH_i to NH_{i+1}); ( ) sequential assignments *via* d_3 (NOE from NH_{i+1} to $C^{\beta}H_i$); (○) assignment in the sequence relied on the identification of the spin system in the COSY spectrum, whereby the amide proton resonances of these residues were not observed. The arrows indicate locations where all the resonances were assigned but the connectivity between 2 neighbouring residues was not established.

$C^{\beta}H_i$ played an important role for the sequential assignments (see Fig. 6 for a survey of the experiments used), even though their inherent reliability is considerably lower than that for the connectivities among amide and C^{α} protons (Billeter *et al.*, 1982). To illustrate this practically important aspect, all three connectivities *via* d_1, d_2 and d_3 are documented in what follows for the C-terminal segment residues 18 to 29 of glucagon (Figs 8 to 11). Only one connectivity for each assignment is documented for the other regions (Figs 11 to 13).

(c) *Sequential connectivities* via d_1, d_2 *and* d_3 *in glucagon residues 18 to 29*

The sequential assignments were unravelled in a H_2O NOESY spectrum of micelle-bound glucagon recorded at 37°C (Fig. 7) and in a corresponding COSY spectrum. In the NOESY spectrum, only the region (ω_1 = 1·4 to 9·0 p.p.m., ω_2 = 7·5 to 9·0 p.p.m.) was needed for the assignments, so that the analysis was not impeded by the strong vertical bands of spurious noise that arose between 1 and 5 p.p.m. from the water irradiation and from the sharp diagonal peaks of residual protons in the deuterated lipid. In Figures 8 to 13, small regions of the NOESY spectrum in Figure 7 and in some cases of the corresponding COSY spectrum are displayed. The chemical shifts of the assigned resonances are listed in Table 2.

As indicated in Figure 6, a continuous line of sequential assignments *via* d_3 was obtained for glucagon residues 18 to 29. These data are documented in Figure 8. The connectivity diagram starts at the chemical shift of the Thr29 amide proton in the lower right corner. A vertical line at this position goes through the centre of only one peak, which was therefore assigned to $C^{\beta}H$ of Asn28 (Billeter *et al.*, 1982). The amide proton of Asn28 was then identified in the COSY spectrum *via* the J-connectivities in the fragment $-C^{\beta}H_2-C^{\alpha}H-NH-$. The COSY connectivity is not shown; instead, the (NH_{29}–$C^{\beta}H_{28}$) cross peak in the NOESY spectrum of Figure 8 is connected with the ($C^{\beta}H_{28}$–NH_{28}) cross peak. A vertical line at the chemical shift of Asn28 NH goes through the centre of only one peak, which was therefore assigned to $C^{\beta}H$ of Met27. The horizontal connection from (NH_{28}–$C^{\beta}H_{27}$) to ($C^{\beta}H_{27}$–NH_{27}) was again based on the identification of the amide proton chemical shift of Met27 in the COSY spectrum. The next step was more ambiguous, since

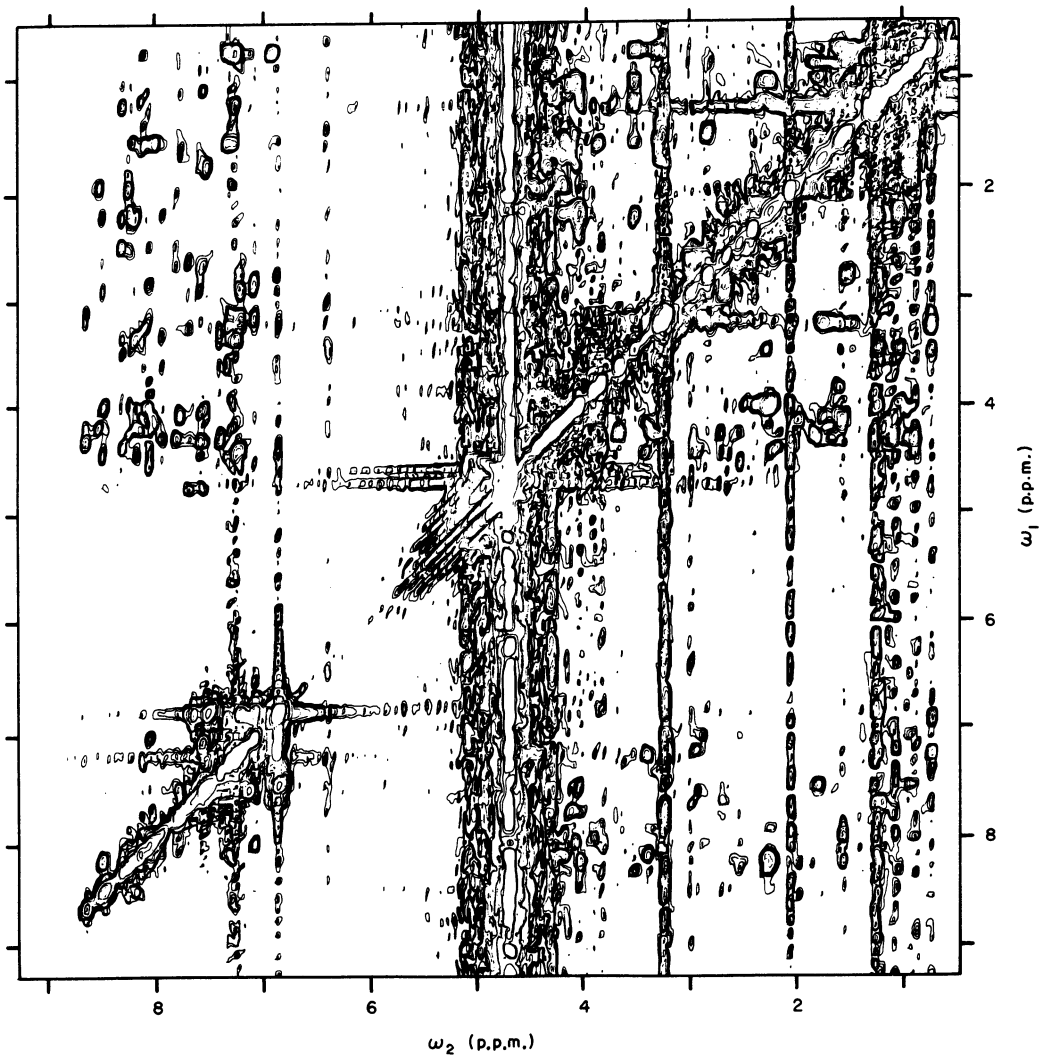


FIG. 7. Contour plot of a 360 MHz ^{1}H NOESY spectrum observed in the same sample of micelle-bound glucagon as the COSY spectrum of Fig. 5. The spectrum was recorded in 26 h, using a mixing time of 100 ms with random modulation of amplitude 6 ms (Macura *et al.*, 1981). The digital resolution is 6·3 Hz/point. An absolute value spectrum is shown.

there are three cross peaks on the vertical line through the chemical shift position of Met27 NH. The correct assignment was obtained on the basis of the amino acid sequence and the identified side-chain spin systems. Continuing in this way, assignments were obtained from Thr29 all the way to $C^\beta H$ of Arg18.

Statistical studies of protein crystal structures have shown that the inherent reliability of sequential assignments *via* d_3 is only ~75%, as compared to ~88% for d_1 or d_2, and between 92 and 99% when two of the three connectivities can be established simultaneously. Therefore, we looked for further connectivities in

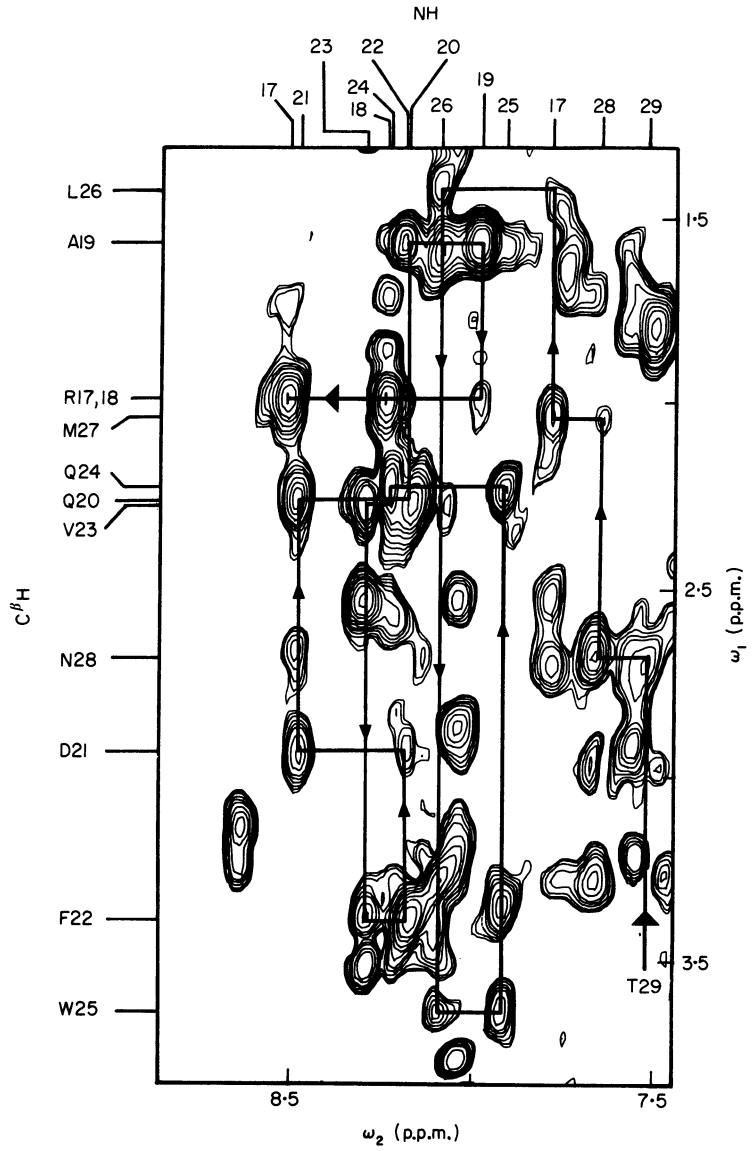


FIG. 8. Spectral region (1·3 to 3·8 p.p.m.) × (7·4 to 8·8 p.p.m.) of the 360 MHz ^{1}H NOESY spectrum of micelle-bound glucagon shown in Fig. 7. The continuous lines with arrows indicate the sequential assignments for the polypeptide segment residues 17 to 29, which were obtained from NOEs between amide protons and the C$^\beta$ protons of the preceding residues (d_3). The numbers at the top of the Figure indicate the amide proton chemical shifts of the corresponding residues, those on the left margin indicate the chemical shifts of 1 C$^\beta$ proton for each residue. The amide proton chemical shifts for Arg17 and Arg18 were identified *via* d_1 and d_2 connectivities (Figs 11 and 12) and have from there been added to the present Figure.

TABLE 2

Chemical shifts, δ†, of the assigned ^1H n.m.r. lines of glucagon bound to perdeuterated dodecylphosphocholine micelles, pH 6·0, t = 37°C

Amino acid residue	δ (±0·01 p.p.m.)†			
	NH	$C^{\alpha}H$	$C^{\beta}H$	Others
His1‡	n.o.	4·66‡	3·07, 3·21‡	$C^{\delta}H$ 7·22, $C^{\epsilon}H$ 8·11
Ser2‡	n.o.	4·51‡	3·86, 3·86‡	
Gln3§	8·75	4·41	2·01, 2·17	$C^{\gamma}H_2$, 2·38, 2·38
Gly4	8·47	4·02		
		4·02		
Thr5	8·06	4·34	4·19	$C^{\gamma}H_3$ 1·07
Phe6	8·63	4·64	3·13, 3·22	Ring 7·27
Thr7	8·11	4·23	4·24	$C^{\gamma}H_3$ 1·17
Ser8	8·04	4·37	3·78, 3·86	
Asp9	8·30	4·57	2·52, 2·52	
Tyr10	8·05	4·56	2·85, 3·18	$C^{\delta}H$ 7·06, 7·06
				$C^{\epsilon}H$ 6·82, 6·82
Ser11	8·03	4·06	3·96, 3·96	
Lys12	7·75	4·09	1·31, 1·59	$C^{\gamma}H_2$ 1·10, 1·52
				$C^{\delta}H_2$ 1·50, 1·50
				$C^{\epsilon}H_2$ 2·80, 2·80
Tyr13	7·55	4·42	2·89, 3·23	$C^{\delta}H$ 7·19, 7·19
				$C^{\epsilon}H$ 6·84, 6·84
Leu14	7·52	4·33	1·61, 1·81	$C^{\gamma}H$ 1·77
				$C^{\delta}H_3$ 0·88, 0·96
Asp15	7·55	4·75	2·63, 2·76	
Ser16	8·45	4·23	3·63, 3·63	
Arg17‖	8·50	4·17	1·96, 1·96	$C^{\gamma}H_2$ 1·73, 1·73
				$C^{\delta}H_2$ 3·25, 3·25
Arg18‖	8·24	4·25	1·86, 1·94	$C^{\gamma}H_2$ 1·71, 1·71
				$C^{\delta}H_2$ 3·27, 3·27
Ala19	7·99	4·17	1·56	
Gln20§¶	8·17	4·02	2·25	
Asp21	8·48	4·47	2·62, 2·92	
Phe22	8·18	4·44	3·38, 3·38	Ring 7·23
Val23	8·28	3·49	2·26	$C^{\gamma}H_3$ 1·02, 1·21
Gln24§¶	8·24	4·00	2·22	
Trp25	7·92	4·30	3·33, 3·62	$C^{\delta}H$ 7·37
				$N^{\epsilon}H$ 10·53
				$C^{\epsilon}H$ 7·33
				$C^{\zeta 2}H$ 7·54
				$C^{\zeta 3}H$ 6·89
				$C^{\eta}H$ 7·11
Leu26	8·09	3·27	1·41, 1·57	$C^{\gamma}H$ 1·56
				$C^{\delta}H_3$ 0·72, 0·72
Met27	7·79	4·30	2·03, 2·13	$C^{\gamma}H_2$ 2·53, 2·70
				$C^{\epsilon}H_3$ 2·04
Asn28§	7·66	4·74	2·68, 2·97	
Thr29	7·53	4·04	4·11	$C^{\gamma}H_3$ 1·07

† The chemical shifts, δ, are relative to external sodium 3-trimethylsilyl-[2,2,3,3-$^2H_4$]propionate. The ε-methyl resonance of Met27 was taken to be at 2·04 p.p.m. and was used as internal reference.

‡ After sequential assignments for the amino acid residues 3 to 29, the remaining resonances were assigned to His1 and Ser2 based on comparison with the corresponding random coil values (Bundi & Wüthrich, 1979*a*). The amide protons for these residues were not observed (n.o.).

glucagon residues 18 to 29. From Figure 8, the chemical shifts of the amide protons of residues 19 to 29 and at least one C^{β} proton of residues 18 to 29 are known, and the $C^{\alpha}H$ chemical shifts were established from the COSY spectrum. Accurate positions for the cross peaks needed to establish d_1 and d_2 connectivities could thus be predicted. In Figures 9 to 11, the predicted sequential d_1 and d_2 connectivities have been drawn into the experimental spectra.

The d_1 connectivities between NH_{i+1} and $C^{\alpha}H_i$ are presented in a combined COSY–NOESY connectivity diagram (Wagner *et al.*, 1981; Wagner & Wüthrich, 1982), which manifests the through-space NOE connectivities between NH_{i+1} and $C^{\alpha}H_i$, as well as the through-bond J-connectivities between $C^{\alpha}H_i$ and NH_i. A record of sequential assignments *via* d_1 then consists of a spiral-like connectivity pattern. In Figure 9, the spectral regions containing the $C^{\alpha}H$–NH cross peaks have been combined in such a way that the connectivities are *via* a virtual diagonal, exactly analogous to the connectivities through the real diagonal peaks when the entire halves of the NOESY and COSY spectra are joined along the diagonal (Wagner *et al.*, 1981). The first step starts with the COSY peak of Thr29, from where a horizontal line leads to the virtual diagonal position of the Thr29 amide proton. A vertical line then leads to the NOESY cross peak position with $C^{\alpha}H$ of Asn28, which was predicted from the NH and $C^{\alpha}H$ chemical shifts known on the basis of the previous assignments (Fig. 8). Next, a horizontal line leads to the virtual, diagonal position of $C^{\alpha}H$ of Asn28 and a vertical line to the COSY peak of Asn28. Continuing on, we follow the d_1 connectivities for residues 29 to 23 in Figure 9 and for residues 23 to 17 in Figure 10. When comparing the predicted positions for the cross peaks with the actual spectra, one finds that the NOESY spectrum provides clear evidence only for the d_1 connectivities between residues 29 to 27, 25 to 23 and 20 to 19. The predicted cross peak positions connecting residues 26 to 25, 23 to 22, 19 to 18 and 18 to 17 do not coincide with cross peaks in the spectrum. For the connectivities 27 to 26, 22 to 21, and 21 to 20, it could not be determined unambiguously whether they are manifested by a cross peak, since they fall into crowded locations.

The d_2 connectivities are indicated in Figure 11. Since d_2 is symmetrical with respect to the direction of the polypeptide chain (Billeter *et al.*, 1982), a non-terminal residue may be connected *via* two NH–NH cross peaks with the residues preceding it and following it in the sequence. The positions for all these cross peaks predicted from the assignments in Figure 8 are connected. The first three steps

---

§ In the NOESY spectrum recorded in H_2O, 3 cross peaks were observed at (ω_1 = 6·83 p.p.m., ω_2 = 7·63 p.p.m.), (6·82 p.p.m., 7·50 p.p.m.) and (6·77 p.p.m., 7·45 p.p.m.), which must come from the side-chain amide groups of Gln3, Gln20, Gln24 and Asn28. They have so far not been assigned individually.

|| Two cross peaks between Arg $C^{\delta}H_2$ and $N^{\epsilon}H$ were observed at (ω_1 = 3·26 p.p.m., ω_2 = 7·67 p.p.m.) and (3·26 p.p.m., 7·49 p.p.m.) in the COSY spectrum (Fig. 5). Since both $C^{\gamma}H_2$ and $C^{\delta}H_2$ have the same chemical shifts for the 2 residues, the resonances of the guanidinium groups could not be assigned individually.

¶ Two cross peaks at (ω_1 = 2·44 p.p.m., ω_2 = 2·59 p.p.m.) and (2·46 p.p.m., 2·52 p.p.m.) in the COSY spectrum recorded in ^2H_2O must correspond to the C^{γ} protons of Gln20 and Gln24. For these 2 residues, the connectivities between the 2 C^{β} protons and between C^{β} and C^{γ} protons could not be established.

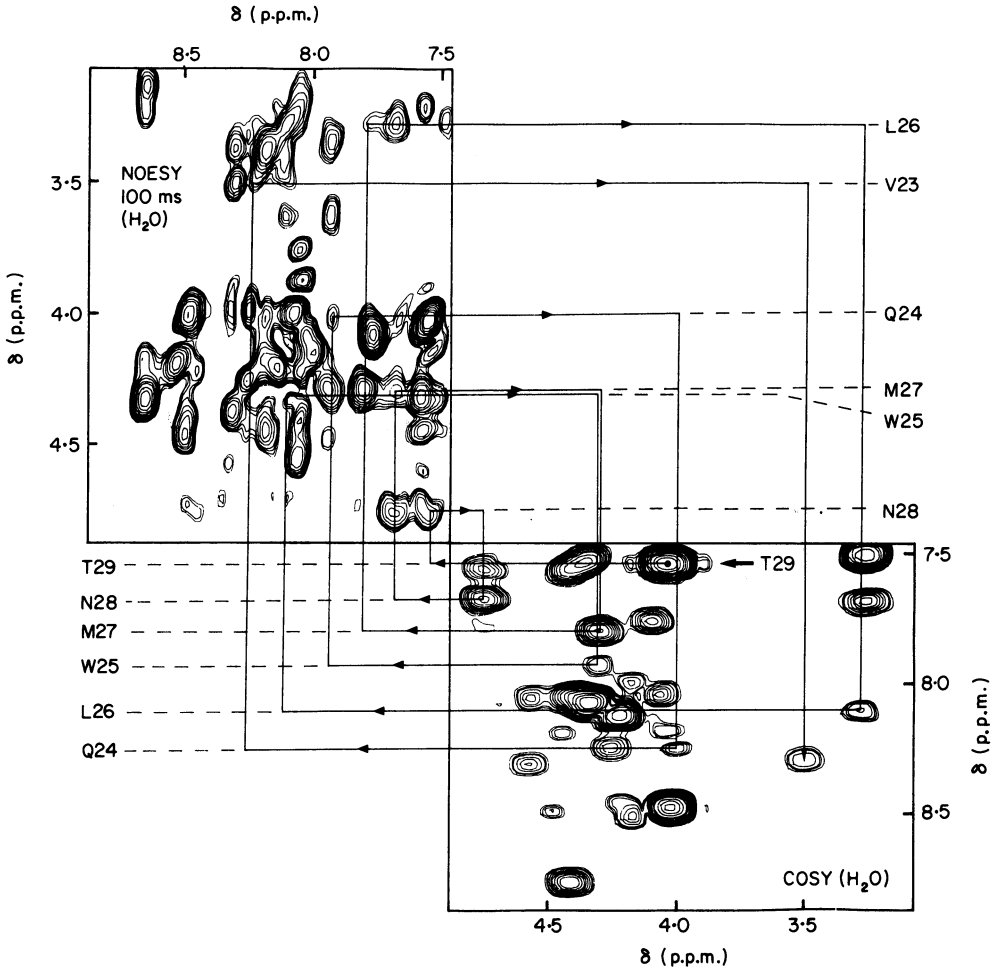


FIG. 9. Combined COSY–NOESY connectivity diagram for sequential resonance assignments *via* NOEs between amide protons and the C^{α} protons of the preceding residue (d_1) (Wagner *et al.*, 1981). In the upper left, the region (3·1 to 4·9 p.p.m.) × (7·4 to 8·9 p.p.m.) from the 360 MHz ^1H NOESY spectrum of micelle-bound glucagon in Fig. 7 is presented. In the lower right, the corresponding region from a COSY spectrum recorded from the same sample under identical conditions is shown. The straight lines and arrows indicate the connectivities between neighbouring residues in the segment residues 23 to 29 of micelle-bound glucagon. "←T29" identifies the starting point of the d_1 connectivity pattern. The amide proton chemical shifts are indicated by the assignments in the lower left corner, those for the C^{α} protons by the assignments in the upper right corner of the Figure.

from Thr29 to Leu26 are easy to follow, since they lead in each case to a lower field position. Leu26 connects to Trp25, which is at higher field. In these somewhat more complex cases, the connectivities to the diagonal peaks are indicated with broken lines. Continuing on, the connectivity pattern ends at Arg18. Comparison of the predicted peak positions with the experiment shows that the two are fully compatible, since none of the predicted cross peaks falls into an empty region of the spectrum. On the other hand, because of the small chemical shift differences

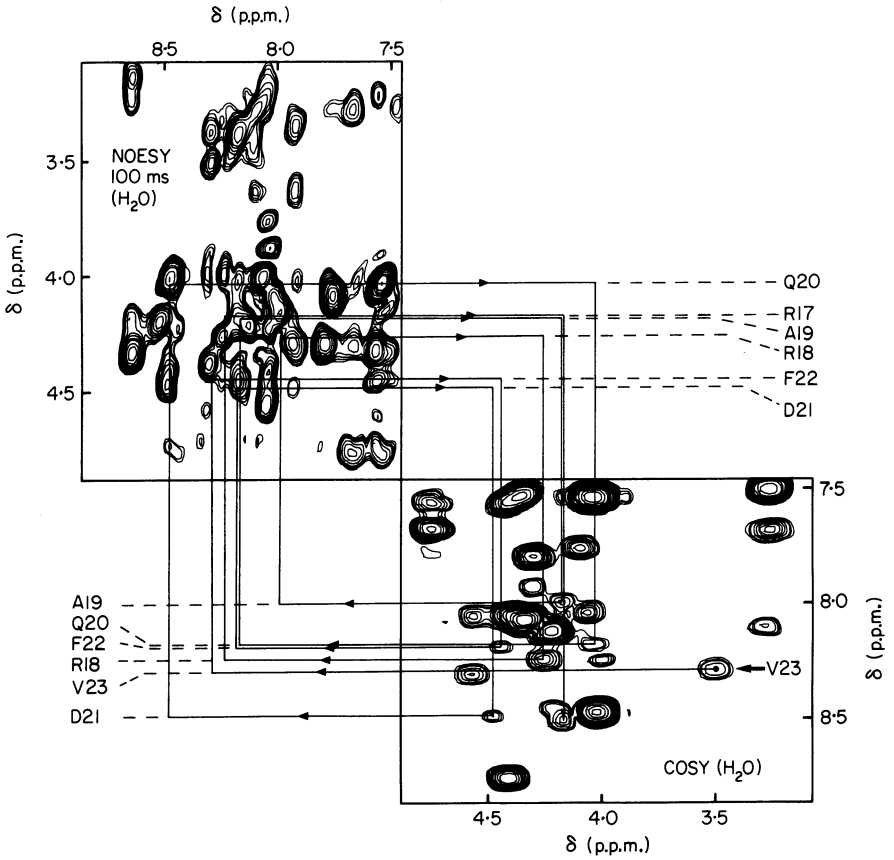


FIG. 10. Combined COSY–NOESY connectivity diagram for sequential resonance assignments *via* NOEs between amide protons and the C^{α} protons of the preceding residues (d_1) (same spectra as in Fig. 9). The straight lines and arrows indicate the connectivities between neighbouring residues in the segment residues 17 to 23 of micelle-bound glucagon. "←V23" identifies the starting point of the d_1 connectivity pattern. The amide proton chemical shifts are indicated by the assignments in the lower left corner, those for the C^{α} protons by the assignments in the upper right corner of the Figure.

between the amide protons of the different residues, most of the predicted NOE cross peak locations are so near the diagonal that it can hardly be seen whether there is indeed a cross peak. Overall, it appears that in the segment residues 18 to 29 only the d_2 connectivity between Met27 and Leu26 could have been established unambiguously from this spectrum.

The comparison of the d_1, d_2 and d_3 connectivities in glucagon residues 18 to 29 provides a valuable complementation of the theoretical and statistical evaluations of the reliability of sequential assignments *via* the different connectivities (Billeter *et al.*, 1982). It is seen that an inherently highly reliable parameter, such as d_2 in the present example, may be virtually uninformative, depending on the spectral characteristics of the protein. The intrinsically least reliable d_3 connectivity may thus become a very valuable parameter in many actual systems.

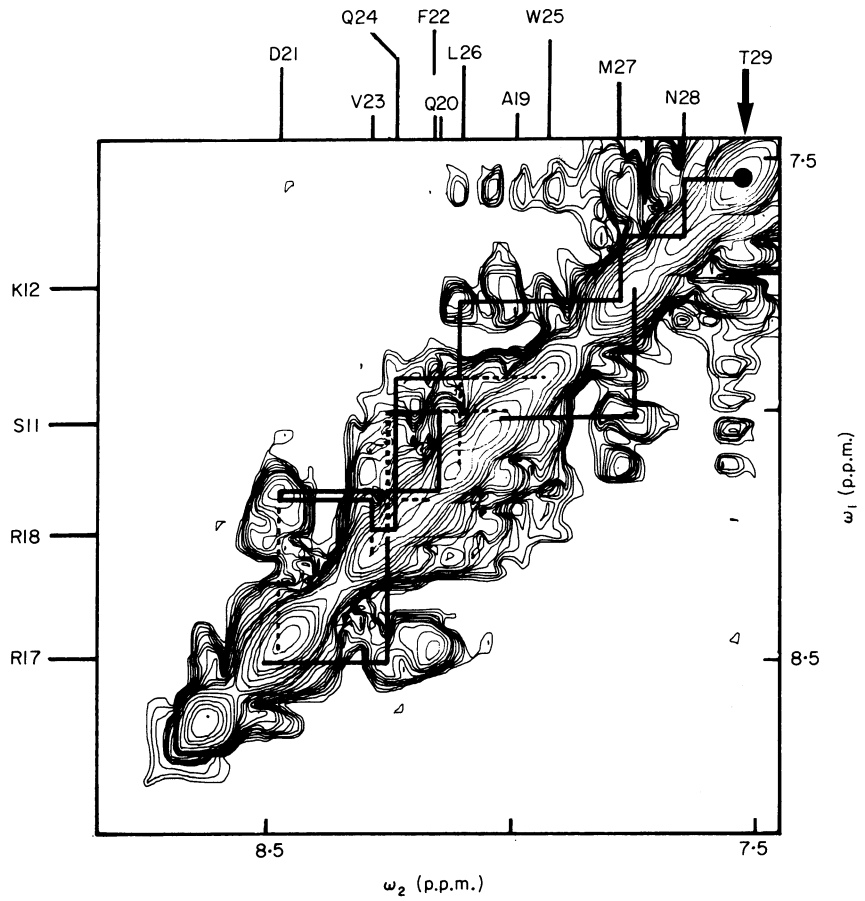


FIG. 11. Region from 7·4 to 8·8 p.p.m. of the 360 MHz ^{1}H NOESY spectrum of micelle-bound glucagon shown in Fig. 7, after symmetrization (Baumann *et al.*, 1981). In this spectral region, the diagonal peaks include all the backbone amide proton resonances, and the cross peaks manifest NOEs between different amide protons. The straight continuous and broken lines in the upper left triangle indicate the NOEs between the amide protons of neighbouring residues (d_2) in the polypeptide residues 18 to 29. In the lower right triangle, the sequential assignments obtained for the segment residues 11 to 12 and 17 to 18 are indicated. The chemical shifts for the amide protons in the segment residues 19 to 29 are indicated at the top of the Figure, those for the 2 short segments on the left margin.

(d) *Individual assignments for glucagon residues 1 to 18*

As outlined in Figure 6, the major part of the sequential assignments in this region were obtained *via* d_1 connectivities, with some missing links filled by d_2 or d_3 connections. The results are documented in Figures 11 to 13 in diagrams similar to those used for glucagon residues 29 to 18. Adding to this previously assigned segment, we first have the d_2 connectivity Arg18–Arg17 in Figure 11. Next we have d_1 connectivities from Arg17 to Leu14, which also resolved the ambiguity in the assignments of Arg18 and Arg17 (see Fig. 8). Connectivities within the tripeptide

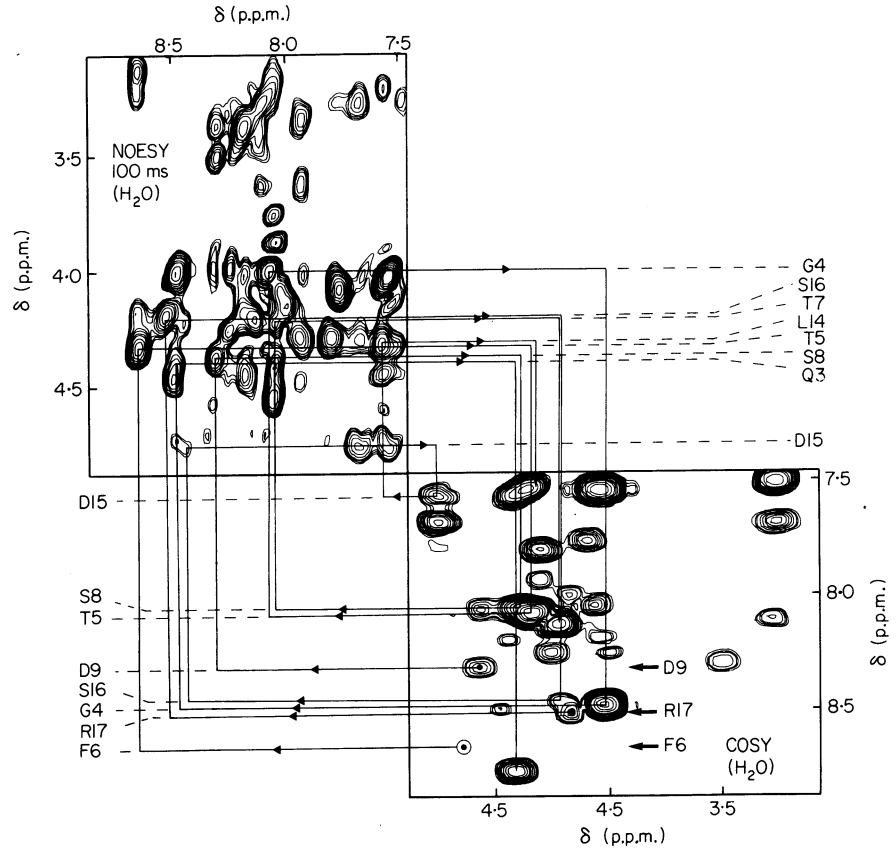


FIG. 12. Combined COSY-NOESY connectivity diagram for sequential resonance assignments *via* NOEs between amide protons and the C^{α} protons of the preceding residues (d_1) (same spectra as in Fig. 9). The straight lines and arrows indicate the sequential resonance assignments obtained for the segment residues 3 to 6, 7 to 9 and 14 to 17. At 37°C, the $C^{\alpha}H$–NH COSY peak of Phe6 coincides with the water resonance and was therefore bleached out by the solvent irradiation. Its location (⊙) was taken from the data in Figs 2 and 13.

Tyr13-Lys12-Ser11 were obtained *via* d_3 (Fig. 13) and *via* d_2 (Fig. 11). A d_3 connection from Tyr10 to Asp9 (Fig. 13) is followed by d_1 connectivities from Asp9 to Tyr7 and from Phe6 to Gln3 (Fig. 12), with a d_3 link between residues 7 and 6 (Fig. 13). Finally, the two N-terminal residues were assigned on the basis of the spin system identifications (Fig. 3) and comparison with the random coil chemical shifts for these residues (Bundi & Wüthrich, 1979*a*). The imidazole ring proton lines of His1 were identified independently from their chemical shifts (Bösch *et al.*, 1980; Table 2).

4. Conclusions

The assignment of the ^{1}H n.m.r. spectrum of micelle-bound glucagon presented in this paper is almost complete (Table 2). All the non-labile protons were observed

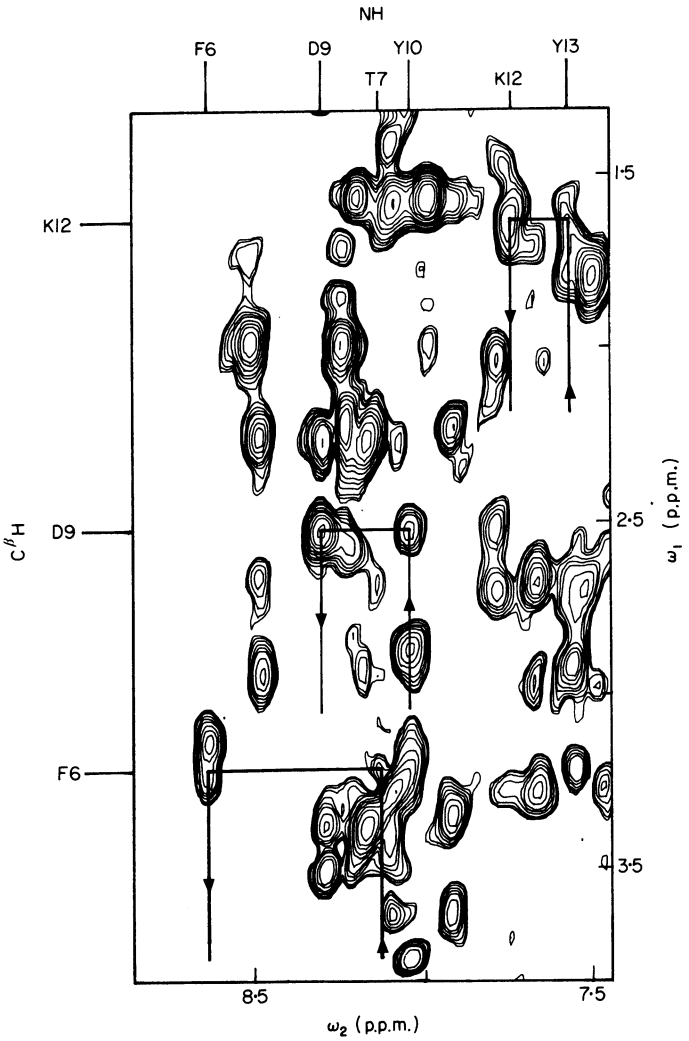


FIG. 13. Spectral region (1·3 to 3·8 p.p.m.) × (7·4 to 8·8 p.p.m.) of the 360 MHz ^{1}H NOESY spectrum of micelle-bound glucagon shown in Fig. 7. The continuous lines with arrows indicate the sequential assignments for the polypeptide segment residues 6 to 7, 9 to 10 and 12 to 13, which were obtained from NOEs between amide protons and the C^{β} protons of the preceding residues (d_3). The numbers at the top of the Figure indicate the amide proton chemical shifts of the corresponding residues, those on the left margin indicate the chemical shifts of one C^{β} proton for each residue.

and assigned, with the exception only of the C^{γ} protons of Gln20 and Gln24. These two residues have nearly identical $C^{\alpha}H$ and $C^{\beta}H$ chemical shifts, which was probably the main reason for failure to establish the connectivities to $C^{\gamma}H_2$. Among the labile protons, those of the two arginine side-chains were observed (Fig. 5), but could not be assigned individually because of the degeneracy of the $C^{\delta}H_2$ resonances. Three NH–NH NOE cross peaks were identified as originating from side-chain amide groups of the single asparagine and the three glutamine residues

(Table 2). So far, no unambiguous individual assignment on the basis of NOE connectivities with $C^{\beta}H_2$ in Asn or $C^{\gamma}H_2$ in Gln (Billeter *et al.*, 1982) could be obtained. The labile backbone protons of His1 and Ser2 and the labile side-chain protons of His1, Lys12 and all the serine, threonine and tyrosine residues were not observed, presumably because they exchange too rapidly with the solvent to be seen as separate lines (Wüthrich, 1976; Bundi & Wüthrich, 1979*a*,*b*).

Previously, a three-dimensional structure for residues 19 to 27 of micelle-bound glucagon was determined on the basis of intramolecular distance constraints between a small number of assigned resonances, which were measured by NOE experiments and analysed with the use of a distance geometry algorithm (Braun *et al.*, 1981). The new assignments described here provide the basis for a much more accurate characterization of the polypeptide conformation. In the earlier work, only very few of the observed NOEs in ^2H_2O solution (Bösch *et al.*, 1981) could be used for the structure determination, since the others had not been assigned (Braun *et al.*, 1981). For the new structure determination, essentially all the previously reported NOEs (Bösch *et al.*, 1981) and, in addition, all those with the backbone amide protons 3 to 29 will be assigned to specific pairs of closely spaced protons. The work on the determination of the three-dimensional structure is in progress.

We thank Drs L. R. Brown and A. Chrzeszczyk for help with the synthesis of perdeuterated dodecylphosphocholine, Dr W. Braun, M. Billeter and Dr G. Wagner for stimulating discussions on sequential resonance assignments, Mrs E. Huber for the careful preparation of the manuscript, Mrs E. H. Hunziker for some of the illustrations, and the Schweizerischer Nationalfonds for financial support (project 3.528.79).

REFERENCES

Anil Kumar, Ernst, R. R. & Wüthrich, K. (1980*a*). *Biochem. Biophys. Res. Commun.* **95**, 1–6.

Anil Kumar, Wagner, G., Ernst, R. R. & Wüthrich, K. (1980*b*). *Biochem. Biophys. Res. Commun.* **96**, 1156–1163.

Aue, W. P., Bartholdi, E. & Ernst, R. R. (1976). *J. Chem. Phys.* **64**, 2229–2246.

Baumann, R., Wider, G., Ernst, R. R. & Wüthrich, K. (1981). *J. Magn. Reson.* **44**, 402–406.

Billeter, M., Braun, W. & Wüthrich, K. (1982). *J. Mol. Biol.* **155**, 321–346.

Bösch, C., Bundi, A., Oppliger, M. & Wüthrich, K. (1978). *Eur. J. Biochem.* **91**, 209–214.

Bösch, C., Brown, L. R. & Wüthrich, K. (1980). *Biochim. Biophys. Acta*, **603**, 298–312.

Bösch, C., Anil Kumar, Baumann, R., Ernst, R. R. & Wüthrich, K. (1981). *J. Magn. Reson.* **42**, 159–163.

Braun, W., Bösch, C., Brown, L. R., Go, N. & Wüthrich, K. (1981). *Biochim. Biophys. Acta*, **667**, 377–396.

Brown, L. R. (1979). *Biochim. Biophys. Acta*, **557**, 135–148.

Brown, L. R., Bösch, C. & Wüthrich, K. (1981). *Biochim. Biophys. Acta*, **642**, 296–312.

Bundi, A. & Wüthrich, K. (1979*a*) *Biopolymers*, **18**, 285–298.

Bundi, A. & Wüthrich, K. (1979*b*). *Biopolymers*, **18**, 299–312.

Dubs, A., Wagner, G. & Wüthrich, K. (1979). *Biochim. Biophys. Acta*, **577**, 177–194.

Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. R. (1979). *J. Chem. Phys.* **71**, 4546–4553.

Kalinichenko, P. (1976). *Stud. Biophys.* **58**, 235–240.

Macura, S., Huang, Y., Suter, D. & Ernst, R. R. (1981). *J. Magn. Reson.* **43**, 259–281.

Nagayama, K. & Wüthrich, K. (1981). *Eur. J. Biochem.* **114**, 365–374.

Nagayama, K., Wüthrich, K. & Ernst, R. R. (1979). *Biochem. Biophys. Res. Commun.* **90**, 305–311.

Nagayama, K., Anil Kumar, Wüthrich, K. & Ernst, R. R. (1980). *J. Magn. Res.* **40**, 321–334.
Panijpan, B. & Gratzer, W. B. (1974). *Eur. J. Biochem.* **45**, 547–553.
Pohl, S. L., Birnbaumer, L. & Rodbell, M. (1969). *Science*, **164**, 566–569.
Rodbell, M., Birnbaumer, L., Pohl, S. L. & Sundby, F. (1971). *Proc. Nat. Acad. Sci., U.S.A.* **68**, 909–913.
Rubalcava, B. & Rodbell, M. (1973). *J. Biol. Chem.* **248**, 3831–3837.
Sasaki, K., Dockerill, S., Adamiak, D. A., Tickle, I. J. & Blundell, T. (1975). *Nature (London)*, **257**, 751–757.
Wagner, G. & Wüthrich, K. (1982). *J. Mol. Biol.* **155**, 347–366.
Wagner, G., Anil Kumar & Wüthrich, K. (1981). *Eur. J. Biochem.* **114**, 375–384.
Wüthrich, K. (1976). *NMR in Biological Research: Peptides and Proteins*, North-Holland Publishing Company, Amsterdam.
Wüthrich, K., Wider, G., Wagner, G. & Braun, W. (1982). *J. Mol. Biol.* **155**, 311–319.

Edited by S. Brenner