Solution Conformation of *N*-Acetyl-L-prolyl-L-glutaminyl-L-prolyl-L-prolyl-L-glutaminamide, a Pentapeptide Fragment of the Type I Collagen α-1 Chain C-Telopeptide, Determined by Phase-Sensitive Two-Dimensional NMR Techniques

Albin Otter,[†] Paul G. Scott,[‡] and George Kotovych*[†]

Contribution from the Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada, and Department of Oral Biology, University of Alberta, Edmonton, Alberta T5N 2N8, Canada. *Received March 16, 1987*

Abstract: The solution conformation of the pentapeptide *N*-acetyl-L-Pro[1]-L-Gln[2]-L-Pro[3]-L-Pro[4]-L-Gln[5]-NH_2 has been studied by circular dichroism and NMR spectroscopy in aqueous solutions. The pentapeptide is a fragment of type I collagen α-1 chain C-telopeptide. By the combined application of various two-dimensional NMR techniques, a complete assignment of all proton as well as most ^{13}C resonances could be achieved. All spectroscopic evidence indicates that the pentapeptide exists in a nonrandom, rigid conformation that might form a stable nucleus around which the rest of the telopeptide could fold.

Type I collagen is the protein which gives mechanical strength to bones, teeth, skin, tendon, and other connective tissues. In the extracellular space collagen monomers secreted by fibroblasts and other cells aggregate in a specific polarized fashion giving rise to fibrils typically 50 to 200 nm in diameter and several microns in length.[1]

The collagen monomer consists of three polypeptide chains, each a little over 1000 amino acid residues long. Two of these chains are identical and are referred to as α-1 chains, and one differs slightly in sequence and is referred to as α-2. For most of their lengths the three α chains are coiled around each other in the polyproline II helix ("collagen triple-helix") which is extensively stabilized by interchain hydrogen bonds.[2] At each end of the α chains are short sequences, called "telopeptides", which do not adopt the triple-helical conformation. The α-1 chain of bovine type I collagen carries a 16-residue N-telopeptide and a 24-residue C-telopeptide.[3]

Warming a neutral solution of monomeric collagen obtained by extraction of connective tissues under nondenaturing conditions causes the precipitation of fibrils which are morphologically very similar to those seen in vivo.[4,5] Collagen fibrillogenesis in vitro has been intensively studied in hopes of elucidating the mechanism of this important step in the biogenesis of connective tissues. From experiments on proteolytically modified collagens, it has been deduced that the C-telopeptide promotes both linear and lateral growth.[6,7] These effects must arise from interactions between telopeptides and the triple-helices of adjacent collagen monomers in the growing fibrils and will almost certainly depend on the adoption by the telopeptides of specific conformations.

With the ultimate aim of furthering understanding of the mechanism of fibrillogenesis, we have begun to study the solution conformations of the isolated telopeptides. The circular dichroism spectrum of the C-telopeptide in aqueous solution at neutral pH was interpreted as indicating the presence of little regular secondary structure.[8] However, the sensitivity of the spectrum to temperature and to the presence of chemical denaturants suggested some restrictions on chain mobility, i.e., that the conformation could not be adequately described as a random coil. The C-telopeptides of type I collagen from cow,[9] chick,[10] and rabbit[11] all contain the sequence Pro-Gln-Pro-Pro-Gln (P^1-Q^2-P^3-P^4-Q^5)[12] approximately in the middle. In view of the strong conformation-directing properties of the proline ring, this sequence is expected to exhibit severely limited flexibility and consequently might form a stable nucleus around which the rest of the telopeptide could fold. The object of the present paper is to examine this hypothesis by means of extensive two-dimensional NMR techniques for both proton and ^{13}C. The results confirm that the pentapeptide exists in a rigid conformation which is unusual for a noncyclic peptide of that size.

Materials and Methods

The peptide (NAC-L-Pro-L-Gln-L-Pro-L-Gln-NH_2) was synthesized by the Alberta Peptide Institute (Department of Biochemistry, University of Alberta) and supplied as a crude product which was purified on a C-4 reverse-phase HPLC column (Vydac 214TP104, Separations Group Inc.) eluted with a linear gradient of acetonitrile in 0.1% (v/v) trifluoroacetic acid.

Circular dichroism spectra of the pentapeptide dissolved at 0.2 mg/mL in water, methanol, or trifluoroethanol were measured on a JASCO J-20A spectropolarimeter in jacketed thermostated cells of 1-mm path length. The instrument had been calibrated with *d*-10-camphorsulfonic acid.[13]

The NMR samples were prepared by dissolving the solid material in 0.5 mL of 99.996% D_2O (Aldrich) or 0.5 mL of 80% H_2O/20% D_2O (referred to as D_2O and H_2O samples, respectively). The initial pH[14] of 2.5 was kept unchanged. Before sealing, the samples were degassed by bubbling with argon for ca. 30 min. The concentration was 4 mM for the D_2O and 16 mM for the H_2O sample.

All NMR experiments were performed on a Bruker AM300 NMR spectrometer at ambient temperature (297 ± 1 K). Data collection and data processing were controlled by an Aspect 3000 computer equipped with an Array Processor using 1986 Bruker DISNMR software. ^1H and ^{13}C chemical shifts were determined relative to external 5% 3-trimethylsilyl-1-propanesulfonate (TSP) measured in a separate sample under the same experimental conditions.

The one-dimensional proton spectra were recorded using a sweep width of 1075 Hz (2200 Hz in H_2O) and 4K (zero-filled to 16K) data

[†]Department of Chemistry.
[‡]Department of Oral Biology.

(1) Parry, D. A. D.; Craig, A. S. *Ultrastructure of the Connective Tissue Matrix*; Ruggeri, A., Motta, P. M., Eds.; Martinus Nijhoff: Boston, 1984; pp 34-64.
(2) Rich, A.; Crick, F. H. C. *Nature (London)* **1955**, *176*, 915-916.
(3) Glanville, R. W.; Kühn, K. *Fibrous Proteins: Scientific, Industrial and Medical Aspects*; Parry, D. A. D., Cramer, L. K., Eds.; Academic Press: London, 1979; Vol. 1, pp 133-150.
(4) Gross, J.; Highberger, J. H.; Schmitt, F. O. *Proc. Natl. Acad. Sci. U.S.A.* **1955**, *41*, 1-7.
(5) Jackson, D. S.; Fessler, J. H. *Nature (London)* **1955**, *176*, 69-70.
(6) Helseth, D. L., Jr.; Veis, A. *J. Biol. Chem.* **1981**, *256*, 7118-7128.
(7) Capaldi, M. J.; Chapman, J. A. *Biopolymers* **1982**, *21*, 2291-2313.
(8) Scott, P. G. *Biochemistry* **1986**, *25*, 974-980.
(9) Rauterberg, J.; Fietzek, P. P.; Rexrodt, F.; Becker, U.; Stark, M.; Kühn, K. *FEBS Lett.* **1972**, *21*, 75-79.
(10) Fuller, F.; Boedtker, H. *Biochemistry* **1981**, *20*, 996-1006.
(11) Becker, U.; Fietzek, P. P.; Furthmayr, H.; Timpl, R. *Eur. J. Biochem.* **1975**, *54*, 359-366.
(12) Amino acids are abbreviated by their one letter symbol throughout this paper.
(13) De Tar, Del. F. *Anal. Chem.* **1969**, *41*, 1406-1408.
(14) pH data are uncorrected meter readings.

0002-7863/87/1509-6995$01.50/0

Table I. Summary of Experimental Parameters Used in the Two-Dimensional NMR Experiments[a]

parameters	COSYPH	RELAY	JRES	NOESYPH[b]	XHDEPT	COLOC
sweep width in F2 (Hz)	920	920	920	920	3125	3846
sweep width in F1 (Hz)	460	460	58	460	460	250[c]
matrix size (F1 × F2) before zero-filling	256 × 1K	128 × 1K	64 × 1K	250 × 1K	24 × 2K	30 × 2K
matrix size (F1 × F2) after zero-filling	1K × 1K	512 × 1K	256 × 2K	1K × 1K	128 × 2K	128 × 4K
evolution time						
initial value (μs)	3	3	3	3	3	3
increment (ms)	0.543	1.086	4.34	0.543	0.543	1.000
number of scans (dummy scans)	112 (2)	240 (2)	64 (2)	112 (2)	1440 (2)	2080 (2)
acquisition time (s)	0.55	0.55	0.55	0.55	0.33	0.27
relaxation delay (s)	1.45	1.45	1.95	1.45	1.50	1.00
other delays (ms)		80[d]		500[e]	4.2[f]	29/30[g]
window functions for 2-D FT (F1/F2)	S/S	Q/Q	S/S	S/S	G/E	E/E
shifts of window functions in fractions of π (F1/F2)	0/0	0/0	0/0	3/4	−20/5.0[h]	4.0/4.0[h]

[a] All spectra were obtained at ambient temperature (297 ± 1 K). The concentration of the D_2O sample (COSYPH, RELAY, JRES, NOESYPH) was 4 mM. For the ^{13}C (XHDEPT, COLOC) a 16 mM H_2O/D_2O (80:20) sample was used. pH in both cases was 2.5.[14] For abbreviations of the various techniques see Materials and Methods. Window functions are abbreviated as follows: S, sine bell; Q, sine bell squared; G, Gaussian multiplication; E, exponential multiplication. [b] For the phase-sensitive NOESY experiment in H_2O, the sweep widths were 2100 and 1050 Hz in F2 and F1, respectively; 256 experiments incremented by 0.238 ms were recorded with 112 scans, for each T1 value. Unshifted sine bell functions were used in both dimensions. [c] Because of the low sample concentration only a limited number of experiments (defining the resolution in F1) were possible. Therefore, the F1 domain was deliberately folded to allow for a reasonable digital resolution. [d] Coherence transfer time. [e] Mixing time. [f] $1/2J$. [g] Δ_1/Δ_2.[35] [h] Line broadening factors (F1/F2).

points in both solvents. Especially in the case of the H_2O sample, a relatively small number of data points (and, consequently, a short acquisition time of 0.9 s) was required to avoid recovery of the intense water signal which was irradiated during a 5-s relaxation delay. By doing this, more sophisticated water-suppression techniques, such as selective excitation by composite pulses,[15,16] were not required thus keeping base-line distortions at a minimum. A relatively large number of scans (344 in D_2O, 176 in H_2O) were accumulated to obtain a high signal-to-noise ratio allowing extensive resolution enhancement by means of a Gaussian multiplication (line broadening factor −2 Hz in D_2O, −5 Hz in H_2O) prior to Fourier transformation. For the assignment of the two NH protons (inset B, Figure 2), a series of homodecoupling experiments with solvent presaturation was performed. The solvent was irradiated during a 5-s relaxation delay, and homodecoupling at a power level appropriate for 10 Hz couplings was carried out during the 0.9-s acquisition time. In the same manner the β protons were selectively decoupled while observing the effect on the corresponding H_α signals. As this was carried out on the D_2O sample, only a 3-s presaturation period was required and the aforementioned power level was used for homodecoupling during the acquisition time.

The nonselective longitudinal ^1H relaxation times were determined in D_2O using a 180°−τ−90° pulse sequence. Eight τ values from 0.001 to 1.5 s and a 5-s relaxation delay were employed.

One-dimensional NOE measurements (D_2O sample) were carried out in the difference mode using multiple irradiation.[17,18] Each line in a multiplet was irradiated 10 times 50 ms (total irradiation time 0.5 s). This procedure allows a very low decoupler power level (typically 10 dB lower than for a standard NOE experiment), making it much more selective in case of close proximity of resonances. A total of 2560 scans for each multiplet was required; the relaxation delay was 2.5 s.

The ^{13}C NMR spectrum (top of Figure 5) was recorded on the 16-mM H_2O sample using a DEPT (distortionless enhancement by polarization transfer)[19] pulse sequence: 8000 scans at a sweep width of 15 kHz (200 ppm) were acquired; 32K data points zero-filled to 64K and a total relaxation time (including the acquisition time) of 1.9 s were used. The evolution delay for the polarization transfer was 4.2 ms. The final ^1H detection pulse was 135°, creating positive signals for CH and CH_3 groups and negative signals for CH_2. During data acquisition proton broad-band decoupling was performed by means of a composite pulse of the WALTZ[20] type. The FID was multiplied with a Gaussian function

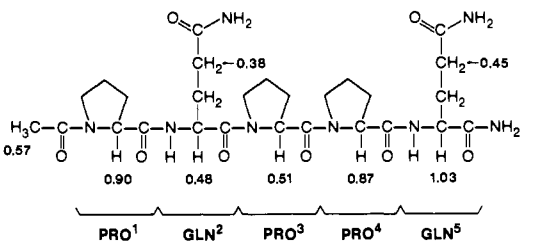


Figure 1. The primary structure of P^1-Q^2-P^3-P^4-Q^5, together with some proton T1 relaxation times (in seconds).

(line broadening factor −1.0, multiply factor 0.15) prior to Fourier transformation. The quaternary carbons were measured by means of a "q-only" sequence.[21] Although this sequence is designed to excite only unprotonated ^{13}C's, residual signals from the proton-bearing carbon resonances may occur due to pulse and delay imperfections. To avoid any folding of such artifacts into the region of interest, a wide sweep width of 13.9 kHz was selected; 32K (zero-filled to 64K) data points, a 5-relaxation delay, and a total spin-echo delay of 8.3 ms were used for a total of 13 300 scans. Composite pulse broad-band decoupling[20] for the ^1H domain was performed during the relaxation period and during data acquisition. The FID was multiplied with an exponential multiplication (line broadening factor 1.0) prior to Fourier transformation.

The parameters used in the various two-dimensional[22,23] NMR techniques are summarized in Table I. Our data are generally zero-filled three times[24] (supplementing N data points by $3N$ zeros) in the t_1 domain, and quadrature detection is used in both dimensions. The COSY (two-dimensional correlated spectroscopy)[25,26] and the NOESY (two-dimensional NOE spectroscopy)[27,28] experiments were both carried out in the phase-sensitive mode.[29,30] Solvent suppression was achieved by contin-

(15) Hore, P. J. *J. Magn. Reson.* **1983**, *55*, 283–300.
(16) Tong, J. P. K.; Kotovych, G. *J. Magn. Reson.* **1986**, *69*, 511–518.
(17) Neuhaus, D. *J. Magn. Reson.* **1983**, *53*, 109–114.
(18) Kinns, M.; Sanders, J. K. M. *J. Magn. Reson.* **1984**, *56*, 518–520.
(19) Doddrell, D. M.; Pegg, D. T.; Bendall, M. R. *J. Magn. Reson.* **1982**, *48*, 323–327.
(20) Shaka, A. J.; Keeler, J.; Freeman, R. *J. Magn. Reson.* **1983**, *53*, 313–340.
(21) Bendall, M. R.; Pegg, D. T. *J. Magn. Reson.* **1983**, *53*, 272–296.
(22) Benn, R.; Günther, H. *Angew. Chem., Int. Ed. Engl.* **1983**, *22*, 350–380.
(23) Bax, A. *Two-Dimensional Nuclear Magnetic Resonance in Liquids*; D. Reidel: Boston, 1982.
(24) Wider, G.; Macura, S.; Kumar, A.; Ernst, R. R.; Wüthrich, K. *J. Magn. Reson.* **1984**, *56*, 207–234.
(25) Aue, W. P.; Bartholdi, E.; Ernst, R. R. *J. Chem. Phys.* **1976**, *64*, 2229–2246.
(26) Bax, A.; Freeman, R. *J. Magn. Reson.* **1981**, *44*, 542–561.
(27) Jeener, J.; Meier, R.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546–4553.
(28) Kumar, A.; Ernst, R. R.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1–6.
(29) Marion, D.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *113*, 967–974.

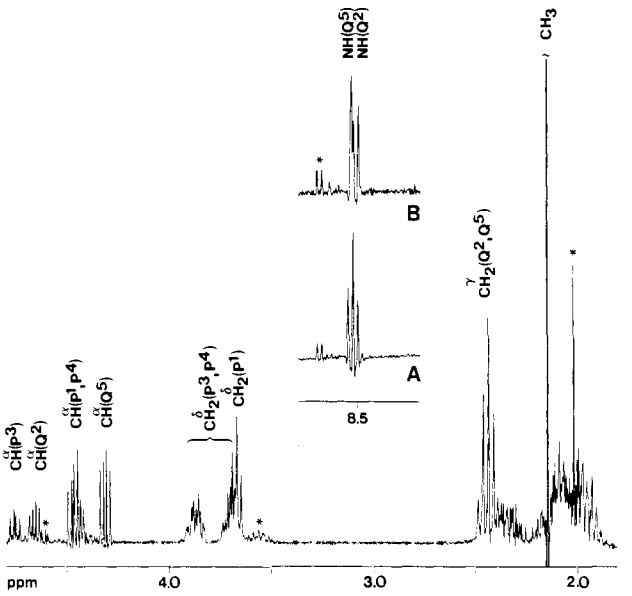


Figure 2. The 300-MHz ^{1}H spectrum of P^1-Q^2-P^3-P^4-Q^5 at 4 mM concentration in D_2O, pH 2.5 at 297 K. Inset A shows the region of the NH(Q^2,Q^5) protons measured in a 16 mM H_2O sample. The result of the homodecoupling experiment with the decoupler irradiating the H_α-(Q^5) proton is depicted in trace B. Resonances arising from the minor (cis) isomer are denoted with a star.

uous irradiation at all times except during data acquisition.[31] The contour plots have been symmetrized with respect to the diagonal. The RELAY (homonuclear relayed coherence transfer spectroscopy)[32,33] experiment was recorded in the absolute-value mode and symmetrized along the diagonal. The same can be said for the *J*-resolved spectrum (abbreviated JRES in Table I) except that the spectrum was first tilted and then symmetrized along F1 = 0. For the $^{13}C/^1H$ shift correlation experiment (XHDEPT)[34] and the COLOC (correlation spectroscopy via long-range couplings),[35] the power spectrum was calculated. It is noteworthy that often assignments cannot be made on the basis of the contour plot alone (especially for the XHDEPT and NOESY experiments) and examination of the appropriate cross sections was required.

Results

The spectral analysis of the P^1-Q^2-P^3-P^4-Q^5 pentapeptide is not straightforward given the relatively short sequence (Figure 1). The fact that only two different amino acids are present (the prolines have a complex spectrum) leads to severe overlapping of many proton resonances, especially in the high-field region of the spectrum (Figure 2). In addition, our previous studies on peptides of similar size[36-38] showed that literature chemical shift data[39] are not applicable for short peptide chains. Therefore, we subjected our sample to a variety of two-dimensional NMR experiments.[22,23] First, a COSY spectrum in the phase-sensitive mode[29] was recorded on the 4 mM D_2O sample. The resulting contour plot is shown in Figure 3. It exhibits the anticipated high spectral resolution, and especially the H_α–H_β connectivities become immediately accessible. However, as the H_α–H_β–$H_{\beta'}$ fragments

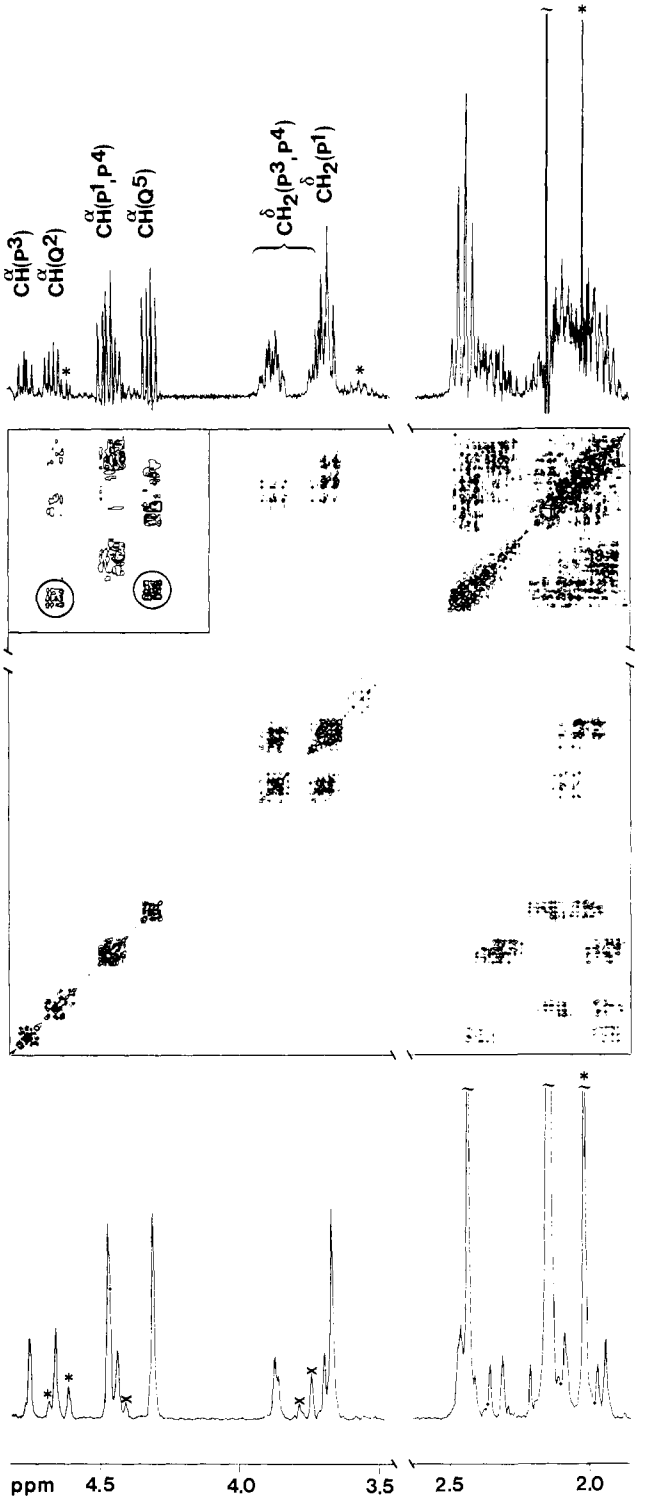
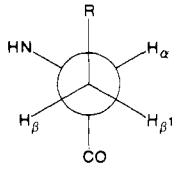


Figure 3. The 300-MHz two-dimensional contour plot of a phase-sensitive COSY experiment of P^1-Q^2-P^3-P^4-Q^5. The inset depicts a part of the RELAY spectrum recorded in the absolute value mode. The encircles cross-peaks reveal the H_α–H_γ connectivities for the two glutamine units. The one-dimensional spectrum is shown at the top, and at the bottom the projection of a *J*-resolved two-dimensional experiment is depicted. All spectra are obtained at 4 mM in D_2O, pH 2.5, 297 K. Minor (cis) isomer peaks are denoted with a star, impurities with an X.

are strongly coupled systems, the coupling constants $J_{\alpha\beta}$ and $J_{\alpha\beta'}$ are not as easily derivable from the cross-peaks as they are, for example, in the "fingerprint"[24] region where NH–H_α connectivities are evident. Nevertheless, the distinction between H_β and $H_{\beta'}$ was feasible because the α protons show well-resolved doublet-of-doublet coupling patterns (Figure 2). Thus, by selectively decoupling each of the β-proton resonances, it can be established

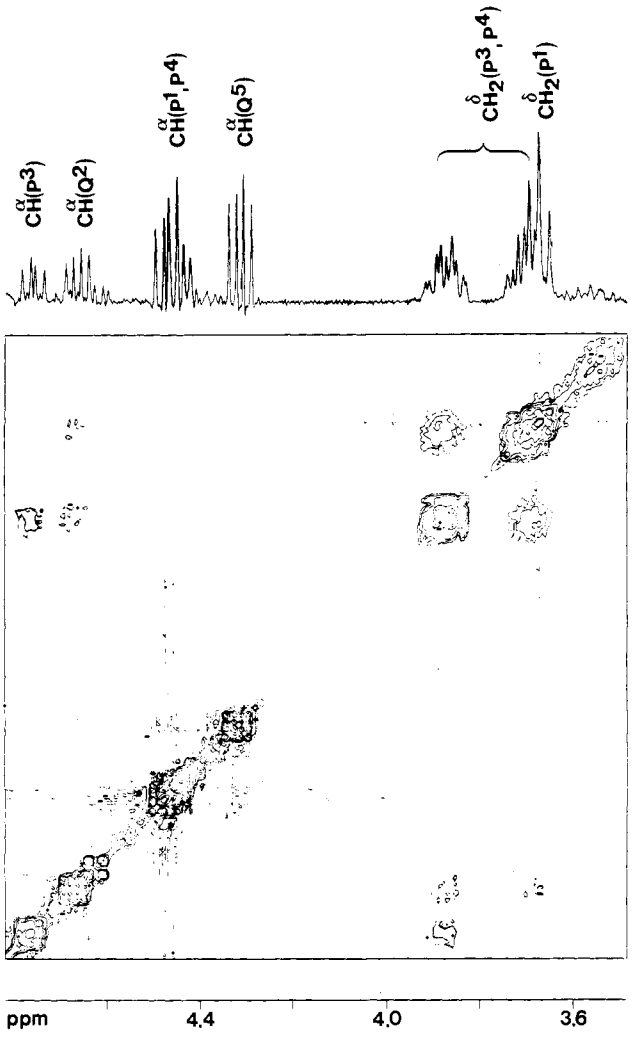
(30) Bodenhausen, G.; Kogler, H.; Ernst, R. R. *J. Magn. Reson.* **1984**, *58*, 370–388.
(31) Wider, G.; Hosur, R. V.; Wüthrich, K. *J. Magn. Reson.* **1983**, *52*, 130–135.
(32) Eich, G.; Bodenhausen, G.; Ernst, R. R. *J. Am. Chem. Soc.* **1982**, *104*, 3731–3732.
(33) Bolton, P. H.; Bodenhausen, G. *Chem. Phys. Lett.* **1982**, *89*, 139–144.
(34) Bendall, M. R.; Pegg, D. T. *J. Magn. Reson.* **1983**, *53*, 144–148.
(35) Kessler, H.; Griesinger, C.; Zarbock, J.; Loosli, H. R. *J. Magn. Reson.* **1984**, *57*, 331–336.
(36) Kawaki, H.; Otter, A.; Beierbeck, H.; Kotovych, G.; Stewart, J. M. *J. Biomol. Struct. Dyn.* **1986**, *3*, 795–803.
(37) Otter, A.; Kotovych, G.; Stewart, J. M. *J. Biomol. Struct. Dyn.* **1986**, *3*, 1121–1131.
(38) Szollosy, A.; Otter, A.; Stewart, J. M.; Kotovych, G. *J. Biomol. Struct. Dyn.* **1986**, *4*, 501–519.
(39) Jardetzky, O.; Roberts, G. C. K. *NMR in Molecular Biology*; Academic Press: New York, 1981; pp 143–186.



which one is connected via a large coupling to H_α (called H_β) and, consequently $H_{\beta'}$ is defined as well (small coupling constant to H_α). In the high-field region the COSY spectrum is naturally very complex and assignments cannot be made unambiguously. In particular, the two H_γ groups of Q^2 and Q^5 cannot be distinguished from each other nor from the various $H_\beta(P)$ lines. This problem could be solved in a very clear-cut way, by applying a RELAY[32,33] experiment. The RELAY is an extension of the COSY experiment, and it allows the observation of cross-peaks between spins that are not mutually coupled but belong to the same spin system. The relevant part of the RELAY contour plot is shown in Figure 3 (inset). Besides the COSY-like cross-peaks, which are always present in such an experiment, the additional H_α–H_γ cross-peaks (encircled) can be seen clearly. They identify the strong resonances at about 2.4 ppm as $H_\gamma(Q^2)$ and $H_\gamma(Q^5)$. Apparently they are exactly separated by 7 Hz, which is precisely the same value as the H_γ–H_β coupling constant in both residues, and therefore the spectrum exhibits the rather unexpected "multiplet" as can be seen in the one-dimensional proton spectrum. In addition, the result of the RELAY experiment also permits the final distinction in the α-proton region between proline and glutamine residues, which otherwise would have to be based on the splitting of chemical shifts between the β and β' protons. Clearly, for a peptide of this size, the RELAY argument is much more reliable than the chemical shifts of the side-chain β protons.

The results of the COSY and RELAY experiments allow the assignments within each amino acid moiety to be established and the distinction between proline and glutamine residues could be made. The remaining question for the complete assignment is the sequential analysis, which depends almost entirely on the observation of nuclear Overhauser effects (NOE)[40,41] between adjacent amino acids. We carried out several one- and two-dimensional NOE experiments[27,28] to obtain the necessary information.

The relevant section of a phase-sensitive two-dimensional NOE experiment[29,30] is shown in Figure 4. There the appearance of cross-peaks between the two α protons at lowest field, previously identified as a proline (4.75 ppm) and a glutamine unit (4.66 ppm), and the $H_\delta(P)$ region of the spectrum can be seen clearly. Close proximity in space, and consequently a relatively strong NOE cross-peak, is only possible for $H_\alpha(P^3)$ and $H_\alpha(Q^2)$ to the $H_\delta/H_{\delta'}$ of P^4 and P^3, respectively, and thus $H_\alpha(P^3)$, $H_\alpha(Q^2)$, and, by exclusion, $H_\alpha(Q^5)$ are identified. It is noteworthy to mention at this point that the observation of these NOE relationships is in perfect agreement with the short T1 relaxation times of $H_\alpha(Q^2)$ and $H_\alpha(P^3)$ with respect to the other α protons (see Figure 1). Furthermore, from the phase of the cross-peaks relative to the diagonal peaks, it can be deduced that the NOE is positive. The distinction between $H_\alpha(P^1)$ and $H_\alpha(P^4)$ can be made by similar arguments although in this case the amide protons $NH(Q^2)$ and $NH(Q^5)$ must be taken into consideration. First, the 1-D spectrum of the pentapeptide at 16 mM concentration in water was recorded (Figure 2A). Then, by selective decoupling of $H_\alpha(Q^5)$, which could be assigned by the previously described experiments, $NH(Q^5)$ could be attributed to the low-field part of the "triplet" at 8.5 ppm (Figure 2B). Finally, the observation of NOE cross-peaks of moderate intensity between $NH(Q^2)$ and $H_\alpha(P^1)$ as well as between $NH(Q^5)$ and $H_\alpha(P^4)$ in a phase-sensitive NOESY experiment on the 16 mM water sample allowed the definitive assignment of the remaining two α protons. Consequently, all H_α resonances and, via the COSY and RELAY experiment, all β and



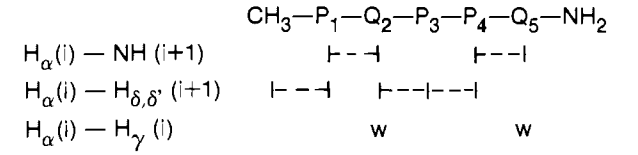


Figure 4. A section of a 300-MHz phase-sensitive NOESY experiment of P^1-Q^2-P^3-P^4-Q^5 in D_2O at 4 mM concentration, pH 2.5, and the corresponding one-dimensional ^1H spectrum. Diagonal peaks have a negative phase; off-diagonal elements are positive indicating positive NOE's. A summary of the observed NOE's is shown on the bottom. A dashed line stands for a medium (5–10%) strength NOE; w denotes weak NOE's (<5%).

γ resonances could be assigned. For the determination of the exact chemical shift values the J-resolved spectrum[42] was recorded. Its F2-projection, revealing the net chemical shift without the coupling constants, is shown in Figure 3 at the bottom, and the chemical shifts together with the coupling constants are summarized in Table II. It should be emphasized that the distinction between the δ protons of the three proline moieties is based (i) on a moderate NOE cross-peak between the methyl group of the N-terminal acetyl unit and the intense, triplet-like structure at 3.67 ppm, and (ii) on the difference in chemical shifts of NOE cross-peaks between $H_\alpha(P^3)$ and $H_\alpha(Q^2)$ with respect to the corresponding H_δ resonances. Because of some limitations in the achievable signal-to-noise ratio in the two-dimensional NOE experiments, we measured the NOE relationship of each α proton by means of a one-dimensional multiple irradiation[17,18] experiment.

(40) Noggle, J. H.; Schirmer, R. E. *The Nuclear Overhauser Effect*; Academic Press: New York, 1971.
(41) Sanders, J. K. M.; Mersh, J. D. *Prog. NMR Spectrosc.* **1982**, *15*, 353–371.
(42) Aue, W. P.; Karhan, J.; Ernst, R. R. *J. Chem. Phys.* **1976**, *64*, 4226–4227.

Table II. ^1H Chemical Shifts,[a] Coupling Constants,[b] and ^{13}C Chemical Shifts[a] of P^1-Q^2-P^3-P^4-Q^5 in Aqueous Solutions[c]

	NH	H_α	H_β	$H_{\beta'}$	H_γ	$H_{\gamma'}$	H_δ	$H_{\delta'}$[f]	NH_2(Z/E)
P^1[d,e]		4.44	2.31	1.94	2.02		3.67		
Q^2	8.51	4.66	1.94	2.14	2.46				7.67/7.16
P^3		4.75	2.41	1.93	2.08		3.87	3.70	
P^4		4.47	2.36	1.97	2.08		3.86	3.64	
Q^5	8.53	4.31	2.01	2.16	2.44				7.67/7.16

	J_{NH-H_α}	$J_{H_\alpha-H_\beta}$	$J_{H_\alpha-H_{\beta'}}$	C_α	C_β	C_γ	C_δ
P^1		8.8	4.7	62.50	32.52	26.94	51.35
Q^2	7.0	9.5	4.8	53.41	29.03	33.38	N/A[g]
P^3		8.4	5.7	61.48	30.73	27.30	50.52
P^4		8.5	5.5	63.04	31.86	27.30	50.39
Q^5	7.0	9.3	5.1	55.57	29.61	33.83	N/A

[a] Chemical shifts in ppm relative to external 5% TSP measured in a separate sample under the same experimental conditions. [b] Coupling constants in Hz. [c] All spectra were measured at ambient temperature (297 ± 1 K). The concentration was 4 mM in D_2O and 16 mM in 80% H_2O/20% D_2O (for ^{13}C measurements and exchangeable protons). In both samples the pH was 2.5.[14] [d] The chemical shifts of *N*-acetyl are: 2.15 (^1H), 24.01, and 175.64 ppm (^{13}C). [e] For the cis isomer the ^1H data (ppm) are: $H_\alpha(P^1)$, 4.61, $J_{\alpha\beta}$ = 8.9 Hz, $J_{\alpha\beta'}$ = 3.4 Hz; $H_{\delta,\delta'}(P^1)$, 3.56; $NH(Q^2)$, 8.68, J_{NH-H_α} = 6.6 Hz; $H_\alpha(Q^2)$, 4.68, $J_{\alpha\beta}$ = 9.4 Hz, $J_{\alpha\beta'}$ = 5.0 Hz; CH_3, 2.02. ^{13}C data (ppm): CH_3, 24.61; $C_\alpha(P^1)$, 63.34; $C_\beta(P^1)$, 34.28; $C_\alpha(Q^2)$, 53.33; C=O (*N*-acetyl), 175.95. [f] H_δ and $H_{\delta'}$ may be interchanged. [g] Because of the low sample concentration these and the following carbonyl resonances could not be assigned without ambiguity: 180.77, 180.69, 178.66, 177.18, 177.01, 174.76, 173.67 ppm.

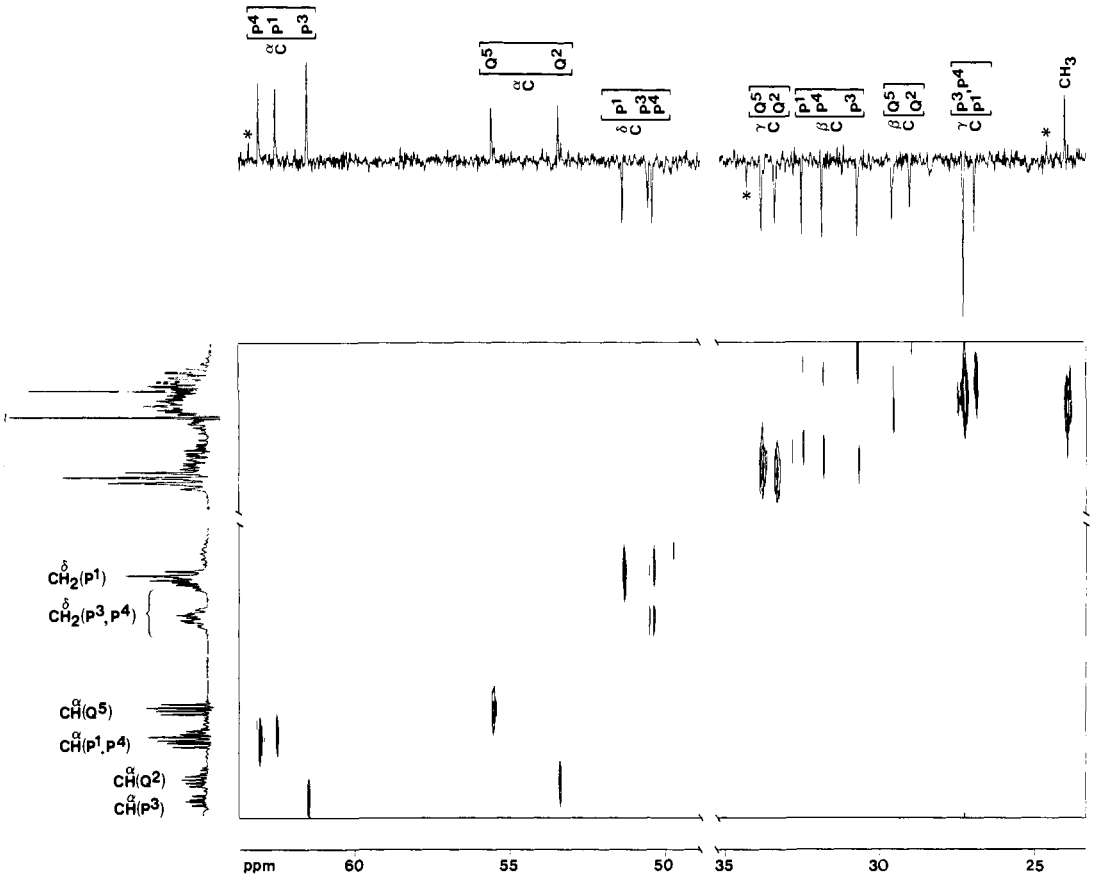


Figure 5. The contour plot of a $^{13}C/^1H$ shift correlated experiment of P^1-Q^2-P^3-P^4-Q^5 at 16 mM concentration in H_2O, pH 2.5. The one-dimensional proton spectrum is shown as well as (on top) the result of a DEPT experiment. CH and CH_3 groups are positive and CH_2 signals are negative in phase.

The results, summarized in Figure 4, show that the magnitudes of the NOE's bearing relevant structural and conformational information are in the range of 5 to 10% based on one proton.

The ^{13}C spectrum of P^1-Q^2-P^3-P^4-Q^5 was first recorded for the protonated carbons by means of a DEPT[19] pulse sequence applied to the 16 mM H_2O sample. The result is shown at the top of Figure 5. CH and CH_3 signals are positive; CH_2 groups are identified by negative peaks. For an unambiguous assignment of the ^{13}C resonances, the sample was then subjected to a $^{13}C/^1H$ two-dimensional shift correlation experiment[34] using the DEPT sequence for the polarization transfer from ^1H to ^{13}C. The experiment was performed with a delay time that was found suitable in a similar experiment on a tetrapeptide.[43] The spectrum is shown in Figure 5. Basically, all ^{13}C resonances can be assigned with ease although for some of the weaker signals an examination of the appropriate cross section parallel to the proton axis was required. The chemical shifts are summarized in Table II. The carbonyl resonances (six in the backbone plus two at the end of the glutamine side chains) were measured by a "q-only" sequence,[21] and an attempt was made for an unambiguous assignment by recording a COLOC[35] experiment. However, because of the low concentration of the sample together with the inherently low sensitivity of carbonyl resonances, only a few signals could be assigned without ambiguity.

The circular dichroism spectrum of the pentapeptide (Figure 6) showed only a single feature, a minimum near 204 nm. Heating the peptide in water (5 to 65 °C), adding methanol (to a final concentration of 10 to 100% (v/v)) or trifluoroethanol (80%), had

(43) Otter, A.; Kotovych, G. *J. Magn. Reson.*, in press.

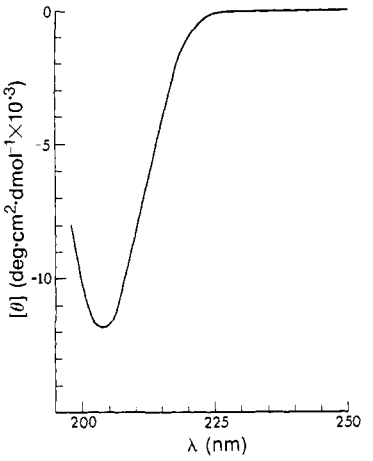


Figure 6. The circular dichroism spectrum (average of two scans) of the pentapeptide in water at room temperature.

no effect on the spectrum. Urea (5 M) likewise had no effect, at least over the accessible wavelength range (205–250 nm). These data indicate that the secondary structure of the peptide is totally insensitive to the environment. Consequently we did not consider it worthwhile to record NMR spectra on samples in nonaqueous solvents.

Discussion

The desirable conformational information for an oligopeptide of the size of P^1-Q^2-P^3-P^4-Q^5 is the determination of torsion angles[44] in the peptide backbone and the evaluation of side-chain rotamer populations. The combination of nuclear Overhauser enhancements and coupling constants allow an almost complete definition of all of the relevant rotational angles.

First we consider the determination of the backbone related angles ϕ, ψ, and ω. The torsion angle ϕ, describing the rotation about the N–C_α bond, is related to the dihedral angle θ, which itself is related to the NH–H_α coupling constant by the Karplus equation.[45] Despite the various sets of constants for this equation,[46] we find a narrow limit for θ by using four different sets of Karplus constants.[39] Therefore, we conclude that ϕ in Q^2 and Q^5 is either $-78° < \phi < -85°$ or $-155° < \phi < -162°$. The former values can be excluded on the base of NOE data as NH(Q) and H_β(Q) would come very close to each other in the case of a ϕ angle of ca. −80° and therefore would exhibit an NOE which was clearly not observed. The ψ angles cannot be derived from ^{1}H coupling constants,[47] but they can be qualitatively well described when considering the $H_\alpha(i)$ – NH(i + 1) and $H_\alpha(i)$ – $H_{\delta,\delta'}(i + 1)$ NOE's (Figure 4). With the help of a CPK model, it can be seen easily that with the existence of the above-mentioned NOE and the assumption that the peptide bonds are trans (ω = 180°), the C_α–H_α bond must be approximately trans to the carbonyl group. Consequently, the angle ψ is about +120° for the amino acid residues P^1 to P^4. ψ of Q^5 is not defined as there is no spectroscopic evidence to prove a fixed conformation about ψ at the C-terminal end of the pentapeptide. Finally, the peptide bond angle ω has to be considered. Because of the partial double-bond character, ω is in general either 180° (trans) or 0° (cis), with the former one being preferred. However, with the presence of three proline moieties, which can undergo cis/trans isomerization quite easily,[38,43,48] the "all-trans" assumption must be examined carefully. The presence of two (or more) cis/trans pairs would be observable by doubling of resonance lines, particularly in the ^{13}C domain.[39,49]

All our NMR data show clearly that the peptide bond linking the N-terminal acetyl group and P^1 exists in a cis/trans equilibrium (ca. 75% trans, 25% cis), but for all the remaining peptide bonds there is no indication of such an isomerization. The effect of cis/trans isomerization on the chemical shifts decreases strongly with increasing distance along the peptide chain. Only for the acetyl group itself and the resonances of P^1 a measurable chemical shift difference in both isomeric forms is observable (Table II). This, together with the 3:1 ratio in favor of the trans isomer, is very similar to the result found in a R^1-P^2-K^3-P^4 tetrapeptide.[37,38]

The side-chain rotamer populations can be derived from the H_α–H_β and H_α–$H_{\beta'}$ coupling constants (Table II) in terms of the populations of the three classical rotamers about the C_α–C_β bond.[39] Using parameters as described in the literature,[50] the relative populations for Q^2 are P_I = 0.09, P_{II} = 0.77, and P_{III} = 0.14. For Q^5 the values are P_I = 0.13, P_{II} = 0.75, and P_{III} = 0.12. In both cases one rotamer is highly dominating but this is not unequivocal evidence for the existence of a single predominant conformation.[39] This is further corroborated by the fact that the pair of coupling constants ($J_{\alpha\beta}$, $J_{\alpha\beta'}$) does not differ by θ = 120° when analyzed by the Karplus equation.[45,46] Therefore, reduced mobility rather than a fixed side-chain conformation reflects the spectroscopic evidence best. Precisely the same can be said with respect to the H_γ(Q) protons where a weak NOE to the H_α signal is observed, indicating that the C_β–C_γ bond is not rotating freely and χ^1 and χ^2 are approximately −120 and 180°, respectively, in the predominating rotamer. With respect to the proline units the situation is more complicated. Because the seven spins (H_α to $H_{\delta,\delta'}$) of each proline ring represent a system of higher order (e.g., $J_{\beta\gamma} \neq J_{\beta\gamma'}$), an analysis would only be possible by a complete spectral simulation for each proline moiety. However, in order to achieve this, a vast majority of the proline lines would have to be accessible for analysis. Owing to the complexity of the spectrum, especially at high field, this is clearly not the case and, consequently, the conformations for the three proline rings remain undefined.

The final consideration with respect to the conformation of the pentapeptide is the evaluation of possible hydrogen bonds between NH and carbonyl groups. From a study of the temperature dependence of NH resonances (data not shown), we conclude that such interaction is not present. The values for the chemical shifts as a function of temperature are all within 8 to 12 × 10^{-3} ppm/T. These are very typical effects for freely accessible protons which are neither hydrogen bonded nor otherwise shielded from an interaction with the solvent.[39,51]

The conformation consistent with the above data can best be described as extended: ϕ and ψ angles for Q^2, the ϕ angle for Q^5, and the ψ angles for P^1, P^3, and P^4 are all close to those of the antiparallel pleated sheet of β-poly(L-alanine).[52] Earlier[53] and more recent[54] applications of secondary structure prediction algorithms have suggested that the region of the C-telopeptide incorporating the pentapeptide studied here would adopt a β-turn conformation. Clearly the results from such probability-based methods must be viewed with caution.

Conclusion

By the combined application of various one- and two-dimensional NMR techniques, all ^{1}H and proton-bearing ^{13}C resonances of the pentapeptide P^1-Q^2-P^3-P^4-Q^5 could be unambiguously assigned. Under the given experimental conditions (aqueous solutions, ph 2.5, room temperature), it exists as a 3:1 mixture of trans:cis isomer with the isomerization site at the acetyl–P_1 amide bond. All other peptide bonds are trans (ω = 180°) and the ϕ angles for both glutamine moieties are ca. −160°. Based on several NOE's the angle ψ can be estimated to be ca. +120° throughout the peptide backbone with the exception of ψ in Q^5 where the angle

(44) IUPAC–IUB Commission on Biochemical Nomenclature *Biochemistry* **1970**, *9*, 3471–3479.
(45) Karplus, M. *J. Am. Chem. Soc.* **1963**, *85*, 2870–2871.
(46) Becker, E. L.; Bleich, H. E.; Day, A. R.; Freer, R. J.; Glasel, J. A.; Visintainer, J. *Biochemistry* **1979**, *18*, 4656–4668.
(47) Bystrov, V. F. *Prog. NMR Spectrosc.* **1976**, *10*, 41–81.
(48) London, R. E.; Matwiyoff, N. A.; Stewart, J. M.; Cann, J. R. *Biochemistry* **1978**, *17*, 2277–2283.
(49) Deslauriers, R.; Smith, I. C. P. *Topics in Carbon-13 NMR Spectroscopy*; Levy, G. C., Ed.; Wiley-Interscience: New York, 1976; pp 2–66.
(50) Feeney, J. *J. Magn. Reson.* **1976**, *21*, 473–478.
(51) Sahal, D.; Balaram, P. *Biochemistry* **1986**, *25*, 6004–6013.
(52) Arnott, S.; Dover, S. D.; Elliott, A. *J. Mol. Biol.* **1967**, *30*, 201–208.
(53) Helseth, D. L., Jr.; Veis, A. *Chemistry and Biology of Mineralized Connective Tissues*; Veis, A., Ed.; Elsevier/North-Holland, New York, 1981; pp 85–92.
(54) Dion, A. S.; Myers, J. C. *J. Mol. Biol.* **1987**, *193*, 127–143.

is not derivable from our data. The evaluation of side-chain rotamer populations in both glutamine units reveals that there is a reduced mobility (one rotamer predominating), but there is no fixed side-chain conformation present. The existence of hydrogen bonds can be excluded based on the high-temperature dependence of the exchangeable protons. In summary, all spectroscopic data (that is NMR and circular dichroism) indicate clearly that the pentapeptide exists in a nonrandom, highly organized conformation that is unusual for a peptide of that size.

Acknowledgment. The authors gratefully acknowledge the financial assistance of the Natural Sciences and Engineering Research Council of Canada, the Canadian Medical Research Council, The Central Research Fund of the University of Alberta, and the Alberta Heritage Foundation for Medical Research (through a fellowship to A.O.).

Assignment of Bacteriochlorophyll *a* Ligation State from Absorption and Resonance Raman Spectra

Patricia M. Callahan and Therese M. Cotton*

Contribution from the Department of Chemistry, University of Nebraska—Lincoln, Lincoln, Nebraska 68588-0304. *Received April 6, 1987*

Abstract: Absorption and Soret excitation resonance Raman (RR) spectra have been obtained for a series of coordination forms of monomeric bacteriochlorophyll *a* (BChl *a*). Strong and moderate intensity bands are observed in the RR spectrum at 1609 and 1530 cm^{-1} for five-coordinate species, which shift to 1595 and 1512 cm^{-1}, respectively, in the six-coordinate form. These coordination-sensitive vibrations are independent of the nature of the axial ligand and are suggested to have significant C_aC_m character, while several other less intense coordination-sensitive bands at 1463, 1444, and 1375 cm^{-1} are considered to arise from C_bC_b and C_aN stretching vibrations. These coordination-sensitive RR bands were used to determine BChl *a* ligation state in the solvents used, and structure correlations based on absorption maxima have been developed. The Q_x absorption band position is sensitive not only to BChl *a* Mg^{2+} coordination number but also to the nature of the axial ligand, i.e., oxygen, sulfur, or nitrogen. Q_x maxima are observed at 570, 575–580, and 582 nm for five-coordinate oxygen, sulfur, and nitrogen ligands, respectively, and at 590–595 and 605–612 nm, for six-coordinate oxygen and nitrogen species, respectively. The Q_y absorption maximum is insensitive to coordination number changes but is dependent on the nature of the axial ligand: 770 nm for oxygen ligand(s) and 775 nm for nitrogen ligand(s). A similar series of absorption and Soret excitation RR spectra were obtained for the demetalated form of BChl *a*, BPheo *a*. The Q_x maxima and RR spectra are insensitive to the solvents used, thereby confirming that coordination changes are responsible for the observed BChl *a* spectral shifts. The structure correlations for monomeric BChl *a* established here, based on the absorption and RR data, are discussed with reference to earlier molecular orbital calculations.

The distribution and function of bacteriochlorophyll *a* (BChl *a*) in bacterial photosynthesis range from a water-soluble antenna protein in the green bacterium *Prosthecochlorus aestuarii* to membrane-bound antenna proteins in the purple photosynthetic bacteria, e.g., *Rhodobacter sphaeroides* and *Rhodospirillum rubrum*, and to the reaction center (RC) pigments of green and purple bacteria.[1-3] The structure of BChl *a*, shown in Figure 1, is a tetrahydroporphyrin with a condensed Vth ring and with conjugated ketone functionalities at position 2 (acetyl) and position 9 (ketyl). The coordination sphere of Mg^{2+} is not satisfied by the four pyrrole nitrogens, and a fifth ligand is always present, either as an exogenous ligand in monomeric systems or as the C_2 acetyl or C_9 ketyl groups of a neighboring BChl *a* molecule in self-aggregates. In these cases the central Mg^{2+} remains out of the plane of the pyrrole nitrogens. A sixth ligand, of sufficient base strength, can be added to the coordination sphere which forces the central metal down into the plane of the bacteriochlorin ring.

The absorption spectrum of BChl *a* is characterized by an intense Soret band in the near-UV composed of B_x and B_y components, a band of moderate intensity in the visible region, Q_x, and another intense band in the near-infrared region of the spectrum, identified as Q_y. Both of the Q bands have vibrational overtones at approximately 1300 cm^{-1} to higher energy of the 0–0 electronic transitions. Several molecular orbital calculations of the porphyrins and (bacterio)chlorins have been carried out[4-7] and their electronic energy levels can be described by using the four-orbital model of Gouterman.[8,9] The Q_y and Q_x transition dipoles are oriented along the unsaturated rings I and III and the reduced rings II and IV,[10,11] respectively, as shown in Figure 1.

In its different protein environments the Q_y band of BChl *a* has been observed to range from 800 to 917 nm as a result of environment and exciton effects.[12,13] However, an exact description of very few systems has been made at this time. In recent calculations of the reaction center pigments of *Rh. viridis* and *Rb. sphaeroides* an ad hoc red shift of the Q_y band of monomeric BChl *a* from 770 to 807 nm for accessory BChl *a* and to 837 nm for the "special pair" P870 was needed in order to fit the exciton calculations to the known crystal structure coordinates of the pigments.[14] In this paper we are concerned with the spectral shifts

(1) Thornber, J. P.; Trosper, T. L.; Strouse, C. E. In *Photosynthetic Bacteria*; Clayton, R. K., Sistrom, W. R., Eds.; Plenum: New York, 1978; pp 133–160.
(2) Wraight, C. A. In *Photosynthesis: Energy Conversion by Plants and* Bacteria; Govindjee, Ed.; Academic: New York, 1982; Vol. 1, pp 19–61.
(3) Okamura, M. Y.; Feher, G.; Nelson, N. In *Photosynthesis: Energy Conversion by Plants and Bacteria*; Govindjee, Ed.; Academic: New York, 1982; Vol. 1, pp 197–272.
(4) Seely, G. R. *J. Chem. Phys.* **1957**, *27*, 125–133.
(5) Weiss, C. Jr. *J. Mol. Spectrosc.* **1972**, *44*, 37–80.
(6) Petke, J. D.; Maggiora, G. M.; Shipman, L.; Christofferson, R. E. *Photochem. Photobiol.* **1979**, *30*, 203–223.
(7) Petke, J. D.; Maggiora, G. M.; Shipman, L. L.; Christofferson, R. E. *Photochem. Photobiol.* **1980**, *32*, 399–414.
(8) Gouterman, M. *J. Chem. Phys.* **1959**, *30*, 1139–1161.
(9) Gouterman, M. *J. Mol. Spectrosc.* **1961**, *6*, 138–163.
(10) Goedheer, J. C. In *The Chlorophylls*; Vernon, L. P., Seely, G. R., Eds.; Academic: New York, 1966; pp 147–184.
(11) Moog, R. S.; Kuki, A.; Fayer, M. D.; Boxer, S. G. *Biochemistry* **1984**, *23*, 1564–1571.
(12) Pearlstein, R. In *Photosynthesis: Energy Conversion by Plants and Bacteria*; Govindjee, Ed.; Academic: New York, 1982; Vol. 1, pp 293–330.
(13) Garcia, D.; Parot, P.; Vermeglio, A.; Madigan, M. T. *Biochim. Biophys. Acta* **1986**, *850*, 390–395.

0002-7863/87/1509-7001$01.50/0