Secondary Structure of Acyl Carrier Protein As Derived from Two-Dimensional ^1H NMR Spectroscopy†

T. A. Holak and J. H. Prestegard*
Department of Chemistry, Yale University, New Haven, Connecticut 06511
Received March 12, 1986; Revised Manuscript Received May 5, 1986

ABSTRACT: Sequence-specific assignments of ^1H NMR resonances were obtained for the backbone protons in acyl carrier protein (ACP) from *Escherichia coli*, a protein of 77 residues. The observations, in the NOESY spectra, of ^1H–^1H sequential and medium-range connectivities indicate the presence of three or four α-helical segments joined by short sequences of mixed conformations. The observations are used to refine a secondary structure model previously proposed on the basis of a Chou–Fasman algorithm [Rock, C. O., & Cronan, J. E., Jr. (1979) *J. Biol. Chem. 254*, 9778–9785].

Acyl carrier proteins (ACPs)[1] are at the heart of a system responsible for fatty acid biosynthesis and, to some extent, the acyl chain distribution found in membrane lipids and fatty deposits. This arises because acyl ACPs function as substrates for the fatty acid synthetase complex, where they carry fatty acids during elongation by successive addition of two carbon fragments from malonyl CoA (Thompson, 1981; Wakil et al., 1983). A possible relationship between the structure of various acyl ACPs and their reactivity in various steps of fatty acid biosynthesis has stimulated an interest in a more complete structural characterization.

Among all known ACPs, ACP from *Escherichia coli* is perhaps the best characterized (Prescott & Vagelos, 1972; Vanaman et al., 1968a,b). ACP from *E. coli* is a soluble protein of 8847 molecular weight. It contains 77 amino acid residues, a large proportion being acidic (ca. 27%) and a small proportion (ca. 8%), predominantly clustered at the NH terminus, being positively charged. The amino acid sequence of strain E-26 is known (Vanaman et al., 1968a,b), and the 4′-phosphopantetheine prosthetic group, through which fatty acids are covalently attached, is known to be linked to Ser-36.

No crystal structure of ACP yet exists, although crystals giving suitable diffraction patterns have been obtained (McRee et al., 1985). From optical rotary dispersion (Takagi & Tanford, 1968; Prescott et al., 1969) and circular dichroism studies (Schultz, 1975), ACP seems to possess a high α-helix content. Rock and Cronan (1979) have applied the predictive algorithm of Chou and Fasman (1978a,b) to ACP, yielding an approximate model of the secondary structure. This model was refined and a tertiary structure of ACP was proposed on the basis of the observed NOE cross relaxation rates from one-dimensional NMR spectra (Mayo et al., 1983). Few of the resonance assignments could, however, be regarded as definitive in this early work, and the model should be viewed as a hypothetical model consistent with a limited set of observations.

In the present study a complete sequence-specific assignment of the ^1H NMR resonances for the backbone protons of ACP is obtained by using two-dimensional NMR methods. These methods were introduced by Wüthrich and co-workers (Wüthrich et al., 1982; Wagner et al., 1981; Billeter et al., 1982; Wüthrich, 1983) and have been applied to a number of protein structure determination problems. ACP, due to its size and large number of nearly degenerate complex spin systems (e.g., 14 Glu), stands at the edge of feasibility for these methods. In most cases applications have been made to smaller proteins (Wagner et al., 1981; Wagner & Wüthrich, 1982; Štrop et al., 1983; Zuiderweg et al., 1983a; Neuhaus et al., 1985; Weber et al., 1985) or to proteins where partial assignments and structure determinations are deemed useful (Wand & Englander, 1985; Boyd et al., 1985). Recently a protein with 74 residues containing extended strands of β-sheets has been analyzed by using such methods (Kline & Wüthrich, 1985).

On the basis of the sequential assignments of ACP resonances it is possible to precisely characterize the secondary structure of this protein. Herein we present this characterization. Since there are some variations from the Chou–Fasman predictions, it also appears that the previously proposed tertiary structure model will undergo significant revision.

Materials and Methods

ACP was isolated from *E. coli* B cells (Grain Processing) by using the method of Rock and Cronan (1980) and reduced with dithiothreitol (DTT) to the free sulfhydryl form. ACP samples were passed through a short Chelex column, followed by a Sephadex G-25 column (5 mM potassium phosphate buffer) and freeze-dried. HPLC and NMR analyses showed these samples to contain approximately 10% minor isostructural forms of ACP. The samples were deemed sufficiently pure for NMR analysis since it was felt that the 10% contaminants were unlikely to give observable resonances. This has proven to be largely correct. There are, however, a few weak cross-peaks present in the amide–$C^\alpha H$ region of the

† This work was supported by a grant from the National Institutes of Health (GM32243) and benefitted from instrumentation provided through shared instrumentation programs of the National Institute of General Medical Science (GM32243S1) and the Division of Research Resources of NIH (RR02379).

[1] Abbreviations: ACP, acyl carrier protein; DTT, dithiothreitol; β-EA-SH, β-mercaptoethylamine; 2D NMR, two-dimensional nuclear magnetic resonance; COSY, two-dimensional *J*-correlated NMR spectroscopy; DQF-COSY, double quantum filtered COSY; RELAYED-COSY (RELAY), relayed coherence transfer spectroscopy; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; ppm, parts per million; $d_{\alpha N}(i,j)$, cross relaxation connectivity between the $C^\alpha H$ proton on residue *i* and the amide proton on residue *j*; $d_{\beta N}(i,j)$, cross relaxation connectivity between a $C^\beta H$ proton on residue *i* and the amide proton on residue *j*; $d_{NN}(i,j)$, cross relaxation connectivity between the amide proton on residue *i* and the amide proton on residue *j*.

0006-2960/86/0425-5766$01.50/0

COSY spectrum to be discussed shortly (Figure 1) that are assignable to minor species. Those peaks close to the very intense peak of the terminal β-mercaptoethylamine (β-EA-SH) of the prosthetic group, for example, could be identified unambiguously as originating from β-EA-SHs of the minor forms because of the unique multiplet structure of the β-EA-SH cross-peaks. The origin of the weak peak at f_2 = 8.69 ppm, f_1 = 4.49 ppm (between Asn-73 and Ala-34) is not specifically known but could arise from minor forms. It is worth mentioning that the NOESY data sets showed no corresponding connectivities in this region, nor any connectivities in parts of spectra which could have been connected to these regions (cf. Figure 2).

Samples for NMR measurements were redissolved in D_2O or H_2O/D_2O (9:1) immediately before the NMR experiments. The samples were run at a final protein concentration of ~10 mM in about 250 mM potassium phosphate buffer. At this concentration samples held at room temperature for the times required for 2D NMR study are susceptible to oxidation and degradation. DTT was added to the samples in order to prevent the oxidation of the SH group. The solubility of ACP in water decreases dramatically at pH values lower than 5.2. Near this value reasonable spectra could only be obtained at a temperature of 30 °C. Above 35 °C ACP degrades too quickly and below 28 °C the line widths were broadened significantly. Thus, all data were acquired at 30 °C. By the use of an ACP sample freshly prepared in D_2O near the low pH limit (5.4), a qualitative identification of slowly exchanging amides was obtained by immediately acquiring a COSY spectrum. The retention of some amide proton signals after approximately 30 h of acquisition was used to give support for the location of helical segments in the secondary structure (Wagner & Wüthrich, 1982b; Zuiderweg et al., 1983b).

For sequential assignments, COSY (Aue et al., 1976; Nagayama et al., 1980; Bax & Freeman, 1981; Wider et al., 1984), NOESY (Jeener et al., 1979; Anil Kumar et al., 1980; Wider et al., 1984), and RELAYED-COSY (Eich et al., 1982; Bax & Drobny, 1985) experiments were performed. The NOESY spectra were acquired in phase-sensitive mode by using the method of States (States et al., 1982). NOE mixing times, normally 140–180 ms, were randomly varied by 10% to eliminate coherent transfers. Recycling times were 1.4 s. There are three regions in the ACP spectrum (Figure 1) for which there are 4 or 5 residues which have similar NH chemical shifts. These regions, at 8.5, 8.2, and 7.8 ppm, were the most difficult to assign, and only a careful comparison of three NOESY spectra run at three different pHs, 5.6, 6.1, and 7.1, allowed the assignment of the NH–C^αH and/or NH–NH cross-peaks.

The double quantum filter (DQF) COSY experiment at pH 6.1 was recorded in the phase-sensitive mode by using the method of States et al. (States et al., 1982; Rance et al., 1983). The COSY experiments at other pHs were obtained as absolute magnitude spectra. RELAY experiments were obtained as absolute magnitude spectra (Bax & Drobny, 1985) with mixing times of 32–50 ms. All spectra were acquired on a Bruker WM-500 spectrometer or at 490 MHz on a home-built spectrometer, using quadrature detection. The majority of the two-dimensional spectra were collected as 450–550 t_1 experiments, each with 2K complex data points over a spectral width of 5.4 kHz in both dimensions, with the carrier placed on the water resonance. For the DQF-COSY in H_2O and D_2O, the frequency range along f_1 was contracted to improve digital resolution (3 kHz), hence part of the data were folded over. For samples dissolved in H_2O, the water signal was irradiated for 1.2 s prior to the start of the pulse sequence and only allowed to recover during signal acquisition in t_2. Signal averaging times varied from 30 h for some COSY spectra to 70 h for the NOESY spectra.

The data were transferred to a VAX 11/750 computer equipped with a CSPI Minimap array processor. Data sets were multiplied in both dimensions by sine-bell functions or 10° phase-shifted sine-bell functions (extending to 500 t_1 points) and zero-filled to 2K in the t_1 dimension before Fourier transformation. Processing, display, and plotting routines were written by Dr. Dennis Hare.

Results

Identification of Spin Systems of Amino Acid Side-Chain Protons. Spin systems for several amino acid types could be unambiguously identified and connected to their appropriate cross-peaks in the amide proton–α proton (NH–C^αH) region of two-dimensional spectra after an inspection of COSY, NOESY, and RELAY spectra run in D_2O and H_2O. These amino acids included seven alanines, seven valines, six threonines, seven isoleucines, four glycines, three serines, two phenylalanines, one tyrosine, one histidine, and five of the seven glutamines/asparagines found in ACP. Beginning with the examination of the RELAY experiment proved particularly useful because it provided direct connectivities between well-resolved methyl peaks and C^αH or amide resonances for many residues. Spin systems assigned on the basis of RELAY experiments included isoleucines [$C^\gamma H_3$–C^βH–C^αH fragment], valines, alanines, and five of the six threonines. In addition, several AMXY systems were identified, mostly those of aspartic acids.

Four glycines could be easily identified from the amide–C^αH region of the COSY spectrum in H_2O, even before the αα connectivities characteristic of glycines were found. The fingerprint region of the COSY spectrum in Figure 1 shows characteristic NH–C^αH patterns for glycines, such as that for Gly-74, along with a number of other assigned cross-peaks. The groups of eight peaks connecting the glycine amide to each of the two C^αH protons are particularly well resolved because of the fortuitous cancellation of minor peaks in the f_2 dimension and the particularly large C^αH–C^αH coupling in the f_1 dimension.

Figure 1 also shows some cross-peaks from one of the unique groups in ACP, the phosphopantetheine prosthetic group. This group contains β-mercaptoethylamine and β-alanine. Amide to methylene cross-peaks appear in the upper part of Figure 1 at 8.11, 3.27 and 7.96, 3.52 ppm. These resonances were actually first assigned from a multiple quantum spectrum taken in H_2O (data not shown) by using a procedure described by Prestegard and Scarsdale (1985). Such spectra give cross-peaks for methylenes coupled to amides. This grouping, of course, occurs for glycines as well, but the prosthetic group resonances can be easily distinguished by their additional coupling to γ methylenes.

Additional assignments of residues in the prosthetic group were facilitated by the fact that these residues exhibit very sharp NOE cross-peaks and very intense and distinctive patterns of couplings in the COSY C^βH–NH and RELAY C^γH–NH cross-peaks. The identification of one of the protons of the CH_2 group of 4-phosphopantoic acid (and of C^βHs of Ser-36) was confirmed with a one-dimensional double-resonance ^{31}P–^1H experiment in which the ^{31}P resonance of 4-phosphopantoic acid was irradiated (data not shown).

Figure 2 shows a region of a NOESY spectrum corresponding to the COSY spectrum in Figure 1. The number of cross-peaks far exceeds the number in the COSY spectrum

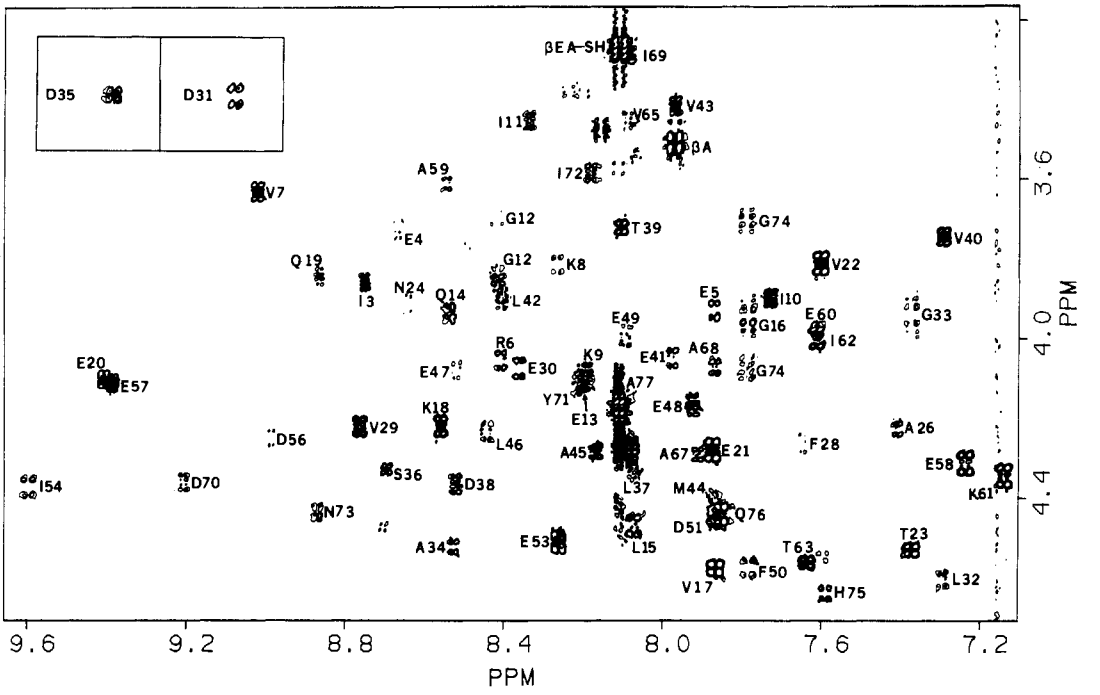

FIGURE 1: Absorption mode ^1H DQF COSY spectrum (490 MHz) of ACP in 90% H_2O and 10% D_2O, pH 6.1. A contour plot of the region shown contains most of the $C^{\alpha}H$–NH proton connectivities; the cross-peaks for Asp-31 and Asp-35 were moved to the inset, and the chemical shifts for these cross-peaks are given in Table I. Spectra are plotted with the negative and positive components of the cross-peaks at a level slightly above noise; as a result a weak cross-peak of Gln-66 is not visible in the plot. In addition, cross-peaks from Thr-52, Asn-25, Ser-27, and Thr-64 were lost because of water irradiation. The cross-peaks for Ser-27, Asn-25, and Thr-52 were observed in a magnitude COSY because slightly lower decoupling power was used to saturate the water resonance.

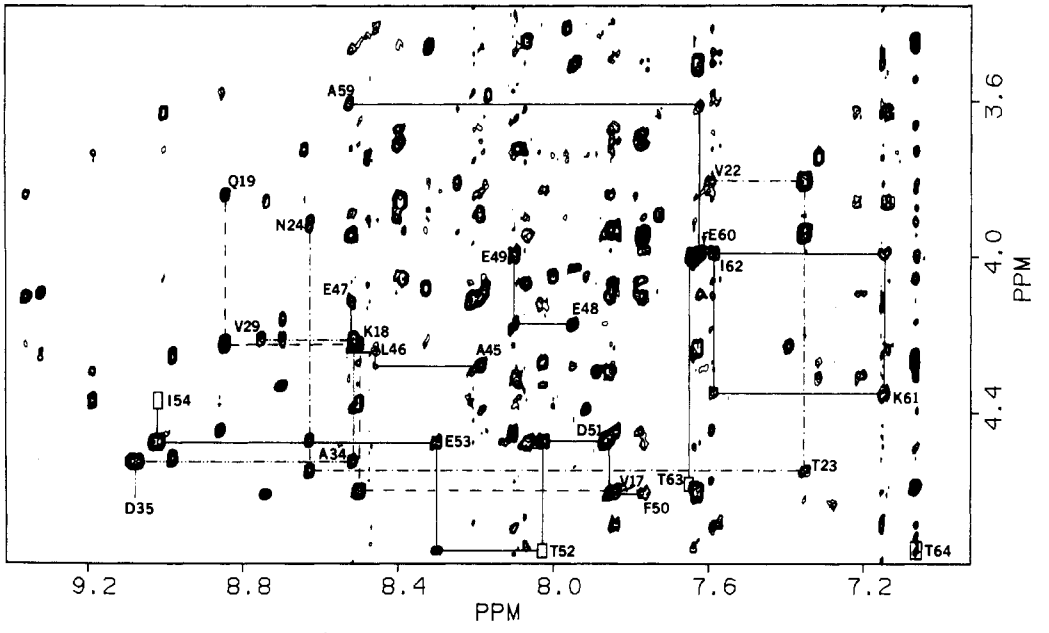

FIGURE 2: Region of a pure absorption 500-MHz ^1H NOE spectrum of ACP, pH 5.6, containing the amide–$C^{\alpha}H$ proton connectivities. The spectrum was recorded with a mixing time of 170 ms. Only intraresidue $C^{\alpha}H$–NH NOEs, as identified from their coincidence with peaks in a COSY spectrum, are labeled. The sequential interresidue $C^{\alpha}H$–NH NOEs are indicated by unlabeled peaks at the intersection of horizontal and vertical lines. Pathways of sequential $C^{\alpha}H$–NH NOEs are indicated for the peptide fragments 17–19 (---), 22–24 (-·-), 45–54 (—), and 59–63 (—). In addition a long-range 29-34-35 connectivity is shown (-··-).

because of the presence of both intra- and interresidue peaks. Comparison of COSY and NOESY spectra, such as those presented in Figures 1 and 2, is useful in distinguishing intra- and interresidue NOEs, as well as in assigning amino acid spin systems in cases where connectivities may be weak or absent in one set but strong and well resolved in the other.

Five glutamine/asparagine spin systems were assigned by comparing NOESY and COSY spectra in H_2O. Glutamine and asparagine are unique in that their side-chain amide protons show no coupling to $C^{\alpha}H$ or other aliphatic protons in a COSY spectrum. In NOESY data sets, however, they show the connectivities between the pairs of side-chain amide resonances and the $C^{\beta}H$ resonances of asparagines or $C^{\gamma}H$ protons of glutamines (Billeter et al., 1982). See, for example, columns labeled 24 and 25 at 7.14 and 6.97 ppm for asparagines 24 and 25, respectively, in Figure 3.

The appearance of ring proton–$C^{\beta}H$ proton connectivities in NOE spectra along with COSY spectra also allowed the

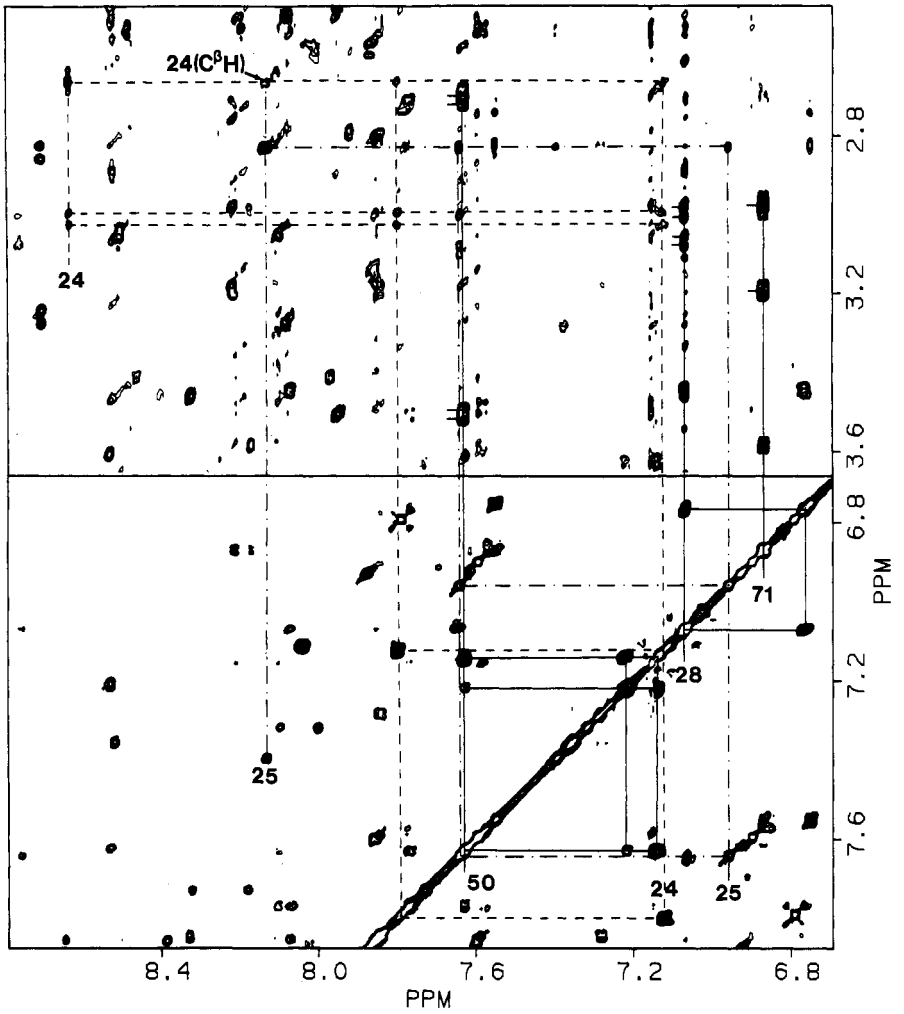

FIGURE 3: NOESY diagram (500 MHz, τ_m = 170 ms, pH 5.6) showing the connections of the side-chain amide resonances to the glutamine $C^\gamma H_2$ resonances or asparagine $C^\beta H_2$ resonances. Vertical lines are labeled according to their amide resonance assignments. NOE cross-peaks are identified for Asn-24 (---) and Asn-25 (-·-). In addition, connectivities between some aromatic proton resonances and side-chain resonances are shown (solid lines).

assignment of all the aromatic spin systems (Billeter et al., 1982). Some of these connectivities are shown in the NOESY data of Figure 3. Tyr-71 C^δHs at 6.88 ppm, for example, show cross-peaks to two resonances in the β region of the spectrum at 2.99 and 3.20 ppm. Examination of a COSY spectrum shows these to be spin-coupled to one another and to C^αH and NH resonances at shifts typical of an aromatic residue. Thus the entire spin system of the single tyrosine is assigned.

Arguments such as those presented above lead to assignment of about 60% of all amide, C^αH, and C^βH resonances to specific amino acid spin systems. At this point, the sequential assignment was started, and the remainder of the backbone resonances, along with additional side-chain resonances, were assigned during the sequential assignment by comparing the NH–C^βH regions of NOESY, RELAY, and COSY spectra.

Sequential Assignments of the Backbone Protons and Further Side-Chain Identification. The NOESY spectra recorded in H_2O with a τ_m value of 170 ms were used for the delineation of sequential backbone proton connectivities. The sequential resonance assignment of proteins by 2D NMR is now well established (Wüthrich et al., 1982; Wüthrich, 1983; Billeter et al., 1982). In this section only the highlights of the sequential assignments will be described. Assignments rely primarily on $d_{NN}(i, i \pm 1)$ and $d_{\alpha N}(i, i + 1)$ connectivities, but they were in a few cases supplemented with $d_{\beta N}(i, i + 1)$, $d_{\alpha N}(i, i + 3)$, and $d_{\alpha N}(i, i + 4)$ connectivities (Wüthrich et al., 1982, 1984; Zuiderweg et al., 1983b). In a few cases $d_{\alpha\beta}$ connectivities were observed, but they were not used in the sequential assignments.

The assignment procedure was usually started from the low-field region of the NOESY spectrum where NH–NH connectivities appear. On average, resonances belonging to segments of 3–5 sequential residues could be connected. When resonances for two or more of the residues had previously been assigned to a specific amino acid type, it was frequently possible to assign the intervening residues on the basis of the assumed analogy between the amino acid sequence for *E. coli* B and the published amino acid sequence for *E. coli* E-26. For example, a peptide segment six residues long, Ile-X-X-X-Val-X (where X stands for the amino acid residue for which the spin system had not been determined), was assigned to Ile-3–Glu-4–Glu-5–Arg-6–Val-7–Lys-8, since a segment of Ile-X-X-X-Val occurs only once in the sequence. Figure 4 illustrates some of the sequential connections established by using the d_{NN} connectivities. Nine regions have been connected by using primarily data of this type. These include Ile-3–Lys-8, Lys-9–Ile-11, Gly-12–Gly-16, Phe-28–Leu-32, Leu-37–Leu-42, Val-43–Glu-47, Glu-57–Asn-60, Ile-69–Asp-70, Ile-72–His-75.

Long stretches of moderately intense NH–NH connectivities, coupled with somewhat weaker NH–α connectivities to the same residues, are characteristic of the α-helical structures believed to dominate the ACP structure (Rock & Cronan, 1979). Examination of columns at the frequencies of NH protons involved in these networks frequently show additional

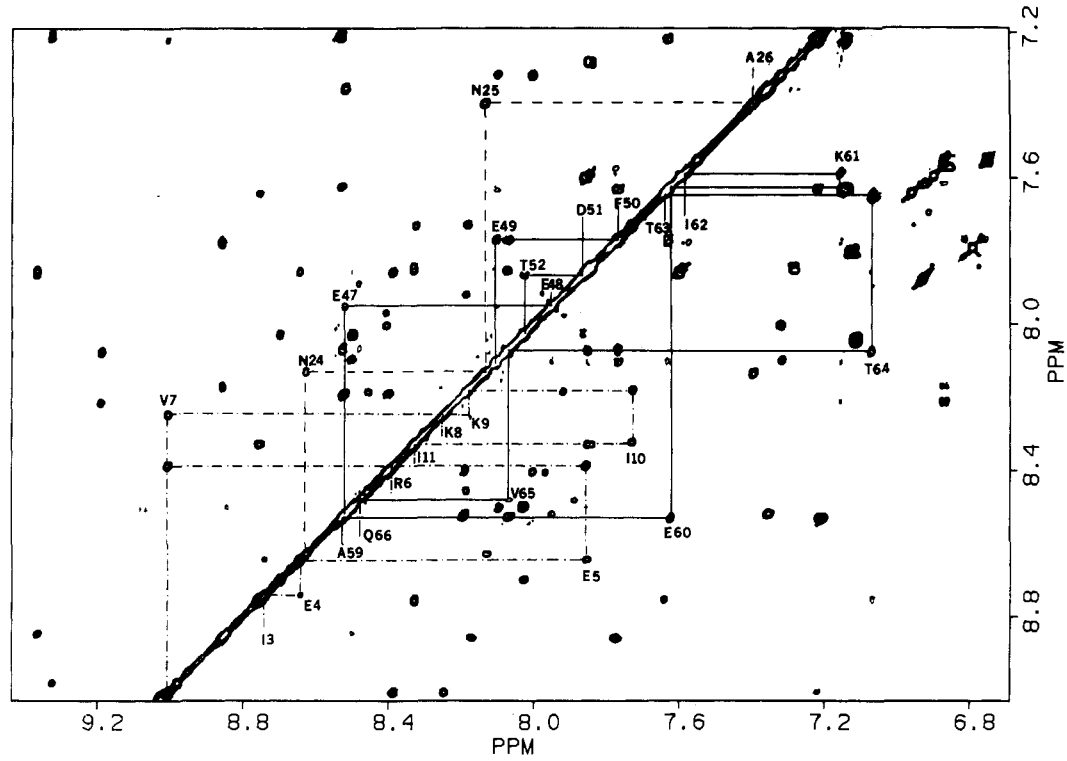

FIGURE 4: NH–NH cross-peak region in ACP; pH 5.6, spectrum symmetrized. The connectivities for fragments 3–11 (–·–), 24–26 (– – –), 47–52 (—, above diagonal), and 59–66 (—, below diagonal) are shown. Cross-peaks are labeled according to their amide resonance position in the f_2 dimension.

$d_{\alpha N}$ cross-peaks. $d_{\alpha N}(i, i + 3)$ cross-peaks are in fact expected for α-helical structures because of the 3.6 residue pitch of the helix. These additional peaks prove useful in confirming assignments, in connecting the above short segments, and in filling small gaps between the segments. For example, although there is no observable NH–NH cross-peak between Lys-8 and Lys-9 at pH 6.1 and 7.1, presumably because the cross-peaks are near the diagonal, the fragments Ile-3–Lys-8 and Lys-9–Ile-11 could be joined by $d_{\alpha N}(i, i + 3)$ connectivities between Val-7 and Ile-10 and between Lys-8 and Ile-11. The resonances for Gly-12 which could not be connected to Ile-11 by NH–NH cross-peaks at pH 5.6 and 7.1 could be assigned by the $d_{\alpha N}(i, i + 3)$ connectivity from Lys-9 to Gly-12 and the entire Lys-9–Ile-11 fragment could be connected to the Gly-12–Gly-16 fragment by an Ile-11 to Gln-14 connectivity. The use of $d_{\alpha N}(i, i + 3)$ cross-peaks is illustrated for the Ile-10 to Val-7 connectivity (most upfield) in Figure 5. A vertical line at the position of the amide resonance (7.7 ppm) shows connectivities to its own $C^{\alpha}H$, the $i - 1$ (Lys-9) and the $i - 3$ (Val-7) residue, strongly supporting the proposed assignments.

In addition, most residues that were assigned by NH–NH NOE connectivities could be connected by use of $d_{\beta N}(i, i + 1)$ connectivities. $d_{\alpha N}(i, i + 4)$ connectivities were also observed, but these cross-peaks were all of lower intensities than the $C^{\alpha}H(i)$–$NH(i + 1,3)$ cross peaks and were only seen in the plots with relatively low contour levels.

There are several regions where the assignments are particularly interesting or challenging. The region from Gly-16 to Asn-25 is interesting both because the structure is less regular and because there seems to be a variation in sequence between *E. coli* B and *E. coli* E-26. The connectivity for the backbone protons in segment Gly-16–Val-17–Lys-18–Gln-19 is characteristic of a β-turn. Strong $d_{\alpha N}(i, i + 1)$ and $d_{\beta N}(i, i + 1)$ connectivities are observed. These very strong $d_{\alpha N}$ type of connectivities are depicted in Figure 2 for the Val-17–Gln-19 segment. Starting with Gln-19, strong $d_{NN}(i, i + 1)$ connectivities allowed the continuation of the assignment to Val-22. The NH–NH connectivity pathway is broken at this point, but a strong NH–$C^{\alpha}H$ cross-peak leads to Thr-23 which in turn can be connected to residue 24 by $d_{\alpha N}(i, i + 1)$ and $d_{\beta N}(i, i + 1)$ connectivities (Figure 2). The connectivity $d_{\alpha N}(i, i + 1)$ between residue 24 and Asn-25 is missing, but a weak $C^{\beta}H$–NH cross-peak is observed (marked in Figure 3); the NH–NH connectivity is also present (Figure 4). The $C^{\alpha}H$ resonance of Asn-25 is close to the water line, and therefore the Asn-25–Ala-26 $C^{\alpha}H$–NH cross-peak is lost because of water irradiation. Nevertheless, $C^{\beta}H$–NH and NH–NH connectivities are present.

This leads to the sequence Gln-19–X-20–X-21–Val-22–Thr-23–X-24–Asn-25–Ala-26, for which identified residues are in agreement with those appearing in the sequence published for *E. coli* strain E-26. Spin systems for residues 20 and 21 agree with glutamic acids which occur in the sequence, but the spin system for X-24, which had been clearly identified as Asn because of connectivity to side-chain amides (Figure 3), does not agree with the aspartic acid appearing in the published sequence. Either the ACP isolated from strain B differs from that of *E. coli* strain E-26, by replacement of Asp-24 with Asn-24, or an Asn was misassigned to Asp.

The stretch from Ala-26 to Ser-36 is distinguished by some unusual long-range contacts between backbone protons. Ordinarily additional long-range contacts of this type could confuse sequential assignments, but in ACP this is the only case where such observations are made, and there are sufficient normal sequential connectivities to eliminate any ambiguity. Ala-26 could be connected to Ser-27 through a very intense $d_{\alpha N}(i, i + 1)$ connectivity, and the NH–NH cross-peak pathway leads the sequential assignment to Ala-34. The NH–NH cross-peak between Leu-32 and Gly-33 is very weak, but the $C^{\alpha}H$–NH connectivity is distinct. From Ala-34, a strong $d_{\alpha N}(i, i + 1)$ connectivity leads to fragment Ala-34–

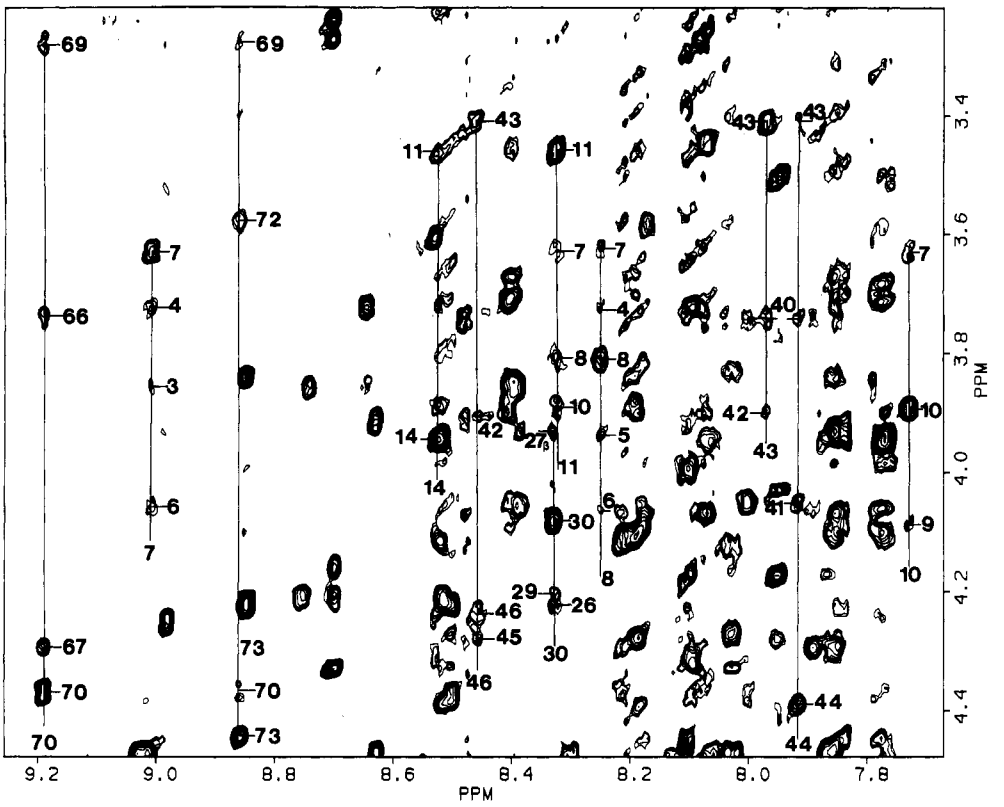

FIGURE 5: Spectral region showing medium-range NH–C$^{\alpha}$H proton cross-peaks from a phase-sensitive NOESY spectrum of ACP, pH 5.6, recorded with a mixing time of 170 ms. Peaks labeled include $d_{\alpha N}(i, i + 1)$, $d_{\alpha N}(i, i + 3)$, and $d_{\alpha N}(i, i + 4)$ interresidue NOEs and $d_{\alpha N}(i,i)$ intraresidue NOEs. Vertical lines are labeled according to their amide resonance position in the f_2 dimension.

Asp-35–Ser-36 (Figure 2). The unusual aspect is that Ala-34 also has a strong C$^{\alpha}$H–NH cross-peak to Val-29 (Figure 2). No C$^{\beta}$H–NH connectivity between these two residues is present. Additional connectivities between C$^{\gamma}$H$_3$ of Val-29 and NH of Ala-34 as well as NH of Glu-30 are exhibited, however. The structural implications of the presence of this rather unusual nonsequential connectivity, between Val-29 and Ala-34, will be deferred to the Discussion section.

Among the more difficult regions in the assignment were stretches from 47–51 and 61–64. These assignments are based on more complex arguments. In order to accurately represent the certainty of our assignments, we present these arguments in detail (Figures 2 and 4). Although the fragment Glu-60–Thr-64 contains three residues for which spin systems were identified (Ile, Thr, Thr), the sequential assignment posed some difficulties because of degeneracies in chemical shifts and the absence of amide cross-peaks to C$^{\alpha}$H resonances underlying water. The chemical shifts of C$^{\alpha}$H and NH protons of Glu-60 and Ile-62 are identical at pH 7.1 and 6.1. However, we were able to resolve the NH resonances at pH 5.6 (Figure 2). Even with this resolution, the amide proton resonance of Thr-63 remains close to the resonance line of the NH for Glu-60. The Thr-64 NH lies exactly on a line of aromatic protons of Phe-28, making it difficult to identify unique connectivities, and the C$^{\alpha}$H–NH cross-peak from Val-65 NH to Thr-64 C$^{\alpha}$H is eliminated due to water irradiation. Thus the normal redundancy of connectivities is not present. One unique pathway, however, was established to fit all the observed connectivities, e.g., C$^{\alpha}$H–NH connectivities from Glu-60 to Thr-63 and NH–NH connectivities from Glu-60 to Ile-62 and from Thr-63 to Val-65 (Figures 2 and 4). This leads to assignment of resonances for the segment Glu-60–Lys-61–Ile-62–Thr-63–Thr-64.

The connectivities in the fragment Glu-47–Asp-51 are also complicated. The NH–NH cross-peak between Leu-46 and Glu-47 is very weak but could be found in the NOESY spectrum at pH 5.6. The NH–NH cross-peak between Phe-50 and Asp-51 is missing in the NOESY spectrum at pH 5.6 and 7.1. At pH 6.1, a cross-peak is present at the right frequency for the NH–NH of Phe-50 and Asp-51. However, at this pH, the NH of residue 16 is degenerate in shift with the NH of Phe-50, and the NH of residue 17 is degenerate in shift with the NH of Asp-51. Thus, this cross-peak could have been assigned to a cross-peak connecting residues 16 and 17. Nevertheless, assuming the Asp-51 amide underlies the amide of Val-17 allows a connectivity for Asp-51–Thr-52 to be unambiguously assigned to the only NH–NH cross-peak not yet assigned, and the entire sequence is completed.

Figure 6 summarizes the sequential NOE connectivities observed in ACP. The chemical shifts of the specific resonance assignments are presented in Table I. These are given for pH 6.1, as at this pH, the highest quality COSY and RELAYED-COSY spectra were acquired. All C$^{\alpha}$H, C$^{\beta}$H, and amide assignments are given. Relatively few spin systems for long side-chain amino acids have been assigned at present. It should be possible to obtain many of these assignments from the existing data. However, NH, C$^{\alpha}$H, and C$^{\beta}$H assignments allow a nearly complete discussion of the secondary structure, and we choose to defer additional assignments to a paper dealing with tertiary structure.

Discussion

While there is no single observable parameter that would lead to a unique identification of a particular type of secondary structure in a protein, the combination of several types of NOE connectivities extending over several residues can be indicative of secondary structure (Billeter et al., 1982; Williamson et al., 1984; Wüthrich et al., 1984). The presence of a continuous succession of $d_{NN}(i, i \pm 1)$, $d_{\alpha N}(i, i + 1)$, and $d_{\alpha N}(i, i + 3)$ connectivities, for example, provides evidence that a peptide

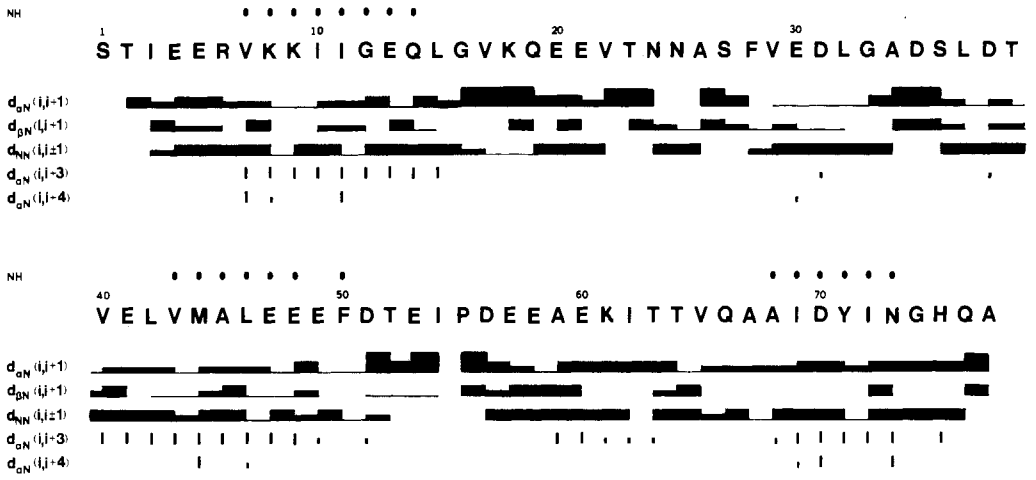

FIGURE 6: Amino acid sequence of ACP and survey of NMR data used for locating secondary structure. The data are compiled from NOESY spectra taken at pH 5.6 and 6.1. The NOEs, classified as weak, medium, strong, and very strong, are represented by the thickness of the bars. The bars for $d_{\alpha N}(i, i + 3,4)$ are positioned at the NHs of the $i + 3$ and $i + 4$ residues and connectivities are to the $C^{\alpha}H$ positions of a residue i. Filled circles above the sequence indicate slowly exchanging NHs.

Table I: Sequence-Specific Resonance Assignments of ACP, pH 6.1[a]

residue	NH	$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$ and others	residue	NH	$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$ and others
Ser-1	n.o.[b]	4.44	3.93, 4.11		Glu-41	7.97	4.07	2.12, 2.17	
Thr-2	9.00[c]	4.61	4.80	1.40	Leu-42	8.40	3.91	1.52, 1.71	
Ile-3	8.74	3.87	2.04	$C^{\gamma}H_3$ 0.73	Val-43	7.96	3.42	2.25	0.91, 1.04
Glu-4	8.66	3.73	1.98, 2.23		Met-44	7.86	4.42	2.18, 2.20	2.80
Glu-5	7.86	3.94	2.30, 2.30		Ala-45	8.16	4.30	1.54	
Arg-6	8.40	4.07	2.00, 2.10		Leu-46	8.43	4.25	2.33, 2.33	2.09, $C^{\delta}H_3$ 0.79, 0.90
Val-7	9.01	3.64	2.21	0.93, 1.02	Glu-47	8.51	4.10	2.52, 2.52	
Lys-8	8.26	3.83	1.95, 1.95		Glu-48	7.92	4.18	2.12, 2.18	
Lys-9	8.18	4.09	2.01, 2.04		Glu-49	8.09	4.01	2.13, 2.13	
Ile-10	7.72	3.91	2.15	$C^{\gamma}H_3$ 0.90	Phe-50	7.78	4.60	2.70, 3.51	$C^{\epsilon}H$ 7.15, $C^{\zeta}H$ 7.23, $C^{\delta}H$ 7.64
Ile-11	8.33	3.46	1.91	$C^{\gamma}H_3$ 0.83	Asp-51	7.87	4.46	2.56, 3.14	
Gly-12	8.41	3.70	3.86 (α)		Thr-52	8.06	4.78	3.85	1.08
Glu-13	8.20	4.11	2.18, 2.24		Glu-53	8.25	4.54	1.95, 1.95	
Gln-14	8.53	3.95	2.11, 2.11	2.12, 2.45; $N^{\epsilon}H_2$ 6.99, 7.90	Ile-54	9.59	4.39	1.65	$C^{\gamma}H_2$ 0.80, 1.59, $C^{\delta}H_3$ 0.62, $C^{\gamma}H_3$ 0.31
Leu-15	8.07	4.49	1.79, 1.79		Pro-55		4.53	1.96, 2.55	
Gly-16	7.78	3.94	3.96 (α)		Asp-56	8.98	4.26	2.63, 2.75	
Val-17	7.86	4.61	2.12	0.79, 0.95	Glu-57	9.38	4.12	2.06, 2.06	2.33
Lys-18	8.55	4.23	1.71, 1.74	1.62	Glu-58	7.23	4.33	1.74, 1.74	1.96, 2.12
Gln-19	8.86	3.85	2.08, 2.17	2.42, 2.46, $N^{\epsilon}H_2$ 6.80, 7.80	Ala-59	8.54	3.62	1.41	
Glu-20	9.40	4.11	2.06, 2.06	2.32	Glu-60	7.60	4.00	2.04, 2.04	
Glu-21	7.86	4.29	2.12, 2.12		Lys-61	7.14	4.35	1.93, 1.97	
Val-22	7.59	3.82	2.40	0.86, 1.02	Ile-62	7.61	4.01	2.09	$C^{\gamma}H_2$ 0.74, 1.76, $C^{\gamma}H_3$ 0.92, $C^{\delta}H_3$ 0.41
Thr-23	7.37	4.56	4.48	1.22	Thr-63	7.64	4.59	4.49	1.17
Asn-24	8.63	3.92	2.69, 3.03	$N^{\delta}H_2$ 7.14, 7.81	Thr-64	7.07	4.79	4.08	1.07
Asn-25	8.14	4.79	2.84	$N^{\delta}H_2$ 6.97, 7.65	Val-65	8.08	3.46	2.55	0.77, 1.16
Ala-26	7.40	4.24	1.45		Gln-66	8.48	3.75	1.82	2.30, 2.42, $N^{\epsilon}H_2$ 7.12, 8.06
Ser-27	10.02	4.89	3.94, 4.04		Ala-67	7.91	4.30	1.71	
Phe-28	7.64	4.29	3.00, 3.08	$C^{\zeta}H$ 6.77, $C^{\delta\epsilon}H$ 7.08	Ala-68	7.86	4.08	1.58	
Val-29	8.75	4.24	2.03	0.97, 1.03	Ile-69	8.08	3.29	1.98	$C^{\gamma}H_3$ 0.90
Glu-30	8.36	4.08	1.91, 1.97	2.34, 2.41	Asp-70	9.20	4.38	2.68, 2.72	
Asp-31	7.85	5.06	2.83, 3.20		Tyr-71	8.21	4.12	2.99, 3.20	$C^{\epsilon}H$ 6.49, $C^{\delta}H$ 6.88
Leu-32	7.29	4.64	2.32, 2.32		Ile-72	8.17	3.59	1.73	$C^{\gamma}H_2$ 0.93, 1.74, $C^{\delta}H_3$ 0.54, $C^{\gamma}H_3$ −0.03
Gly-33	7.37	3.93	3.97 (α)		Asn-73	8.86	4.45	2.74, 2.84	
Ala-34	8.52	4.54	1.23		Gly-74	7.78	3.71	4.08 (α)	
Asp-35	9.16	5.00	2.82, 3.23		His-75	7.58	4.69	2.55, 3.49	$C^{\delta}H$ 7.15, $C^{\epsilon}H$ 8.21
Ser-36	8.69	4.34	4.17, 4.22		Gln-76	7.84	4.46	2.08, 2.24	2.44, 2.52, $N^{\epsilon}H_2$ 6.92, 7.60
Leu-37	8.08	4.29	1.82, 1.82		Ala-77	8.10	4.18	1.42	
Asp-38	8.51	4.38	2.44, 3.07		Pant.[d]			CHOH 4.04	CH_2 3.43, 3.81, CH_3 0.88, 0.98
Thr-39	8.10	3.73	4.33	1.22	β-Ala	7.96		3.52	2.50
Val-40	7.28	3.75	2.29	1.04, 1.15	β-EASH[e]	8.11		3.27	2.56

[a] Chemical shifts are referenced to internal DSS and are accurate to ±0.01 ppm. [b] n.o., not observed. [c] pH 5.4. [d] Pantoic acid. [e] β-Mercaptoethylamine.

segment adopts a helical structure. The occurrence of a series of very strong $d_{\alpha N}(i, i + 1)$ connectivities is, on the other hand, indicative of an extended chain conformation. The presence of short stretches of relatively strong $d_{\alpha N}(i, i + 1)$ or d_{NN} connectivities, perhaps with an isolated residue showing both $d_{\alpha N}(i, i + 1)$ and d_{NN} connectivities, indicates turns or a non regular secondary structure.

In the case of ACP the presence of helical structure is most

easily recognized from strings of $d_{\alpha N}(i, i + 3)$ and $d_{\alpha N}(i, i + 4)$ connectivities. This is shown graphically in Figure 6, where the intensities of various types of connectivities are indicated by the heights of the bars under the appropriate amino acid residue. When isolated absences of $d_{\alpha N}(i, i + 3,4)$ connectivities in otherwise continuous stretches are ignored, helical segments are indicated to be present from 3 to 15, 37 to 51, 56 to 63, and 65 to 75. Except for 56 to 63, these segments coincide with regions having slowly exchanging amide resonances (indicated as circular dots above the sequence in Figure 6). Exchange of amide protons for deuterons when proteins are dissolved in D_2O is known to be slow for regions of well-defined secondary structures (half-times ca. 300 h at pH 4) (Wagner & Wüthrich, 1982b; Zuiderweg et al., 1983b). Since our data result from a single COSY set accumulated over 30 h, it is difficult to speculate on the time scale of exchange or the reason for the absence of amide peaks in the 56–63 segment. The higher pH (5.4) may accelerate exchange to the point where amides are lost from even relatively stable secondary structural elements in this time period.

The precise ends of helical segments identified above are difficult to specify since cross-peaks may be absent for spectroscopic as well as structural reasons. It is also possible that residues 64 and 65, which separate two proposed helical regions, are part of one long helix from 56 to 75. There are no unusually strong $d_{\alpha N}$ connectivities in this region to indicate turns, and the moderate $d_{NN}(i, i \pm 1)$ connectivities present are characteristic of helical segments.

The use of NOEs for distinction between an α-helix and a 3_{10} helix is also possible based on NOE data. Observation of many $d_{\alpha N}(i, i + 4)$ connectivities is indicative of an α-helical structure, whereas the presence of $d_{\alpha N}(i, i + 2)$ contacts is indicative of a 3_{10} helix (Wüthrich et al., 1984). Many $d_{\alpha N}(i, i + 4)$ connectivities are observed for the residues of ACP in the helical fragments (Figures 5 and 6), suggesting an α-helix.

Segments of ACP near residues 16–19, 23–27, 34–36, and 52–55 show short stretches of strong $d_{\alpha N}(i, i + 1)$ connectivities. They are most consistent with turns in the largely helical structure. The segment from residues 28 to 34, which shows the strong long-range contact between Val-29 and Ala-34 mentioned in the Results section, is also most consistent with a sharp turn. Fragment 28–33, which shows the moderate NH–NH connectivities of a helical-like structure but is missing the usual $C^{\alpha}H$–NH short- and medium-range connectivities for an α-helix, could produce observed connectivities if residues 31–34 are in a tight turn type I. This turn would also bring Val-29 and Ala-34 into proximity and allow the intense long-range NOE cross-peaks we observe.

It is interesting to compare the four α-helical, 5-turn secondary structure observed, to that previously proposed by using the Chou–Fasman algorithm. The latter showed α-helical segments between residues 3 and 21, in reasonable agreement with our observation of an α-helix from 3 to 15. It further showed an α-helix from 37 to 53, in excellent agreement with our determination of one from 37 to 51, and an α-helix from 58 to 69, in good agreement with the two segments we see from 56 to 63 and 65 to 75. β-Turns predicted at 21–26, 32–37, and 53–58 agree well with turns we see at 23–27, 34–36, and 52–55. The major deviations are that the evidence is weak for the predicted short helix from 26 to 32 and the first helix terminates sooner than predicted with an additional turn or segment of irregular structure appearing at 16–19.

We had previously proposed a tertiary structure based on the Chou–Fasman secondary structure and a small number of NOE measurements from one-dimensional spectra (Mayo et al., 1983). Despite the substantial agreement between observation and the Chou–Fasman model, it now appears that this model will undergo significant revision. Data to be used in revising the structure will come largely from long-range connectivities and will come after assignment of additional side-chain resonances and stringent exclusion of spin diffusion as the cause for the observed NOEs.

ACKNOWLEDGMENTS

We acknowledge a number of useful discussions with P. J. Jones during the course of this work. We thank P. Demou for assisting with the instrumental aspects of the work.

REFERENCES

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

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

Bax, A., & Freeman, R. (1981) *J. Magn. Reson. 44*, 542–561.

Bax, A., & Drobny, G. (1985) *J. Magn. Reson. 61*, 306–320.

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

Boyd, J., Dobson, C. M., & Redfield, C. (1985) *Eur. J. Biochem. 153*, 383–396.

Chou, P. Y., & Fasman, G. D. (1978a) *Annu. Rev. Biochem. 47*, 251–276.

Chou, P. Y., & Fasman, G. D. (1978b) *Adv. Enzymol. Relat. Areas Mol. Biol. 48*, 45–148.

Eich, G., Bodenhausen, G., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 3731–3732.

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

Kline, A. D., & Wüthrich, K. (1985) *J. Mol. Biol. 183*, 503–507.

Mayo, K. H., Tyrell, P. M., & Prestegard, J. H. (1983) *Biochemistry 22*, 4485–4493.

McRee, D. E., Richardson, J. S., & Richardson, D. C. (1985) *J. Mol. Biol. 182*, 467–468.

Nagayama, K., Anil Kumar, Wüthrich, K., & Ernst, R. R. (1980) *J. Magn. Reson. 40*, 321–334.

Neuhaus, D., Wagner, G., Vašák, M., Kägi, J. H. R., & Wüthrich, K. (1985) *Eur. J. Biochem. 151*, 257–273.

Prescott, D. J., & Vagelos, P. R. (1972) *Adv. Enzymol. Relat. Areas Mol. Biol. 36*, 269–311.

Prescott, D. J., Elovson, J., & Vagelos, P. R. (1969) *J. Biol. Chem. 244*, 4517–4521.

Prestegard, J. H., & Scarsdale, J. N. (1985) *J. Magn. Reson. 61*, 136–140.

Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.

Rock, C. O., & Cronan, J. E., Jr. (1979) *J. Biol. Chem. 254*, 9778–9785.

Rock, C. O., & Cronan, J. E., Jr. (1980) *Anal. Biochem. 102*, 362–364.

Schultz, H. (1975) *J. Biol. Chem. 250*, 2299–2304.

States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.

Štrop, P., Wider, G., & Wüthrich, K. (1983) *J. Mol. Biol. 166*, 641–667.

Takagi, T., & Tanford, C. (1968) *J. Biol. Chem. 243*, 6432–6435.

Thompson, G. A. (1981) *The Regulation of Membrane Lipid Metabolism*, pp 20–23, 33, CRC, Boca Raton, FL.

Vanaman, T. C., Wakil, S. J., & Hill, R. L. (1968a) *J. Biol. Chem. 243*, 6409–6419.

Vanaman, T. C., Wakil, S. J., & Hill, R. L. (1968b) *J. Biol. Chem. 243*, 6420-6431.
Wagner, G., & Wüthrich, K. (1982a) *J. Mol. Biol. 155*, 347-366.
Wagner, G., & Wüthrich, K. (1982b) *J. Mol. Biol. 160*, 343-361.
Wagner, G., Anil Kumar, & Wüthrich, K. (1981) *Eur. J. Biochem. 114*, 375-384.
Wakil, S. J., Stoops, J. K., & Joshi, V. C. (1983) *Annu. Rev. Biochem. 52*, 537-579.
Wand, A. J., & Englander, S. W. (1985) *Biochemistry 24*, 5290-5294.
Weber, P. L., Wemmer, D. E., & Reid, B. R. (1985) *Biochemistry 24*, 4553-4562.
Wider, G., Macura, S., Anil Kumar, Ernst, R. R., & Wüthrich, K. (1984) *J. Magn. Reson. 56*, 207-234.
Williamson, M. P., Marion, D., & Wüthrich, K. (1984) *J. Mol. Biol. 173*, 341-359.
Wüthrich, K. (1983) *Biopolymers 22*, 131-138.
Wüthrich, K., Wider, G., Wagner, G., & Braun, W. (1982) *J. Mol. Biol. 155*, 311-319.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715-740.
Zuiderweg, E. R. P., Kaptein, R., & Wüthrich, K. (1983a) *Eur. J. Biochem. 137*, 279-292.
Zuiderweg, E. R. P., Kaptein, R., & Wüthrich, K. (1983b) *Proc. Natl. Acad. Sci. U.S.A. 80*, 5837-5841.

Preparation of Protein Conjugates via Intermolecular Hydrazone Linkage†

Te Piao King,* Shu Wei Zhao, and Terence Lam
The Rockefeller University, New York, New York 10021-6399
Received March 18, 1986; Revised Manuscript Received May 27, 1986

ABSTRACT: Proteins can be modified at their amino groups under gentle conditions to contain an average of three to six aryl aldehyde or acyl hydrazide groups. These two types of modified proteins at about 10 μM concentration condense with each other at pH ~5 to form conjugates linked by hydrazone bonds. Under proper conditions conjugates mainly of dimers and trimers in size or, if desired, higher oligomers can be obtained. The conjugates can be dissociated to their individual protein components by an exchange reaction with an excess of acetyl hydrazide. The reversible hydrazone bonds of conjugates can be reduced with $NaCNBH_3$ to give stable hydrazide bonds. The stability of protein-hydrazone conjugates was found to be significantly greater than that of the model compound, the *N*-acetylhydrazone of *p*-carboxybenzaldehyde. This difference is believed to result from the presence of multiple hydrazone linkages in protein conjugates.

For various biochemical studies it will be useful to have simple methods for preparing conjugates of one protein with another protein or peptide. The most widely used method for this purpose involves the coupling of two components by reaction with glutaraldehyde (Reichlin, 1980) to give a mixture of conjugates of like and unlike components. To avoid formation of conjugates of like components, one uses two components with different reactive groups for the coupling reaction. Three reported methods utilizing this principle involve the reaction of the sulfhydryl group of one protein with the haloacetyl group (Rector et al., 1978; Eberle et al., 1977), the maleimido group (Kitagawa & Akikawa, 1976), or the 4-dithiopyridyl group (Carlsson et al., 1978; King et al., 1978) of another protein or peptide.

Hydrazone formation has been used to prepare conjugates of proteins with various ligands (Heitzmann & Richards, 1974; Itaya et al., 1975; Rando et al., 1979). In those studies, glycoproteins were either chemically or enzymatically oxidized to generate aldehyde groups and then coupled to low molecular weight ligands containing hydrazide groups.

In this work, methods were devised so that proteins can be readily modified to contain aryl aldehyde or acyl hydrazide groups. Studies were then made to establish optimal conditions for coupling of aldehyde- and hydrazide-containing proteins to form conjugates by hydrazone bond formation and for reduction of hydrazone linkages with $NaCNBH_3$. The chemical reactions involved are given in eq 1-5 for Figure 1. Model experiments were also made on the formation and reduction of the hydrazone of *p*-carboxybenzaldehyde with acetyl hydrazide.

Materials and Methods

Unless noted otherwise, all organic chemicals were obtained from Aldrich Chemical Co., and all protein samples were obtained from Worthington Biochemical Corp. or Sigma Chemical Co. *N*-Hydroxysuccinimide esters of *N*-(bromoacetyl)-β-alanine (Santi & Cunnion, 1974), *p*-carboxybenzaldehyde (Kraehenbuhl et al., 1974), 2-nitro-5-thiobenzoic acid (Degani & Patchornik, 1971), and *N*-acetylhomocysteinyl hydrazide (Taylor & Wu, 1980) were prepared according to published procedures.

The proteins, which contain free sulfhydryl groups, were treated at a concentration of about 5 mg/mL with 1.5 mM iodoacetamide in 0.1 M phosphate of pH 7.8 for 1 h at room temperature and dialyzed to remove excess reagent prior to modifications.

Synthesis of N-Acetylhydrazones of Different Aldehydes. These were prepared by bringing to boil a 0.1-0.5 M solution of aldehyde and acetyl hydrazide (10% molar excess) in ethanol for a minute or less. On cooling, the crystalline product was collected by filtration and then recrystallized from ethanol. The hydrazones of acetaldehyde, benzaldehyde, and *p*-nitrobenaldehyde had the reported values for their melting points (Lindegren & Nieman, 1949; Tisler, 1957; Gutman et al., 1961). The hydrazone of *p*-carboxybenzaldehyde melted

† This research was supported in part by Research Grant AI-17021 from the USPHS.

0006-2960/86/0425-5774$01.50/0