Two- and Three-Dimensional ^{1}H NMR Studies of a Wheat Phospholipid Transfer Protein: Sequential Resonance Assignments and Secondary Structure†

Jean-Pierre Simorre,*,‡ Anita Caille,‡ Dominique Marion,§ Didier Marion,∥ and Marius Ptak‡,⊥

Centre de Biophysique Moléculaire (CNRS) and Université d'Orléans, 1A, Avenue de la Recherche Scientifique, 45071 Orléans Cedex 2, France, Laboratoire de Biochimie et Technologie des Protéines, INRA, BP 527-44026 Nantes Cedex 03, France, and Laboratoire de Résonance Magnétique en Biologie et Médecine, CEA/DSV/DBMS and CNRS/SDI/IMABIO, 38041 Grenoble Cedex, France

Received May 31, 1991; Revised Manuscript Received August 23, 1991

ABSTRACT: Two- and three-dimensional ^{1}H NMR experiments have been used to sequentially assign nearly all proton resonances of the 90 residues of wheat phospholipid transfer protein. Only a few side-chain protons were not identified because of degeneracy or overlapping. The identification of spin systems and the sequential assignment were made at the same time by combining the data of the two- and three-dimensional experiments. The classical two-dimensional COSY, HOHAHA, and NOESY experiments benefit from both good resolution and high sensitivity, allowing the detection of long-range dipolar connectivities. The three-dimensional HOHAHA-NOESY experiment offers the advantage of a faster and unambiguous assignment. As a matter of fact, homonuclear three-dimensional NMR spectroscopy proved to be a very efficient method for resonance assignments of protein ^{1}H NMR spectra which cannot be unraveled by 2D methods. An assignment strategy which overcomes most of the ambiguities has been proposed, in which each individual assignment toward the C-terminal end is supported by another in the opposite direction originating from a completely different part of the spectrum. Location of secondary structures of the phospholipid transfer protein was determined by using the method of analysis introduced here and was confirmed by $^3J_{\alpha NH}$ coupling and NH exchange rates. Except for the C-terminal part, the polypeptide chain appears to be organized mainly as helical fragments connected by disulfide bridges. Further modeling will display the overall folding of the protein and should provide a better understanding of its interactions with lipids.

Membrane biogenesis requires the transport of phospholipids from their site of synthesis to other cell membranes. It is postulated that a part of this lipid movement is facilitated by proteins which are characterized as phospholipid transfer proteins (PLTP).[1] PLTP are ubiquitous proteins found in microorganisms (Daum & Paltauf, 1984; Tai & Kaplan, 1985) and mammalian (Kamp et al., 1973; Lutton & Zilversmit, 1976; Whitlow et al., 1980) and plant cells (Douady et al., 1982; Watanabe & Yamada, 1986; Kader et al., 1984). Some PLTP are very specific, such as the phosphatidylcholine transfer protein of bovine liver (Kamp et al., 1973) or the phosphatidylinositol–phosphatidylcholine transfer proteins from rat brain (Venuti & Helmkamp, 1988) and yeast (Daum & Paltauf, 1984). The others exhibit a wide range of specificity, especially the plant PLTP which transport PC, PE, PI, and glycolipids (Douady et al., 1982; Watanabe & Yamada, 1986). The primary structures of nonspecific plant PLTP (Takishima et al., 1986; Bouillon et al., 1987; Tchang et al., 1988) are quite different from the sequence of the animal ns-PLTP (Westerman & Wirtz, 1985). In particular, the plant proteins have four disulfide bridges, while the animal proteins have only one cysteine residue. Furthermore, the plant protein synthesized as a precursor of higher molecular weight which is in fact composed of a peptide signal and of the mature protein (Tchang et al., 1988). These structural and biosynthetic data are usually characteristic of exported proteins and might suggest that plant ns-PLTP's are not involved in the cytosolic transport of phospholipid. To understand the biological significance of plant ns-PLTP, the mechanisms involved in in vitro lipid transfer have to be elucidated. Many data have been obtained for animal proteins, especially on lipid-binding structure [for a review, see Wirts et al. (1986) and Rueckert and Schmidt (1990)], while only predictive structure studies have been carried out on plant ns-PLTP. It is worthwhile to mention that no X-ray structure of any member of this family is yet available.

Therefore, we have started an investigation of the solution three-dimensional (3D) structure of wheat PLTP using homonuclear two- and three-dimensional (2D and 3D) nuclear magnetic resonance (NMR) techniques. Results obtained over the last 5 years have shown that NMR is a leading technique for obtaining structural information at the atomic level for small proteins in solution [for a review, see Wüthrich (1986), Clore and Gronenborn (1987), and Bax (1989)]. The great power of multidimensional NMR techniques for protein spectroscopy stems from their ability to solve the twin problem of assignment (spin correlation) and resolution (two-dimensional display). The key for the determination of a 3D

† This work was supported by the Centre National de la Recherche Scientifique, the Région Centre, the University of Orléans, and INRA (A.I.P. 90/4749). It is a part of J.P.S's Ph.D. thesis, supported by a Grant from MRT.

* To whom correspondence should be addressed.

‡ Centre de Biophysique Moléculaire, CNRS, Orléans.

§ Laboratoire de Résonance Magnétique en Biologie et Médecine, Grenoble.

∥ Laboratoire de Biochimie et Technologie des Protéines, INRA, Nantes.

⊥ Université d'Orléans.

[1] Abbreviations: PLTP, phospholipid transfer protein; ns, nonspecific; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PI, phosphatidylinositol; NMR, nuclear magnetic resonance; 2D, 3D, and 4D, two, three, and four dimensional; NOE, nuclear Overhauser effect; COSY, *J*-correlated spectroscopy; DQF, double-quantum filtered; HOHAHA, homonuclear Hartmann–Hahn spectroscopy; NOESY, NOE spectroscopy; SCUBA, stimulated cross peaks under bleached alphas.

0006-2960/91/0430-11600$02.50/0

structure lies in the assignment of the maximum number of interproton NOE's and of three-bond coupling constants (^3J) which yield, respectively, approximate internuclear distances and torsion angles, defining a basis of geometric constraints for modeling the 3D structures. Several procedures are used for generating structures either in distance space or in torsion angle space: *distance geometry* (Havel et al., 1983; Braun, 1987), *restrained molecular mechanics* (Fesik et al., 1986; Holak et al., 1987), *restrained molecular dynamics* (including simulated annealing) (Kaptein et al., 1985), and *restrained Monte-Carlo simulations* (Bassolino et al., 1988). In most cases, modeling including constraints derived from NMR parameters yields families of conformers corresponding to multiple minima formed in the optimization process or which actually reflect some internal dynamics. One way of validating the proposed structures is to back-calculate the NOE intensities from the relaxation matrix (Lefèvre et al., 1987).

Conventional ^{1}H 2D NMR techniques have been successfully used for determining 3D structures of proteins of 40 to about 100 residues [for a review, see Markley (1989)]. For larger proteins up to 150 residues, or when the spread of chemical shifts is narrow, the assignment of resonances using standard 2D NMR techniques becomes a laborious and sometimes impossible task. A means to extend resolution relies on the increase of dimensionality of acquired spectra going from 2D to 3D NMR (Griesinger et al., 1987). It overcomes the severe problem of resonance overlap, and the range of proteins amenable to NMR is extended. In heteronuclear 3D NMR, the third dimension brought by the ^{15}N or ^{13}C chemical shifts resolves overlaps encountered in 2D correlation spectra and can thus be described as X-resolved ^{1}H–^{1}H correlation experiments (Fesik & Zuiderweg, 1988; Kay et al., 1989). On the other hand, homonuclear 3D spectra yield extra cross peaks due to two different transfers (for instance, a HOHAHA transfer followed by a NOESY one); with these data, the two basic steps of the assignment (spin-system identification and interresidue connectivities) are performed on a single data set (Oschkinat et al., 1988; Vuister et al., 1990). A more reliable assignment is thus obtained because an NOE cross peak also bears a HOHAHA pattern which is a clear signature of a residue. However, the occurrence of two transfer steps is a major sensitivity limitation which has to be overcome by an optimized design of the pulse sequence.

On the wheat PLTP, the assignment procedure has been carried out by jointly using high-resolution 2D spectra (for the identification of the cross-peak fine structure) and HOHAHA-NOESY 3D spectra (for a faster and unambiguous assignment). This combined strategy is likely to be extensively used in the future for any protein, when easily usable software packages for 3D NMR processing become generally available.

Materials and Methods

Wheat PLTP was isolated according to a modification of the procedure described by Douady et al. (1985). Wheat seeds (4 kg) were soaked for 72 h at 4 °C in 5 L of 100 mM Tris-HCl, pH 7.8, containing 0.1 M KCl, 5 mM EDTA, and 8 mM EDTA. Seeds were ground at 4 °C in Tris buffer, and the homogenate was centrifuged at 5000g for 30 min. The supernatant was brought to pH 5 with glacial acetic acid and centrifuged at 10000g for 30 min. The pH of the resulting supernatant was increased up to 7.8 with 2 M Tris, and solid ammonium sulfate was added to 40% saturation. The mixture was stirred for 2 h and then centrifuged at 1500g for 15 min. The supernatant was brought to 80% saturation ammonium sulfate and stirred for 2 h. After centrifugation at 5000g for 15 min, the precipitate was suspended in 10 mM ammonium acetate, pH 6.8, and dialyzed 24 h at 4 °C against acetate buffer.

The protein solution was applied on a column (95 cm × 10 cm) packed with Sephadex G75 (Pharmacia, France) and equilibrated with 10 mM ammonium acetate buffer. Elution was carried out at a flow rate of 480 mL/h. The fractions containing PLTP were pooled and passed through a column (15 cm × 5 cm) packed with a cation exchanger (Sepharose S Fast Flow, Pharmacia, France). Proteins were eluted with a linear gradient from 10 to 250 mM ammonium acetate, pH 6.8, (40 mM/h) at a flow rate of 360 mL/h. The fractions containing the pure PLTP were pooled and dialyzed overnight deionized water and freeze-dried. The chemical determination of the primary sequence was concomitant with NMR assignments and is used in this work without further comment. This will be the subject of another paper (not yet published).

The protein samples were dissolved in ^2H_2O at a concentration of 2 mM or in a 90% ^1H_2O/10% ^2H_2O mixture at a concentration of 4 mM. The solutions were buffered with 50 mM acetate buffer, adjusted to a pH meter reading of pH 5.4, and finally sealed under argon. ^{1}H NMR spectra were recorded on a Bruker AM 500-X spectrometer, and 2D data were processed on a Bruker X32 data station using the UXNMR program. All these spectra were acquired at three temperatures, 293, 300, and 308 K, and the chemical shifts were calibrated on the water signal, respectively, at 4.82, 4.75, and 4.68 ppm. The mixing times were set, respectively, to 80, 150, 200, and 250 ms for NOESY experiments (Jeener et al., 1979) and to 30, 50, and 80 ms for HOHAHA experiments (Braunschweiler & Ernst, 1983; Davis & Bax, 1985). To reduce the first-order phase correction in NOESY experiments, the first $90°_x$ pulse was replaced by a $[90_x]^{-1} = 360°_x - 270°_{-x}$ composite pulse (Freeman et al., 1988). In both experiments, the water resonance was suppressed by combination of coherent low-power (γB_2 = 50 Hz) presaturation during the relaxation delay period and a "jump and return" read pulse. This combination eliminates any danger of receiver overflow due to radiation damping and yields a flat baseline near the water signal. Four double-quantum filtered COSY (Shaka & Freeman, 1983; Rance et al., 1983), where no selective excitation can be implemented, a stronger water suspension was used in conjunction with the SCUBA technique to restore the magnetization at the bleached $C_\alpha H$ (Brown et al., 1988). Further elimination of the dispersive component of H_2O was done by the time-domain convolution technique (Marion et al., 1989a). For all spectra, the hypercomplex method (Müller & Ernst, 1979; States et al., 1982) has been used to achieve quadrature detection along the F_1 dimension. The data set was collected as a (2 × 400, 2 × 1024) (t_1, t_2) matrix and transformed into a (2 × 1024, 2 × 1024) frequency domain spectrum. The DQF-COSY data set was collected as a (2 × 400, 2 × 2048) (t_1, t_2) matrix. Additional information was obtained from a homonuclear HOHAHA-NOESY experiment at 308 K (HOHAHA mixing = 80 ms, NOESY mixing = 150 ms) (Simorre & Marion, 1991). The Waltz-17 mixing scheme was optimized for balancing longitudinal and rotating frame cross-relaxation according to the technique known as clean TOCSY (Griesinger et al., 1988). The 180° pulses of the Waltz were sandwiched by two delays set to twice the 90° pulse length. The water signal was suppressed in a manner similar to that for 2D experiments (presaturation and jump and return). Further elimination of the dispersive component of H_2O was done by the time-domain convolution technique. Simultaneous acquisitions were used in all three dimensions. The 3D data set was transformed on a SUN 3-260 station

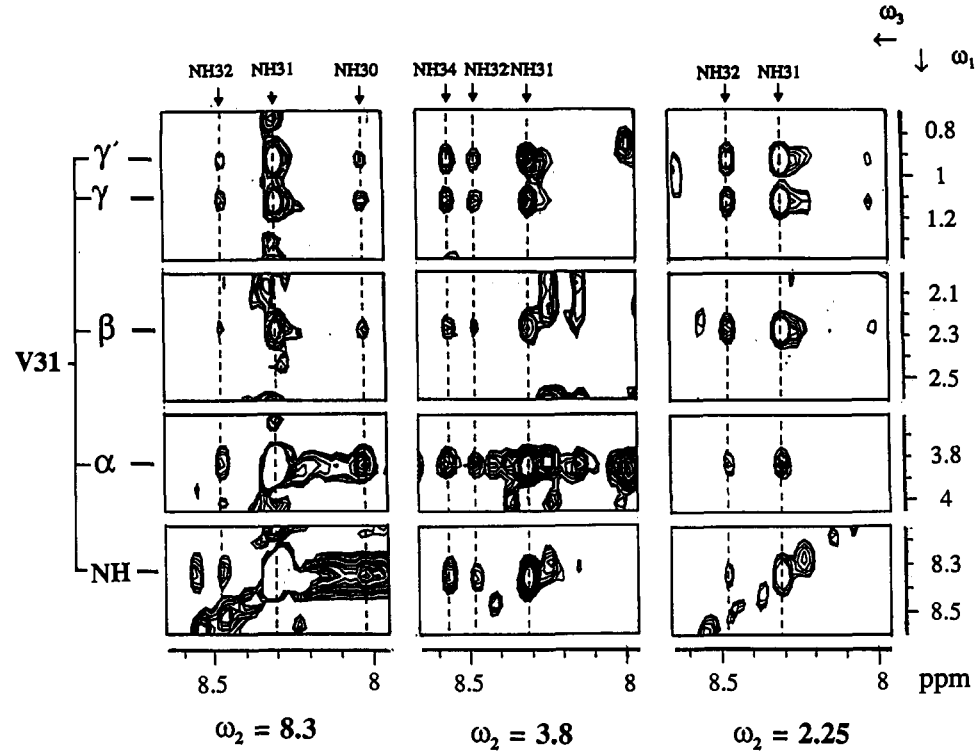

FIGURE 1: Planes extracted from the HOHAHA-NOESY 3D spectrum recorded at 308 K (HOHAHA mixing = 50 ms, NOESY mixing = 150 ms). The left inset shows the plane taken at $F_2 = \delta$ (Val 31 NH), the central one at $F_2 = \delta$ (Val 31 $C_\alpha H$), and the right one at $F_2 = \delta$ (Val 31 $C_\beta H$). The same pattern displayed in all insets is characteristic of the Val spin systems and is transferred at various location of the spectrum by an NOE involving the protons of Val 31 (see text). The HOHAHA pattern originates from several protons of this residue, and this feature is valuable for longer side chains such as Lys, where the HOHAHA transfer from the NH to the end of the side chain vanishes due the transverse relaxation.

using the NMR2 software (NMRi Syracuse, NY) and additional software written at the NIH by L. E. Kay and D. Marion (Marion et al., 1989b).

Results

Assignment Strategy. Two distinct, but in fact complementary, procedures have been proposed for obtaining ^1H NMR resonance assignments of peptides and proteins. The first one, suggested by Wüthrich and co-workers, first attempts to identify the complete spin systems of the amino acid residues before their sequential assignment via interresidue NOESY connectivities (Wüthrich et al., 1982). The other method, *main-chain-directed* (MCD) analysis, skips the spin-system identification step and identifies repeating NOE patterns of secondary structure elements (Englander & Wand, 1987). As a matter of fact, these two strategies are in no way exclusive, and in practice we used them jointly.

All reported assignments were made by combining the analysis of 2D and 3D NMR experiments. They were doubly checked on three 2D data sets acquired at different temperatures (293, 300, and 308 K) and on a homonuclear 3D data set. In some sense, temperature can be conceived as a third dimension parameter instead of a chemical shift, leading to a pseudo-3D data set (versus temperature) complementary to conventional homonuclear 3D data sets, but with higher sensitivity. It must be considered more as a means to authenticate connectivities than to reduce the crowding of 2D spectra since, over the temperature range acceptable for proteins, only a part of overlaps and degeneracies are lifted. Furthermore, monitoring the temperature is not free of danger since both the conformation and the NMR parameters are temperature dependent.

As noted before, the cross-cross-peaks of interest in a 3D spectrum (i.e., those which have been transferred first via isotropic mixing and then via NOE) exhibit a rather weak intensity (typically less than 1% of the main diagonal). Therefore most of the detected sequential connectivities correspond to strong spin–spin interactions whereas weak connectivities associated with either large distances (≥4 Å) and/or small three-bond couplings (≤3 Hz) are missing.

Nevertheless, we show below that 3D homonuclear NMR techniques save a lot of time in the assignment procedure by resolving most of the difficulties due to the complexity of the spectra. 3D NMR makes use of the same spin interactions as classical 2D NMR (NOESY and HOHAHA), but the information is combined differently: in frequency space for 2D NMR (the frequency of a cross peak is transferred from one spectrum to the other) but in the spin space for 3D NMR. As a result, a HOHAHA-NOESY spectrum allows one to ascertain the assignment of a NOESY cross peak to a given residue, not only on the basis of its frequency but also from the more informative HOHAHA pattern connected to this peak.

Let us consider two $C_\alpha H$'s belonging to residues A and B with almost exactly the same chemical shifts. From each of these $C_\alpha H$, each HOHAHA pattern (to the NH but also to the side chain) is transferred via NOE to several NH($i \rightarrow i$, $i \rightarrow i+1$, ...) which are likely to be at different frequencies.

This very powerful approach is illustrated in Figure 1 for Val 31 spin system. The middle panel shows that several $C_\alpha H$'s overlap at $\delta = 3.80$ ppm on top of that of Val 31, but the HOHAHA pattern (NH ← $C_\alpha H$, $C_\alpha H \rightarrow C_\beta H$, $C_\alpha H \rightarrow C_\gamma H_3$, $C_\alpha H \rightarrow C_\gamma H_3$) is NOE-transferred to three distinct NH's, which to not overlap. Even when severe overlap occurs, interresidue connectivities can be safely established, as shown by considering the NOESY peak between $C_\alpha H$ of Gln 45 and NH of Cys 48. It accidently overlaps with the sequential (Ala 47 $C_\alpha H \rightarrow$ Cys 48 NH) NOESY cross peak, which will be

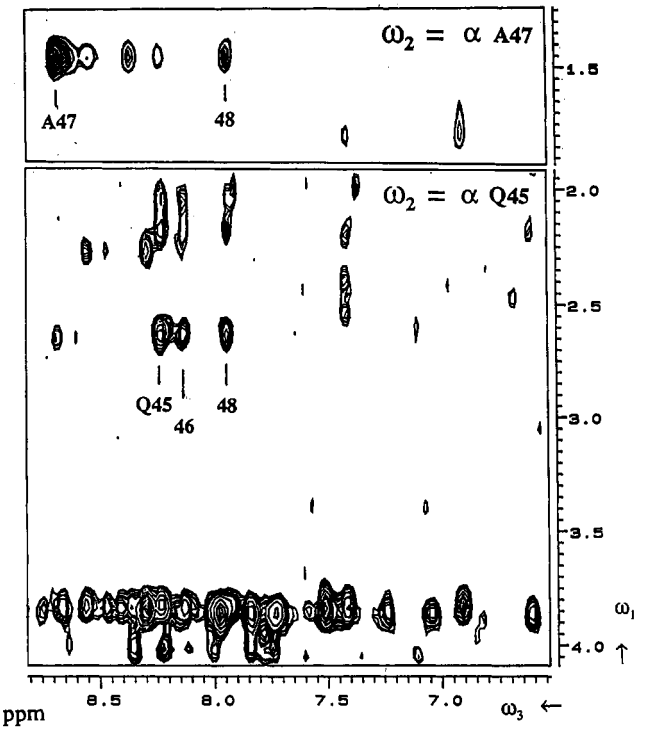

FIGURE 2: Use of 3D NMR in order to assign overlapping NOESY cross peaks. Because of the short distance corresponding to sequential NOE's, the (C_αH 47 → NH 48) cross peak exhibits a sizeable intensity in any case. As a result, the close proximity of the chemical shift of CαH of Gln 45 and that of Ala 47 (see Table I) prevents us from assigning without any doubt the possible (C_αH 45 → NH 48) cross peak in the 2D spectrum shown in Figure 3. The entire spin-system pattern of Gln 45 is NOE transferred at the frequency of NH 48. Combination of the two 2D spectra (NOESY and HOHAHA) does not provide such a piece of information.

present anyway, whatever the secondary structure is (Figure 3). In Figure 2, the HOHAHA pattern from C_αH of Gln 45 is NOE-transferred among others to the NH belonging to Cys 48 (δ = 7.99 ppm), definite evidence of a ($i \rightarrow i+3$) connectivity.

Spin System Identification. The identification of characteristic spin systems of Gly, Ala, Thr, Val, Leu, Ile, and Val was achieved mainly by analysis of scalar correlated experiments recorded in D_2O. The four Ala residues were unambiguously identified by their two COSY peaks. The single Thr (Thr 80) was recognized by comparing its COSY and HOHAHA patterns. Frequent overlaps in the high-field spectral region (between 0.0 and 1.0 ppm) are a major source of ambiguities for discriminating between Leu and Ile residues (spectra not shown). The efficient HOHAHA transfers (slow transverse relaxation) permitted a complete correlation from the NH up to the γCH_3 and the δCH_3 of Ile and Leu. From the C_αH's fingerprint region in COSY spectra, Gly residues were easily identified by their two cross peaks that line up at the NH frequency with a peculiar fine structure. From the comparison of two spectra at different temperatures (300 and 308 K), proof was obtained that both NH–C_αH cross peaks have exactly the same temperature dependence. The only trap in Gly identification was the presence of δCH_2–ϵNH resonances corresponding to Arg residues; when the two δCH_2's showed well distinct chemical shifts (for instance Arg 11, 44, and 56), and two cross peaks analogous to those of Gly appeared, corresponding to the rather slowly exchanging ϵNH protons. Assignment of a Gly residue relies on the observation of two direct connectivities with one NH, and, here again, 3D NMR is a valuable tool for identification of overlapping Gly spin systems, but the limited digital resolution of the 3D data set prevents us from discriminating Gly from Ser with degenerated chemical shifts. Discrimination between Arg and Lys is not straightforward at this stage because they only differ by a CH_2 group. The ring protons of the aromatic residues show connectivities with the remaining protons of these residues in long-mixing NOESY spectra in D_2O. Slight pH differences between the H_2O and D_2O sample has been a real problem in the case of His 59, where an uncertain chemical shift has deliberately not been reported in Table I, in which all the assignments are collected.

Sequential Resonance Assignments. As noted earlier, this step strongly overlaps with the previous one and has been set aside for sake of clarity. As reported in Figure 3 most of the NH(i) → NH(i+1) NOE connectivities were established and confirmed by the C_αH(i) → NH(i+1) connectivities. The first four residues in the amino acid sequence exhibited a small number of very weak connectivities: this feature obviously originates from conformational averaging and already supports a random-coil structure for this N-terminal part of the molecule. Gaps were introduced by Pro residues (numbers 12, 21, 23, 70, 71, and 78), which lack a NH proton. Nevertheless, spare connectivities with the neighbors were provided by their $C_\delta H_2$ protons. Large sequences of C_βH(i) → NH(i+1) connectivities were also identified up to the 70–71 Pro doublet. C_αH(i) → NH(i+3) and C_αH(i) → C_βH(i+3) connectivities complete the interproton dipolar interactions.

Secondary Structures. From the NOE connectivity pattern, one can conclude that the major part of the protein (residues ranging from 5 to 69) exhibits a helical conformation. Figure 4 and 5 are illustrative examples of the observed connectivities in two arbitrarily chosen fragments: 13–20 (Figure 5) and 42–49 (Figure 4). Four types of information are in agreement with an α-helix: short NH(i) → NH(i+1) and C_βH(i) → NH(i+1) distances, small backbone J coupling, and above all a long-range NOE involving residues i and i+3. 3D NMR was of great help by speeding up this stage of secondary structure assignments, and it can be presented as a new method. Helical fragments are easily assigned because of the symmetry of the NH–NH connectivity: in fact, one can either move upstream [NH(i) → NH(i+1)] or downstream [NH(i) → NH(i+1)] (Simorre & Marion, 1991). Figure 6 shows the assignment of a stretch of helix from Tyr 16 to Gly 20. Each plane has been taken at the F_2 frequency of the corresponding NH, and, along F_1 dimension, the HOHAHA pattern of this residue shows up. This pattern is transferred via NOE from the starting NH to both the following and preceding NH. In each plane, the same pattern is shown at three locations [δ NH(i–1), δ NH(i), and δ NH(i+1)]: for example, the peculiar pattern of Val 17 is observed at δ = 7.30 ppm (Tyr 16), δ = 7.97 ppm (Val 17), and δ = 7.61 ppm (Gln 18). Cross-cross-peaks will be detected at the same frequencies in two other planes, but the HOHAHA pattern will be different owing to the unlike nature of the central residues. The recognition of an entire pattern, which counterpart in 2D NMR spectra is a single NOESY cross peak, leaves no room for doubt, and this task can be devoted to a computer program for computer-assisted assignment. The HOHAHA pattern highlighted in Figure 6 is also displayed at other locations of the 3D spectrum. As previously noted, the signature of Val 31 in Figure 1 is shown at various pairs of chemical shifts which correspond to all NOESY connectivities in a 2D spectrum: NH(31) → NH(32), NH(31) → NH(30) (left panel), C_αH(31) → NH(32), C_αH(31) → NH(34) (center panel), and C_βH(31) → NH(32) (right panel) as well as

Table I: ^{1}H Chemical Shifts (ppm) for PLTP at pH 5.4 and 35 °C

residue	NH	αH	βH	others
Ile 1	7.13	4.02	1.85	γCH_2 1.68, 1.30; γCH_3 0.89; δCH_3 0.84
Asp 2	8.51	4.87	2.83, 2.79	OH 9.29
Cys 3	8.23	4.94	2.80, 2.61	
Gly 4	8.86	3.97, –[a]		
His 5	8.00	4.58	3.43, 3.36	δ^2H 7.10; ϵ^1H 8.55
Val 6	7.88	3.39	2.29	γCH_3 1.06, 0.94
Asp 7	8.64	4.14	2.84, 2.78	
Ser 8	7.53	4.15	4.05, –	
Leu 9	7.66	4.20	2.10, 1.45	γH 1.68; δCH_3 0.94, 0.71
Val 10	7.51	4.64	2.31	γCH, 0.92, 0.87
Arg 11	7.92	4.25	2.24, 1.97	γCH_2 1.79; δCH_2 3.31, 3.22; ϵNH 7.65
Pro 12		4.64	2.54, 1.31	γCH_2 1.96, –; δCH_2 3.76, 3.58
Cys 13	8.78	4.79	3.31, 3.21	
Leu 14	8.24	4.01	1.51, –	γH 2.03; δCH_3 0.97, 0.76
Ser 15	8.66	4.24	4.06, 3.96	
Tyr 16	7.30	4.65	3.51, 3.15	δH 7.04; ϵH 6.61
Val 17	7.97	3.75	2.45	γCH_3, 1.10, 1.03
Gln 18	7.61	4.69	2.08, 2.08	γCH_2 2.43, 2.37; ϵNH_2 7.15, 6.86
Gly 19	7.74	4.61, 3.50		
Gly 20	8.33	4.46, 3.82		
Pro 21		4.73	2.35, –	γCH_2 2.03, 1.92; δCH_2 3.66
Gly 22	8.18	3.92, 3.21		
Pro 23		3.82	1.83, 1.79	γCH_2 1.56, –; δCH_2 3.35, 2.86
Ser 24	8.99	4.44	4.18, 3.82	
Gly 25	9.04	3.96, 3.82		
Gln 26	8.37	4.20	2.10, 2.05	γCH_2 2.60; ϵNH_2 –, –
Cys 27	8.00	4.37	3.36, 2.87	
Cys 28	8.23	4.73	3.13, 2.89	
Asp 29	8.90	4.45	2.77, 2.69	
Gly 30	8.03	4.34, 3.95		
Val 31	8.31	3.80	2.24	γCH_3 1.10, 0.91
Lys 32	8.48	3.95	1.99, 1.95	γCH_2 1.56, 1.37; δCH_2 1.72; ϵCH_2 2.95; ζNH_3^+ –
Asn 33	8.26	4.56	3.07, 2.78	δNH_2 7.23, 6.98
Leu 34	8.56	3.89	2.18, 1.55	γH 1.83; δCH_3 0.99
His 35	8.30	4.33	3.35, 3.03	δ^2H 7.10; ϵ^1H 8.05
Asn 36	7.91	4.40	3.03, 2.95	δNH_2 7.75, 6.96
Gln 37	7.97	4.32	2.19, 1.99	γCH_2 2.65, 2.49; ϵNH_2 7.39, 6.93
Ala 38	7.99	4.65	1.18	
Arg 39	8.30	4.40	1.88, 1.81	γCH_2 1.60, 1.53; δCH_2 3.21; ϵNH 7.08
Ser 40	8.80	4.56	4.32, 4.08	
Gln 41	9.15	3.88	2.19, –	γCH_2 2.53, 2.33; ϵNH_2 7.77, 6.87
Ser 42	8.56	4.20	3.92, 3.88	
Asp 43	7.87	4.54	2.88, 2.43	
Arg 44	8.68	4.00	2.10, 1.98	γCH_2 –, 1.70; δCH_2 3.48, 3.18; ϵNH 8.45
Gln 45	8.24	3.78	2.14, 2.04	γCH_2 2.60, –; ϵNH_2 7.58, 6.66
Ser 46	8.14	4.34	4.12, 4.01	
Ala 47	8.70	3.75	1.42	
Cys 48	7.97	3.58	3.29, 3.27	
Asn 49	8.26	4.50	2.91, 2.77	δNH_2 7.31, 7.15
Cys 50	8.66	4.60	3.22, 3.08	
Leu 51	8.38	4.03	2.12, 1.90	γH 1.50; δCH_3 0.82, 0.76
Lys 52	8.37	4.28	2.06, 1.99	γCH_2 1.91, 1.89; δCH_2 1.63; ϵCH_2 3.15; ζNH_3^+ –
Gly 53	7.86	3.97, 3.83		
Ile 54	7.77	3.74	1.94	γCH_2 1.21, 1.13; γCH_3 0.84; δCH_3 0.79
Ala 55	8.14	3.69	1.36	
Arg 56	7.74	3.97	1.96, –	γCH_2 1.75; δCH_2 3.29, 3.25; ϵNH 7.33
Gly 57	7.53	4.29, 3.80		
Ile 58	7.25	4.17	1.95	γCH_2 1.62, 1.19; γCH_3 0.79; δCH_3 0.70
His 59	8.74	4.62	3.22, 2.91	δ^2H –; ϵ^1H –
Asn 60	8.65	4.51	2.88, 2.81	δNH_2 7.52, 6.82
Leu 61	7.62	3.87	1.76, 1.48	γH 1.42; δCH_3 0.91
Asn 62	9.04	4.80	2.86, 2.50	δNH_2 7.49, 6.87
Glu 63	8.67	3.98	2.10, 1.99	γCH_2 2.38, 2.31
Asp 64	8.08	4.40	2.81, 2.69	
Asn 65	8.22	4.16	2.40, 2.45	δNH_2 6.97, 6.83
Ala 66	8.32	3.88	1.62	
Arg 67	8.75	4.10	2.07, 1.98	γCH_2 1.90, 1.68; δCH_2 3.25; ϵNH 7.45
Ser 68	7.72	4.53	4.38, 4.10	
Ile 69	7.43	3.82	2.16	γCH_2 1.10, 1.01; γCH_3 0.97; δCH_3 0.97
Pro 70		4.36	2.49, 2.36	γCH_2 1.76, –; δCH_2 3.77, 3.83
Pro 71		4.45	2.31, 2.24	γCH_2 2.17, 2.00; δCH_2 3.66, 3.63
Lys 72	8.27	4.17	1.99, 1.97	γCH_2 1.53, 1.79; δCH_2 1.68; ϵCH_2 3.12, 3.06; ζNH_3^+ –
Cys 73	8.29	4.87	2.95, 2.85	
Gly 74	7.98	3.97, 3.84		
Val 75	8.03	4.22	1.52	γCH_3 0.68
Asn 76	8.55	4.66	2.82, 2.58	δNH_2 –, –
Leu 77	7.27	4.49	1.63, 1.61	γH 1.00; δCH_2 0.83, 0.69

Table I (Continued)

residue	NH	αH	βH	others
Pro 78		4.66	2.07, –	γCH_2 1.92, –; δCH_2 3.99, 3.36
Tyr 79	6.67	4.69	3.16, 3.14	δH 6.72; εH 6.60
Thr 80	6.80	4.33	4.21	γCH_3 0.92
Ile 81	7.99	3.69	1.66	γCH_2 0.97, 0.91; γCH_3 0.84; δCH_3 0.84
Ser 82	6.93	4.56	3.89, 3.12	
Leu 83	9.10	4.42	1.69, 1.69	γH 1.61; δCH_3 0.97, 0.90
Asn 84	8.64	4.89	2.89, 2.55	δNH_2 7.52, 6.88
Ile 85	7.33	4.19	1.84	γCH_2 1.44, 1.21; γCH_3 0.93; δCH_3 1.08
Asp 86	8.40	4.86	2.94, 2.61	
Cys 87	8.67	4.70	2.86, 3.30	
Ser 88	8.59	4.38	4.06, 3.95	
Arg 89	7.40	4.64	2.17, –	γCH_2 1.68, 1.55; δCH_2 3.21; εNH 7.13
Val 90	6.84	4.30	2.44	γCH_3 1.10, 0.94

[a] Unassigned resonances.

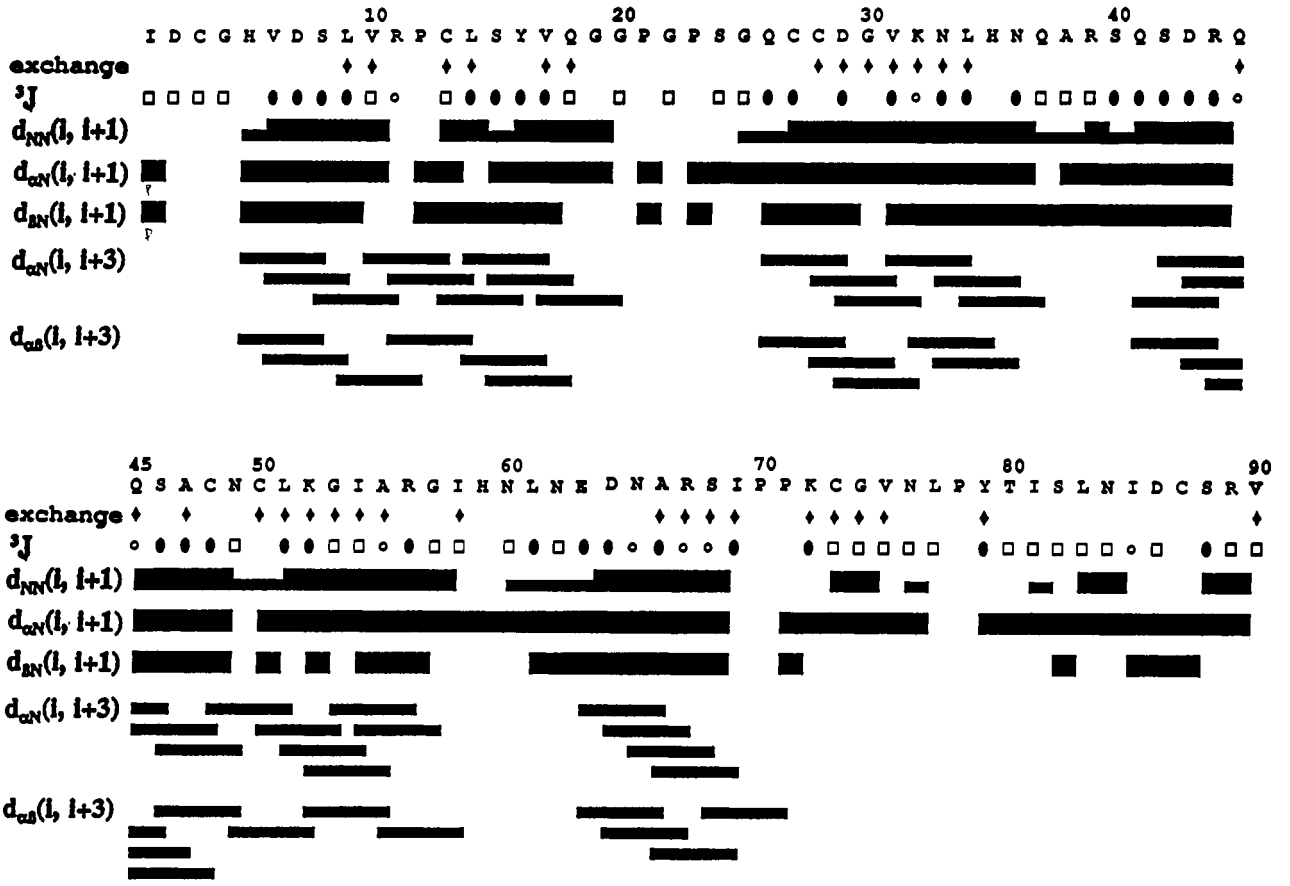

FIGURE 3: Sequence of PLTP together with a schematic summary of the sequential connectivities which supports the resonance assignment. The sequential NOESY connectivities have not been interpreted quantitatively [except for NH(i) → NH(i+1)] at this stage of the study, because our sequential cross peaks always evidence sizeable intensities, whatever the secondary structure is (for instance $d_{\alpha N}$ is always <3.5 Å). Definite evidence of the helical regions is given by the four stretches of $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$. Values of $^3J_{HN\alpha}$ coupling constants are classified as small (≤6 Hz), medium (>6 Hz and <9 Hz), or large (≥9 Hz) as indicated by filled oval (●), open circles (○), and open squares (□), respectively. The black diamond shapes (♦) on the row labeled "exchange" denote residues that contain slowly exchanging amide protons.

intraresidue connectivities. A long-range NOE between residues 31 and 34 is clear evidence of the helical organization of this region.

Within an accuracy of ±1 residue, the helical regions of PLTP can be defined as follows: 5 → 20, 26 → 36, 41 → 58, and 63 → 69. From an NMR point of view, a helix is characterized by the simultaneous observation of several NMR parameters, including short NH(i) → NH(i+1) and C_βH(i) → NH(i+1) distances, a small backbone J coupling, and a (i → i+3) long-range NOE and slowly exchanging protons. From inspection of the data summarized in Figure 3, it clearly appears that some of the parameters (described in textbooks) are missing for a few residues, and this deviation from a theoretical behavior will be discussed next.

Let us first remark that the first helix (5 → 20) is actually broken by the presence of Pro 12. Although prolines are known to disrupt helices, the almost continuous stretch of (i → i+3) connectivities clearly shows that Pro 12 only induces a slight bend in the 5 → 20 helix.

NH exchange rates have been gathered from two HOHAHA spectra recorded on a sample freshly dissolved in ^2H_2O. Data acquisition was started only half an hour after dissolution, and two HOHAHA spectra were recorded in, respectively, 4 and 16 h. Roughly 30 NH protons (out of 83 NH) were still present in the very first spectrum and identified as slowly exchanging protons in Figure 3. Obviously, this set is by far smaller than what one could expect, in agreement with the four long stretches of helices mentioned above. The slowest NH exchange rate in proteins is observed at about pH 3 when both acid- and base-catalyzed mechanisms become inefficient; thus,

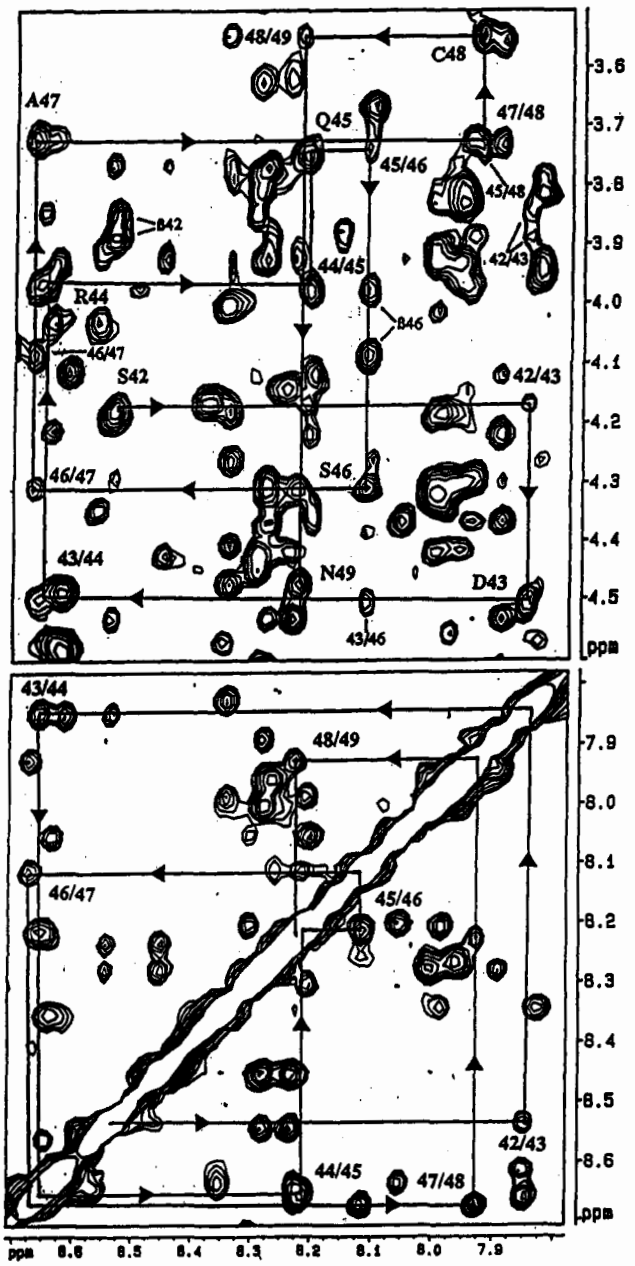

FIGURE 4: Evidence of an α-helical region at residues 42–49. In the lower part, the NH(i) → NH(i+1) NOESY connectivities are displayed and confirmed by the C_αH(i) → NH(i+1) peaks. Some long-range NOE peaks (i → i+3) are labeled, particularly the (C_αH 45 → NH 48) peaks which have been assigned by 3D NMR (see Figure 2).

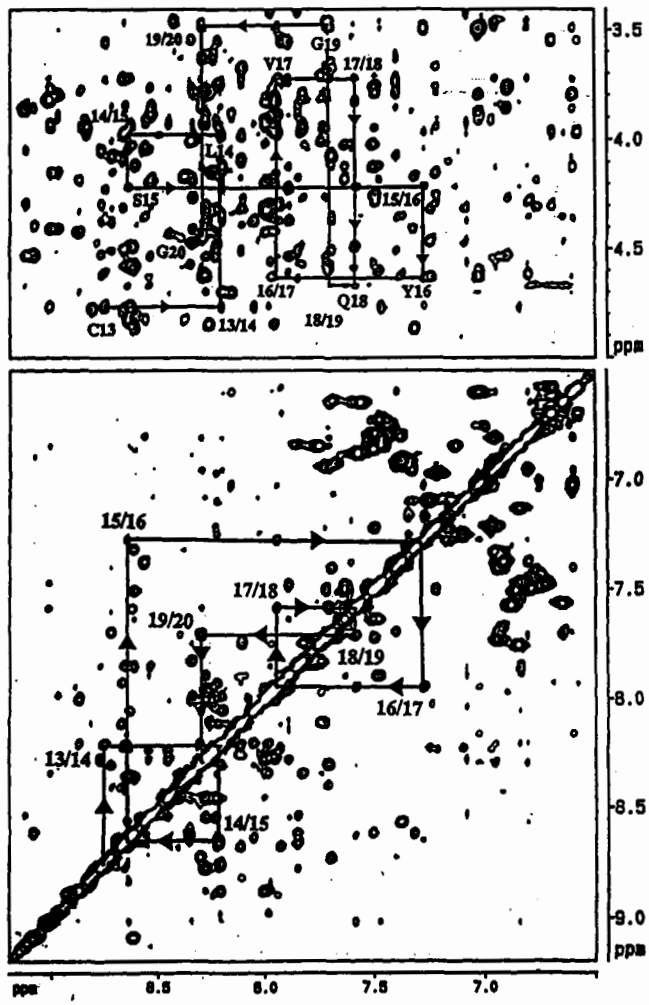

FIGURE 5: Evidence of an α-helical region at residues 13–20 by NOESY connectivities (see Figure 3).

at pH 5.4, the rate is speed up more than 100 times, and a larger fraction of NH's is likely to be already exchanged after 4 h (Wüthrich & Wagner, 1979). It should be remembered that, in an α-helix, the very first four NH which are forming some kind of turn are not hydrogen bonded. This is the case for the 5 → 20 helice, where the NH's from 5 to 8 are not seen in ^2H_2O. In the second helix (26 → 36), the residues from 28 to 34 have slowly exchanging NH's, and in the third helix (41 → 58) all but five residues from 44 to 58 have the same behavior. Finally, in the last helix (63 → 69), the C-terminal NH exchanges slowly with the solvent. At this stage of the data interpretation, it is difficult to explain why some NH's in helices have been reported missing; the tertiary structure of PLTP will probably explain some of these striking features. Some of the side chains are also able to modify the exchange behavior of the backbone NH, such as the OH group of Ser 15, which could speed up the exchange rate of both Ser 15 and Tyr 16 NH.

A third kind of NMR spectral parameters, the vicinal coupling constants $^3J_{HN\alpha}$, have not been used so far as evidence of secondary structure. In Figure 3, the values of these constants determined on the COSY maps are reported and roughly arranged according to their low, intermediate, and high values. Two sets of low values that appear in the 5 → 20 fragment were evidence of helical organization and were previously given. Another set is found for the 26 → 36 fragment, which exhibits a rather characteristic NOE pattern usually assigned to a helical structure. The presence of a third helical fragment was previously identified in the 42 → 48 region by considering the NOE connectivities, and it is likely confirmed by a continuous series of small coupling constants. Two other fragments, 50 → 58 and 61 → 69, could also involve some helical organization. Clearly, the last C-terminal fragment following the 70 → 71 Pro doublet corresponds to extended structure. The overall folding of the polypeptide chain should involve several tight turns connecting the different secondary structure elements, and strong additional constraints are introduced by the four S–S bridges. It is then that some severe distortions can affect the helical structures previously proposed and can make a very precise definition of the global helicity of PLTP difficult.

Concluding Remarks.

We have demonstrated the value of 3D NMR HOHAHA-NOESY spectra for assigning a 9-kDa protein such as PLTP.

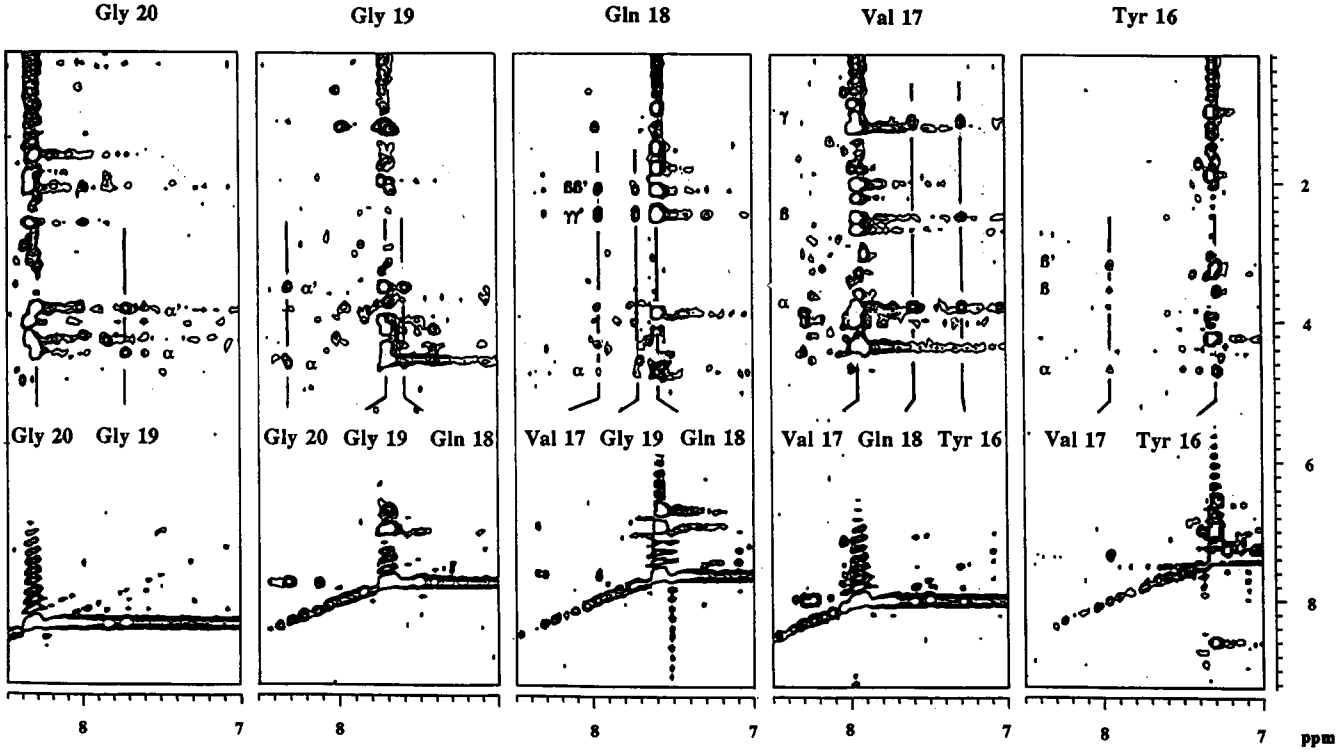

FIGURE 6: Assignment of helices by 3D homonuclear NMR. This figure displays the same assignment process as Figure 4, but with 3D NMR. In each plane taken at the F_2 frequencies of several NH's, a HOHAHA pattern corresponding to this residue is transferred to the adjacent residues. Definite evidence of a connectivity is given in two planes: in the plane of Val 17, its spin system moves to the NH of Gln 18 ($i \rightarrow i+1$), while, in the plane of Gln 18, the latter system moves to the NH of Val 17 ($i+1 \rightarrow i$).

Our method for the interpretation of such spectra is based on the downstream and upstream transfer of magnetization and on an identification using the entire HOHAHA pattern appearing at several frequencies of neighboring protons (Figure 2). This method allows us to overcome a number of problems due to resonance overlaps and considerably increases the reliability of residue assignment and the detection of secondary structure elements.

Obviously, this does not prevent the use of 2D NMR, which offers the advantage of good resolution and optimal sensitivity, allowing the detection of long-range interactions. Having at our disposal an efficient and reliable method of assignment was specially important in the case of wheat PLTP, inasmuch as the entire peptide sequence was not available at the beginning of the NMR investigation. NMR provides useful information on those fragments whose sequence was not fully determined by chemical analysis and certainly contributes to the definitive establishment of the entire sequence of the protein. Up to now, no results have been gathered on the disulfide bridges and on how the Cys residues match together. Once a sequential resonance assignment is obtained, several additional NOE's in the neighborhood of the cysteines were observed. For instance, weak NOE's corresponding to long-range connectivities are detected between the following pairs of protons: Cys 30 C_αH–Cys 50 C_βH, Cys 27 NH–Cys 13 C_αH, Cys 73 C_αH–Cys 28 C_βH, and Cys 27 C_αH–Pro 12 C_βH, etc. This network is in full agreement with the S–S bridges determined in a homologous protein (PLTP from germinated castor bean seeds) (Takishima et al., 1988) by using chemical methods. Note that only chemical determination provides definite evidence for disulfide bridges; in the same way as for hydrogen bonds (where an NOE cannot directly connect the hydrogen-bonded atoms), interpretation of NOE's in terms of S–S bridges can be uncertain.

In the present state of our investigation, it can be assumed that the main part of the PLTP chain is organized in secondary structure elements separated by short gaps involving mainly Pro residues. Several helical regions were well characterized, but a more precise picture of the whole organization of the polypeptide chain will require a complete modeling based on the NMR data. In this regard, homonuclear 3D NMR techniques are of a great help by providing a number of NOE connectivities which cannot easily be detected in 2D spectra. It has been largely proved that the reliability of a protein structure deduced from NMR data sharply increases with the number of definitively assigned NOESY cross peaks. The next step of our structure determination is now to build up and to refine a model of the 3D structure of PLTP. This 3D structure should enable us to establish some basis for the interaction of the proteins with lipids, and we will investigate this shortly by using 2D and 3D NMR techniques.

References

Bassolino, D. A., Hirata, F., Kitchan, D., Pardi, A., & Levy, R. M. (1988) *Int. J. Supercomput. Appl. 2*, 41–61.

Bax, A. (1989) *Annu. Rev. Biochem. 58*, 223–256.

Bouillon, P., Drischel, C., Vergnolle, C., Duranton, H., & Kader, J. C. (1987) *Eur. J. Biochem. 166*, 387–391.

Braun, W. (1987) *Q. Rev. Biophys. 19*, 115–157.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.

Brown, S. C., Weber, P. L., & Mueller, L. (1988) *J. Magn. Reson. 77*, 166–169.

Clore, G. M., & Gronenborn, A. M. (1987) *Protein Eng. 1*, 275–288.

Daum, G., & Paltauf, F. (1984) *Biochim. Biophys. Acta 794*, 385–391.

Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820–2821.

Douady, D., Grosbois, M., Guerbette, F., & Kader, J. C. (1982) *Biochim. Biophys. Acta 710*, 143–153.

Douady, D., Grosbois, M., Guerbette, F., & Kader, J. C.

(1985) *Physiol. Veg. 23*, 373–380.
Englander, S. W., & Wand, A. J. (1987) *Biochemistry 26*, 5953–5958.
Fesik, S. W., O'Donnell, T. J., Gampe, R. T., Jr., & Olejniczak, E. T. (1986) *J. Am. Chem. Soc. 108*, 3165–3170.
Fesik, S. W., & Zuiderweg, E. R. P. (1988) *J. Magn. Reson. 78*, 588–593.
Freeman, R., Friedrich, J., & Wu, Xi-Li (1988) *J. Magn. Reson. 79*, 561–567.
Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870–872.
Griesinger, C., Sørensen, O. W., & Ernst, R. R. (1987) *J. Magn. Reson. 84*, 14–63.
Havel, T. F., Kuntz, I. D., & Grippen, G. M. (1983) *Bull. Math. Biol. 45*, 665–720.
Holak, T. A., Prestegard, J. H., & Forman, J. D. (1987) *Biochemistry 26*, 4652–4660.
Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4556.
Kader, J. C., Julienne, M., & Vergnolle, C. (1984) *Eur. J. Biochem. 139*, 411–416.
Kamp, H. H., Wirtz, K. W. A., & van Deenen, L. L. M. (1973) *Biochim. Biophys. Acta 318*, 313–325.
Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R., & van Gunsteren, W. F. (1985) *J. Mol. Biol. 182*, 179–182.
Kay, L. E., Marion, D., & Bax, A. (1989) *J. Magn. Reson. 84*, 72–84.
Lefèvre, J. F., Lane, T. L., & Jardetzky, O. (1987) *Biochemistry 26*, 5076–5090.
Lutton, C., & Zilversmit, D. B. (1976) *Biochim. Biophys. Acta 441*, 370–379.
Markley, J. L. (1989) *Methods Enzymol. 176*, 12–64.
Marion, D., Ikura, M., & Bax, A. (1989a) *J. Magn. Reson. 84*, 425–430.
Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., & Bax, A. (1989b) *J. Am. Chem. Soc. 111*, 1515–1517.
Müller, L., & Ernst, R. R. (1979) *Mol. Phys. 38*, 963–992.
Oschkinat, H., Griesinger, C., Kraulis, P. J., Sørensen, O. W., Ernst, R. R., Gronenborn, A. M., & Clore, G. M. (1988) *Nature (London) 332*, 374–376.
Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.
Ruecker, D. G., & Schmidt, K. (1990) *Chem. Phys. Lipids. 56*, 1–20.
Shaka, A. J., & Freeman, R. (1983) *J. Magn. Reson. 51*, 169–173.
Simorre, J. P., & Marion, D. (1991) *J. Magn. Reson. 94*, 426–432.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Takishima, K., Watanabe, S., Yamada, M., & Mamiya, G. (1986) *Biochim. Biophys. Acta 870*, 248–255.
Takishima, K., Watanabe, S., Yamada, M., & Mamiya, G. (1988) *Eur. J. Biochem. 177*, 241–249.
Taï, S. P., & Kaplan, S. (1985) *Chem. Phys. Lipids 38*, 41–50.
Tchang, F., This, P., Stiefel, V., Arondel, V., Morch, M.-D., Pages, M., Puigdomenech, P., Grellet, F., Delseny, M., Bouillon, P., Huet, J. C., Guerbette, F., Beauvais-Cante, F., Duranton, H., Pernollet, J.-C., & Kader, J.-C. (1988) *J. Biol. Chem. 263*, 16849–16855.
Venuti, S. E., & Helmkamp, G. M., Jr. (1988) *Biochim. Biophys. Acta 646*, 119–128.
Vuister, G. W., Boelens, R., Padilla, A., Kleywegt, G. J., & Kaptein, R. (1990) *Biochemistry 29*, 1829–1839.
Watanabe, S., & Yamada, M. (1986) *Biochim. Biophys. Acta 876*, 116–123.
Westerman, J., & Wirtz, W. A. (1985) *Biochem. Biophys. Res. Commun. 127*, 333–338.
Whitlow, C. D., Pool, G. L., Brumley, G. W., & Lumb, R. H. (1980) *FEBS Lett. 113*, 221–224.
Wirtz, K. W. A., Op den Kamp, J. A. F., & Roelofsen, B. (1986) in *Progress in Lipid–Protein Interactions* (Watts, A., & De Pont, J. J. H. H. M., Eds.) pp 221–265, Elsevier, Amsterdam.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., & Wagner, G. (1979) *J. Mol. Biol. 130*, 1–18.
Wüthrich, K., Wieder, G., & Wagner, G., & Braun, W. (1982) *J. Mol. Biol. 155*, 311–319.