Sequence-Specific ^{1}H NMR Assignments and Solution Structure of Bovine Pancreatic Polypeptide

Xiang Li,[‡] Michael J. Sutcliffe,[§] Thue W. Schwartz,[∥] and Christopher M. Dobson*[,‡]

Oxford Centre for Molecular Sciences and Inorganic Chemistry Laboratory, University of Oxford, Oxford, OX1 3QR, U.K., Biological NMR Centre, P.O. Box 138, Leicester University, Leicester, LE1 9HN, U.K., and Laboratory for Molecular Endocrinology, Rigshospitalet 6321, Blegdamsvej 9, DK-2100 Copenhagen, Denmark

Received August 14, 1991; Revised Manuscript Received October 22, 1991

ABSTRACT: Sequence-specific ^{1}H NMR assignments for the 36 residue bovine pancreatic polypeptide (bPP) have been completed. The secondary and tertiary structure of bPP in solution has been determined from experimental NMR data. It is shown that bPP has a very well-defined C-terminal α-helix involving residues 15–32. Although regular secondary structure cannot be clearly defined in the N-terminal region, residues 4–8 maintain a rather ordered conformation in solution. This is attributed primarily to the hydrophobic interactions between this region and the C-terminal helix. The two segments of the structure are joined by a turn which is poorly defined. The four end residues both at the N-terminus and the C-terminus are highly disordered in solution. The overall fold of the bPP molecule is very closely similar to that found in the crystal structure of avian pancreatic polypeptide (aPP). The RMS deviation for backbone atoms of residues 4–8 and 15–32 between the bPP mean structure and the aPP crystal structure is 0.65 Å, although there is only 39% identity of the residues. Furthermore, the average conformations of some (mostly from the α-helix) side chains of bPP in solution are closely similar to those of aPP in the crystal structure. A large number of side chains of bPP, however, show significant conformational averaging in solution.

Pancreatic polypeptides (PPs)[1] are single-chain peptides of 36 amino acid residues. Since the first member of the family was isolated from avian pancreas (Kimmel et al., 1968), homologous peptides have been extracted from a number of mammalian as well as many nonmammalian species (Schwartz et al., 1989). PPs are localized in a single endocrine cell type in the pancreas where they are synthesized as part of a larger precursor (Schwartz & Tager, 1981). They appear to play an important role in the physiological feedback inhibition that regulates pancreatic secretion after a protein meal (Schwartz, 1983). Specific receptors for PP have been characterized on neuronal cell lines (Schwartz et al., 1987) and recently also in areas of the central nervous system which are accessible to circulating hormones (Whitecomb et al., 1990). It is now believed that PPs belong to a large group of regulatory peptides which also includes the intestinal hormone peptides YY (PYYs) and the important and abundant neuropeptides Y (NPYs). A common feature to the PP group of molecules is a blocked amidated C-terminal residue, Tyr36, which is believed to play a role in stabilizing the molecule in solution (Tonan et al., 1990).

During the last decade, numerous studies have been performed on the biological functions of the PP-fold molecules. However, the structure of only one member of the family, avian (a)PP, has been determined by X-ray crystallography (Blundell et al., 1981; Glover et al., 1983). Despite the fact that it is such a small molecule, aPP assumes a globular conformation which comprises a polyproline type II helix (residues 1–8) and an α-helix (residues 14–32). The two helices, joined by a type II β turn, interact via hydrophobic contacts to form a compact and stable tertiary structure. Further hydrophobic interactions are achieved through dimerization, which is preserved in solution. PPs and their relatives from mammalian species show high sequence identity, but they differ from aPP at more than half of the 36 positions (Table I). Nevertheless, a similar fold to that found for aPP has been assumed to prevail among all the members of the PP family. This is supported by molecular modeling studies (Glover et al., 1985) and is widely referred to in investigations of their biological functions. Until very recently, most of the direct information on the conformations of PPs in solution has been provided by CD studies, many involving the use of synthetic peptide fragments as well as the intact molecules (Tonan et al., 1990; Fuhlendorff et al., 1990; Minakata & Iwashita, 1990). A recent NMR study of pNPY (Saudek & Pelton, 1990), however, suggests a very different picture of the N-terminal conformation of the molecule, possibly arising as a consequence of cis–trans isomerization of the 3 N-terminal prolines under the experimental conditions used.

In this study, we present the sequence-specific ^{1}H NMR assignments of bovine (b)PP, together with its solution structure determined from NMR data.

Materials and Methods

bPP Peptide. The bPP used in the present study was natural peptide purified from side fractions in insulin production. The peptide was purified through multiple ion-exchange chromatography followed by reverse-phase HPLC as described previously (Hoffmann et al., 1983). The peptide was kindly

* To whom correspondence should be addressed.
‡ University of Oxford.
§ Leicester University.
∥ Laboratory for Molecular Endocrinology, Copenhagen.

[1] Abbreviations: bPP, bovine pancreatic polypeptide; aPP, avian pancreatic polypeptide; hPP, human pancreatic polypeptide; pPP, porcine pancreatic polypeptide; NPY, neuropeptide Y; PYY, intestinal hormone peptide YY; NMR, nuclear magnetic resonance; 2-D, two dimensional; COSY, two-dimensional *J*-correlated spectroscopy; NOE, nuclear Overhauser enhancement; NOESY, two-dimensional NOE spectroscopy; TOCSY, two-dimensional total coherence spectroscopy; E-COSY, exclusive correlation spectroscopy; $d_{NN}(i,j)$, NOE connectivity between the NH proton on residue i and the NH proton on residue j; $d_{\alpha N}(i,j)$, NOE connectivity between the αCH proton on residue i and the NH proton on residue j; $d_{\beta N}(i,j)$, NOE connectivity between the βCH proton on residue i and the NH proton on residue j; d_{NN}, $d_{NN}(i,i+1)$; $d_{\alpha N}$, $d_{\alpha N}(i,i+1)$; $d_{\beta N}$, $d_{\beta N}(i,i+1)$; J, coupling constant; RMS, root mean square; RMSD, root mean square deviation.

0006-2960/92/0431-1245$03.00/0

Table I: Amino Acid Sequences of Some Members of the Pancreatic Polypeptide Family

	5	10	15	20	25	30	35	
bPP	APLEP	EYPGD	NATPE	QMAQY	AAELR	RYINM	LTRPR	Y[a]
aPP	GPSQP	TYPGD	DAPYE	DLIRF	YDDLQ	QYLNV	VTRHR	Y[a]
hPP	APLEP	VYPGD	NATPE	QMAQY	AADLR	RYINM	LTRPR	Y[a]
pPP	APLEP	VYPGD	DATPE	QMAQY	AAELR	RYINM	LTRPR	Y[a]
nPY	YPSKP	DNPGE	DAPAE	DLARY	YSALR	HYINL	ITRQR	Y[a]
pYY	YPAKP	EAPGE	DASPE	ELSRY	YASLR	HYLNL	VTRQR	Y[a]

provided by Ronald E. Chance, Eli Lilly, Indianapolis, IN, and Leo Snell, Novo Nordisk, Bagsvaerd, Denmark. All NMR samples were dissolved in either 90% H_2O/10% D_2O or 99.99% D_2O at a concentration of ca. 7 mM and a pH of 4.6 (direct pH meter reading without correction for isotope effects). All spectra were recorded at 28 °C, except for the hydrogen exchange experiment which was performed at 14 °C in order to reduce the exchange rates.

NMR Spectroscopy. All NMR experiments were performed on a home-built 500-MHz spectrometer equipped with an Oxford Instruments Co. magnet, a GE/Nicolet 1280 computer and 293D pulse programmer, and a Bruker probe. All spectra were recorded in the pure-phase absorption mode according to the method of States et al. (1982). TOCSY (Braunschweiler & Ernst, 1983; Davis & Bax, 1985) spectra were recorded with mixing times of 50 ms in D_2O and 58 ms in H_2O. RF pulses used in the TOCSY experiment were obtained via the attenuated low-power transmitter channel (90° pulse ~23–26 μs). Spin-locking was achieved with an MLEV17 pulse scheme sandwiched between two 2.5-ms trim pulses. Short delays of approximately 2.2 times the 90° pulse length were inserted into the 180° composite pulse in the MLEV17 sequence to eliminate rotating-frame cross-relaxation effects (Griesinger et al., 1988). This windowed TOCSY showed better performance, especially for tracing long-range through-bond connectivities, than the standard experiment. COSY (Aue et al., 1976; Bax & Freeman, 1981) and NOESY (Jeener et al., 1979; Anil Kumar et al., 1980) spectra were recorded using standard phase-cycling schemes. Mixing times of 100, 150, and 200 ms were used in the NOESY experiments and zero-quantum effects were suppressed by 10% random variation of the mixing time (Macura et al., 1981). Pre-TOCSY COSY and NOESY spectra (Otting & Wüthrich, 1987) were also collected; in these experiments, a short (~25 ms) MLEV17 mixing sequence was inserted immediately after the first 90° pulse in order to recover αCH resonances bleached out by the strong irradiation of the H_2O peak during the relaxation delays for solvent suppression. Data sets consisting of 512 t_1 increments of 32 or 48 transients with 2048 data points in the t_2 domain were accumulated for the COSY and pre-TOCSY COSY spectra. For TOCSY, NOESY, and pre-TOCSY NOESY spectra, typically 256 t_1 increments of 32 transients were collected. The transmitter offset was placed on the water resonance, and a spectral width of 5405.4 Hz was used in both dimensions. All spectra were resolution enhanced in t_2 by trapezoidal and double-exponential weighting functions and in t_1 by squared sine-bell and double-exponential functions. After zero-filling, the digital resolution was 2.6 Hz/point in both t_1 and t_2 dimensions. E-COSY (Griesinger et al., 1987) spectra were recorded in D_2O to measure $^3J_{\alpha\beta}$ coupling constants. A spectral width of 4608.29 Hz was used; 512 t_1 increments of 32 transients with 4K data points in t_2 were collected in order to achieve the required high digital resolution.

Structure Calculations. Three-dimensional structures were calculated on the basis of the NMR data using a hybrid metric matrix distance geometry–dynamic simulated annealing approach (Nilges et al., 1988) as described previously (Sutcliffe & Dobson, 1991). The metric matrix distance geometry calculations (Havel & Wüthrich, 1984) were carried out using a program supplied by Dr. R. Scheek, Groningen. The dynamic simulated annealing calculations were achieved using the program XPLOR (Brunger et al., 1987). In the distance geometry stage of the procedure, a set of distance values between a subset of heavy atoms was generated within the distance bounds imposed by all the NOE constraints, the complete covalent geometry, and the hard-sphere atomic van der Waals radii. The substructures were then "embedded" into Cartesian coordinate space. The resulting structures were optimized using a program from Dr. A. Smellie, Oxford. Full side chains were then positioned on to the substructures, using a least-squares procedure to fit the N, C_α, and C atoms from the substructure to the corresponding dihedral angles, where they existed in the substructure. The remaining atoms were built using the most probable conformer for the residue in question (Sutcliffe et al., 1987).

The second stage of the calculations starts with 1.5 ps of dynamics at 1000 K. Large values of NOE and ϕ angle restraining force constants (K_{NOE} and K_ϕ) of 50 kcal/(mol·Å^2) and 200 kcal/(mol·rad^2), respectively, were used, while keeping the REPEL force constant (K_{REPEL}) at 0.01 kcal/(mol·Å^4). During the following 3 ps of dynamics, K_{NOE} was set to 10 kcal/(mol·Å^2) and K_ϕ to 20 kcal/(mol·rad^2), while K_{REPEL} was increased from 0.01 to 10 kcal/(mol·Å^4) by a factor of $1000^{1/40}$ in 40 steps of 0.75-ps dynamics. 1.4 ps of steady cooling from 1000 to 300 K was then performed, followed by 200 steps of Powell minimization. A further 8.8 ps of dynamics was run in the presence of the NMR restraints with a full potential force field (i.e., full van der Waals, electrostatic, and hydrogen-bonding potentials). The first 3 ps of the trajectories was run at 1000 K, and the temperature was reduced over the next 2.8 ps to 300 K in 25 K steps, followed by 3 ps of dynamics run at 300 K. The final structure was then obtained by averaging the last 2 ps of the trajectories at 300 K and optimized with 200 steps of restrained Powell minimization.

Results

Assignment of the ^1H NMR Spectra. The ^1H NMR assignments of bPP presented here were achieved by means of standard sequential assignment procedures (Wüthrich, 1986) and extend preliminary studies carried out in this laboratory (Howarth, 1986; Rehse, 1989). The first stage of assignment involved the identification of spin systems. This was established primarily by using TOCSY spectra, which show both direct and multiple-relayed correlation. COSY spectra were used to complement this to obtain unambiguous identification of direct scalar connectivities. In the TOCSY spectra (Figure 1a,b), the 2 δ protons from leucine residues gave pairs of cross-peaks to γ, β, α, and even to NH protons of the same residue, forming a unique pattern in the upfield region of the spectra. The complete assignment of all three leucine spin systems was thus fairly straightforward and unambiguous. The γCH_3 resonance of the single isoleucine residue showed a strong cross-peak to its βH and αH resonances in the TOCSY

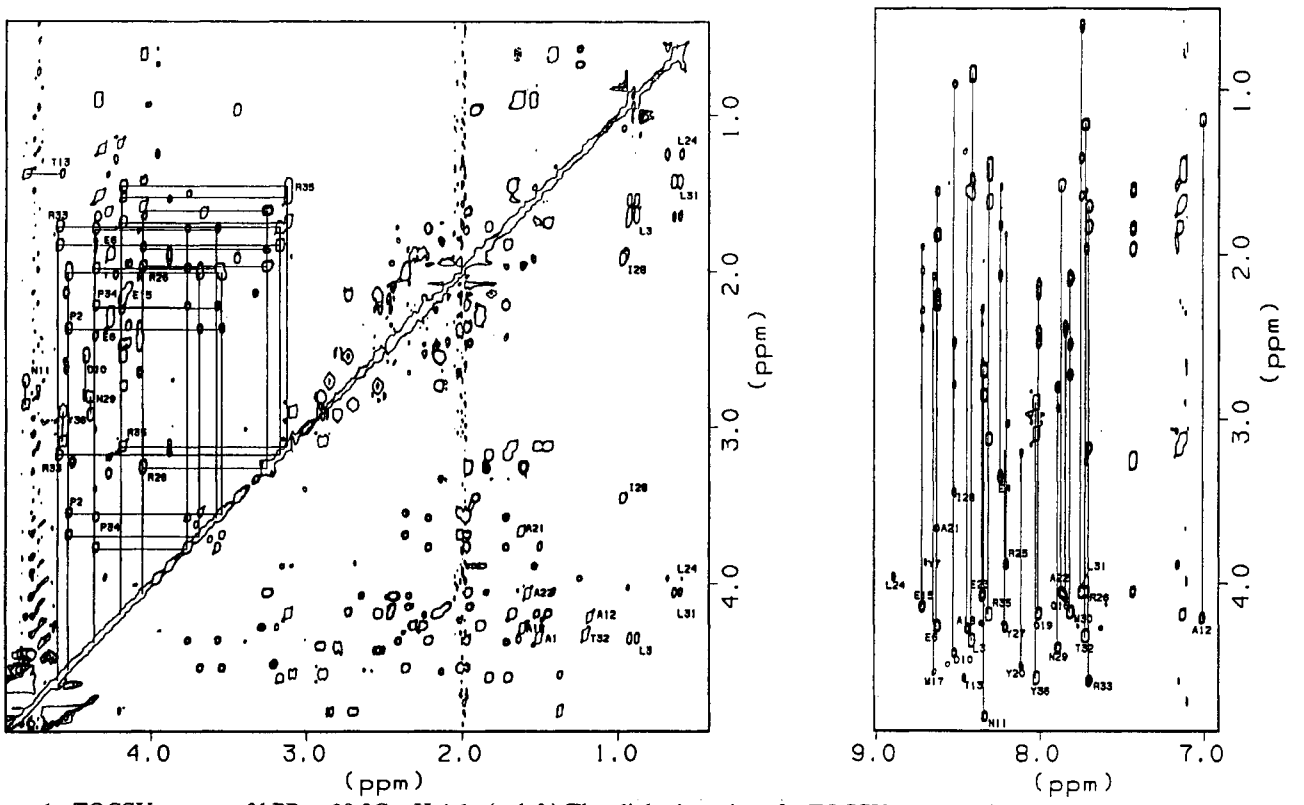


FIGURE 1: TOCSY spectra of bPP at 28 °C, pH 4.6. (a, left) The aliphatic region of a TOCSY spectrum in D_2O, with a mixing time of 50 ms. Some long side-chain spin systems (e.g., arginine, proline) are shown by solid lines with sequence numbers near the corresponding cross-peaks. (b, right) The NH–aliphatic region of a TOCSY spectrum, recorded in H_2O, with a mixing time of 55 ms. Cross-peaks from the same spin system are linked by solid lines with the sequence numbers at the positions of the direct NH–C_αH peaks.

spectra, and the cross-peak between the δCH_3 and γCH_2 protons formed a unique pattern in the upfield region of the COSY spectra. All five alanine residues in bPP gave strong αH–βH cross-peaks in the characteristic region of the TOCSY and COSY spectra, and the corresponding cross-peaks from NH to α,β protons were easily identified in the amide region of the TOCSY spectra, except for the alanine at the N-terminus. Assignment of the resonances from the threonine residues was achieved in a similar manner.

The four tyrosines are the only aromatic residues in bPP. Correlation of the ring protons to the backbone protons of the corresponding residues was achieved from NOESY spectra. The five arginines were assigned in a similar manner to the leucine residues, because of the complete and characteristic spin networks found in the TOCSY spectra for these residues. Three of the five remaining proline residues were successfully assigned by tracing their spin networks in the TOCSY experiments. A fourth proline residue showed only very weak signals in the TOCSY spectra but stronger intensities in the NOESY spectra; this residue was therefore identified by both intra- and interresidue NOEs. The fifth proline residue has not been unambiguously assigned, because it failed to give any cross-peaks in either TOCSY or COSY spectra. Some NOEs were, however, found from its adjacent residue, which could not be assigned to other spin systems. The reason for the absence of the peaks from this proline residue in TOCSY and COSY spectra is not known. None of the assigned prolines showed cis–trans isomerization as reported for pNPY (Saudek & Pelton, 1990). Inspection of their sequential NOEs suggested that the trans conformation predominates throughout our experimental conditions.

All the cross-peaks from the remaining residues (i.e., Asn, Asp, and Gln, Glu, and Met) were present in the TOCSY spectra and were initially grouped into AMX or AM(PT)X

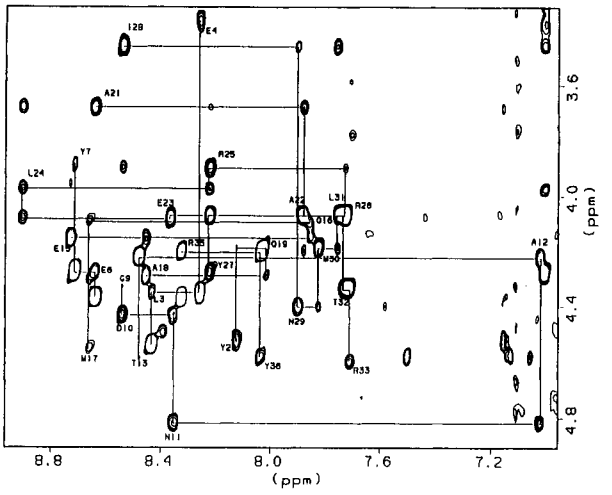


FIGURE 2: Fingerprint region of the pre-TOCSY NOESY spectrum of bPP, recorded in H_2O at 28 °C, with a mixing time of 200 ms. Sequential links through $d_{\alpha N}$ connectivities are shown in solid lines. Sequence numbers are indicated at the positions of direct intraresidue $C_\alpha H(i)$–$NH(i)$ cross-peaks.

spin systems. Assignment to individual residues was only achieved at the later sequence-specific assignment stage.

The second stage of the assignment involved alignment of the spin systems with specific amino acid residues in the protein sequence, using short sequential NOE (<5 Å) connectivities (i.e., NH, C_αH, C_βH, and, for proline, C_δH). This procedure was initiated at the single isoleucine residue, Ile 28. Because of ample sequential NOE (i.e., NH–NH, NH–C_αH) signals observed in the C-terminal region of the sequence (Figures 2 and 3), the assignment of residues 14–36 was straightforward. The backbone connectivities through NH resonances were broken in the N-terminal region at the proline residues

Table II: ^{1}H Chemical Shifts of bPP at 28 °C, pH 4.6

residue	NH	αH	βH	others
Ala1		4.35	1.51	
Pro2		4.53	2.38, 1.99	$^{\gamma}CH_2$ 2.04; $^{\delta}CH_2$ 3.69, 3.56
Leu3	8.42	4.35	1.55, 1.66	$^{\gamma}H$ 1.63; $^{\delta}CH_3$ 0.94, 0.89
Glu4	8.25	3.36	1.60, 1.73	$^{\gamma}CH_2$ 1.83, 2.12
Pro5		4.37	2.42	$^{\gamma}CH_2$ 1.79; $^{\delta}CH_2$ 3.39, 3.03
Glu6	8.63	4.27	1.90	$^{\gamma}CH_2$ 2.25, 2.32
Tyr7	8.70	3.89	3.00, 2.70	(2,6)H 6.86; (3,5)H 6.69
Pro8				
Gly9	8.20	4.36, 3.43		
Asp10	8.53	4.42	2.81, 2.55	
Asn11	8.35	4.81	2.87, 2.71	$^{\gamma}NH_2$ 7.53, 6.86
Ala12	7.02	4.22	1.19	
Thr13	8.47	4.58	4.78	$^{\gamma}CH_3$ 1.38
Pro14		4.29	2.38, 2.28	$^{\gamma}CH_2$ 2.03; $^{\delta}CH_2$ 3.96
Glu15	8.72	4.15	2.10, 1.96	$^{\gamma}CH_2$ 2.45, 2.34
Gln16	7.85	4.07	2.45, 2.49	$^{\gamma}CH_2$ 1.75; $^{\delta}NH_2$ 7.56, 6.83
Met17	8.65	4.55	2.14	$^{\gamma}CH_2$ 2.59, 2.62
Ala18	8.45	4.29	1.62	
Gln19	8.01	4.19	2.18, 2.24	$^{\gamma}CH_2$ 2.48, 2.55; $^{\delta}NH_2$ 7.59, 6.79
Tyr20	8.12	4.51	3.22	(2,6)H 7.14; (3,5)H 6.61
Ala21	8.63	3.69	1.63	
Ala22	7.87	4.07	1.59	
Glu23	8.35	4.08	2.33, 2.43	$^{\gamma}CH_2$ 2.64, 2.05
Leu24	8.89	3.97	1.59	$^{\gamma}H$ 1.25; $^{\delta}CH_3$ 0.69, 0.60
Arg25	8.22	3.89	1.89, 1.96	$^{\gamma}CH_2$ 1.55; $^{\delta}CH_2$ 3.17; $^{\epsilon}NH$ 7.17
Arg26	7.72	4.06	1.97, 1.63	$^{\gamma}CH_2$ 1.85; $^{\delta}CH_2$ 3.26; $^{\epsilon}NH$ 7.43
Tyr27	8.22	4.28	3.29, 3.19	(2,6)H 7.00; (3,5)H 6.65
Ile28	8.54	3.46	1.93	$^{\gamma}CH_3$ 0.99; $^{\gamma}CH_2$ 1.88, 1.09; $^{\delta}CH_3$ 0.95
Asn29	7.90	4.39	2.94, 2.81	$^{\gamma}NH_2$ 7.58, 6.86
Met30	7.82	4.19	2.16	$^{\gamma}CH_2$ 2.55, 2.74
Leu31	7.75	4.07	1.65	$^{\gamma}H$ 1.43; $^{\delta}CH_3$ 0.65, 0.61
Thr32	7.73	4.32	4.32	$^{\gamma}CH_3$ 1.21
Arg33	7.70	4.59	1.84	$^{\gamma}CH_2$ 1.72; $^{\delta}CH_2$ 3.18; $^{\epsilon}NH$ 7.16
Pro34		4.36	2.23, 1.73	$^{\gamma}CH_2$ 1.98; $^{\delta}CH_2$ 3.77, 3.59
Arg35	8.32	4.19	1.69	$^{\gamma}CH_2$ 1.53, 1.47; $^{\delta}CH_2$ 3.13; $^{\epsilon}NH$ 7.14
Tyr36	8.03	4.57	3.09, 2.92	(2,6)H 7.13; (3,5)H 6.80
$^{\alpha}NH_2$	7.49			

Pro2, Pro5, Pro8, and Pro14. The sequential information in each segment was, however, complete. In NOESY spectra, each proline residue with resolved resonances had NOEs between its two δH and the αH protons of the previous residue, and the NH proton of the following residue had an NOE cross-peak linking it to its $C_\alpha H$ resonance. These segments could therefore be joined together and coordinated into the entire sequence. Only one residue, Pro 8, was not identified in this manner; this is the proline whose resonances could not be resolved. A summary of the short- and medium-range NOEs involving NH, $C_\alpha H$, and $C_\beta H$, as well as the $C_\delta H$ of proline is shown in Figure 4. A complete list of assignments is given in Table II.

A diagonal plot of the NOEs observed in bPP is shown in Figure 5a. It is evident from examining this, together with Figure 4, that a stretch of d_{NN} connectivities is observed for residues 15–28. In addition, medium-range $d_{\alpha N}(i,i+3)$ connectivities, as well as many $(i,i+3)$ and $(i,i+4)$ NOEs involving side-chain protons, are present, linking residues 15–32. This indicates the formation of an α-helix in this region of the molecule. The N-terminal segment, containing the three proline residues at the 2, 5, and 8 positions, has virtually no detectable nonsequential NOEs. A number of long-range NOEs were, however, detected between this region and the C-terminal α-helix. Hydrogen exchange data for the amide protons of bPP are also given in Figure 4. Residues running from 17 to 31 were found to have slow to medium exchange rates (see Figure 4 for more details); NH protons from residues 21 to 28 were persistent in D_2O for more than 12 h at pH 4.6, 14 °C. Cross-peaks from these latter residues could be unequivocally identified in NOESY spectra under these conditions. The slow hydrogen exchange rates for these residues are

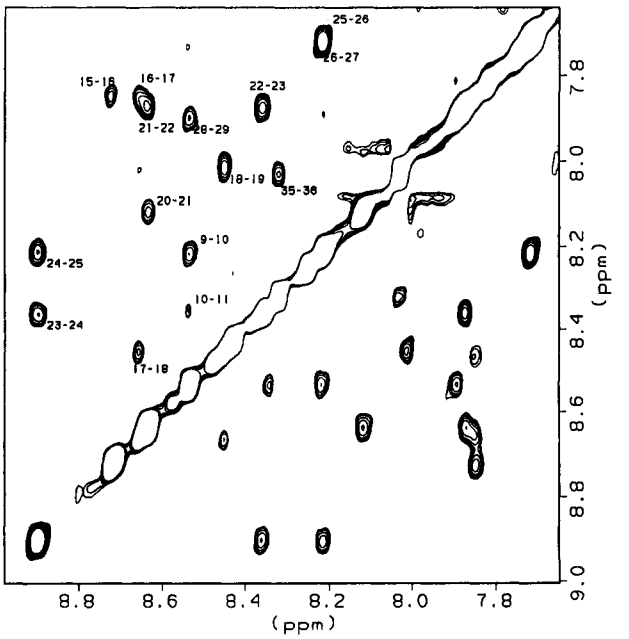


FIGURE 3: NH–NH region of the 200-ms pre-TOCSY NOESY spectrum of bPP, recorded in H_2O at 28 °C. Sequential NH(i)–NH(i+1) NOEs are indicated.

consistent with the formation of hydrogen bonds present in the α-helix. As the amide protons of the first four residues of the α-helix do not participate in hydrogen bonding, the higher exchange rates from residues 15 to 18 are not unexpected. The fact that the residues 29–32 of the α-helix show higher exchange rates may indicate that this end of the helix

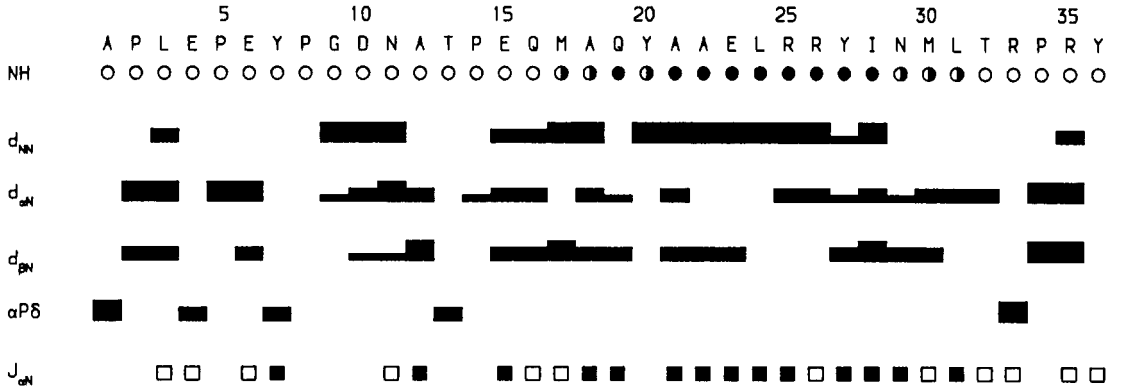


FIGURE 4: Amino acid sequence of bPP, together with a summary of all the short-range NOEs involving the NH, $C_\alpha H$, and $C_\beta H$, as well as the $C_\delta H$ protons of proline residues. Filled, half-filled, and empty circles indicate amides with slow (≥3 h), medium (≥30 min) and fast (<10 min) exchange rates, respectively. The NOEs are classified as strong, medium, or weak by the thickness of the lines. Filled and open boxes indicate, respectively, $J_{NH\alpha}$ coupling constants of less than 6 Hz or larger than 7 Hz.

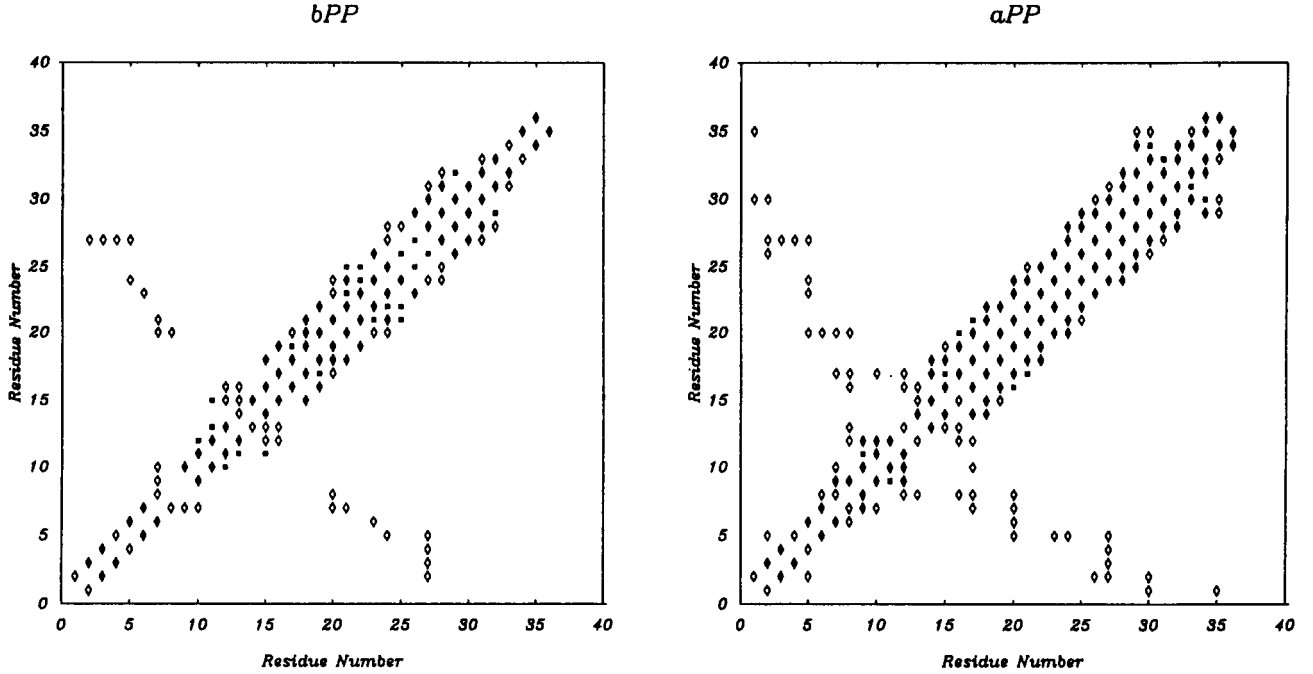


FIGURE 5: NOE pattern of (a) bPP (left) and (b) aPP (right). NOEs of aPP were generated from the crystal structure, with cut-off distances of 4.5 Å for side-chain protons and 4.0 Å for backbone protons, respectively. Solid boxes denote NOEs between backbone protons, and open diamonds denote NOEs between side-chain protons or between side-chain and backbone protons.

is less stable and may be associated with the high mobility of the last four residues of bPP (see later).

Further information about the conformations of proteins in solution can be extracted from measurements of coupling constants. The $^3J_{\alpha\beta}$ coupling constants are related to side-chain conformation (χ^1 torsion angles), and the $^3J_{NH\alpha}$ coupling constants are related to backbone conformation (ϕ torsion angles). $^3J_{\alpha\beta}$ coupling constants for 18 out of 36 residues of bPP have been determined. The values (Table III) show that most of these are averaged to a significant extent (Wagner et al., 1987; Smith et al., 1991) and hence that, except for Asp10 and probably Asn11, no single side-chain rotamer can be defined from the data. It was therefore decided not to attempt stereospecific assignments for the residues of bPP. $^3J_{NH\alpha}$ coupling constant data were obtained by fitting the cross sections of the NH–C_αH cross-peaks from a 4K COSY spectrum (Redfield & Dobson, 1990). The results, shown in Figure 4, are consistent with the α-helical conformation in the C-terminal region; the fact that some of the $^3J_{NH\alpha}$ values are slightly larger than would be expected from a standard α-helical conformation may partly result from the curvature of the helix present in bPP (see later).

Table III: $J_{\alpha\beta}$ Coupling Constants of bPP Measured at 28 °C, pH 4.6

residue	$J_{\beta\beta}$	$J_{\alpha\beta_1}$	$J_{\alpha\beta_2}$
Pro2	−13	5.4	8.0
Leu3	−15	7.0	
Pro5	−13.7	7.26	8.0
Glu6		5.5	
Tyr7	−15.3	11.7	6.5
Asp10	−18.0	10.9	4.7
Asn11	−15.9	8.4	5.4
Ala12		8	
Thr13		5	
Glu15	−14.7	8.2	9.0
Met17		6	
Gln19	−17	8	8
Glu23	−14	6	12
Tyr27	−15.2	6.25	10.2
Asn29	−16.0	6.0	9.5
Arg33	−15.5	12.1	
Pro34	−13.5	6.8	7.8
Tyr36	−13.8	8.8	6.5

Tertiary Structure Determination. For the structure calculations a set of 352 NOEs was used. This consisted of 204 short-range ($|i - j| < 5$) and 23 long-range ($|i - j| \geq 5$) in-

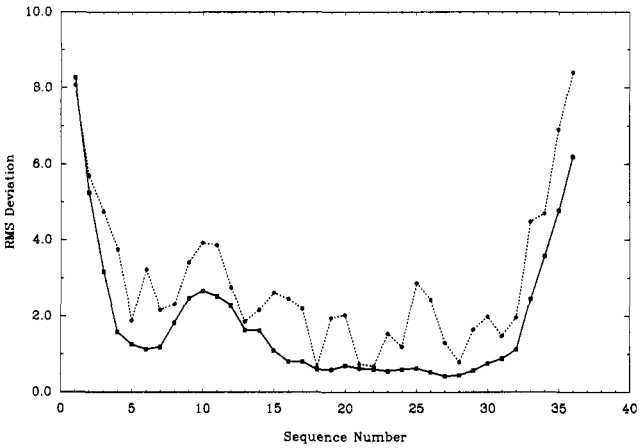


FIGURE 6: Average atomic RMS deviations of individual residues versus sequence number. Backbone (N, C_α, C) RMSD values are indicated by the solid line with filled squares at each sequence position; side-chain RMSD values are indicated by the dashed line with filled circles.

terresidue NOEs and 125 intraresidue NOEs. These NOEs were first identified in 200-ms NOESY spectra and then, from the shorter mixing time NOESY spectra, classified into four distance ranges, namely, 1.8–2.5, 1.8–3.0, 1.8–3.5, and 1.8–5.0 Å, corresponding to strong, medium strong, medium, and weak NOEs. As no stereospecific assignments were defined, the upper limits of distances involving a group of two or more protons, as well as the aromatic protons, were modeled by using pseudoatoms (Wüthrich, 1986). In the case of NOEs involving methyl groups, a further 0.5 Å was added to account for the higher apparent intensity of the methyl resonances (Wagner et al., 1987). A list of all the NOE information used in the structure calculations is given in the Supplementary Material. In addition to the NOE distance restraints, 26 distance restraints for 13 backbone NH–CO hydrogen bonds were included in the input distance restraints list at the dynamic simulated annealing phase. Restraints were set to 1.8 ± 0.5 Å for NH–O and 2.8 ± 0.5 Å for N–O distances, respectively, on the basis of the secondary structure and hydrogen exchange data presented above.

A batch of 25 structures was generated in the distance geometry phase of the calculation. In order to overcome the local image problem, these structures were inverted to produce their global images; 50 structures were therefore processed in the subsequent dynamic annealing phase. Some of the structures were of high energy at the end of this procedure, probably because they were the wrong image; these structures were picked out and discarded before the full dynamics calculation was carried out. A total of 19 structures resulted from this calculation. These structures all had only small NOE violations, as indicated by the low E_{NOE} term, and nonbonded contacts were very good (Table IV). The average RMS deviation for C_α atoms of residues 4–8 and 15–32 between the individual structures and the mean structure (obtained by averaging the coordinates of individual structures), is 0.95 ± 0.29 Å. A plot showing the average RMS deviation of the individual residues versus sequence is presented in Figure 6. It is clearly seen from this and from Figure 7 that the backbone conformation of the C-terminal helix (15–32) is very well defined and that the N-terminal segment (4–8) maintains a relatively stable position with respect to this helix. It is evident, however, that the N-terminal region is less well defined than the C-terminal helical region. The turn region (9–14) is poorly defined relative to the helix, as indicated by its large RMS deviations (Figure 6). The four residues both at the N-terminus (1–4) and at the C-terminus (33–36) are highly disordered; that these residues of the molecule are mobile in solution, rather than simply undefined by the experimental NMR data, is supported by their near random-coil chemical

Table IV: Statistics of the Structure Calculations[a]

structural statistics	⟨SA⟩	structural statistics	⟨SA⟩
RMS deviations		E_{NOE} (kcal/mol)	28.4 ± 5.4
NOEs (Å)	0.080 ± 0.008	E_{bond} (kcal/mol)	5.7 ± 0.8
bonds (Å)	0.0075 ± 0.0005	E_{angle} (kcal/mol)	127.6 ± 7.6
angles (deg)	2.85 ± 0.09	E_{impr} (kcal/mol)	2.6 ± 0.6
impropers (deg)	0.30 ± 0.04	E_{VDW} (kcal/mol)	1.6 ± 0.3

[a] ⟨SA⟩ are the final 19 structures. The values listed in the right-hand column are the average over the ensemble. The RMS deviations of NOEs from the experimental restraints are calculated with respect to the upper and lower bounds of the distance restraints. RMS deviations of bond lengths, angles, and improper dihedrals are calculated with respect to idealized geometries. E_{NOE} is calculated with a force constant of 10 kcal/(mol·Å^2) and E_{VDW} with a force constant of 10 kcal/(mol·Å^4). Energy terms from bonds, angles and improper dihedrals are calculated with standard XPLOR topology files.

FIGURE 7: Overlap of the C_α atom positions of the 19 bPP structures calculated from the NMR data. The structures are aligned to give a best fit for residues 4–8 and 15–32 to the mean structure. The average RMS deviation for these C_α atoms is 0.95 ± 0.29 Å.

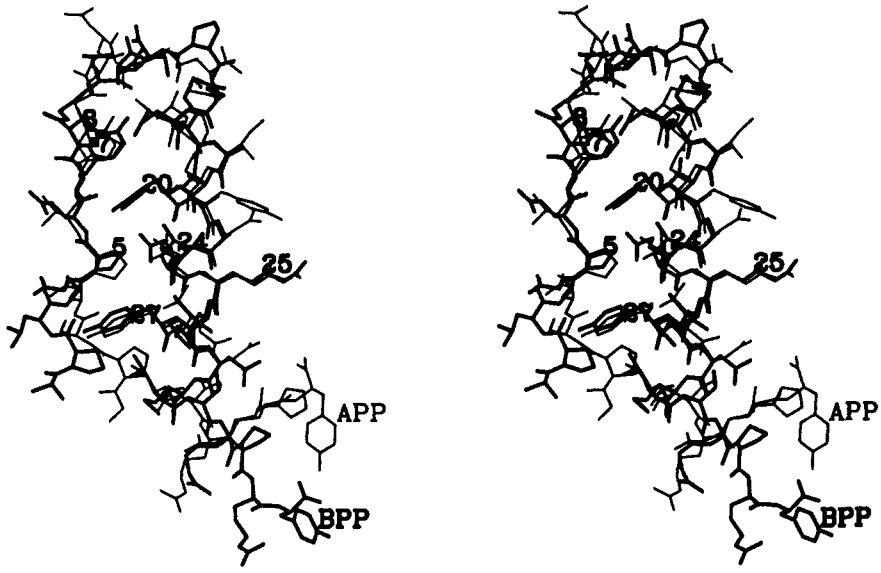
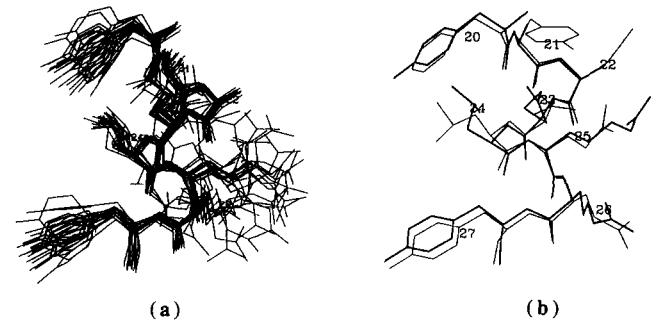


FIGURE 8: Backbone overlap of the bPP mean structure in solution with the aPP crystal structure, superimposed at residues 4–8 and 15–32. The RMS deviation between the two structures is 0.65 Å for the atoms superimposed. Side chains from helical residues (20–28) and the structurally important residues in the N-terminal region, Pro5, Tyr7, and Pro8, are also shown.

shifts, averaged J values, and relatively narrow line widths. The line widths of the NH resonances from these terminal residues, for example, are 7.3 ± 1.2 Hz, the narrowest (5.4 Hz) being from Y36, compared to values of 11.6 ± 1.7 Hz for the helical residues.

Discussion

Generation of the structure of bPP in solution has enabled us to make a detailed comparison between it and the crystal structure of aPP, which has been refined to very high resolution (0.98 Å) (Glover et al., 1983). First, it is of interest to see what kind of NOE pattern would be expected from the aPP structure. Figure 5b shows an NOE diagram of aPP, calculated from the crystal structure with a cut-off distance of 4.5 Å; this indicates the NOEs which would be expected for aPP were its structure to be maintained in solution. The overall resemblance between this figure and Figure 5a, which shows the experimental results for bPP, is remarkable. It is interesting to note that at this cut-off distance no nonsequential NOEs involving residues 2–8, which form a polyproline type II helix, are predicted from the aPP crystal structure. This implies that inspection of NOEs from these residues is insufficient to detect this type of secondary structure in solution. The close resemblance of long-range NOE patterns observed for bPP and predicted from the aPP structure suggests strongly that the main-chain folds of the two molecules must be closely similar. This is clearly seen when the mean NMR structure calculated for bPP is compared with the X-ray structure of aPP (Figure 8). The α-helical region (residues 15–32) can be seen to be essentially identical. The helical curvature reported for aPP (Glover et al., 1983) can be observed here for bPP as well; this has been attributed to the effect of water molecules participating in hydrogen bonding both to the side chains and also to the carbonyl atoms along the hydrophilic face of the helix. The presence of the helical curvature in aPP makes the predicted $^3J_{NH\alpha}$ values for residues 14–32 about 1 Hz greater than would be expected from a standard α-helical conformation. A larger deviation of the bPP solution structure from that of aPP in the crystalline state occurs in the N-terminal segment of the molecule. In the crystal structure of aPP, a polyproline helix is identified between residues 2 and 8; although the present NMR data are unable to define directly the existence of this secondary element in the solution structure



(a) (b)

FIGURE 9: (a, left) Overlap of residues 20–27 of 19 bPP structures, aligned to give a best fit to the mean structure in this region; only backbone atoms were used in the superposition (RMS deviation 0.37 ± 0.11 Å). (b, right) Overlap of residues 20–27 of the bPP mean structure (thicker line) with the aPP crystal structure; again only backbone atoms were used in the superposition (RMS deviation 0.22 Å).

of bPP, the conformation of the mean bPP structure is similar to that of the aPP crystal structure in this region. Despite the fact that a large number of side chains of bPP undergo conformational averaging in solution and that the sequence identity between bPP and aPP is rather low (39%), it is interesting to see (Figure 9a,b) that the averaged side chains from the α-helical region of the bPP molecule in solution have very similar orientations to those of the respective side chains in the aPP crystal structure. It is also apparent that side chains from residues such as Tyr20, Leu24, and Tyr27, which are located in the region of the molecule which makes contacts with the N-terminal segment, are better defined than those from residues such as Arg25 and Arg26, which are directed away from the interface. This phenomenon is reminiscent of the greater ordering of internal residues compared to surface residues found for globular proteins (Smith et al., 1991). The side chains from the N-terminal segment of bPP are more disordered; this is due at least in part to the limited extent to which the side chains from this part of the bPP molecule could be defined.

Secondary, or conformational dependent, shifts were calculated by subtracting the experimental chemical shifts from values anticipated for a random-coil structure; a positive value therefore indicates an upfield shift and a negative value a downfield shift. In order to show the trend of these values,

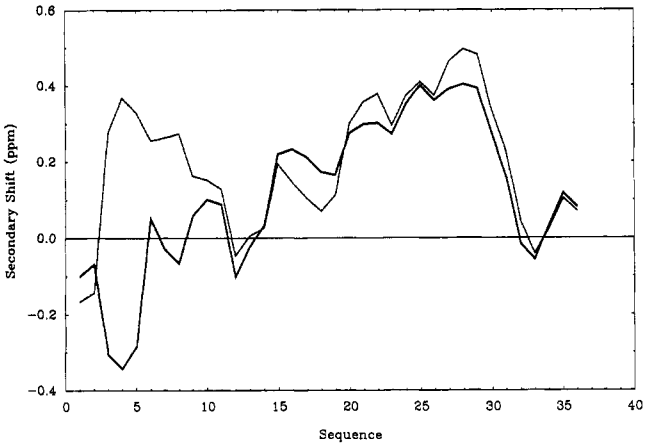


FIGURE 10: Secondary shifts of the C_α protons of bPP. Each data point is averaged with Gaussian weighting factors over the two neighboring residues as described in the text. Shifts with corrections for ring current effects are indicated by the solid line; shifts without corrections for ring current effects by the dashed line. Ring current shifts were calculated as described previously (Hoch et al., 1982; Hoch, 1983).

the secondary shifts were averaged over their two neighboring data points after applying a Gaussian weighting function; secondary shifts for $C_\alpha H$ protons of bPP calculated in this way are plotted as a function of sequence number in Figure 10. It is known (Jardetzky & Roberts, 1981; Szilágyi & Jardetzky, 1989) that, among other effects, α-helix formation will result in upfield shifts of $C_\alpha H$ resonances. This is clearly seen here for residues from the α-helix of bPP. The experimental upfield shifts of resonances from residues in the N-terminal part of the molecule, however, appear to result from their proximity to aromatic rings. This is particularly noticeable for Glu4, with a $C_\alpha H$ proton resonance upfield shift of 0.93 ppm from its random-coil position. This is supported by ring current shift calculations based on the bPP mean structure, which suggest that the shift arises predominately from contacts with Tyr27. That the calculated shift (1.95 ppm) exceeds the experimentally observed value may well be a consequence of the terminal residues of the bPP structure in solution being highly disordered.

In the calculation of the bPP structure, we have considered so far that the molecule exists as a single monomeric species in solution. In fact, it is more likely that it exists as a dimer; dimerization in solution has been described for mammalian PPs (Noelken et al., 1980; Chang et al., 1980) and for NPYs (Minakata et al., 1989). It has not, however, been possible to identify for bPP any NOEs that can be unequivocally identified as coming from dimeric interactions. In the crystalline state, aPP has close dimeric interactions involving both the polyproline and α-helices as well as the turn region. These involve side-chain contacts only, and several residues, notably Tyr7 and Phe20, are particularly important in this regard. In bPP, residue 20 is tyrosine, not phenylalanine as in aPP, and the following residue, Ala21, is tyrosine in aPP. Tyr7 shows NOEs to Tyr20 in our study, but these are expected on the basis of the fold of the monomeric structure. There are, however, some weak NOEs involving both Tyr7 and Tyr20 which have not been assigned and could arise from the dimeric contacts. This aspect of the structure has not, therefore, been resolved. Furthermore, some nonpolar side chains (e.g., Leu24) from the α-helical segment of bPP point away from the helix, in the same manner as is seen in the aPP crystal structure; this feature of the structure is likely to be more favorable in a dimeric molecule. Analysis of the aPP dimer contacts suggests, however, that failure to distinguish dimer and monomer NOEs is unlikely to have significant influence on the experimentally determined overall fold of the molecule.

The finding in the present NMR study that the bPP molecule in solution exists in a folded structure which is similar to the aPP crystal structure has important implications for the biological function of PP-fold peptides. Structure–function studies, mainly performed on the NPY molecule, have shown that the receptors require both ends of the PP-fold molecules. Substitutions which prevent the formation of, for example, the α-helix have dramatic effects on receptor binding of the peptide (Fuhlendorff et al., 1990). Furthermore the PP-fold can be exchanged among different PP-fold peptides with only minor effects on receptor binding, as long as the terminal sequences are preserved. The PP-fold, or at least up to 20 amino acids of it, has even been exchanged by synthetic nonpeptide bridging constructs without loss of biological activity (Fuhlendorff et al., 1990; Krstenansky & Buck, 1987). The PP-fold has been suggested, therefore, to function mainly as a structural scaffold which brings the C- and N-terminal segments of the molecule in close proximity and presents the combined moiety to the receptor (Schwartz et al., 1990; Krstenansky & Buck, 1987). The present finding that this folded structural arrangement is stable and well defined in solution is consistent with this proposal. It is of particular interest in this regard, however, that the residues presented to the receptor are, in contrast to those defining the overall fold, in the highly flexible region of the bPP molecule in solution. Whether or not these dynamic properties are significant for the specificity of the interaction remains to be established.

ACKNOWLEDGMENTS

We are very grateful to Dr. S. P. Wood and Prof. T. L. Blundell for suggesting this project, for encouragement of the work, and for providing samples for preliminary studies and the coordinates of the aPP crystal structure. We acknowledge valuable contributions to preliminary studies of aPP and bPP from Drs. M. J. Bogusky, M. Howarth, J. E. Pitts, and P. H. Rehse. We thank Dr. J. Boyd, Dr. C. Redfield, and N. Soffe for their assistance with the NMR experiments; Dr. C. Refield also kindly provided and helped with the fitting program for extracting $^3J_{NH\alpha}$ coupling constants. X.L. thanks the British Council for a Sino–British Friendship Scholarship. C.M.D. is a member of the Oxford Centre for Molecular Sciences. M.J.S. is a Royal Society University Research Fellow.

SUPPLEMENTARY MATERIAL AVAILABLE

A table listing the NOEs identified for bPP and used in the structure calculations along with coordinates for the calculated mean structure and the RMSD values for each atom over the 19 structures (18 pages). Ordering information is given on any current masthead page.

Registry No. bPP, 76415-72-0.

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.

Blundell, T. L., Pitts, J. E., Tickle, I. L., Wood, S. P., & Wu, C.-W. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4175–4179.

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

Brunger, A. T., Kurian, J., & Karplus, M. (1987) *Science 235*, 458–460.

Chang, P. J., Noelken, M. E., & Kimmel, J. R. (1980) *Biochemistry 19*, 1844–1849.
Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2821–2822.
Fuhlendorff, I., Johansen, N. L., Melberg, S. G., Thoegersen, H., & Schwartz, T. W. (1990) *J. Biol. Chem. 265*, 11706–11712.
Glover, I. D., Haneef, I., Pitts, J. E., Wood, S. P., Moss, D., Tickle, I. J., & Blundell, T. L. (1983) *Biopolymers 22*, 293–304.
Glover, I. D., Barlow, D. J., Pitts, J. E., Wood, S. P., Tickle, I. J., Blundell, T. L., Tatemoto, K., Kimmel, J. R., Wollmer, A., Strassburger, W., & Zhang, Y. S. (1985) *Eur. J. Biochem. 142*, 379–385.
Griesinger, C., Sørensen, O. W., & Ernst, R. R. (1987) *J. Magn. Reson. 75*, 474–492.
Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870–7872.
Havel, T. F., & Wüthrich, K. (1984) *Bull. Math. Biol. 46*, 673–698.
Hoch, J. C. (1983) Ph.D. Thesis, Harvard University.
Hoch, J. C., Dobson, C. M., & Karplus, M. (1982) *Biochemistry 21*, 1118–1125.
Hoffmann, J. A., & Chance, R. E. (1983) *Biochem. Biophys. Res. Commun. 116*, 830–835.
Howarth, M. A. (1986) D.Phil. Thesis, University of Oxford.
Jardetzky, O., & Roberts, G. C. K. (1981) *NMR in Molecular Biology*, Academic Press, New York.
Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.
Kimmel, J. R., Pollock, H. G., & Hazelwood, R. L. (1968) *Endocrinology 83*, 1323–1330.
Krstenansky, J. L., & Buck, S. H. (1987) *Neuropeptides 10*, 77–85.
Krstenansky, J. L., Owen, T. J., Buck, S. H., Hagaman, K. A., & Mclean, L. R. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4377–4381.
Macura, S., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259–281.
Minakata, H., & Iwashita, T. (1990) *Biopolymers 29*, 61–67.
Minakata, H., Taylor, J. W., Walker, M. W., Miller, R. J., & Kaiser, E. T. (1989) *J. Biol. Chem. 264*, 7907–7913.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317–324.
Noelken, M. E., Chang, P. J., & Kimmel, J. R. (1980) *Biochemistry 19*, 1838–1843.
Otting, G., & Wüthrich, K. (1987) *J. Magn. Reson. 75*, 546–549.
Redfield, C., & Dobson, C. M. (1990) *Biochemistry 29*, 7201–7214.
Rehse, P. H. (1989) Ph.D. Thesis, University of London.
Saudek, V., & Pelton, J. T. (1990) *Biochemistry 29*, 4509–4515.
Schwartz, T. W. (1983) *Gastroenterology 85*, 1411–1425.
Schwartz, T. W., & Tager, H. S. (1981) *Nature 294*, 589–591.
Schwartz, T. W., Sheikh, S. P., & O'Hare, M. M. T. (1987) *FEBS Lett. 225*, 209–214.
Schwartz, T. W., Fuhlendorff, J., Langeland, N., Thøgersen, H., Jørgensen, J. C., & Sheikh, S. P. (1989) *Neuropeptide Y, XIV Nobel Symposium. V.* (Mutt, T., Hökfelt, Fuxe, K., & Lundberg, J., Eds.) pp 143–151, Raven Press, New York.
Schwartz, T. W., Fuhlendorff, J., Kjems, L. L., Kristensen, M. S., Vervelde, M., O'Hare, M. M. T., Krstenansky, J. L., & Bjørnholmn, B. (1990) *Ann. N.Y. Acad. Sci. 611*, 35–47.
Smith, L. J., Sutcliffe, M. J., Redfield, C., & Dobson, C. M. (1991) *Biochemistry 30*, 987–996.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Sutcliffe, M. J., & Dobson, C. M. (1991) *Proteins 10*, 117–129.
Sutcliffe, M. J., Hayes, F. R. F., & Blundell, T. L. (1987) *Protein Eng. 1*, 385–392.
Szilágyi, L., & Jardetzky, O. (1989) *J. Magn. Reson. 83*, 441–449.
Tonan, K., Kawata, Y., & Hamaguchi, K. (1990) *Biochemistry 29*, 4424–4429.
Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Gō, N., & Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611–639.
Whitecomb, D. C., Taylor, I. L., & Vigna, S. R. (1990) *Am. J. Physiol. 259*, G687–G693.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.