Rosowsky, A., & Yu, C.-S. (1978) *J. Med. Chem. 21*, 170–175.
Rosowsky, A., Forsch, R., Uren, J., & Wick, M. (1981) *J. Med. Chem. 24*, 1450–1455.
Schweisfurth, H., Reinhard, E., Heinrich, J., & Brugger, E. (1983) *J. Clin. Chem. Clin. Biochem. 21*, 605–609.
Seeger, D. R., Cosulich, D. B., Smith, J. M., & Hultquist, M. E. (1949) *J. Am. Chem. Soc. 71*, 1753–1758.
Sherwood, R. F., Melton, R. G., Alwan, S. M., & Hughes, P. (1985) *Eur. J. Biochem. 148*, 447–453.
Sirotnak, F. M., Chello, P. L., Piper, J. R., Montgomery, J. A., & DeGraw, J. I. (1979) in *Chemistry and Biology of Pteridines* (Kisliuk, R. L., & Brown, G. M., Eds.) pp 597–608, Elsevier North-Holland, New York.
Vitols, K. S., Montejano, Y. D., Kuefner, U., & Huennekens, F. M. (1989) *Pteridines 1*, 65–69.

Toward the Solution Structure of Human Insulin: Sequential 2D ^{1}H NMR Assignment of a Des-pentapeptide Analogue and Comparison with Crystal Structure†

Qingxin Hua‡,§,⊥ and Michael A. Weiss*,‡,§,∥

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, Francis Bitter National Magnet Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts 02114

Received June 22, 1990; Revised Manuscript Received August 15, 1990

ABSTRACT: 2D ^{1}H NMR studies are presented of des-pentapeptide-insulin, an analogue of human insulin lacking the C-terminal five residues of the B chain. Removal of these residues, which are not required for function, is shown to reduce conformational broadening previously described in the spectrum of intact insulin [Weiss et al. (1989) *Biochemistry 28*, 9855–9873]. This difference presumably reflects more rapid internal motions in the fragment, which lead to more complete averaging of chemical shifts on the NMR time scale. Sequential ^{1}H NMR assignment and preliminary structural analysis demonstrate retention in solution of the three α-helices observed in the crystal state and the relative orientation of the receptor-binding surfaces. These studies provide a foundation for determining the solution structure of insulin.

Insulin, a small polypeptide hormone, provides a model system for biophysical studies of protein folding, dynamics, and recognition. The insulin fold is highly conserved among vertebrate and invertebrate kingdoms, defining an ancestral family of insulin-like regulatory proteins (Blundell & Humbel, 1980). Insulin is composed of two polypeptide chains, the A chain (21 residues) and the B chain (30 residues) linked by two disulfide bonds. As a foundation for comparative studies of insulin-related peptides, we describe complete 2D NMR resonance assignment of an analogue, des-pentapeptide[B26–B30]-insulin (DPI)[1] (Gattner, 1975; Danho et al., 1975; Rieman et al., 1983; Wang & Tsou, 1986). The deleted residues B26–B30 are not generally conserved, and in their absence an amidated analogue exhibits full receptor-binding activity (Fisher et al., 1985; Casaretto et al., 1987; Mirmira & Tager, 1989). An analysis of the aromatic spectrum of DPI at pH 10 has previously been reported (Hua et al., 1989).

The present experiments are based on previous ^{1}H NMR studies of human insulin and proinsulin in 80% H_2O/20% acetic acid–sodium acetate, pH 3 (Cheshnovsky et al., 1983; Weiss et al., 1989, 1990). This solvent system was found to weaken insulin self-association, enabling the monomer to be studied under conditions in which native structure is retained. A remarkable feature of the ^{1}H NMR spectrum of the insulin monomer under these conditions is the extensive variation observed in the linewidths of amide resonances. Such variation is related to exchange among conformational substates. The existence of substates is an intrinsic feature of the insulin fold and is not related to the particular solvent system used in these studies (Weiss et al., 1990). Although of intrinsic interest in relation to insulin dynamics (Chothia et al., 1983; Kruger et al., 1987), exchange-mediated line broadening was found in the earlier studies to limit the application of established sequential ^{1}H NMR assignment methods. This limitation may be overcome by using less aqueous solvent conditions; sequential assignment of insulin has recently been described in 35% acetonitrile (Kline & Justice, 1990). In this paper an alternative strategy is described, based on systematic screening of insulin analogues.

Our results are presented in three parts. In part I the structural effects of successive C-terminal deletions in the B

† This work was supported in part by grants from the National Institutes of Health, American Diabetes Association, and Juvenile Diabetes Foundation International to M.A.W. The Harvard NMR Center was funded by a National Institutes of Health shared instrumentation grant (1 S10 RR04862-01) and an award from the National Health Resources Foundation. The High-Field NMR Resource of the Francis Bitter National Magnet Laboratory at MIT is supported by the National Institutes of Health (RR-00995). M.A.W. is supported in part by the Pfizer Scholars Program for New Faculty and an American Cancer Society Junior Faculty Research Award.

* Address correspondence to this author at Harvard Medical School.
‡ Harvard Medical School.
§ Massachusetts Institute of Technology.
⊥ Permanent address: Institute of Biophysics, Academia Sinica, Beijing, China.
∥ Massachusetts General Hospital.

[1] Abbreviations: DOI, des-octapeptide-insulin; DPI, des-pentapeptide insulin; DPA, des-pentapeptide-insulin amide; NOE, nuclear Overhauser enhancement; NOESY, 2D NOE spectroscopy; photo-CIDNP, photochemically induced dynamic nuclear polarization; TOCSY, 2D isotropic mixing (Hartmann–Hahn) spectroscopy; DQF-COSY, double-quantum filtered 2D correlation spectroscopy.

0006-2960/90/0429-10545$02.50/0

chain are examined. It is shown that overall features of the insulin fold are retained following deletion of five to eight residues. Because these analogues exhibit less extensive conformational broadening than does native insulin, they provide tractable systems for 2D NMR study. In part II we focus on one such analogue, des-pentapeptide-insulin (DPI). DPI retains the structural information required for high-affinity receptor recognition (Gattner, 1975; Danho et al., 1975; Rieman et al., 1983; Wang & Tsou, 1986) and has been analyzed at high resolution by X-ray crystallography (Bi et al., 1984; Dai et al., 1987). Sequence-specific ^{1}H NMR assignment and secondary-structure determination are described in part III. This analysis demonstrates that the three α-helices observed in the crystal state (Blundell et al., 1971; Bi et al., 1984; Dai et al., 1987) are retained in solution. Long-range NOEs are observed involving residues in the receptor-binding surface; these are also in general accord with the crystal model. These results validate the solvent conditions used in this previous NMR studies (Weiss et al., 1989, 1990) and provide a foundation for modeling the structure and dynamics of insulin in solution.

Materials and Methods

Biosynthetic human insulin was provided as zinc crystals by Eli Lilly and Co. (Indianapolis, IN). Zinc was removed by gel filtration (Sephadex G-25) in 1% acetic acid. The solution was lyophilized and redissolved in NMR buffer (below); the protein concentration was determined by UV absorbance at 278 nm, with the assumption that a 1 mg/mL solution has an absorbance of 1.05/cm (Frank & Veros, 1968). Insulin analogues lacking the C-terminal five to eight residues of the B chain were generously provided by Dr. R. E. Chance and Dr. B. H. Frank (Eli Lilly and Co., Indianapolis, IN) and used without further purification. NMR buffer consists of 20% CD_3COOD and 80% D_2O or H_2O. The pD (pH) was adjusted in the range 1.9–3.0 (direct meter reading) with aliquots of 20% CD_3COOD/80% D_2O (H_2O) containing 1 M NaOD (NaOH). 2D NMR spectra were recorded at 500 MHz at the Harvard Medical School NMR Center; photo-CIDNP spectra were recorded at the Francis Bitter National Magnet Laboratory (MIT) as previously described (Weiss et al., 1989). Two-dimensional experiments were performed by the pure-phase method of States et al. (1982) and processed as described in the figure captions. Isotropic mixing spectra (TOCSY) were obtained by the method of Davis and Bax (1985). Spectra were obtained in H_2O by solvent presaturation.

Results

Our results are presented in three parts. In part I the structural effects of successive C-terminal deletions in the B chain are evaluated in a series of insulin analogues. We then focus on one such analogue, des-pentapeptide-insulin (DPI), for detailed study. ^{1}H NMR spin system identification of DPI is presented in part II; sequential assignment and initial structural characterization is presented in part III.

(I) Des-pentapeptide-insulin as Model System

(i) Overview of C-Terminal Deletions in the B Chain. In the crystal state the C-terminal region of the B chain forms an extended β-strand as shown in Figure 1. Residues B24–B30 interact with dimer-related residues B24′–B30′ to form a antiparallel β-sheet (Adams et al., 1969; Blundell et al., 1971). Although residues B26–B30 may be removed without loss of biological activity (Fisher et al., 1985; Casaretto et al., 1987; Mirmira & Tager, 1989), residues B23–B25 (Gly-Phe-Phe) are highly conserved and appear to form part of the receptor-binding surface (Pullen et al., 1976; De Meyts et al., 1978; Baker et al., 1988). Mutations at positions B24 and B25 in general decrease receptor affinity and are associated with diabetes mellitus in man (Tager et al., 1979, 1980; Keefer et al., 1981; Wollmer et al., 1981; Kobayashi et al., 1982; Shoelson et al., 1983a,b; Haneda et al., 1984, 1985; Nakagawa & Tager, 1986, 1987).

^{1}H NMR spectra of insulin analogues containing successive C-terminal B-chain deletions are shown in Figure 2. The spectrum of intact human insulin is shown in panel A; analogues lacking five to eight residues are shown in panels B–E, respectively. As expected, C-terminal resonances are absent in the spectra of the deletion mutants, as indicated in the figure (T, the methyl resonances of ThrB28 and ThrB30; arrow, the meta resonance of TyrB26; asterisk, the ortho resonance of PheB24). Remarkably, each of these spectra exhibits similar dispersion of chemical shifts, indicating a stably folded structure. Corresponding long-range nuclear Overhauser enhancements (NOEs) are also observed; in the case of DPI these are discussed in detail below.

(ii) Photo-CIDNP. Maintenance of tertiary structure may also be investigated by photochemically induced dynamic nuclear polarization (photo-CIDNP). The photo-CIDNP effect requires collision between the triplet-state dye and an accessible tyrosine hydroxyl group, which leads in turn to formation of a radical pair (Kaptein, 1980). Hence, relative tyrosine accessibilities (and inaccessibility) following C-terminal deletion may be compared with those obtained in previous photo-CIDNP studies of intact insulin (Muszkat et al., 1984; Weiss et al., 1989). The photo-CIDNP spectrum of the insulin monomer exhibits nearly equal enhancement of three accessible tyrosines (B16, B26, and A14). No enhancement is observed for TyrA19, reflecting its position in an internal crevice as illustrated in Figure 1 (Blundell et al., 1971; Baker et al., 1988). However, A19 enhancement is observed following thermal unfolding (Weiss et al., 1989), indicating that its inaccessibility in the native state provides a probe for tertiary structure.

The photo-CIDNP spectra of des-pentapeptide-insulin (DPI) and des-octapeptide-insulin (DOI) are shown in Figure 3. "Dark" spectra of DPI and DOI are shown in panels A and A′, respectively; spectra following laser irradiation are shown in panels B and B′; and respective difference spectra are shown in panels C and C′. In the absence of B26 two major enhancements are observed; these are assigned to TyrA14 (labeled a in the figure) and TyrB16 (labeled b) by analogy to the spectrum of insulin. Although the meta resonance of A19 is not resolved from that of B16 in either spectrum, integration of the difference spectrum indicates that its photo-CIDNP enhancement (c) is less than 10% relative to A14 or B16.

These ^{1}H NMR and photo-CIDNP studies indicate that the hydrophobic core of insulin is stably folded in the absence of the C-terminal portion of the B chain. This result is in accord with the divergence of B-chain sequences and B-chain length in the evolution of the insulin-related gene family (Blundell & Humbel, 1980); the B chain of porcine relaxin, for example, appears to end at position B23 (Bedarkar et al., 1977). Thus, these truncated insulin analogues provide valid models for defining the general features of insulin as a motif of protein folding (Wang & Tsou, 1986). We focus below on one such analogue, DPI, which also retains structural information required for receptor recognition (Riemen et al., 1983; Fisher et al., 1985; Casaretto et al., 1987; Mirmira & Tager, 1989).

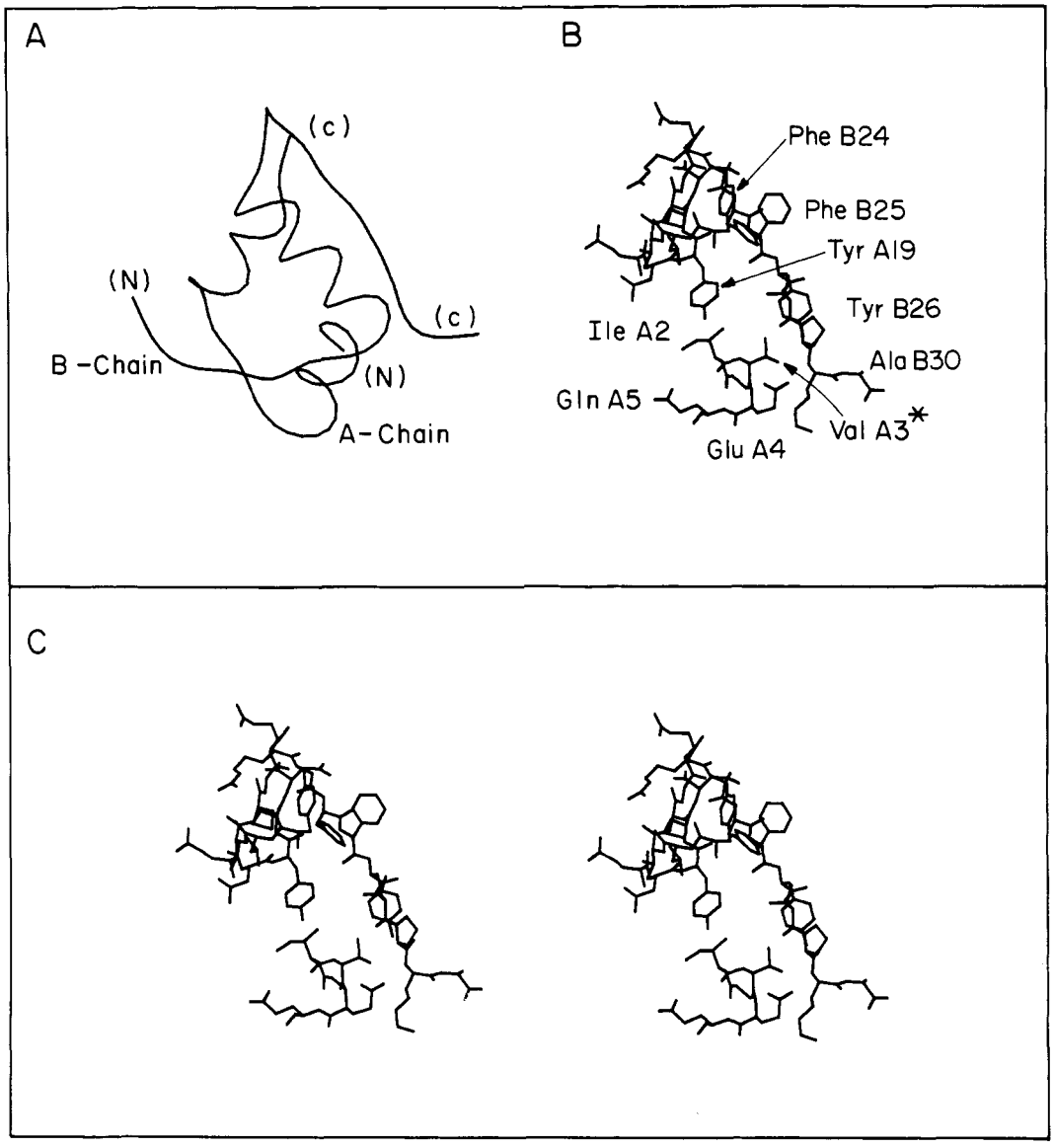

FIGURE 1: (A) Ribbon representation of molecule 2 in the two-Zn structure of porcine insulin (Blundell et al., 1971). (B) α-Carbon representation with selected side chains illustrating the C-terminal region of the B chain and neighboring portions of the A chain. (C) Stereorepresentation of panel B. Diabetes-associated mutations have been identified at positions A3 (asterisk), B24, and B25 (Haneda et al., 1984).

(iii) Comparison of DPI and Intact Human Insulin. The amide and aromatic resonances of DPI (panel A) and intact human insulin (panel B) are shown in Figure 4. These spectra exhibit an overall similarity, including four well-resolved amide resonances that are shifted downfield (labeled a–d in the figure). However, the resonances in the spectrum of DPI are in general sharper than those in the spectrum of intact insulin. This difference (i.e., less rapid T_2 relaxation in the fragment) is more pronounced following a Carr–Purcell–Meiboom–Gill pulse train with 20-ms transverse relaxation (data not shown).

There are two possible mechanisms for more significant line broadening in intact insulin. At high protein concentrations (>2 mM) partial dimerization of insulin (but not DPI) occurs under the conditions of study, as monitored by UV difference and CD spectroscopy (Weiss et al., 1989). These optical perturbations reflect ordering of the B-chain tyrosines (B16 and B26) in the dimer interface (Rupley et al., 1967; Wood et al., 1975; Strickland & Mercola, 1976); in the absence of residues B26-B30 dimerization is not observed (Liang et al., 1985). In addition, monomer-specific broadening of selected amide resonances is less extensive in the spectrum of DPI than in the spectrum of insulin at low protein concentrations (Weiss et al., 1989, 1990). This second mechanism appears to be related to intermediate exchange among conformational substates. Presumably, more rapid motions in the fragment lead to more complete averaging of chemical shifts on the NMR time scale. The overall similarity in chemical shifts between DPI and insulin is consistent with the overall similarities in their crystal structures (Bi et al., 1984; Dai et al., 1987; Baker et al., 1988).

(II) DPI Spin System Identification

The primary structure of DPI is shown in Figure 5; the A chain contains 21 residues, and the B chain contains 25 residues. Their NMR spin systems may be classified as (i) methyl containing (A, I, L, T, and V), (ii) long methylene chains (E, Q, R), (iii) aromatic residues (F, H, and Y), (iv) the remaining AMX spin systems (C, N, and S), and (v) glycine. These groups of spin systems are discussed in turn. DPI contains no aspartic acid, lysine, methionine, or proline residues.

In the following analysis a single spin system was observed for each residue in the protein, consistent with the presence of a well-defined oligomeric state (the monomer). In the case of insulin and its derivatives this is not a trivial observation.

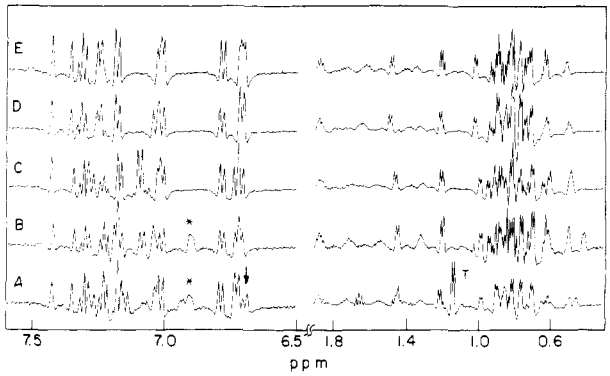

FIGURE 2: (A) ^1H NMR spectrum of 1 mM human insulin in 20% acetic acid–sodium acetate (pD 3) and 30 °C. Corresponding spectra of analogues lacking residues B26–B30 (DPI; panel B), residues B25–B30 (panel C), residues B24–B30 (panel D), and residues B23–B30 (DOI; panel E). Asterisk indicates ortho resonance of PheB24; arrow indicates meta resonance of TyrB26; and T indicates threonine methyl resonances of B27 and B30. A convolution difference with exponential parameters 2 and 4 Hz and subtraction ratio 1.0 was applied.

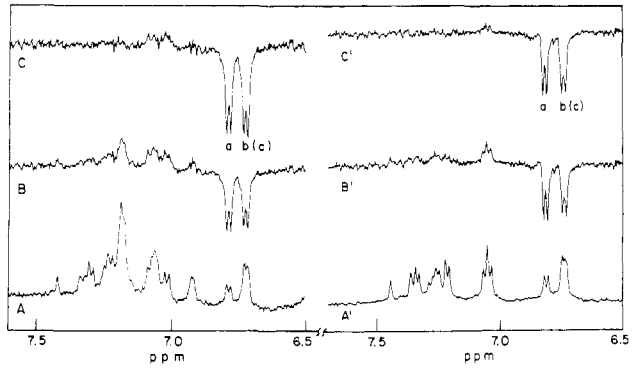

FIGURE 3: Photo-CIDNP spectra of DPI (panels A–C, analogue lacking residues B26–B30) and DOI (panels A′–C′, analogue lacking residues B23–B30). (A, A′) "Dark" spectra observed in the absence of laser irradiation. (B, B′) Spectra observed following laser irradiation. (C, C′) Calculated difference spectra. Difference signals assigned to TyrA14 are labeled a; those of TyrB16 are labeled b; and those of TyrA19 are labeled c. Photo-CIDNP spectra were obtained at pD 3.0 and processed as previously described (Weiss et al., 1989).

In fact, under alternative solvent conditions (e.g., 6% mM protein in 10% deuterated DMSO, 90% H_2O) DPI exhibits independent spin systems arising from distinct oligomeric species in slow exchange [data not shown; see also Jeffrey (1986)].

(*i*) *Methyl-Containing Spin Systems.* DPI contains one alanine (B14), one threonine (A8), two isoleucines (A2 and A10), six leucines (A13, A16, B6, B11, B15, and B17), and four valines (A3, B2, B12, and B18). Their characteristic spin systems are readily distinguished in DQF-COSY, RCT, and TOCSY spectra. These spin systems are illustrated in Figure 6.

(*ii*) *Long Side Chains.* ArgB22 is distinguished by an NOE in H_2O from its side-chain NH to H_δ (indicated by an asterisk in panel D of Figure 9). Seven Glx spin systems are apparent in the DQF-COSY and TOCSY spectra in D_2O. These are illustrated in Figure 7. The three glutamine spin systems are distinguished from the four glutamic acid spin systems by strong H_γ–$NH_{2\epsilon}$ corsspeaks in NOESY spectra recorded in H_2O.

(*iii*) *Aromatic Residues.* DPI contains three phenylalanines (*B1*, B24, and B25), two tyrosines (A14 and B16), and two histidines (B5 and B10). The aromatic ring resonances in insulin have previously been assigned by mutagenesis and

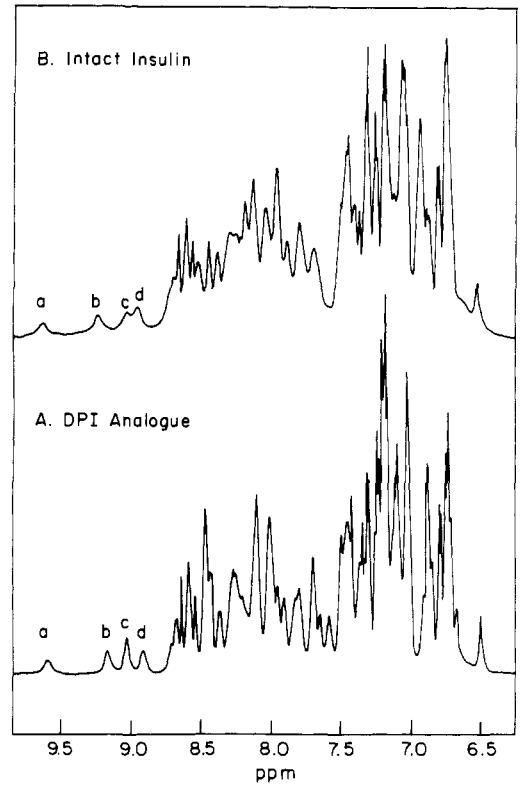

FIGURE 4: 1D NMR spectra of DPI (panel A) and native human insulin (panel B) in 20% acetic acid/80% H_2O solution at pH 1.9 and 25 °C. Four broad amide resonances are resolved downfield; these are assigned to A11 (labeled a), B8 (labeled b), B9 (labeled c), and B6 (labeled d). A 3-Hz exponential multiplication was applied prior to Fourier transformation. The protein concentration was 3.4 mM. The solvent resonance was eliminated by presaturation.

chemical modification (Weiss et al., 1989); these assignments extend to DPI and are consistent with the sequence-specific assignments given below. These resonances are linked to corresponding $\alpha\beta\beta'$ AMX spin systems by NOESY crosspeaks in D_2O. These are shown in Figure 8. The assumption that these represent intraresidue effects is verified by sequential analysis (part III; below). Additional interresidue NOEs are observed between ring protons and H_α and H_β resonances; these provide structural information and are discussed below.

(*iv*) *Other AMX Spin Systems.* Twelve remaining AMX spin systems are identified in DQF-COSY and TOCSY spectra in D_2O, corresponding to three asparagines (A18, A21, and B3), six cysteines (A6, A7, A11, A20, B7, and B19), and three serines (A9, A12, and B9). These are also shown in Figure 8. The asparagine spin systems are distinguished by strong H_β–$NH_{2\delta}$ crosspeaks in NOESY spectra in H_2O.

(*v*) *DPI Contains Four Glycines at Positions A1, B8, B20, and B23.* Two glycine-specific H_N–H_α/H_N–$H_{\alpha'}$ fingerprint crosspeaks are observed in TOCSY and NOESY spectra in H_2O, as illustrated in panel D of Figure 9. One glycine spin system contains degenerate H_α resonances but may be defined by sequential connectivities (below). As expected, no crosspeaks are observed from the NH_2-terminal GlyA1; its AB spin system is assigned by elimination and confirmed by a sequential A1(H_α)–A2(H_N) NOE.

(III) Sequence-Specific Assignment and Structural Features

(*i*) *Sequential Assignment.* The A and B chains may be considered independently, and their assignment pathways are separately illustrated in Figures 9 and 10. The H_N–H_α NOESY fingerprint region is shown in Figure 9; A-chain

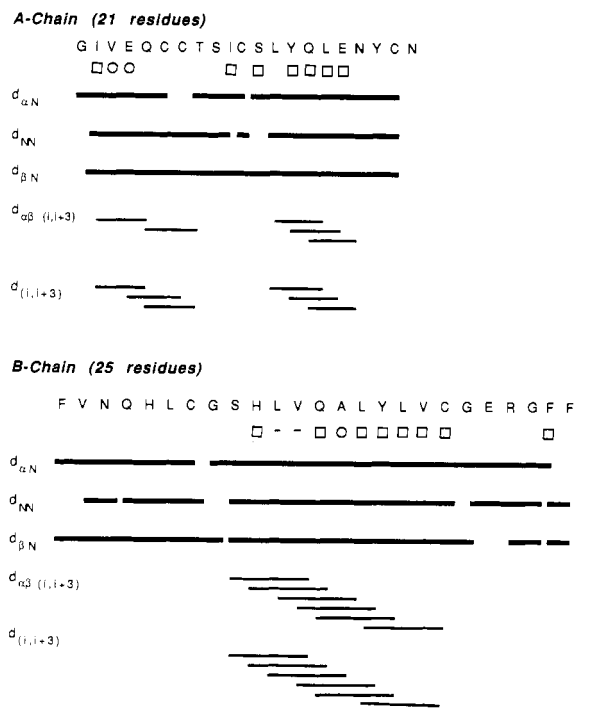

FIGURE 5: Schematic representation of sequential connectivities in the A chain (top panel) and B chain (bottom panel). The primary structures of the two chains are given in one-letter code. Symbols $d_{\alpha N}$, d_{NN}, etc. are as defined by Wuthrich (1986). Two classes of slowly exchanging amide resonances are indicated: □ (slower) and O (faster). The amide resonances of LeuB11 and ValB12 (indicated by –) could not be evaluated in the 1D spectrum due to resonance overlap. $d_{(i,i+3)}$ denotes NOEs between the α proton of residue i and the amide proton of residue $i + 3$. α-Helices (A-chain residues 2–8 and 13–18; B-chain residues 9–18) are defined by patterns of $(i,i+3)$ NOEs and by continuous strings of d_{NN} connectivities that are stronger than those generally observed in nonhelical regions; strong d_{NN} connectivities are also observed in the turn between residues B20 and B23. In this and the following figures, residue B13 should be identified as glutamic acid (E) rather than glutamine (Q) as indicated; see Table I.

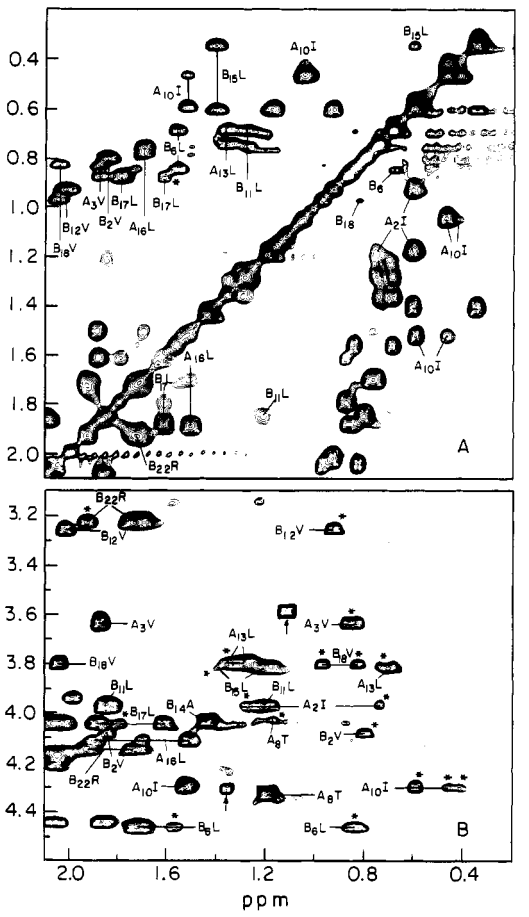

FIGURE 6: TOCSY spectrum of DPI showing identification of spin systems containing methyl groups. Conditions: mixing time 30 ms; protein concentration 3 mM, and 25 °C. Prior to measurement amide protons were exchanged for deuterium. Four hundred t_1 values were observed and zero filled to 4096 points prior to Fourier transformation. In each dimension the digital resolution was 1.34 Hz/point, and exponential and sine shifted window functions were applied. Relevant crosspeaks have been labeled with the sequence position and the amino acid type. Hartmann–Hahn relay crosspeaks are denoted by asterisks. Impurity peaks are indicated by arrows.

residues are labeled in panel B, and the B-chain residues are labeled in panel D. Corresponding d_{NN} connectivities are shown in panels A (A chain) and C (B chain); sequential $d_{\beta N}$ connectivities are outlined in Figure 10. In each case a connected string of sequential NOESY-COSY connectivities may be established for the entire chain (Wuthrich et al., 1983). Sequential connectivities are illustrated in schematic form in Figure 5, and ^{1}H NMR assignments are summarized in Table I.

These assignments are in overall accord with those reported by Kline and Justice (1990). Significant changes in chemical shifts are observed, however, in the amide resonances of B8 ($\Delta\delta$ 0.36 ppm), B21 ($\Delta\delta$ 0.35 ppm), and A21 ($\Delta\delta$ 0.34 ppm) and also in the α resonance of PheB24 ($\Delta\delta$ 0.34 ppm). These differences in chemical shift may reflect the absence of residues B26–B30 or the effects of different solvent compositions in the two studies.

DPI contains three disulfide bonds (A6–A11, A7–B7, and A20–B19). Although no homonuclear coupling is established by these bonds, the spatial proximity of the amide protons, α-protons, and β-protons involved provides a NOESY connection between the spin networks of the two chains. The A6–A11 disulfide gives rise to A6(H_α)–A11(H_β) and A6-(H_N)–A11($H_{\beta'}$) NOEs, the A7–B7 disulfide gives rise to A7(H_β)–B7(H_α) and A7($H_{\beta'}$)–B7(H_N) NOEs, and A20–B19 disulfide gives rise to an A20(H_N)–B19(H_β) NOE. No other NOEs are observed between cysteines. These NOEs, which reflect the local geometry of each disulfide, provide a consistency check for the sequential assignments. Similar disulfide-specific NOEs have been described by Kline and Justice (1990); however, the A20 and B19 spin systems could not be resolved under their solvent conditions.

Interestingly, disproportionate broadening of the A11 spin system is observed under both sets of solvent conditions (see H_N resonance a in Figure 4; 9.61 ppm) and may relate to millisecond motions in the protein. Disproportionate broadening is also observed among the amide resonances of B8 (resonance b; 9.18 ppm), B9 (resonance c; 9.03 ppm), B6 (resonance d; 8.92 ppm), A5 (8.23 ppm), and A9 (7.43) and the H_α–H_β crosspeak of A2 (not shown). However, several crosspeaks described as absent in the spectrum of insulin in 35% acetonitrile (e.g., H_α–H_β crosspeaks for A13, B15 and B17; Kline & Justice, 1990) are observed in the spectrum of DPI in 20% acetic acid. The origins of these differences are presently unclear.

(ii) Secondary Structure. The presence and location of α-helices in DPI may be established by observation of characteristic $(i,i+3)$ sequential connectivities. These connectivities are outlined in Figure 5 and confirm the presence of three α-helices in solution: two in the A chain (residues A2–A8 and A12–A17) and one in the B chain (B9–B18). As expected, the three helical regions exhibit (i) stronger d_{NN} connectivities

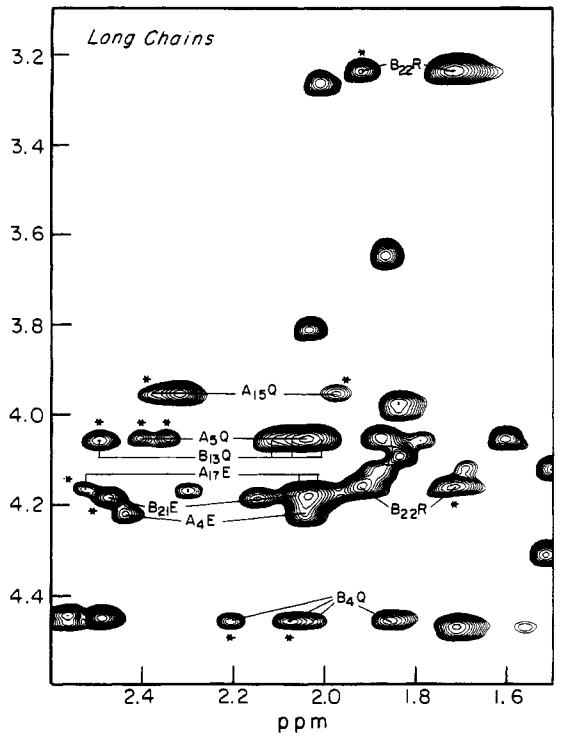

FIGURE 7: TOCSY spectrum showing long-side-chain residues of DPI in regions $\omega 1$ = 1.50–2.60 ppm and $\omega 2$ = 3.10–4.60 ppm. Asterisks represent relay crosspeaks. Conditions are as described in the caption to Figure 6.

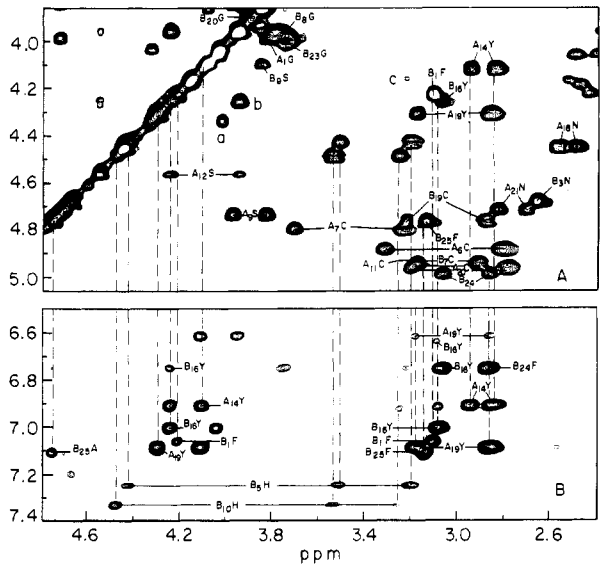

FIGURE 8: (A) Region of TOCSY spectrum showing AMX and glycine spin systems. (B) NOESY spectrum in the region showing corresponding aromatic NOEs. Others are the same as in Figure 3. Vertical lines link $C^{\alpha}H$–$C^{\beta}H$ crosspeaks in (A) with NOESY crosspeaks in (B), establishing the connectivities with the corresponding aromatic spin systems. Crosspeaks a, b, and c in panel A are $C^{\alpha}H$–$C^{\beta}H$ crosspeak of ThrA8, the $C^{\beta 1}H$–$C^{\beta 2}H$ crosspeak of A12 Ser, and $C^{\alpha}H$–$C^{\delta}H$ crosspeak of B22 Arg, respectively. The TOCSY spectrum was recorded as described in the caption to Figure 6; the NOESY spectrum was recorded with a mixing time of 200 ms in 20% acetic acid-d_4/80% D20. The NOESY spectrum was processed with the same window functions as described for the TOCSY spectrum in Figure 6.

than are generally observed elsewhere and (ii) slowly exchanging amide resonances (see below). These results are in accord with the crystal structure of DPI (Bi et al., 1984; Dai et al., 1987) and with the two-Zn structure of porcine insulin (Blundell et al., 1971; Baker et al., 1988). The participation

Table I: Chemical Shifts of the Assigned ^1H NMR Lines of DPI at pH 1.9 and 25 °C

	chemical shifts at 25 °C			
residue	NH	$C^{\alpha}H$	$C^{\beta}H$	others
A1 Gly		3.97, 3.84		
A2 Ile	8.47	3.96	1.17	$C^{\gamma}H_2$ 1.18, 0.93
				$C^{\gamma}H_3$ 0.73, $C^{\delta}H_3$ 0.60
A3 Val	8.09	3.64	1.87	$C^{\gamma}H_3$ 0.88, 0.82
A4 Glu	8.10	4.22	2.05, 2.05	$C^{\gamma}H_2$ 2.45, 2.45
A5 Gln	8.23	4.04	2.07, 2.01	$C^{\gamma}H_2$ 2.42, 2.36
				$N^{\epsilon}H_2$ 6.85, 7.46
A6 Cys	8.28	4.87	3.32, 2.82	
A7 Cys	8.23	4.79	3.71, 3.26	
A8 Thr	8.18	4.04	4.35	$C^{\gamma}H_3$ 1.20
A9 Ser	7.43	4.74	3.97, 3.84	
A10 Ile	7.81	4.31	1.54	$C^{\gamma}H_2$ 0.48, 0.46
				$C^{\gamma}H_3$ 0.61, $C^{\delta}H_3$ 1.04
A11 Cys	9.61	4.94	3.17, 2.88	
A12 Ser	8.72	4.56	4.27, 3.97	
A13 Leu	8.58	3.82	1.37, 1.31	$C^{\gamma}H$ 1.31, $C^{\delta}H_3$ 0.77, 0.70
A14Tyr	7.43	4.12	2.96, 2.87	$C^{2,6}H$ 7.02, $C^{3,5}H$ 6.79
A15 Gln	7.51	3.95	2.33, 1.99	$C^{\gamma}H$ 2.38, $N^{\epsilon}H_2$ 6.90, 7.45
A16 Leu	8.02	4.08	1.85, 1.50	$C^{\gamma}H$ 1.70, $C^{\delta}H_3$ 0.75, 0.75
A17 Glu	8.01	4.16	2.07, 2.01	$C^{\gamma}H_2$ 2.53, 2.28
A18 Asn	7.38	4.44	2.58, 2.52	$N^{\delta}H_2$ 6.50, 7.12
A19 Tyr	7.84	4.31	3.20, 2.88	$C^{2,6}H$ 7.20, $C^{3,5}H$ 6.72
A20 Cys	7.37	4.95	3.21, 2.80	
A21 Asn	8.20	4.71	2.83, 2.70	$N^{\delta}H_2$ 7.47, 6.67
B1 Phe		4.23	3.13, 3.13	$C^{2,6}H$ 7.17, $C^{3,5}H$ 7.31
				$C^{4}H$ 7.24
B2 Val	8.11	4.09	1.85	$C^{\gamma}H_3$ 0.81, 0.81
B3 Asn	8.43	4.68	2.68, 2.68	$N^{\delta}H_2$ 7.48, 6.87
B4 Gln	8.37	4.45	2.08, 1.87	$C^{\gamma}H_2$ 2.22, 2.22
				$N^{\epsilon}H_2$ 7.29, 6.76
B5 His	8.60	4.44	3.51, 3.22	$C^{2}H$ 8.56, $C^{4}H$ 7.35
B6 Leu	8.92	4.47	1.72, 0.84	$C^{\gamma}H$ 1.56, $C^{\delta}H_3$ 0.86, 0.69
B7 Cys	8.30	4.92	3.18, 2.93	
B8 Gly	9.18	3.97, 3.81		
B9 Ser	9.03	4.09	3.93, 3.87	
B10 His	7.98	4.48	3.53, 3.27	$C^{2}H$ 8.66, $C^{4}H$ 7.44
B11 Leu	7.07	3.97	1.84, 1.21	$C^{\gamma}H$ 1.28, $C^{\delta}H_3$ 0.78, 0.70
B12 Val	7.13	3.27	2.03	$C^{\gamma}H_3$ 0.94, 0.94
B13 Glu	7.90	4.05	2.07, 2.02	$C^{\gamma}H_2$ 2.50, 2.14
B14 Ala	7.70	4.04		$C^{\beta}H_3$ 1.44
B15 Leu	8.00	3.83	1.24, 0.77	$C^{\gamma}H$ 1.43, $C^{\delta}H_3$ 0.62, 0.38
B16 Tyr	8.13	4.25	3.10, 3.10	$C^{2,6}H$ 7.10, $C^{3,5}H$ 6.74
B17 Leu	7.79	4.05	1.90, 1.62	$C^{\gamma}H$ 1.80, $C^{\delta}H_3$ 0.88, 0.88
B18 Val	8.47	3.82	2.05	$C^{\gamma}H_3$ 0.98, 0.84
B19 Cys	8.68	4.75	3.22, 2.88	
B20 Gly	7.70	3.90, 3.90		
B21 Glu	8.48	4.18	2.16, 2.05	$C^{\gamma}H_2$ 2.48, 2.48
B22 Arg	7.95	4.17	1.93, 1.93	$C^{\gamma}H_2$ 1.74, 1.71
				$C^{\delta}H_2$ 3.23, 3.23, $N^{\epsilon}H$ 7.08
B23 Gly	7.59	3.98, 3.77		
B24 Phe	7.65	4.99	3.07, 2.88	$C^{2,6}H$ 6.88, $C^{3,5}H$ 7.02
B25 Phe	8.26	4.73	3.14, 3.14	$C^{2,6}H$ 7.21, $C^{3,5}H$ 7.28

of GlyA1 in the N-terminal helix of the A chain is unclear.

(iii) Slowly Exchanging Amides. Several slowly exchanging H_N resonances are observed over 24 h in D_2O solution, reflecting the presence of a stable structure. The time course of exchange is presented at 25 °C in Figure 11. The parent spectrum in H_2O is shown in panel B and upon initial dissolution in D_2O in panel C. Spectra obtained at successive time points between 20 min and 24 h are shown in panels D–F. Slowly exchanging H_N resonances from residues are well-resolved in the 1D NMR spectrum and may be assigned unambiguously; an envelope of resonances is observed at 8.0 ppm, which contain several overlapping H_N resonances (A16, A17, B10, and B15). In addition, exchange of A3, A4, and B14 amide protons occurs within minutes and is initially observable in D_2O solution. The assigned resonances map to the B-chain helix (B10, B13, and B15–B18), the N-terminal portion of the second A-chain helix (residues A15–A17), and nonhelical residues PheB24 and IleA10. These nonhelical residues participate in packing of the hydrophobic core (section iv) and apparently do not exhibit slow-exchange features in 35%

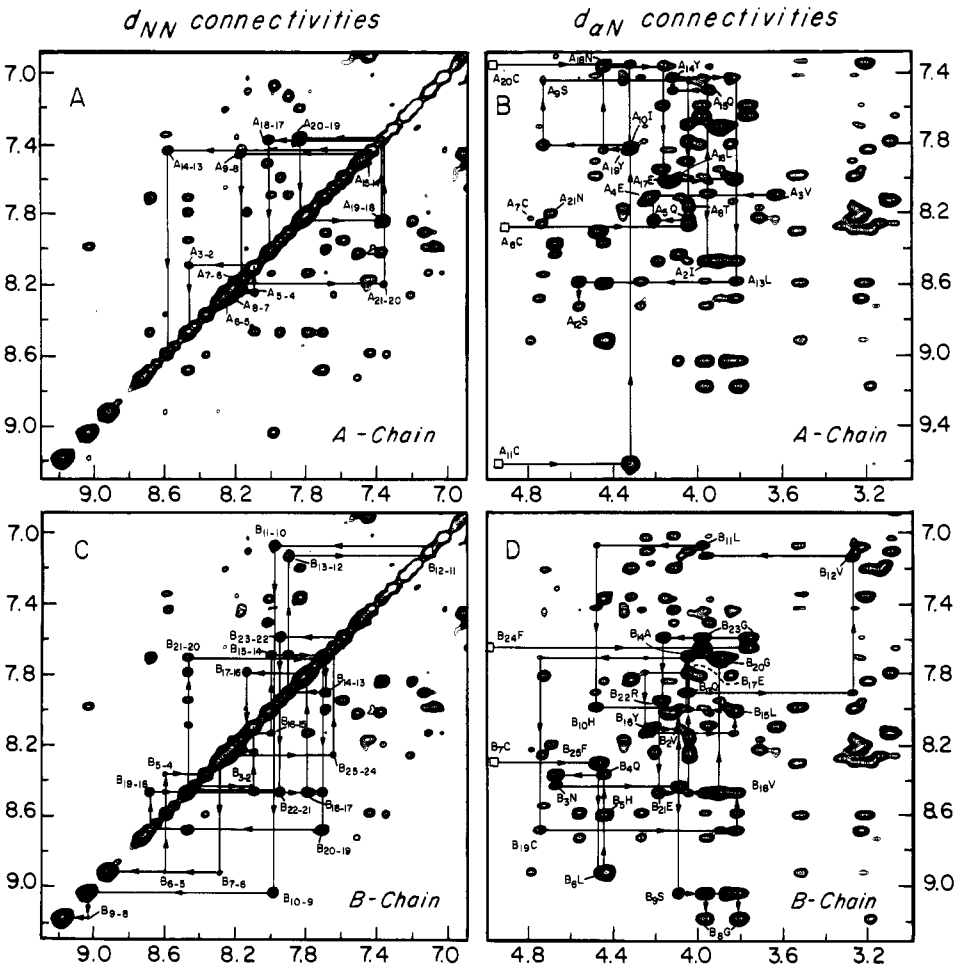

FIGURE 9: d_{NN} and $d_{\alpha N}$ connectivities of the A chain (panels A and B, respectively) and of the B chain (panels C and D, respectively) in the NOESY spectrum in 20% deuterated acetic acid/80% H_2O and a protein concentration of 3.4 mM. The mixing time was 200 ms. Open squares (□) represent $C^{\alpha}H$–NH crosspeaks bleached under these conditions by solvent presaturation. Asterisk in panel D represents an intraresidue NOE from the ArgB22 side-chain NH to H_{δ}.

acetonitrile (Kline & Justice, 1990).

(*iv*) *Long-Range NOEs.* DPI in solution exhibits three segments of α-helix (section ii above) whose relative orientation gives rise to many long-range NOEs. NOEs involving aromatic rings and methyl resonances are shown in Figure 12 and may be qualitatively compared to those expected on the basis of the crystal model (Blundell et al., 1971; Bi et al., 1984; Dai et al., 1987). These may be classified as intrachain or interchain and are discussed in turn.

Within the B chain the ortho and meta protons of PheB24 exhibit NOEs involving ValB12 and LeuB15 (Figure 12). These NOEs are expected from the crystal models, as illustrated in panel A of Figure 13. Additional packing interactions within the A chain are observed between TyrA19 and IleA2, as expected from the crystal model shown in Figure 1. These interactions define part of the hydrophobic core. The absence of significant NOEs involving PheB25 is consistent with an outward configuration as observed in the DPI crystal (Bi et al., 1984; Dai et al., 1987). Such a configuration was previously inferred in NMR studies of intact insulin (Weiss et al., 1989) in accord with molecule 2 of the two-Zn structure (Blundell et al., 1971). However, weak NOEs are observed between the ortho resonance of PheB25 and LeuB15 ($CH_{2\beta}$) and between the B25 meta resonance and LeuB15 ($C^{\delta 2}H_3$). These effects are not predicted by either ring orientation in the two-Zn crystal structure and are likely to represent transient interactions between a flexible side chain (PheB25) and a well-ordered side chain (LeuB15). This interpretation is consistent with molecular mechanics calculations (Weiss et al., 1989) and may be explored further by molecular dynamics simulations (Torda et al., 1990).

Long-range NOEs between the A chain and B chain reflect the relative orientation of the three helices. As predicted by the crystal model, NOEs are observed between the first helix of the A chain and the N-terminal portion of the B-chain helix [IleA2 side chain and B7,8(H_{β}) and B11(H_{α}); ValA3 and B11(H_{δ})] and between the second helix of the A chain and the C-terminal portion of the B-chain helix [AsnA21 H_N and B22(H_{β}); TyrA19 meta and B15 δ-methyl; A20–B19 disulfide as above]. Previous studies of insulin analogues have mapped the putative receptor-binding surface to the two A-chain helices, the C-terminal portion of the B-chain helix, and nonhelical residues B23–B25 (Pullen et al., 1976). The qualitative agreement between the crystal model and the observed NOE pattern indicates that the overall configuration of the receptor-binding surface is retained under these solution conditions.

A prominent NOE is observed between the H_{ϵ} resonance of HisB5 and IleA10; in contrast, no NOEs are observed in this region involving HisB10. This is in accord with the participation of HisB5 in a local pocket between the A and B chains and the absence of intramolecular interactions expected in the monomer for HisB10. These contrasting histidine environments have been described in NMR studies of intact insulin and proinsulin and analyzed by molecular mechanics calculations (Weiss et al., 1989, 1990).

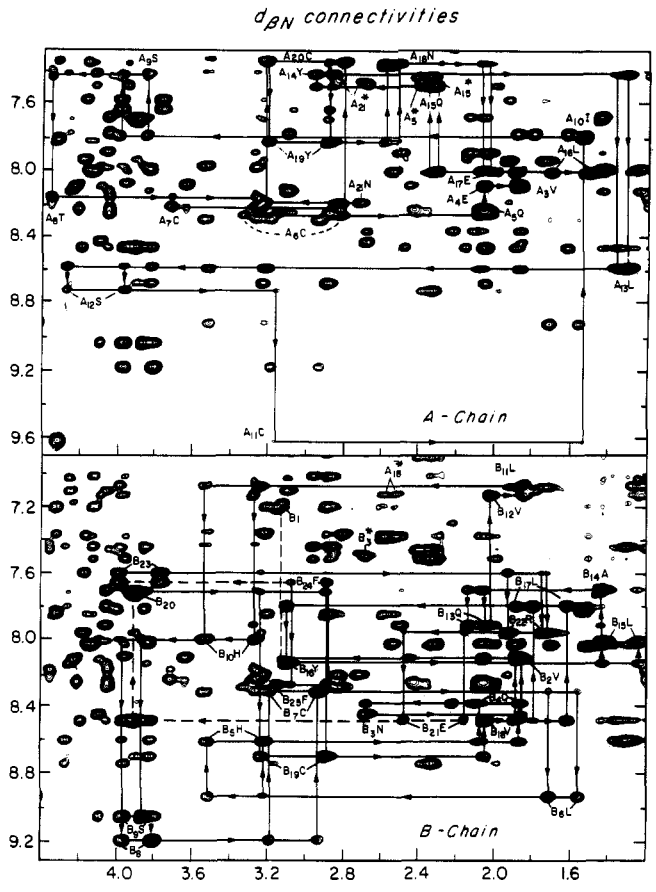

FIGURE 10: $d_{\beta N}$ connectivities of the A chain (region A) and the B chain (region B) in NOESY spectrum observed as described in the caption to Figure 9. In the lower panel horizontal dashed lines represent glycine B20 and B23 connectivities. Since N-terminal NH_2 resonances are not observed, a vertical dashed line connects a PheB1 aromatic H_β crosspeak to ValB2 H_N.

No significant nonlocal NOEs are observed involving residues B1–B3. Their absence is in accord with flexibility in the N-terminal region of the B chain, as previously inferred from CD studies (Wollmer et al., 1979) and molecular dynamics calculations (Kruger et al., 1987). This region of insulin is not thought to be involved in receptor binding (Pullen et al., 1977; Baker et al., 1988). However, a weak NOE is observed between the ortho resonance of PheB1 and LeuA13 ($C^{\delta 2}H_3$). Such an NOE is predicted by the two-Zn structure of porcine insulin, as shown in panel B of Figure 13 (asterisk); this interaction is not observed in the crystal structure of DPI (Bi et al., 1984; Dai et al., 1987), however, nor in other crystal forms of insulin (Dodson et al., 1979). It is likely that the various crystal models are represented in an ensemble of solution structures and that the observed B1–A13 NOE reflects a transient excursion of the B-chain in some members of this ensemble. Quantitative interpretation of such NOEs may be pursued in the future by an analysis of $^{13}C_\alpha$ spin–lattice relaxation times and molecular dynamics simulations (Torda et al., 1990).

Discussion

The primary objective of this paper is the complete sequential assignment of des-pentapeptide-insulin (DPI), a monomeric insulin analogue lacking the five C-terminal residues of the B chain (Gattner, 1975; Danho et al., 1975; Wang & Tsou, 1986). These studies are conducted in 20% acetic acid, conditions in which insulin is monomeric and stably folded (Weiss et al., 1989). Control experiments demonstrate

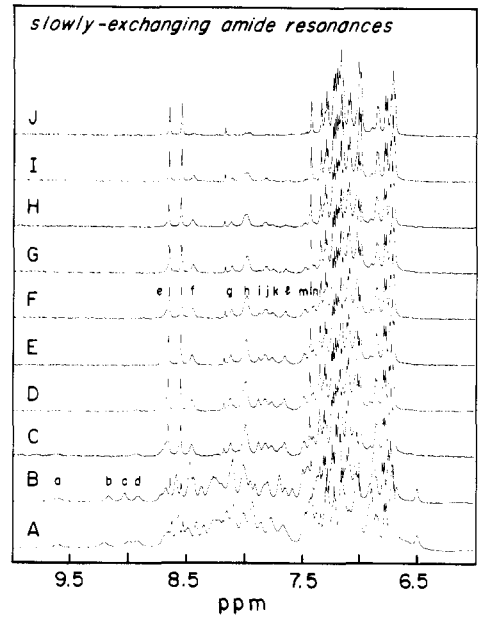

FIGURE 11: 1D NMR spectra of DPI freshly dissolved in a 80% deuterated acetic acid/20% D_2O solution. Slow exchange of selected amide resonances is observed at successive time points; panels C–J were observed at exchange times of respectively 10, 20, 40, 60, 260, 400, and 14 440 min (24 h). Corresponding spectra of intact human insulin and DPI in 80% deuterated acetic acid/20% H_2O solution are shown in panels A and B, respectively. Resonances a–d are defined in the caption to Figure 4. Resonances e–n are as follows: (e) A12 and B19; (f) A2 and B18; (g) B16, A3, and A4; (h) A16, A17, B10, and B15; (i) B13; (j) A10; (k) B17; (l) B14 and B24; (m) A15; (n) A14.

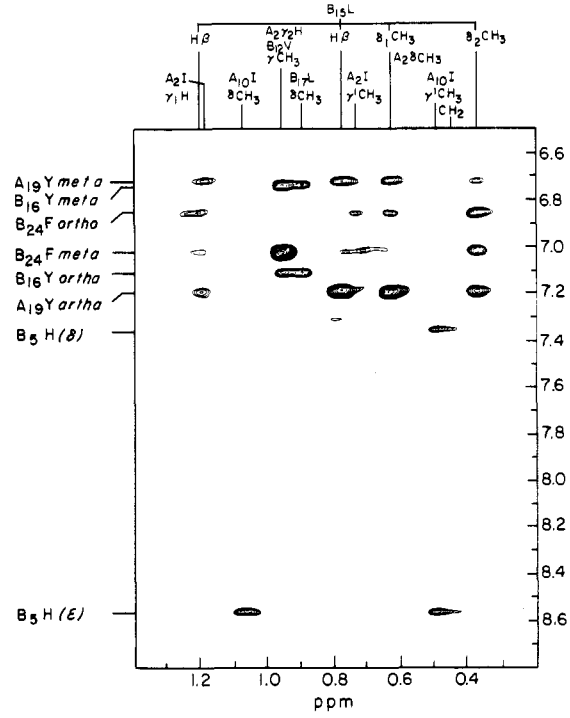

FIGURE 12: Long-range NOEs between aromatic and methyl resonances demonstrate tertiary structure. Very weak NOEs are observed at lower contours (not shown) between PheB25 and LeuB15 and between PheB1 and LeuA13 (see text); these are likely to reflect transient interactions involving flexible residues B1 and B25. The NOESY spectrum was recorded in 80% deuterated acetic acid/20% D_2O with a 200-ms mixing time.

that acetic acid in the range 0–20% does not significantly perturb the insulin fold (Weiss et al., 1990). DPI was chosen as a model system following systematic screening of insulin

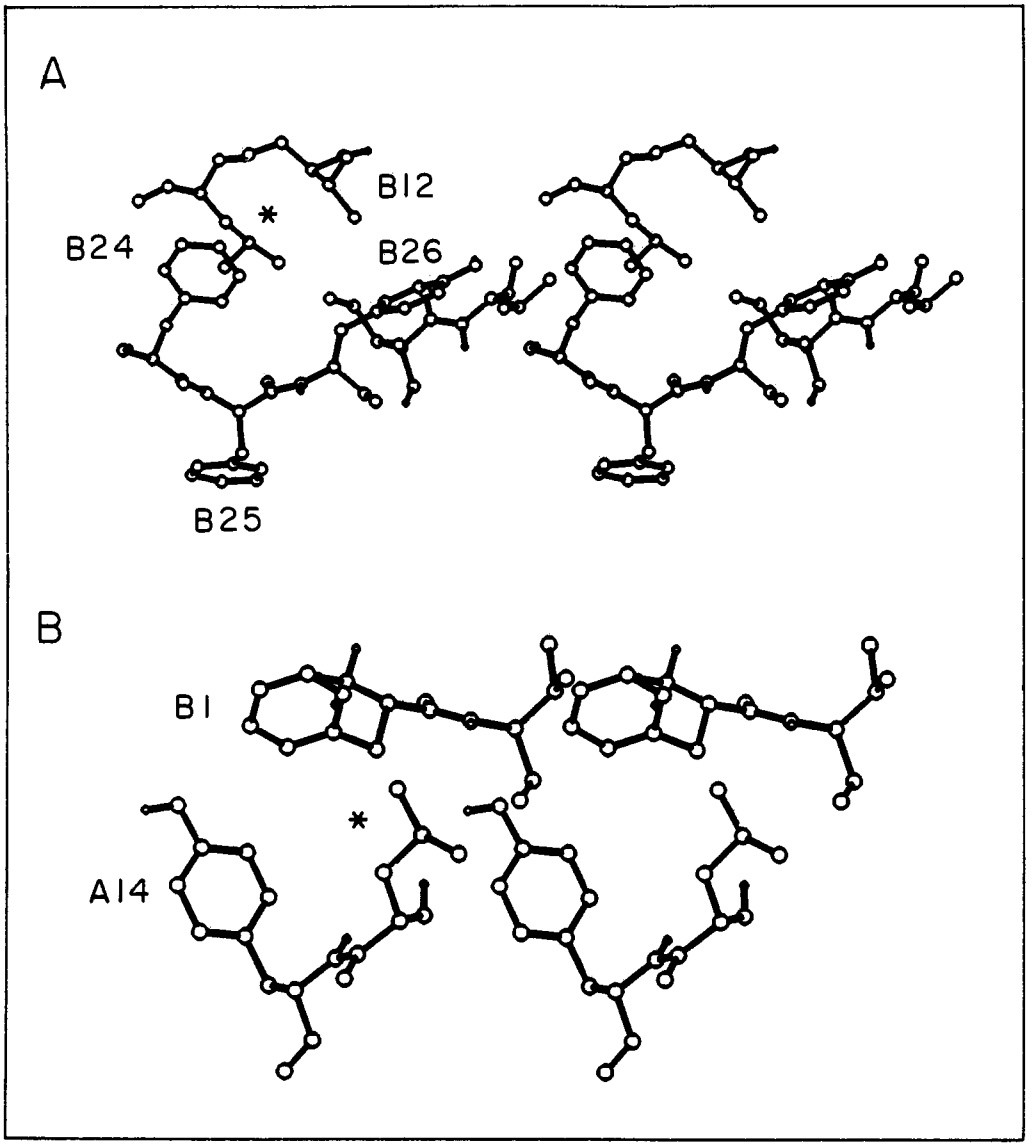

FIGURE 13: (A) Stereorepresentation of long-range interactions between PheB24 and LeuB15 (asterisk) and ValB12 in the crystal structure of molecule 2 of the two-Zn crystal structure of porcine insulin (Blundell et al., 1971); a similar interaction is observed in the crystal structure of DPI (Bi et al., 1984; Dai et al., 1986). (B) Stereorepresentation of long-range interactions between PheB1 and LeuA13 (asterisk) in the crystal structure of molecule 2 of the two-Zn crystal structure of porcine insulin; this interaction is not observed in the crystal structure of DPI (Bi et al., 1984; Dai et al., 1986) and in other crystal forms of porcine insulin (Dodson et al., 1979; G. Dodson, personal communication).

analogues. An NMR analysis of DPI at pH 10 has previously been described (Hua et al., 1989).

The C-terminal five residues of the insulin B chain are not conserved (Pullen et al., 1976) and in the amide form may be removed with retention of biological activity (Fisher et al., 1985). For this reason shortened insulin analogues have been used as a model for protein design (Casaretto et al., 1987; Mirmira & Tager, 1989). DPI also offers two technical advantages as a model for NMR studies. First, partial removal of the dimerization surface permits higher protein concentrations (>2 mM) to be used without aggregation. Second, amide resonances are less broad in the spectrum of DPI than in the spectrum of the insulin monomer (Weiss et al., 1989); such broadening reflects intermediate exchange among conformational substates and appears to be a general feature of the insulin fold (Weiss et al., 1990). Presumably, more rapid motions in the fragment lead to more complete averaging of chemical shifts on the NMR time scale. Accordingly, sequential assignment strategies (Wuthrich et al., 1983), previously found to be difficult to employ in the case of insulin, may successfully be applied to DPI.

An alternative strategy to obtain an interpretable spectrum of the insulin monomer has recently been described (Kline & Justice, 1990). This strategy uses a more organic solvent (35% acetonitrile) to obtain sharper amide resonances. Presumably, these conditions also alter the time scale of internal motions. Comparison of the structure of insulin under different solution conditions is of intrinsic interest and may complement our understanding of alternative subunit conformations observed in different crystal forms (Blundell et al., 1971; Dodson et al., 1979; Chothia et al., 1983). Differences among crystal forms emphasize the importance of nonlocal displacement of elements of secondary structure. Such displacement, which requires adjustment in multiple side-chain configurations, provides a model for the transmission of conformational change in proteins. 2D NMR analyses under different solvent conditions may enable the effects of crystal packing interactions to be distinguished from the intrinsic conformational repertoire of the monomer.

Sequential assignment of DPI indicates that the N- and C-terminal helices of the A chain and the central α-helix of the B chain are retained in solution. The existence of a com-

pact and stably folded tertiary structure is indicated by the protection of TyrA19 from photo-CIDNP enhancement, pattern of slowly exchanging amide protons, and long-range nuclear Overhauser enhancements. The latter NOEs suggest that the receptor-binding surface is similar in solution and in the crystal state. The present results provide a foundation for further investigation of the structure of DPI in solution by computer-based modeling methods (Wuthrich, 1989). Complementary interpretation of nonlocal NOEs in flexible regions of the protein is likely to require analysis of transient interactions by molecular-dynamics simulations (Torda et al., 1990). Such flexibility has been proposed to play a key role in receptor recognition (Dodson et al., 1979; Baker et al., 1988; Mirmira & Tager, 1989). In the future this hypothesis may be tested through comparative NMR studies of mutant insulins, including those associated with diabetes mellitus in man.

Added in Proof

Sequential ^{1}H NMR assignment of DPI at pH 1.7–1.9 has also been described in a recent paper by Kaptein and colleagues (Boelens et al., 1990). This study was conducted in aqueous solution; the assignments obtained are in general accord with those reported here.

Acknowledgments

We thank L. J. Neuringer (MIT) for advice and support, R. E. Chance and B. F. Frank (Eli Lilly & Co.) for DPI and biosynthetic human insulin, I. Khait for assistance with photo-CIDNP measurements, D. Nguyen for helpful discussion and computer graphics, J. P. Lee for assistance with NMR measurements, A. Kline for communication of results prior to publication, A. Bax and D. Torchia for helpful discussion, and Jai Wen-Hua for assistance with NMR processing. M.A.W. thanks Professors Martin Karplus, Christopher T. Walsh, Elkan Blout, Eugene Braunwald, and John T. Potts, Jr., for their encouragement.

Registry No. DPI, 55599-09-2; DOI, 39416-70-1; deshexapeptide B25–30 insulin, 39471-22-2; desheptapeptide B24–30 insulin, 51798-49-3; insulin, 11061-68-0.

References

Adams, M. J., Blundell, T. L., Dodson, E. J., Dodson, G. G., Vijayan, M., Baker, E. N., Hardine, M. M., Hodgkin, D. C., Rimer, B., & Sheet, S. (1969) *Nature (London) 224*, 491–495.

Baker, E. N., Blundell, T. E., Cutfield, G. S., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M. C., Hubbard, R. E., Iassac, M. W., Reynolds, D. C., Sakabe, K. S., Sakabe, N., & Vjayan, N. M. (1988) *Philos. Trans. R. Soc. (London) B319*, 389–456.

Bi, R. C., Dauter, Z., Dodson, E., Dodson, G., Giordano, F., & Reynolds, C. (1984) *Biopolymers 23*, 391–395.

Blundell, T. L., & Humbel, R. E. (1980) *Nature (London) 287*, 781–787.

Blundell, T. E., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dondson, G. G., Hodgkin, D. C., Mercola, D. A., & Vijayan, M. (1971) *Nature (London) 231*, 506–511.

Boelens, R., Ganadu, M. L., Verheyden, P., & Kaptein, R. (1990) *Eur. J. Biochem. 191*, 147–153.

Casareto, M., Spoden, M., Diaconescu, C., Gattner, H.-G., Zahn, H., Brandenburg, D., & Wollmer, A. (1987) *Biol. Chem. Hoppe-Seyler 368*, 709–716.

Cheshnovsky, D., Neuringer, L. J., & Williamson, K. L. (1983) *J. Protein Chem. 1*, 335–339.

Chothia, C., Lesk, A. M., Dodson, G. G., & Hodgkin, D. C. (1983) *Nature 302*, 500–505.

Dai, J.-B., Lou, M.-Z., You, J.-M., & Liang, D.-C. (1987) *Sci. Sin. 30* (1), 55–65.

Danho, W. O., Gattner, H.-G., Nissen, D., & Zarn, H. (1975) *Hoppe-Seyler's Z. Physiol. Chem. 356*, 1406–1412.

Davis, D. G., & Bax, A. (1985) *J. Magn. Reson. 64*, 533–535.

De Meyts, P., Van Obberghen, E., Roth, J., Brandenburg, D., & Wollmer, A. (1978) *Nature (London) 273*, 504–509.

Dodson, E. J., Dodson, G. G., & Hodgkin, D. C. (1979) *Can. J. Biochem. 57*, 469–479.

Fischer, W. H., Saunders, D., Brandenburg, D., Wollmer, A., & Zahn, H. (1985) *Hoppe-Seyler's Z. Biol. Chem. 366*, 521–525.

Gattner, H. G. (1975) *Hoppe-Seyler's Z. Physiol. Chem. 356*, 1397–1404.

Haneda, M., Polonsky, K. S., Bergenstal, R. M., Jaspan, J. B., Shoelson, S. E., Blix, P. M., Chan, S. J., Kwok, S. C. M., Wishner, W. B., Zeidler, A., Olefsky, J. M., Freidenberg, G., Tager, H. S., Steiner, D. F., & Rubenstein, A. H. (1984) *N. Engl. J. Med. 310*, 1288–1294.

Haneda, M., Kobayashi, M., Maegawa, H., Watanabe, N., Takata, Y., Ishibashi, O., Shigeta, Y., & Inoute, K. (1985) *Diabetes 34*, 568–573.

Hua, Q.-X., Chen, Y.-J., Wang, C.-C., Wang, D.-C., & Roberts, G. C. K. (1989) *Biochim. Biophys. Acta 994*, 114–120.

Jeffrey, P. D. (1986) *Biol. Chem. Hoppe-Seyler 367*, 363–369.

Kaptein, R. (1980) in *Photo-CIDNP Studies of Proteins in Biological Magnetic Resonance* (Berliner, W., & Rueben, J., Eds.) Vol. 4, pp 145–191, Plenum Press, New York.

Keefer, L. M., Piron, M.-A., DeMeyts, P., Gattner, H.-G., Diaconescu, C., Saunders, D., & Brandenburg, D. (1981) *Biochem. Biophys. Res. Commun. 100*, 1229–1236.

Kline, A. D., & Justice, R. M., Jr. (1990) *Biochemistry 29*, 2906–2913.

Kobayashi, M., Ohgaku, S., Iwasaki, M., Maegawa, H., Shigeta, Y., & Inouye, K. (1982) *Biochem. J. 206*, 597–603.

Kruger, P., Strassburger, W., Wollmer, A., van Gunsteren, W. F., & Dodson, G. G. (1987) *Eur. Biophys. J. 14*, 449–459.

Liang, D.-C., Stuart, D., Dai, J.-B., Todd, R., You, J.-M., & Luo, M.-Z. (1985) *Sci. Sin. 28*, 472–484.

Mirmira, R., & Tager, H. S. (1989) *J. Biol. Chem. 264*, 6349–6354.

Muszkat, K. A., Khait, I., & Weinstein, S. (1984) *Biochemistry 23*, 5–10.

Nakagawa, S. H., & Tager, H. S. (1986) *J. Biol. Chem. 261*, 7332–7341.

Nakagawa, S. H., & Tager, H. S. (1987) *J. Biol. Chem. 262*, 1204–12058.

Pullen, R. A., Lindsay, D. G., Wood, S. P., Tickle, I. J., Blundell, T. L., Wollmer, A., Krail, A., Brandenburg, D., Zahn, H., Gliemann, J., & Gammeltoft, S. (1976) *Nature (London) 259*, 369–373.

Riemen, M. W., Pon, L. A., & Carpenter, F. H. (1983) *Biochemistry 22*, 1507–1515.

Rupley, J. A., Renthal, R. D., & Praissman, M. (1967) *Biochim. Biophys. Acta 140*, 185.

Shoelson, S. E., Haneda, M., Blix, P., Nanjo, K., Sanke, T., Inouye, K., Steiner, D., Rubenstein, A., & Tager, H. (1983a) *Nature (London) 302*, 540–542.

Shoelson, S. E., Fickova, M., Haneda, M., Nahum, A., Musso, G., Kaiser, E. T., Rubenstein, A. H., & Tager, H. (1983b) *Proc. Natl. Acad. Sci. U.S.A. 80*, 7390–7394.

Smith, G. D., Swenson, D. C., Dodson, G. G., & Reynolds, C. D. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7093–7097.

States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Strickland, E. H., & Mercola, D. A. (1976) *Biochemistry 15*, 3875–3884.
Tager, H., Given, B., Baldwin, D., Mako, M., Markese, J., Rubenstein, A., Olefsky, J., Kobayashi, M., Kolterman, O., & Poucher, R. (1979) *Nature (London) 281*, 122–125.
Tager, H., Thamoas, N., Assoian, R., Rubenstein, A., Saekow, M., Olefsky, J., & Kaiser, E. T. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 3181–3185.
Torda, A. E., Sheek, R. M., & van Gunsteren, W. F. (1990) *J. Mol. Biol. 14*, 223–235.
Wang, C.-C., & Tsou, C.-L. (1986) *Biochemistry 25*, 5336–5340.
Weiss, M. A., Nguyen, D. T., Khait, I., Inouye, K., Frank, B. H., Beckage, M., O'Shea, E. K., Shoelson, S. E., Karplus, M., & Neuringer, L. J. (1989) *Biochemistry 29*, 9855–9873.
Weiss, M. A., Frank, B. H., Khait, I., Pekar, A., Heiney, R., Shoelson, S. E., & Neuringer, L. J. (1990) *Biochemistry 29*, 8389–8401.
Wollmer, A., Fleischhauer, J., Strassburger, W., Thiele, H., Bradenburg, D., Dodson, G., & Mercola, D. (1979) *Biophys. J. 20*, 233–243.
Wommer, A., Strassburger, W., Glatter, U., Dodson, G. G., & McRittel, W. (1981) *Z. Physiol. Chem. 362*, 581–591.
Wood, S. P., Blundell, T. L., Wollmer, A., Lazarus, N. R., & Neville, R. W. J. (1975) *Eur. J. Biochem. 55*, 531–542.
Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
Wuthrich, K. (1989) in *Methods of Enzymology* (Oppenheimer, N. J., & James, T. L., Eds.) Academic Press, New York.
Wuthrich, K., Wider, G., Wagner, G., & Braun, W. (1983) *J. Magn. Reson. 155*, 311.

Physicochemical Properties of Amaranthin, the Lectin from *Amaranthus caudatus* Seeds†

Stephen J. Rinderle,‡,§ Irwin J. Goldstein,*,‡ and Edward E. Remsen‖

Department of Biological Chemistry, University of Michigan, Ann Arbor, Michigan 48109-0606, and Monsanto Company, St. Louis, Missouri 63198

Received March 8, 1990; Revised Manuscript Received July 17, 1990

ABSTRACT: Amaranthin is the lectin present in the seeds of *Amaranthus caudatus*, which specifically binds the T-disaccharide (Galβ1,3GalNAcα-*O*-). The lectin is composed of a single type of subunit with M_r = 33 000–36 000 (Rinderle et al., 1989). Equilibrium sedimentation (M_r = 62 900) and low-angle laser light scattering (M_r = 61 400) methods have been used to unambiguously establish the native multimeric structure of amaranthin as a homodimer. These absolute molecular weight methods and the calculated Stokes radius (27.2 Å) indicate that the amaranthin dimer is highly compact relative to typical globular proteins, and thus, anomalous molecular weight values are obtained when simple size exclusion chromatography is used to determine the molecular weight of amaranthin. Studies with a homobifunctional cross-linking reagent and amaranthin further support the existence of a lectin homodimer. The stoichiometry of carbohydrate binding was determined to be one T-disaccharide-binding site per amaranthin subunit ($K_a = 3.6 \times 10^5$ M^{-1}). Amaranthin exhibits hydrophobic-binding properties as indicated by binding of 8-anilino-1-naphthalenesulfonate ($K_a = 3.6 \times 10^3$ M^{-1}) and 6-toluidinyl-2-naphthalenesulfonate ($K_a = 2 \times 10^4$ M^{-1}). Serological studies suggest that amaranthin does not appear to be present in the stems or leaves of the *A. caudatus* plant, nor were there any indications for the presence of cross-reactive material.

The lectin present in the seeds of *Amaranthus caudatus*, amaranthin, is very specific for the Thomsen–Friedenreich antigen (T-antigen; Galβ1,3GalNAcα-*O*-)[1] and selected derivatives of this disaccharide (Rinderle et al., 1989). The carbohydrate-binding specificity of amaranthin was found to be significantly different from that of peanut (*Arachis hypogaea*) agglutinin, explaining differences in reactivity with glycoproteins and red blood cells for these two T-antigen-specific lectins. Suggestions that the T-antigen may be a carcinoma-associated antigen (Springer et al., 1975; Anglin et al., 1977; Springer, 1984; Yuan et al., 1986) indicate the importance of developing well-defined T-antigen probes.

A recent report concerning the isolation and characterization of amaranthin (Rinderle et al., 1989) indicated that the lectin is composed of a single type of subunit with M_r = 33 000–36 000. The native multimeric structure of amaranthin, however, could not be accurately defined since the lectin migrated with a molecular weight (M_r = 54 000) between that expected for a homodimer and that expected for a monomer

† This work was supported by Grant GM 29470 (to I.J.G.) from the National Institutes of Health.
‡ University of Michigan.
§ Predoctoral Fellow supported by Pharmacological Sciences Training Grant 2 T32-GM07767-11, National Institutes of Health, during the course of this work. Current address: College of Pharmacy, The Ohio State University, Columbus, OH 43210.
‖ Monsanto Co.

[1] Abbreviations: T-antigen, Galβ1,3GalNAcα-*O*-; GalNAc, 2-acetamido-2-deoxy-D-galactose; ANS, 8-anilino-1-naphthalenesulfonate; TNS, 6-toluidinyl-2-naphthalenesulfonate; DSS, disuccinimidyl suberate; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; PBS, phosphate-buffered saline, pH 7.2; LALLS, low-angle laser light scattering.

0006-2960/90/0429-10555$02.50/0