NMR and Photo-CIDNP Studies of Human Proinsulin and Prohormone Processing Intermediates with Application to Endopeptidase Recognition†

Michael A. Weiss,*,‡,§ Bruce H. Frank,∥ Igor Khait,⊥ Allen Pekar,∥ Richard Heiney,∥ Steven E. Shoelson,# and Leo J. Neuringer⊥

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts 02114, Francis Bitter National Magnet Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, Indiana 46285, and Joslin Diabetes Center and Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115

Received March 27, 1990; Revised Manuscript Received May 1, 1990

ABSTRACT: The proinsulin–insulin system provides a general model for the proteolytic processing of polypeptide hormones. Two proinsulin-specific endopeptidases have been defined, a type I activity that cleaves the B-chain/C-peptide junction (Arg^{31}–Arg^{32}) and a type II activity that cleaves the C-peptide/A-chain junction (Lys^{64}–Arg^{65}). These endopeptidases are specific for their respective dibasic target sites; not all such dibasic sites are cleaved, however, and studies of mutant proinsulins have demonstrated that additional sequence or structural features are involved in determining substrate specificity. To define structural elements required for endopeptidase recognition, we have undertaken comparative ^1H NMR and photochemical dynamic nuclear polarization (photo-CIDNP) studies of human proinsulin, insulin, and split proinsulin analogues as models of prohormone processing intermediates. The overall conformation of proinsulin is observed to be similar to that of insulin, and the connecting peptide is largely unstructured. In the ^1H NMR spectrum of proinsulin significant variation is observed in the line widths of insulin-specific amide resonances, reflecting exchange among conformational substates; similar exchange is observed in insulin and is not damped by the connecting peptide. The aromatic ^1H NMR resonances of proinsulin are assigned by analogy to the spectrum of insulin, and assignments are verified by chemical modification. Unexpectedly, nonlocal perturbations are observed in the insulin moiety of proinsulin, as monitored by the resonances of internal aromatic groups. Remarkably, these perturbations are reverted by site-specific cleavage of the connecting peptide at the CA junction but not the BC junction. These results suggest that a stable local structure is formed at the CA junction, which influences insulin-specific packing interactions. We propose that this structure (designated the "CA knuckle") provides a recognition element for type II proinsulin endopeptidase.

Insulin is synthesized as a single polypeptide (preproinsulin) in which the C-terminus of the B-chain is linked to the N-terminus of the A-chain by a connecting peptide (Steiner & Oyer, 1967; Steiner et al., 1967, 1971). Following cleavage of an N-terminal signal sequence in the endoplasmic reticulum, the nascent polypeptide folds and is packaged into secretory granules as a prohormone. Proinsulin is converted to insulin in the B-cell granules by a specific set of proteases (Steiner & Tager, 1979; Docherty & Hutton, 1983; Davidson & Hutton, 1987; Davidson et al., 1988). This pattern of biosynthesis and prepro- and prohormone processing exemplifies a general motif in cell biology (Docherty & Steiner, 1982; Fisher & Scheller, 1988).

Intracellular conversion of proinsulin to insulin is catalyzed by two different Ca^{2+}-dependent endopeptidase activities (designated types I and II) which differ in their pH optima and substrate specificities (Davidson et al., 1988). Type I activity is directed against the B-chain/C-peptide junction (BC junction; Arg^{31}–Arg^{32}), and type II activity cleaves the C-peptide/A-chain junction (CA junction; Lys^{64}–Arg^{65}). Rules governing selection of target sites in this or other processing systems have not been defined. Naturally occurring mutations within the dibasic CA junction impair proinsulin processing and are associated with diabetes mellitus in man (Gabbay et al., 1976; Robbins et al., 1981, 1984). Proinsulin analogues retaining normal BC and CA dibasic junctions but containing deletions elsewhere in the connecting peptide are not properly processed (Docherty et al., 1989; Gross et al., 1989), suggesting that additional features are required for endopeptidase recognition. A highly conserved tetrapeptide in the connecting peptide (residues 33–36) has been proposed to constitute part of the type I recognition site at the BC junction (Gross et al., 1989).

As a first step toward defining conformational requirements of endopeptidase recognition, comparative ^1H NMR and photo-CIDNP[1] studies are presented of human proinsulin, insulin, and split proinsulin analogues. The ^1H NMR aromatic

† This work was supported in part by grants from the National Institutes of Health to M.A.W. and to the MIT NMR Resource (L.J.N.). This work was also supported in part by the Markey Charitable Trust at Harvard Medical School and by the Pfizer Scholars Program for New Faculty (M.A.W.). S.E.S. and M.A.W. were supported by the American Diabetes Association and the Juvenile Diabetes Foundation International. S.E.S. is a Capps Scholar in Diabetes at Harvard Medical School.

* Address correspondence to this author at the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School.

‡ Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School.

§ Massachusetts General Hospital.

∥ Eli Lilly and Co.

⊥ Massachusetts Institute of Techology.

Joslin Diabetes Center, The Brigham and Women's Hospital, Harvard Medical School.

[1] Abbreviations: (25,26) split human proinsulin, proinsulin analogue containing a single proteolytic cleavage between residues 25 and 26 (B25 and B26 in insulin nomenclature); des-(24,25) split human proinsulin, two-chain proinsulin analogue lacking residues 24 and 25 (B24 and B25 in insulin nomenclature); (32,33) split human proinsulin, proinsulin analogue containing a single proteolytic cleavage between residues 32 and 33; (65,66) split human proinsulin, proinsulin analogue containing a single proteolytic cleavage between residues 65 and 66 (A_0 and A_{-1} in insulin nomenclature); Arg-A_0-insulin, insulin analogue containing arginine N-terminal to GlyA1; CD, circular dichroism; COSY, correlated spectroscopy; NOE, nuclear Overhauser enhancement; photo-CIDNP, photochemically induced dynamic nuclear polarization; TOCSY, total correlation spectroscopy based on isotropic mixing.

0006-2960/90/0429-8389$02.50/0 © 1990 American Chemical Society

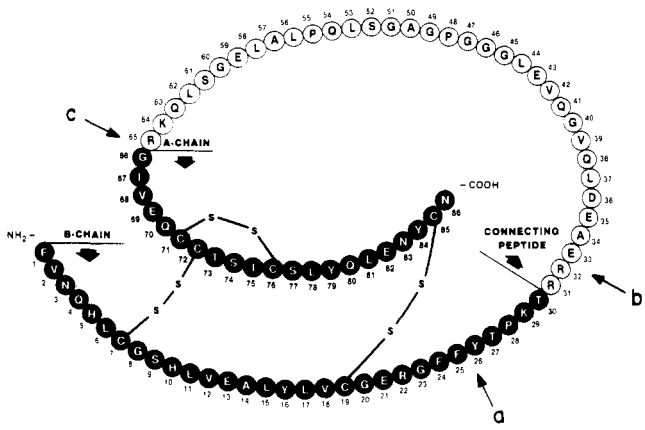

FIGURE 1: The amino acid sequence of human proinsulin is shown in one-letter code. The solid circles represent the insulin moiety (A- and B-chains), and the open circles represent the connecting peptide (residues 31–65). Arrows indicate positions of site-specific cleavage: (a) (25,26) split proinsulin; (b) (32,33) split proinsulin; and (c) (65,66) split proinsulin.

spectrum of human proinsulin is assigned and analyzed in reference to that of insulin and insulin analogues (Bradbury & Brown, 1977; Bradbury et al., 1981; Cheshnovsky et al., 1983; Bradbury & Ramesh, 1985; Hua et al., 1989; Weiss et al., 1989). This is the first NMR analysis of the proinsulin monomer.

The sequence of human proinsulin is shown in Figure 1 (Oyer et al., 1971). Although its three-dimensional structure has not been determined by X-ray crystallography (see Added in Proof), proinsulin appears to consist of a stably folded region (the insulin moiety; shaded circles in Figure 1) and an unstructured region (the connecting peptide; open circles in Figure 1). This model is in accord with (i) formation of proinsulin dimers and hexamers similar to those of insulin (Pekar & Frank, 1972); (ii) specific coordination of Zn^{2+} in the proinsulin hexamer (Grant et al., 1972); (iii) the presence of shared epitopes between insulin and proinsulin; (iv) a pattern of tyrosine reactivities to chemical modification similar to that of insulin (Frank et al., 1983); and (v) the similarity between the far-UV circular dichroism spectrum of proinsulin and the sum of the spectra of insulin and the C-peptide (Frank et al., 1972). In addition, proinsulin binds specifically to the insulin receptor, indicating that a portion of the receptor-binding surface is retained and is accessible (Chance et al., 1968).

In this paper conditions are described in which insulin, proinsulin, and proinsulin analogues are studied as monomers; these are 20% acetic acid–sodium acetate (pH 3). A mixed solvent system is required to weaken insulin self-association (Cheshnovsky et al., 1983; Weiss et al., 1989). Control experiments indicate that the structure of proinsulin is not appreciably perturbed under these conditions. Comparative NMR and photo-CIDNP studies of human insulin and proinsulin demonstrate that the insulin moiety of proinsulin is similar to insulin and that the connecting peptide is largely unstructured. However, perturbations are observed in 2D NMR resonances assigned to the hydrophobic core of the insulin moiety. The structural origins of these differences are investigated in studies of split proinsulin analogues as models of prohormone processing intermediates. Remarkably, proinsulin-specific perturbations are reverted by cleavage of the CA junction, but not by cleavage of the BC junction. These experiments suggest the existence of a stable local structure at the CA junction (designated the "CA knuckle"), which we propose as a possible recognition element for type II endopeptidase.

Materials and Methods

Bovine insulin and biosynthetic human insulin were 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). Zinc-free bovine proinsulin and biosynthetic human proinsulin was provided by Eli Lilly and used without further purification; their concentration was determined by UV absorbance at 278 nm, assuming that a 1 mg/mL solution has an absorbance of 0.70/cm for bovine proinsulin (Pekar & Frank, 1972) and 0.65/cm for human proinsulin. NMR buffer consists of 20% CD_3COOD/80% H_2O or D_2O (pH/pD 3.0, direct meter reading). The pH/pD was adjusted with aliquots of 20% CD_3COOD/80% D_2O containing 1 M NaOD.

Preparation of Proinsulin Analogues. Monoiodinated bovine proinsulin derivatives were prepared under limiting iodination conditions followed by preparative HPLC purification of the unreacted and singly reacted products as described (Frank et al., 1982, 1983). (32,33) split human proinsulin and (65,66) split human proinsulin were prepared as described (Peavy et al., 1985). (25,26) split human proinsulin was prepared by limited pepsin treatment of human proinsulin and isolated by preparative HPLC. Des-(24,25) split human proinsulin was prepared by limited carboxypeptidase digestion of (25,26) split proinsulin; successive removal of Phe24 and Phe25 using carboxypeptidase A was monitored by HPLC. An insulin derivative containing an arginine residue N-terminal to GlyA1[2] (A_0-insulin) was kindly provided by Dr. R. E. Chance (Eli Lilly and Co., Indianapolis, IN). Analogues were characterized by peptide mapping, amino acid analysis, and N-terminal sequencing.

NMR. Two-dimensional experiments were performed by the pure-phase method of States et al. (1982). TOCSY spectra were obtained as described (Davis & Bax, 1985). A total of 1024 complex points were sampled in t_2; 256 t_1 values were obtained, and the data matrix was zero-filled to 1K × 1K. Spectra were obtained in H_2O by solvent presaturation or by selective excitation (Hore, 1983).

Photo-CIDNP. Experiments were performed with riboflavin (0.4 mM). Solutions were prepared immediately before use from a saturated stock solution of riboflavin in NMR buffer. A 500-MHz probe was constructed at the Francis Bitter National Magnet Laboratory to permit laser irradiation of the sample under computer control, as previously described (Weiss et al., 1989). The light source was an Innova-6 continuous argon laser at 488 nm (Coherent, Inc., Palo Alto, CA). Light-minus-dark difference spectra were calculated in subsequent data processing.

Circular Dichroism and UV Difference Spectra. CD data were obtained at 37, 22, and 13 °C on a Jasco 600 spectropolarimeter using quartz cuvettes with path lengths from 0.02 to 5 cm. UV difference spectra at insulin concentrations from 1 to 40 mg/mL were recorded as described by Rupley et al. (1967).

Aggregation State. Protein self-association was monitored by the concentration dependence of near-UV absorption (Rupley et al., 1967) and circular dichroism (Wood et al.,

[2] The numbering scheme for insulin differs from that of proinsulin (Figure 1). Proinsulin residues 1–30 (corresponding to the insulin B-chain) are designated B1–B30; similarly, proinsulin residues 66–86 (A-chain) are designed A1–A21.

1975; Strickland & Mercola, 1976; Haruk et al., 1980) spectra, whose features are sensitive to ordering of aromatic rings in the dimer interface.

Results

The ^{1}H NMR characteristics of aromatic residues in insulin have been extensively investigated, providing a foundation for the present study (Bradbury et al., 1981; Cheshnovsky et al., 1983; Bradbury & Ramesh, 1985; Palmieri et al., 1989; Hua et al., 1989; Weiss et al., 1989). Mammalian insulins and proinsulins usually contain four tyrosines, three phenylalanines, and two histidines (Figure 1; Brown et al., 1955). In the crystal state PheB1 and TyrA14 are surface residues, whereas TyrA19 and PheB24 participate in the packing of the hydrophobic core of the protein (Adams et al., 1969; Blundell et al., 1971; Peking Insulin Structure Group, 1971). TyrB16, PheB25, and TyrB26 are on or near the surface of the monomer and are involved in dimerization. Both histidines are on the surface of the monomer and engage in subunit interactions in the crystal. HisB10 participates in zinc coordination, and HisB5 is involved in hexamer–hexamer interactions. PheB25 and HisB5 are observed to be in different configurations in the two crystallographic protomers in the 2Zn structure. There are no additional aromatic residues in the connecting peptide of human proinsulin (Oyer et al., 1971; Figure 1).

Our results are presented in three parts. In part I conditions of study are characterized by monitoring dimerization using UV difference spectroscopy and circular dichroism. Control experiments are presented indicating that the native structure of proinsulin is retained under these conditions. In part II the aromatic assignments of human insulin (Weiss et al., 1989) are extended to proinsulin, and overall aspects of the solution structure and dynamics are described. Remarkably, proinsulin-specific perturbations in the structure of the insulin moiety are observed that relate to tethering by the connecting peptide. The structural origins of these perturbations are investigated by comparative study of split proinsulin analogues; these positional isomers contain single strand breaks at the junctions between the insulin chains and the connecting peptide, providing models of prohormone processing intermediates. Part III contains a comparative analysis of the local environments of aromatic residues in proinsulin and insulin.

(I) Characterization of Conditions. *(A) Dimerization.* Zinc-free insulin and proinsulin ordinarily exist in solution as an equilibrium mixture of monomers, dimers, tetramers, hexamers, and higher order oligomers (Jeffrey & Coates, 1966a,b; Grant et al., 1972; Pekar & Frank, 1972; Goldman & Carpenter, 1974). To define conditions suitable for NMR study of the insulin monomer, we have previously described the use of an acetic acid mixed solvent system (20% acetic acid–sodium acetate, pH 3; Weiss et al., 1989). These conditions may be extended to proinsulin.

Dimerization also may be monitored by changes in molar UV absorbance, which reflects involvement of the B-chain tyrosines (B16 and B26; see below) in the dimer interface. For insulin the difference in molar UV absorbance at 287 nm between monomer and dimer states is 415–720/cm (Rupley et al., 1967; Frank & Veros, 1970). In 20% acetic acid–sodium acetate (pH 3.0), a 0.5 mM solution of insulin contains 2–5% dimer at 37 °C (Weiss et al., 1989). Remarkably, and in contrast to insulin, concentration-dependent changes are not observed in the molar UV absorbance of proinsulin (0.1–5 mM), suggesting that monomer–dimer–hexamer transitions are not present under these conditions.

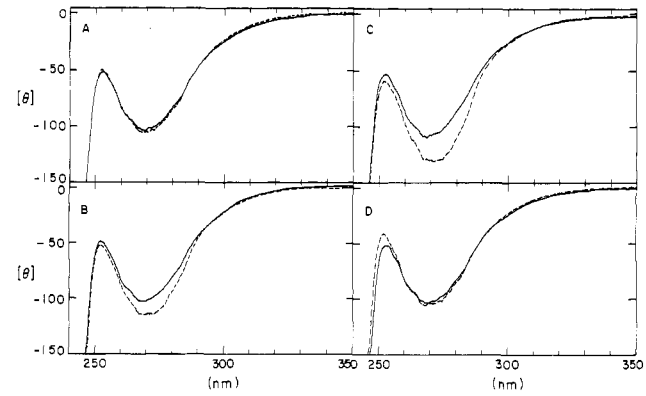

FIGURE 2: Near-UV CD spectra of human proinsulin: (A) 20% acetic acid–sodium acetate (pH 3.1) at 22 °C and at protein concentrations of 0.2 mM (solid line) and 2 mM (dotted line); (B) 5% acetic acid at 22 °C and at protein concentrations of 0.15 mM (solid line) and 1.5 mM (dotted line); (C) pH 3.1 in the absence of acetic acid at 22 °C and at protein concentrations of 0.2 mM (solid line) and 2 mM (dotted line); and (D) 20% acetic acid–sodium acetate (pD 3.1) at a protein concentration of 0.2 mM and at 22 °C (solid line) and 10 °C (dotted line).

In the near-UV CD spectrum, dissociation of the insulin or proinsulin dimer leads to attenuation of the negative ellipticity near 273 nm and a shift in the spectral minimum to shorter wavelengths (Wood et al., 1975; Strickland & Mercola, 1976; Horuk et al., 1980). These perturbations also may be used to determine the fraction of dimer in solutions used for NMR study. The observed features of the near-UV CD spectra of zinc-free human proinsulin in 20% acetic acid–sodium acetate (pH 3.1 and 22 °C) are characteristic of the insulin monomer (panel A of Figure 2); no significant changes are observed between proinsulin concentrations of 0.2 mM (solid line) and 2.0 mM (dashed line), consistent with the above UV absorbance studies. Corresponding data (pH 3.1 and 22 °C) are shown in 5% acetic acid–sodium acetate in panel B and in the absence of acetic acid in panel C. As expected, progressive changes in the concentration-dependent CD bands indicate that the fraction of dimer in solution increases as the percentage of acetic acid is reduced.

(B) NMR Characterization of Conditions. The NMR spectrum of proinsulin, like that of insulin (Weiss et al., 1989), exhibits changes with solvent composition, pH, and temperature which are consistent with those expected from the UV and CD results. As the concentration of acetic acid is increased to 20% (5 mM protein concentration and 22 °C), the resonances become sharper, and the spin systems which are broadened by intermediate exchange between monomer and dimer (B16, B24, B25, B25) become better defined. The spectrum is well resolved under these conditions, and the dispersion of chemical shifts, which is essentially unaffected by the percentage of acetic acid, is consistent with the retention of folded structure. This is illustrated in the TOCSY spectrum of human proinsulin shown in Figure 3. At pD 3.0 and 37 °C no perturbations are seen with increasing protein concentration in the range 0.1–5 mM, consistent with the optical data. Similarly, no concentration-dependent changes are observed in the photo-CIDNP enhancement of the B-chain tyrosines (see part III), which are sensitive to dimerization (Muszkat et al., 1984; Weiss et al., 1989). The use of 35% acetonitrile as an alternative solvent to obtain high-resolution spectra of insulin has also been described (Kline & Justice, 1990).

(C) Control for Solvent Perturbation. Unlike insulin, proinsulin may be observed as a monomer in the *absence* of

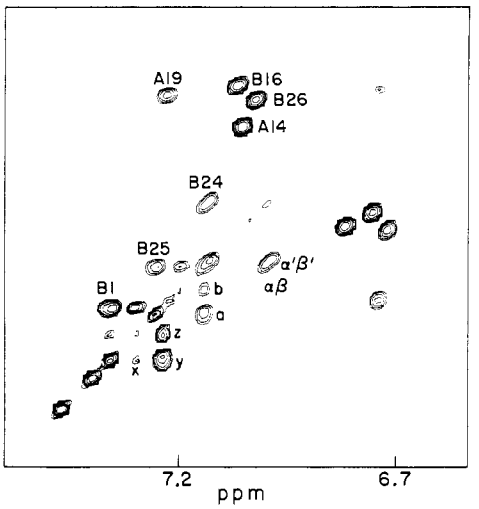

FIGURE 3: Aromatic region of the TOCSY spectrum of human proinsulin under conditions as described in the legend to Figure 4. The mixing time was 30 ms. The assignments of the four tyrosine and three phenylalanine spin systems are indicated. The ortho–metal cross-peak of PheB1 is labeled y; the corresponding meta–para cross-peak is labeled x; and the indirect ortho–para cross-peak is labeled z. The ortho–meta cross-peak is labeled b; the meta–para cross-peak is not observed at these contour levels. The ortho and meta resonances of PheB24 are broadened and the para resonance is not observed; the elongation of the ortho–metal cross-peak parallel to the diagonal reflects intermediate exchange between states $\alpha\beta$ and $\alpha'\beta'$.

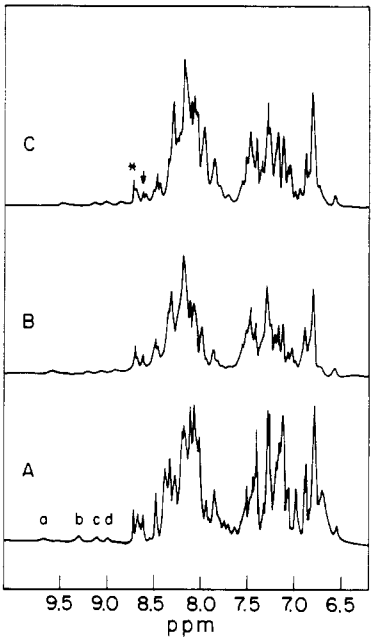

FIGURE 4: Exchangeable and aromatic ^1H NMR resonances of human proinsulin are shown at 600 MHz and 37 °C under the following conditions: (A) 0.5 mM proinsulin at pH 3.1 in 85% H_2O/15% D_2O; (B) 5 mM proinsulin in 10% acetic acid–sodium acetate/90% H_2O at pH 3.1; and (C) 5 mM proinsulin in 20% acetic acid–sodium acetate/80% H_2O at pH 3.1. The vertical scale in spectrum A is enlarged 1.5-fold relative to that in spectra B and C. The asterisk in spectrum C indicates the C_2H resonance of HisB10; the arrow indicates the C_2H resonance of HisB5. Resonances a, b, c, and d, are assigned to A11, B8, B9, and B6, respectively, by analogy to the sequential assignment of des-pentapeptide insulin (Hua & Weiss, 1990). Spectrum A is the sum of 1024 transients; spectra B and C are the sum of 128 transients. An exponential line broadening of 1 Hz was applied before Fourier transformation. The H_2O resonance was eliminated by presaturation.

acetic acid at 0.5 mM protein concentration and 37 °C. Spectra recorded at successive concentrations of acetic acid–sodium acetate (pH 3) thus provide a control for the effects of the mixed solvent system. The spectrum of human proinsulin in 90% H_2O/10% D_2O (pH 3) is shown in panel A of Figure 4. The amide resonances are notable for a large variation in line width; resonances a–d are shifted downfield and broadened. These features were previously observed in the spectrum of human insulin (Weiss et al., 1989) and are discussed further below. Corresponding spectra at 10% and 20% acetic acid–sodium acetate (pH 3) are shown in panels B and C, respectively. A similar pattern of aromatic and amide resonances are observed, including the C_2H resonances of HisB5 and HisB10 (indicated by an arrow and asterisk, respectively) and 2D NMR aromatic spin systems (data not shown). Remarkably, downfield resonances a–d are maintained in the presence of 20% acetic acid–sodium acetate. The small chemical shift differences observed with increasing concentrations of acetic acid in the range 0–20% are similar to the effects of increasing temperature in the range 25–45 °C and are likely to be mediated by changes in protein dynamics.

(II) Global Structure and Dynamics. (A) Amide Resonances and Conformational Substates. The amide resonances of the human proinsulin monomer exhibit large variations in line width (above; see also panel A of Figure 5). This is not due to solvent exchange, since similar amplitudes are observed following selective excitation (Hore, 1983; data not shown). The corresponding amide resonances of human insulin in 20% acetic acid–sodium acetate (pH 3) are shown in panel B of Figure 5. Remarkably, a similar pattern of chemical shift dispersion and line width variation is observed, including downfield resonances a–d. These resonances are assigned to CysA11 (resonance a), GlyB8 (b), SerB9 (c), and LeuB6 (d) following the sequential assignment of des-pentapeptide insulin [Hua & Weiss, 1990; see also Kline and Justice (1990)]. In addition to insulin-like resonances, the spectrum of proinsulin contains a tightly grouped set of sharp resonances between 7.8 and 8.5 ppm. These are similar to the spectrum of the isolated C-peptide (data not shown) and presumably arise from the connecting peptide (residues 31–65; Figure 1). Their narrow line widths and limited chemical shift dispersion suggest an absence of stably folded structure (see Added in Proof). The amide spectrum of human proinsulin thus appears as the superposition of insulin-specific and connecting-peptide resonances. A similar pattern is observed in the aliphatic region.

We attribute differential broadening to intermediate exchange among conformational substates of the protein. Such exchange presumably reflects millisecond motions within the protein which are not completely averaged on the NMR time scale. The correspondence between the pattern of broadening in proinsulin and insulin suggests that these motions occur primarily in the insulin moiety and are not damped by the connecting peptide. Presumably, more rapid fluctuations in the latter [expected to be on a submicrosecond time scale (Karplus, 1986)] result in a subset of resonances with narrow line widths.

Human and bovine insulin differ by three conservative substitutions, whereas human and bovine C-peptide sequences have more significantly diverged (Oyer et al., 1971; Nolan et al., 1971; see Table I). Interestingly, a correspondence between insulin-specific and C-peptide-specific resonances is also observed in bovine insulin and bovine proinsulin (panels C and D of Figure 5, respectively). As in the spectra of the corresponding human proteins, broad resonances are observed between 8.7 and 9.5 ppm (labeled a′–d′). Separation of time scales between the internal motions of the insulin moiety and

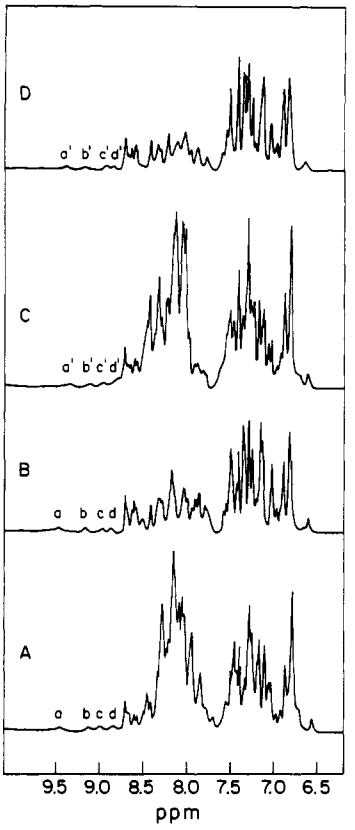

FIGURE 5: Exchangeable and aromatic ^1H NMR resonances of human proinsulin (panel A), human insulin (panel B), bovine proinsulin (panel C), and bovine insulin (panel D) are shown at 600 MHz and 37 °C in 20% acetic acid–sodium acetate/80% H_2O (pH 3.1). In each spectrum four exchange-broadened amide resonances are resolved downfield (labeled a–d in panels A and B, and labeled a′–d′ in panels C and D); accordingly, these are assigned to the insulin moiety of proinsulin. Human and bovine proinsulin were observed at 5 mM protein concentration; human and bovine insulin were observed at 1 mM protein concentration. Water elimination was accomplished by presaturation; an exponential broadening of 1 Hz was applied.

fluctuations in the connecting peptide thus appears to be general, i.e., not dependent on the divergent features of human and bovine connecting peptides (Table I). Motions on distinct time scales appear to be an intrinsic feature of the insulin fold and are not a consequence of a mixed solvent system.

(*B*) *Aromatic Assignment.* Systematic assignment of protein NMR resonances may be accomplished in certain proteins by the sequential analysis of amide–H_α connectivities (Wuthrich et al., 1983). However, sequential assignment techniques are difficult to apply to proinsulin due to variations in amide line widths, and so alternative methods of resonance assignment are required under these conditions. Aromatic assignments obtained in an earlier study of human insulin (Weiss et al., 1989) were based on comparative analysis of mutant and chemically modified insulin analogues. An extension of these assignments to proinsulin is verified by the study of related proinsulin analogues. These assignments are also consistent with the sequential assignment of des-pentapeptide insulin under these conditions (Hua & Weiss, 1990) and of human insulin in 35% acetonitrile (Kline & Justice, 1990).

The aromatic spin systems of human proinsulin are well resolved in the 2D NMR isotropic mixing spectrum (TOCSY), as shown in Figure 3. Histidine resonances are distinct and have been assigned previously in insulin (Bradbury et al., 1981; Cheshnovsky et al., 1983). The tyrosine and phenylalanine spin systems are readily distinguished, and their positions are

Table I: Connecting Peptide Sequences of Human and Bovine Proinsulin[a]

Human	Arg-Arg-*Glu-Ala-Glu-Asp*-Leu-Gln-Val-Gly-	(10)
Bovine	Arg-Arg-*Glu-Val-Glu-Gly*-Pro-Gln-Val-Gly-	
Human	Gln-Val-Glu-Leu-Gly-Gly-Gly-Pro-Gly-Ala-	(20)
Bovine	Ala-Leu-Glu-Leu-Ala-Gly-Gly-Pro-Gly-Ala-	
Human	Gly-Ser-Leu-Gln-Pro-Leu-Ala-Leu-Glu-Gly-	(30)
Bovine	Gly -- -- -- -- -- Gly-Leu-Glu-Gly-	
Human	Ser-Leu-Gln-Lys-Arg	
Bovine	Pro-Pro-Gln-Lys-Arg	

[a] Italics indicates conserved tetrapeptide at BC junction proposed by Gross et al. (1989) to be required for type I endopeptidase recognition.

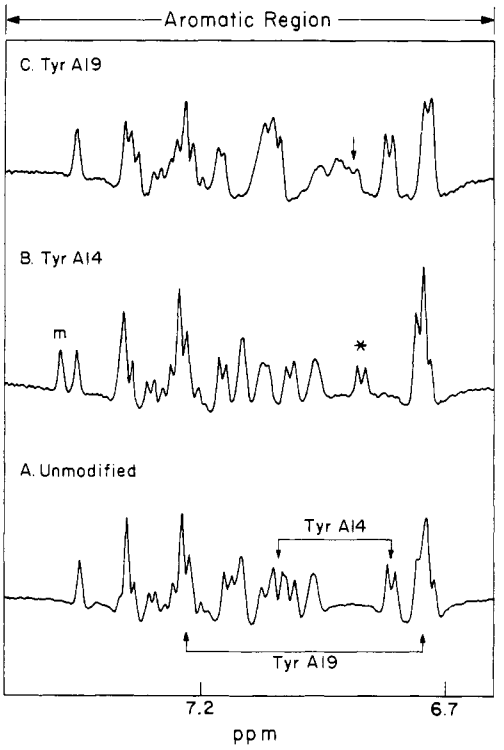

FIGURE 6: Aromatic region of the ^1H NMR spectra of bovine proinsulin (panel A), A14-I-bovine proinsulin (panel B), and A19-I-bovine proinsulin (panel C). The TyrA14 and TyrA19 spin systems are indicated in panel A. The H_5 resonance of I-3-TyrA14 is indicated by an asterisk in panel B; the H_5 resonance of I-3-TyrA19 is indicated by an arrow in panel C. The H_2 singlet resonance of I-3-TyrA14 is labeled m in panel B; the corresponding resonance in the spectrum of A19-I-proinsulin occurs at 7.8 ppm and is not shown. The spectra were acquired at 37 °C in 20% acetic acid–sodium acetate/80% D_2 (pD 3.0). Resolution was enhanced by convolution difference with exponential parameters 4 and 20 Hz and subtraction ratio 1.

similar to but not identical with their positions in corresponding spectra of human insulin (below). To confirm assignments based on analogy, we have independently assigned several of the aromatic spin systems in proinsulin. The assignment of the A-chain tyrosines (A14 and A19) is verified by selective monoiodination (Figure 6). The upfield portion of the aromatic spectrum of A14-$I_{3'}$-bovine proinsulin is shown in panel B. Minimal nonlocal perturbations are observed relative to

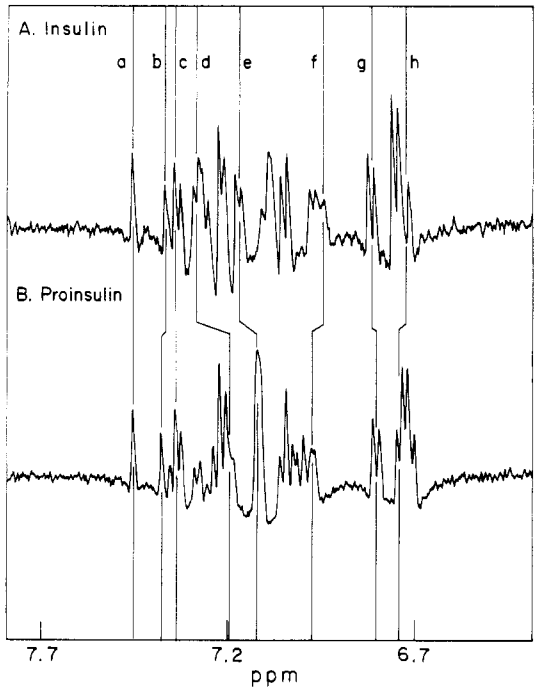

FIGURE 7: Aromatic ^1H NMR resonances of human insulin (panel A) and human proinsulin (panel B) are shown at 500 MHz and 37 °C. The proteins were 0.5 mM in 20% acetic acid–sodium acetate (pD 3.0, direct meter reading). Corresponding resonances are indicated by vertical lines labeled as follows: (a) HisB10 C_4H; (b) HisB5 C_4H; (c) PheB1 meta; (d) TyrA19 ortho; (e) PheB25 ortho; (f) PheB24 ortho; (g) TyrA14 meta; and (h) TyrB26 meta. Resolution was enhanced by convolution difference with exponential parameters 2 and 4 Hz and subtraction ratio 1. The spectra are the sum of 1024 transients; the residual HOD resonance was presaturated for 1 s.

the unmodified spectrum (panel A), consistent with the location of this ring on the surface of the A-chain loop (Blundell et al., 1971; Peking Insulin Structure Group, 1971). The remaining A14 H_5 resonance is indicated by an asterisk; the corresponding uncoupled H_2 singlet resonance is labeled m. In contrast, extensive changes are observed in multiple spin systems in the spectrum of A19-I_3-bovine proinsulin (panel C of Figure 6), reflecting the internal position of this residue. The remaining A19 H_5 resonance is identified in the COSY spectrum (not shown) and indicated by an arrow in panel C. The pattern of perturbations effected by A14 and A19 modification in proinsulin is similar to that described previously in modified insulins (Weiss et al., 1989).

The assignment of TyrB26 is verified by comparison of human proinsulin and (25,26) split proinsulin. The aromatic COSY spectrum of native human proinsulin and the split proinsulin analogue are shown in panel A of Figure 9; cross-peaks (positive contours) from proinsulin are shown in the upper left, and cross-peaks from (25,26) split proinsulin are shown in the lower right. The B26 cross-peak is significantly perturbed (arrow); its chemical shift and T_2 relaxation [observed in the COSY spectrum as relative cross-peak intensity (Weiss et al., 1984)] suggest a random-coil environment. The spectrum of this analogue is discussed further below. The remaining tyrosine spin system (B16) is assigned by elimination. Assignment of PheB1 is verified by inspection of the spectrum of des-(24,25) human proinsulin, in which B1 is the only phenylalanine spin system. Assignment of PheB24 and PheB25 in proinsulin is made by analogy to insulin (below) and has not been independently verified.

(*C*) *Proinsulin-Specific Perturbations.* The aromatic spectra of human insulin (panel A) and proinsulin (panel B) are shown in Figure 7. As indicated by vertical lines, significant changes are observed in the chemical shifts of corresponding residues. These changes are delineated by the spin systems in the COSY spectra shown in panel A of Figure 8; cross-peaks (positive contours) from human insulin are shown in the upper left, and cross-peaks from human proinsulin are shown in the lower right. In this format perturbations are observed as asymmetries in cross-peak position. Changes in chemical shift are observed in each of the spin systems except those of PheB1 and HisB10, which are disordered in solution (Weiss et al., 1989).

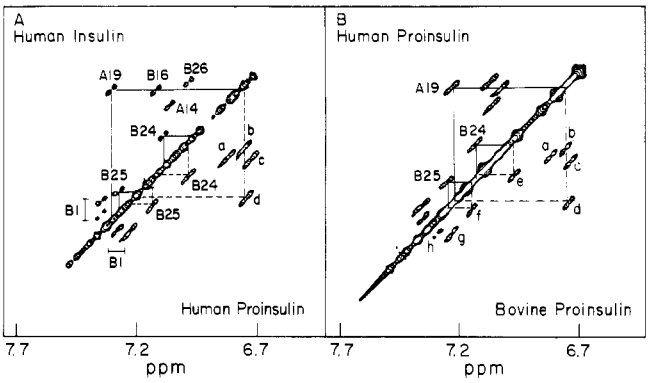

FIGURE 8: Aromatic regions of the COSY spectra in D_2O at 500 MHz and 37 °C of human insulin versus human proinsulin (panel A) and bovine proinsulin versus human proinsulin (panel B). In this format perturbations are seen as asymmetries in cross-peak position (spin systems A19, B16, B24, B25, and B26 in panel A). Only positive components of 2D antiphase multiplets are displayed. The assignments shown for human insulin are as described (Weiss et al., 1989); the tyrosine spin systems in the spectrum of human and bovine proinsulin are labeled a–d, and the phenylalanine cross-peaks are labeled e–h. The concentration of human insulin was 1 mM in 20% acetic acid–sodium acetate (pD 3.0); the concentration of human and bovine proinsulin was 5 mM in the same solvent. 1024 points were collected in t_2. 64 scans were acquired per t_1 value; 256 t_1 values were observed and zero-filled to 1024. A convolution difference with exponential parameters 4 and 20 Hz and subtraction ratio 1 was applied in both dimensions.

The generality of these perturbations may be evaluated by comparison of human and bovine proinsulin. Human and bovine insulins differ by three amino acid substitutions (ThrA8 → Ala, IleA10 → Val, and ThrB30 → Ala) (Brown et al., 1957; Oyer et al., 1971). Despite these differences, the overall spectra of human and bovine insulins are nearly identical. Human and bovine C-peptides differ substantially (Table I); the bovine C-peptide is five residues shorter than the human and contains nine substitutions among the remaining 26 residues (Oyer et al., 1971; Nolan et al., 1971). Remarkably, similar perturbations are observed in both human and bovine proinsulins (panel B of Figure 8), despite these differences in their connecting peptides. Thus, proinsulin-specific perturbations appear to be independent of the exact length or composition of the connecting peptide. However, connecting peptide sequences are not random; mammalian C-peptides are identical at 9/31 positions and exhibit similarities near the BC and CA junctions (Chan et al., 1983, 1984). Functional constraints affecting the evolution of these sequences are proposed to relate to prohormone processing (see Discussion).

(*D*) *Split Proinsulin Analogues.* The origins of proinsulin-specific perturbations may be explored by site-specific proteolysis, generating two-strand analogues as models for prohormone processing intermediates (split proinsulins; Peavy et al., 1985). With appropriate strand cleavage, the effects of tethering at either end of the connecting peptide may be distinguished. Three split proinsulin analogues have been prepared (Figure 1), containing breaks between residues 25

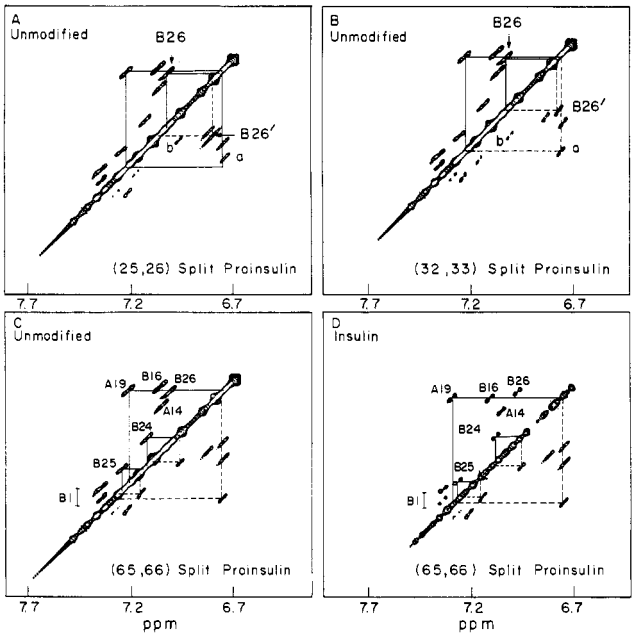

FIGURE 9: Aromatic regions of the COSY spectra in D_2O at 500 MHz and 37 °C of human proinsulin and various split proinsulins. The format and conditions are as described in the legend to Figure 8; only positive components of 2D antiphase multiplets are displayed. Panel A: Spectra of native human proinsulin (upper left) and (25,26) split proinsulin (lower right). The TyrB26 spin system in native proinsulin is outlined in the solid line (arrow in upper left); the random-coil TyrB26′ spin system in (25,26) split proinsulin is outlined in the dashed line (arrow in lower right). Perturbations are observed in the resonances of TyrA19 (a), PheB24 (b), and PheB25 (not labeled). Panel B: Spectra of native human proinsulin (upper left) and (32,33) split proinsulin (lower right). Spin systems are labeled as in panel A. Panel C: Spectra of native human proinsulin (upper left) and (65,66) split proinsulin (lower right). The spin systems of native proinsulin are indicated (A19, B24, and B25 in solid lines in upper left); the corresponding spin systems in (65,66) split proinsulin are outlined in dashed lines in lower right. Panel D: Spectra of human insulin (upper left) and (65,66) split proinsulin (lower right). Spin systems are labeled as in panel C.

and 26 (within the B-chain; arrow a), residues 32 and 33 (adjacent to the BC junction; arrow b) and residues 65 and 66 (the CA junction; arrow c).

The COSY spectrum of (32,33) split proinsulin is shown in panel B of Figure 9 (positive contours). The aromatic spectrum of unmodified human proinsulin is shown in the upper left, and the spectrum of the analogue is shown in the lower right. Remarkably, no significant perturbations are introduced by cleavage at this position (the BC junction). A small shift (<0.05 ppm) is observed in the PheB24 spin system (cross-peak b) in a direction (downfield) opposite to that observed in insulin (Figure 8A); the remaining aromatic resonances, including TyrA19 (cross-peak b), are not affected. In striking constrast, however, scission of the CA junction reverts the A19 and B24 spin systems to their positions in the spectrum of insulin (panels C and D in Figure 9). In addition, the B25 and B26 spin systems shift partially in the direction of the insulin spectrum. Thus, the CA junction, unlike the BC junction, alters the packing of insulin's hydrophobic core.

These data suggest the existence of stable structure at the CA junction, which we designate the "CA knuckle". This term is meant to suggest a structure that is accessible to solvent (and presumably to the type II endopeptidase), involves nonstandard secondary structure (in according with previous CD studies of proinsulin and the C-peptide), and is local (i.e., does not involve distant regions of the connecting peptide; see Added in Proof). Nonlocal effects of the CA knuckle are likely to be transmitted by the N-terminal A-chain α-helix, which packs against the C-terminal regions of the A- and B-chains in the crystal state (Blundell et al., 1971; Peking Insulin Structure Group, 1971). This transmitted conformational change is not due simply to modification of the N-terminal amino group, since the ^1H NMR spectrum of Arg-A_0-insulin is similar to that of native insulin under these conditions (data not shown). Complementary evidence for local structure in the connecting peptide has recently been obtained in CD studies of proinsulin unfolding (D. Brems, personal communication); a multiphasic transition is observed with a structural element in the connecting peptide unfolding at lower concentrations of denaturant.

In (25,26) split proinsulin the PheB24 and TyrA19 spin systems (cross-peaks a and b, respectively, in panel A of Figure 9) are also shifted upfield, i.e., in the direction of insulin. These changes in chemical shift may reflect the local absence of the TyrB26 ring current (see above). Alternatively, the putative CA knuckle may be stabilized by interactions involving the C-terminal region of the B-chain and are indirectly affected in this analogue. The spectrum of the related insulin analogue, des-pentapeptide insulin (DPI; Danho et al., 1975), is essentially unchanged from that of native insulin (Hua & Weiss, 1990).

(III) Local Structure and Dynamics. (A) Phenylalanine Ring Dynamics. The resonances of PheB1 in proinsulin are sharper than those of PheB24 and PheB25 and exhibit fewer NOEs, indicating that the N-terminus of the B-chain is flexible in solution. There is no evidence for a stable packing interaction with TyrA14, as observed in the 2Zn insulin structure (Blundell et al., 1971; Peking Insulin Structure Group, 1971). Similarly, PheB25 appears to be flexible; no NOE is observed with TyrA19, suggesting that B25 adopts an outward configuration in solution (molecule II; Blundell et al., 1971; Peking Insulin Structure Group, 1971). These features of proinsulin are similar to those previously described in the spectrum of insulin (Weiss et al., 1989) and are consistent with the position of PheB25 in the crystal structure of a des-pentapeptide insulin analogue (Bi et al., 1984; Jinbi et al., 1987).

The resonances of PheB24 are broadened; the para resonance is not observed and may overlap the ortho resonance. Distortion of the TOCSY cross-peak (Figure 3) suggests that it is in intermediate exchange among related ring configurations $\alpha\beta$ and $\alpha'\beta'$; such exchange may involve constrained ring rotation about χ_2 or more extended motions including χ_1 and χ_2 (e.g., ring sliding). The apparent absence of a sharp para resonance suggests the latter, since it would not be broadened by ring rotation. Evidence for a second type of motion is provided by the temperature dependence of the PheB24 resonances (asterisk in Figure 10). Ordinarily one would expect intermediate to slow exchange broadening to be "frozen" in slow exchange as the temperature is lowered. However, in the range 25–0 °C disproportionate broadening of the B24 resonances is observed, suggesting a supervening exchange process on an intermediate time scale. In the crystal state electron density for PheB24 is well-defined in a single configuration, which is also observed in the structure of the monomeric analogue des-pentapeptide insulin (Bi et al., 1984; Jinbi et al., 1987).

Disporportionate but less extensive broadening is observed in the TyrA19 and TyrB26 spin systems in the temperature range 25–0 °C (data not shown). Such broadening suggests the existence of barriers to ring rotation and is consistent with molecular mechanics calculations (Gelin & Karplus, 1975) based on the crystal structures (Weiss et al., 1989). Corre-

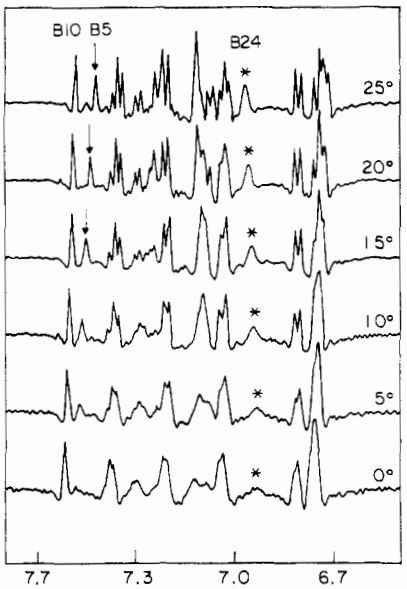

FIGURE 10: Aromatic resonances of human proinsulin are shown as a function of temperature in the range 0–25 °C. The cold-induced broadening of the HisB5 C_4H resonance and PheB24 meta resonance is indicated by an arrow and an asterisk, respectively. Resolution was enhanced by convolution difference with exponential parameters 2 and 4 Hz and subtraction ratio 1.

sponding changes in ring dynamics are also suggested by temperature-dependent perturbations in the near-UV CD spectrum of proinsulin under these conditions. As shown in panel D of Figure 2, a change is observed in the relative maximum near 254 nm between 22 and 10 °C which is opposite to that associated with dimerization (panels B and C; Wood et al., 1975; Horuk et al., 1980). However, disulfide bonds also contribute to the observed ellipticity in this region and cannot be resolved.

(*B*) *The Histidine Resonances.* The C_4H resonance of HisB5 exhibits an NOE to an isoleucine spin system in the NOESY spectrum of human proinsulin and to a valine spin system in the NOESY spectrum of bovine proinsulin (data not shown). Interestingly, human and bovine insulin differ in three sites, including IleA10 (human) and ValA10 (bovine). In the crystal structures of human and bovine insulin, HisB5 is observed in a local pocket bounded in part by IleA10 and ValA10, respectively. This pocket is shown in panel A of Figure 11. Although the isoleucine and valine spin systems involved in the B5 NOE have not been independently assigned, the bovine- and human-specific NOE patterns suggest that the HisB5 pocket is retained in solution. In the crystal state HisB5 is observed in two distict orientations within the pocket (molecule II is shown in Figure 11; Blundell et al., 1971); similar orientations are observed in the crystal structure of a monomeric insulin analogue (Bi et al., 1984; Jinbi et al., 1987). Exchange between these states may account for the extensive broadening exhibited by HisB5 at low temperatures (arrow in Figure 10). Similar B5 exchange features have been observed in the spectrum of human insulin (Weiss et al., 1989); the absence of nonlocal damping by the connecting peptide is consistent with observations of comparable conformational

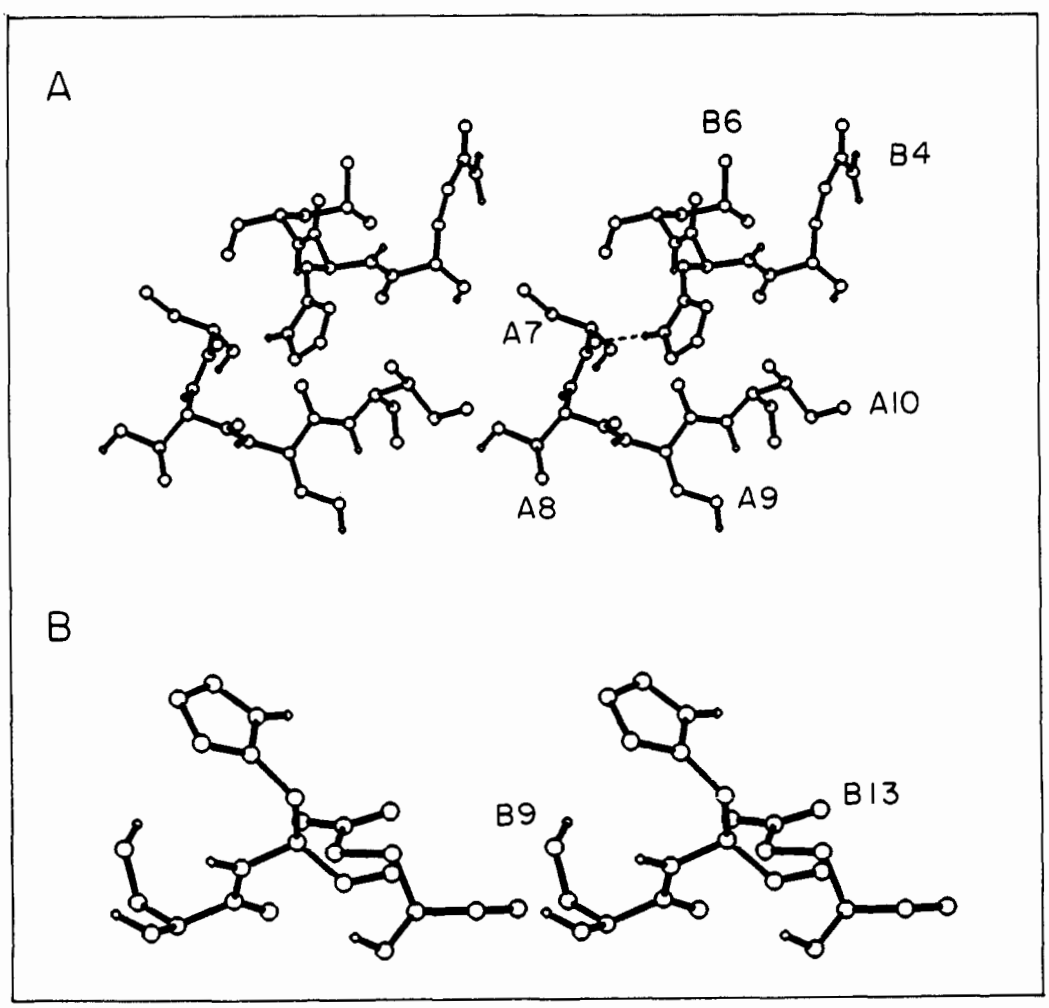

FIGURE 11: Stereoviews of local environments of HisB5 (panel A) and HisB10 (panel B) in molecule II of the 2Zn crystal structure (Blundell et al., 1971; Peking Insulin Structure Group, 1971).

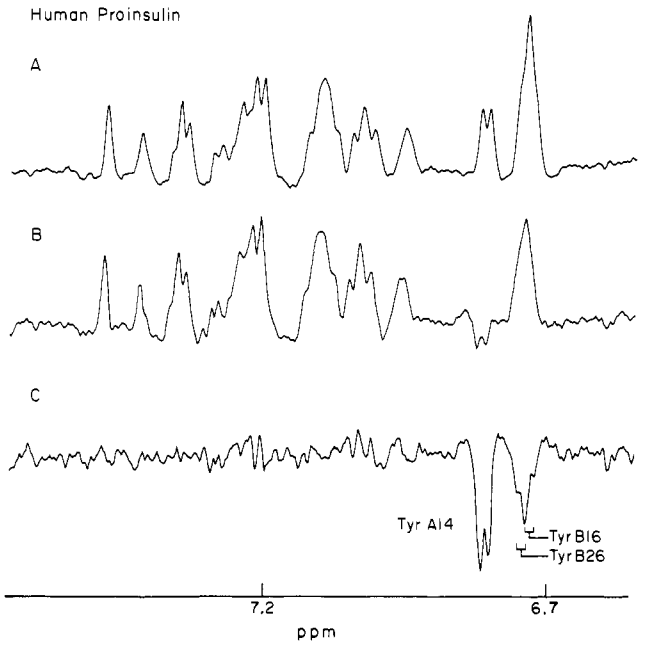

FIGURE 12: Photo-CIDNP spectrum of 4 mM human proinsulin at 500 MHz and 30 °C in 20% acetic acid–sodium acetate (pD 3.0): (A) dark spectrum; (B) spectrum following photoexcitation; and (C) difference spectrum. The photo-CIDNP enhancements of TyrA14, TyrB16, and TyrB26 are indicated in panel C. Resolution was enhanced in spectra A and B by convolution difference with exponential parameters 4 and 20 Hz and subtraction ratio 1. An exponential broadening of 2 Hz was applied in the difference spectrum (panel C).

broadening of insulin-specific amide resonances in insulin and proinsulin (see part I). In contrast, the ring resonances of HisB10 are sharper than those of B5, exhibit fewer NOESY cross-peaks, and do not undergo exchange broadening. These results are in accord with the presumed position of B10 on the surface of the zinc-free insulin monomer (panel B of Figure 11). The chemical shifts of the HisB10 ring resonances are identical in insulin and proinsulin under these conditions.

(*C*) *Tyrosine Accessibility.* The environment of tyrosine residues can be investigated by photochemically induced dynamic nuclear polarization (photo-CIDNP). This technique is based on a reversible photochemical reaction of protein side chains (tyrosine and histidine) with a dye excited to the triplet state (Kaptein, 1980; Berliner & Kaptein, 1981; Muszkat & Gilon, 1978; Muszkat et al., 1984). The insulin monomer exhibits a simple pattern of photo-CIDNP enhancements under acidic conditions (Weiss et al., 1989). Three enhanced signals of equal amplitude are observed, assigned to TyrA14, TyrB16, and TyrB26. The inaccessibility of TyrA19 to photo-CIDNP enhancement reflects its location in a hydrophobic pocket (Blundell et al., 1971; Peking Insulin Structure Group, 1971) and provides a marker for tertiary structure.

The photo-CIDNP spectra of human proinsulin are shown in Figure 12; the "dark" spectrum is shown in panel A, the spectrum following laser irradiation is shown in panel B, and the calculated difference spectrum is shown in panel C. The photo-CIDNP enhancement of TyrA14 is similar in insulin and proinsulin, indicating that the connecting peptide does not shield this portion of the A-chain; this result is in accord with molecular modeling by Blundell et al. (1978). Unlike insulin, however, only partial enhancements are observed for TyrB16 and TyrB26. This may reflect steric interference by the connecting peptide, limiting accessibility of the B-chain tyrosines (but not A14) to riboflavin (mechanism I). Alternatively, tethering by the connecting peptide may stabilize local

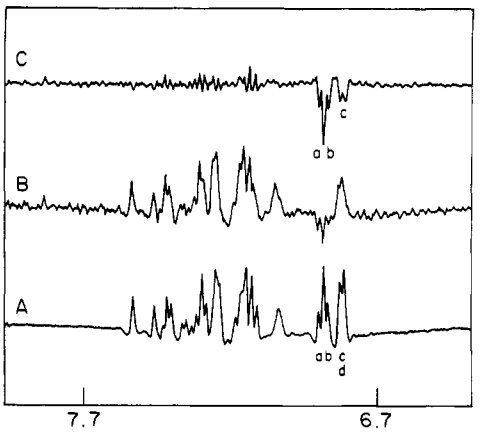

FIGURE 13: Photo-CIDNP spectrum of 2 mM (25,26) split human proinsulin is shown at 500 MHz and 30 °C in 20% acetic acid–sodium acetate (pD 3.0): (A) dark spectrum; (B) spectrum following photoexcitation; and (C) difference spectrum. The photo-CIDNP enhancements of random-coil TyrB26 (labeled a), TyrA14 (labeled b), and TyrB16 (labeled c) are indicated in panel C. Resolution was enhanced in spectra A and B by convolution difference with exponential parameters 2 and 4 Hz and subtraction ratio 1. An exponential broadening of 2 Hz was applied in the difference spectrum (panel C).

packing interactions involving the α-helix and C-terminal β-strand of the B-chain (mechanism II; Dodson et al., 1979; Baker et al., 1988).

Corresponding photo-CIDNP spectra of (25,26) split proinsulin are shown in Figure 13. Complete enhancement of TyrB16 (resonace c in panel C) and TyrB26 (resonance b) is observed in this analogue relative to TyrA14 (resonance a). TyrA19 (resonance d) is not enhanced, indicating that its position in a hydrophobic cleft is maintained. In (32,33) and (65,66) split proinsulins and photo-CIDNP enhancement of B16 and B26 is intermediate between those of insulin and proinsulin (data not shown). These results are consistent with either mechanism I or mechanism II.

Discussion

Prohormone processing by compartment-specific endopeptidases is a general motif in eukaryotic cell biology (Docherty & Steiner, 1982; Fisher & Scheller, 1988). Recognition by such endopeptidases appears to involve sequence or structural features in addition to the local target site (Docherty et al., 1989; Gross et al., 1989). In the case of proinsulin proteolysis two distinct endopeptidases (types I and II) have been identified (Davidson et al., 1988); type I is specific for the BC junction, and type II is specific for the CA junction. To define potential recognition elements at these junctions, we have undertaken comparative ^{1}H NMR studies of human proinsulin and insulin. The overall proinsulin spectrum is largely the sum of the component subspectra: an insulin-specific subspectrum whose resonances are well dispersed and of varying line width and a C-peptide-specific subspectrum whose resonances are sharp and clustered. Proinsulin thus may be approximately modeled as a folded moiety (the insulin core) and an unfolded moiety (the connecting peptide), as previously suggested by comparative physicochemical studies of insulin, proinsulin, and the C-peptide (Frank et al., 1972). This model is also supported by recent X-ray diffraction studies by Blundell and co-workers (see Added in Proof).

A biological role for the connecting peptide has not been defined. Although its sequence has more significantly diverged than those of the A- and B-chains, vertebrate connecting peptides are of similar length and exhibit nonrandom sequence elements (Chan et al., 1981, 1984). A variety of functional

roles have been proposed, including the orientation of sulfhydryl groups in the pathway of protein folding (Steiner, 1978); the provision of a minimal peptide length for transmembrane transport in the rough endoplasmic reticulum (Steiner, 1978, 1984); or the regulation of intracellular trafficking and export (Powell et al., 1988; Parham, 1988). These proposals are not supported by studies of mutant proinsulins containing shortened connecting peptides (Thim, 1986; Markussen, 1986a,b; Powell, 1988). Recently Halban and colleagues (Gross et al., 1989) have suggested that conserved elements of the connecting peptide are involved in specific endopeptidase recognition. This suggestion has motivated the present structural examination of split proinsulin analogues as models for prohormone processing intermediates.

The present studies have been performed under acidic conditions to reduce higher order aggregation. The increased net positive charge of insulin and proinsulin under acidic conditions reduces the dimerization constant (Jeffrey & Coates, 1966a,b; Pekar & Frank, 1972; Goldman & Carpenter, 1975). We have previously described NMR studies of the insulin monomer in an acetic acid mixed solvent (20% acetic acid–sodium acetate; Weiss et al., 1989). These conditions were found to weaken hydrophobic interactions between insulin monomers and yet preserve native-like structure. Dimerization of proinsulin is significantly weaker than that of insulin in the range 0–20% acetic acid–sodium acetate (pH 3). Accordingly, the proinsulin monomer may be studied by one-dimensional NMR methods at low protein concentrations in the *absence* of acetic acid, providing a direct control for the influence of solvent composition (part I). The present studies of human proinsulin have been conducted under the same conditions as previously described for human insulin (20% acetic acid–sodium acetate, pH 3) to facilitate comparative analyses. NMR assignments of the aromatic residues have been obtained by comparison with insulin and verified by studies of chemically modified analogues (part II). Since there are no aromatic residues in the human C-peptide (Figure 1), these resonances provide probes for the solution structure and dynamics of the insulin moiety (part III). This is the first NMR study of the proinsulin monomer to be described.

Conformational Substates of Insulin and Proinsulin. Conformational flexibility of insulin has previously been inferred from structural differences observed in various crystal states (Blundell et al., 1971; Peking Insulin Structure Group, 1971; Bentley, 1976; Chothia et al., 1983; Dodson et al., 1984). Evidence for an equilibrium between conformational substates in both insulin and proinsulin is provided by the observation of large variations in amide line widths. Since these variations are not due to solvent exchange, a structural mechanism is suggested, i.e., the existence of a range of conformations with different chemical shifts that are incompletely averaged on the NMR time scale. The similarity of these resonances in insulin and proinsulin implies that such motions are not damped by the connecting peptide. Because these motions occur on a millisecond time scale, they must be distinguished from the nanosecond fluctuations expected for large-scale displacement of a random-coil segment and subnanosecond motions probed by molecular dynamics simulations (Karplus, 1987). Similar features are observed in the spectra of human and bovine insulins and proinsulins.

Local Dynamics of Insulin and Proinsulin. The insulin moiety of proinsulin exhibits local structural and dynamic features which are similar to those observed previously for insulin (Weiss et al., 1989). In both proteins, for example, HisB10 is flexible and exhibits few packing interactions, consistent with its position on the surface of the B-chain helix in the crystal state (Figure 11; Blundell et al., 1971; Peking Insulin Structure Group, 1971). Although HisB5 also lies on the surface of the protein, it is stably packed in a pocket formed by CysA7 and IleA10 (ValA10 in bovine insulin; Nolan et al., 1971). The two promoters in the 2Zn crytal exhibit alternate hydrogen-bonding patterns (Blundell et al., 1971; Peking Insulin Structure Group, 1971), which alter the environments of the H_2 and H_4 ring protons. Our results demonstrate that this pocket is similar in insulin and proinsulin, and in each protein the B5 ring exhibits exchange phenomena with decreasing temperature.

In the crystal PheB24 projects into the hydrophobic interior and packs against LeuB15 on one side and the A20–B19 disulfide on the other (Blundell et al., 1971; Peking Insulin Structure Group, 1971). TyrA19 projects into a hydrophobic crevice, where it packs against IleA2, LeuA6, and LeuB15. In solution these residues exhibit prominent NOEs with methyl resonances and, accordingly, provide sensitive probes for the ordering of the hydrophobic core. The exchange features of PheB24, which suggest alternative configurations in solution, are in accord with the anomalous bioactivities of D-amino acids at this position (Kobayashi et al., 1982; Casaretto et al., 1987; Mirmira & Tager, 1989). The inaccessibility of TyrA19 in a local cleft is demonstrated by an absence of photo-CIDNP accessibility. Broadening of the A19 resonances is observed at lower temperature (0–20 °C), consistent with constrained ring rotation or more extended motions (e.g., ring swinging). The latter mechanism is suggested by the asymmetric pattern of A19 *B* values in the 2Zn molecule II (Baker et al., 1988).

PheB1 and TyrA14 are flexible in solution in both insulin and proinsulin. No perturbation in B1 is seen in either protein following chemical iodination of TyrA14, and no NOE is observed between these rings. These results indicate that the A14–B1 interaction in the 2Zn crystal is unlikely to be maintained in zinc-free solution (Blundell et al., 1971; Peking Insulin Structure Group, 1971; Wollmer et al., 1979). The flexibility of TyrA14 is consistent with its photo-CIDNP enhancement (which is unchanged from that in insulin) and with the variety of side-chain configurations observed in different crystal forms (Baker et al., 1988). PheB25 also appears to be a surface residue and is not observed in two configurations as in the crystallographic dimer.

Photo-CIDNP spectra indicate that the acessibility of TyrA14 is similar in insulin and proinsulin, as predicted by Blundell et al. (1978). However, the B-chain tyrosines (B16 and B26) are less accessible in proinsulin than in insulin. These data are consistent with either of two mechanisms. Either the connecting peptide directly shields these sites from collision with the dye or decreased accessibility results from nonlocal stabilization of structure in the B-chain. In the crystal state TyrB16 and TyrB26 participate in an interaction between the B-chain α-helix and C-terminal β-strand, stabilizing a U-turn which also packs against the N-terminal α-helix of the A-chain (Blundell et al., 1971; Peking Insulin Structure Group, 1971). Either or both mechanisms are consistent with the partial increase in photo-CIDNP enhancement that is observed following cleavage of either junction.

Proinsulin-Specific Perturbations and Split Proinsulin Analogues. As monitored by ^{1}H NMR chemical shifts, changes are observed in the packing of TyrA19 and PheB24 in the hydrophobic interior of human proinsulin relative to human insulin. A similar pattern of perturbations is observed in bovine proinsulin relative to bovine insulin, despite differences in the connecting peptides of the human and bovine

proinsulins (Oyer et al., 1971; Nolan et al., 1971). Corresponding changes are observed in neighboring methyl resonances, assigned by NOEs to LeuB15, IleA2, and ValB12 on the basis of the crystal structure (Blundell et al., 1971; Peking Insulin Structure Group, 1971). This is the first observation of a structural interaction between the insulin moiety of proinsulin and the connecting peptide. Remarkably, these perturbations are reverted by cleavage of the CA junction but not by cleavage of the BC junction (arrows c and b in Figure 1, respectively). This effect is not due simply to modification of the N-terminal amino group of the A-chain, since the aromatic resonances of Arg-A_0-insulin are similar to those of insulin rather than proinsulin.

We propose that a local structure is formed at the CA junction involving a portion of the connecting peptide, which induces a small change in the packing of the insulin core. This proposal is consistent with the results of recent CD unfolding studies of proinsulin in which an autonomous unfolding transition is observed in the connecting peptide (D. Brems, personal communication). As determined by NMR, this structure is maintained following cleavage of the BC junction; significantly, the corresponding processing intermediate is likely to be the physiological substrate for type II proinsulin endopeptidase (Davidson et al., 1988). We propose that this local structure (designated the CA knuckle) forms part of the recognition element for type II endopeptidase. This hypothesis is consistent with the conservation of C-peptide residues near the CA junction (Gross et al., 1989), which may reflect functional constraints related to the regulated pathway of prohormone processing. The absence of a comparable structure at the BC junction is consistent with its presumed flexibility (Dodson et al., 1979; Baker et al., 1988) and does not rule out context-dependent recognition of Arg^{31}–Arg^{32} by type I endopeptidase as proposed by Gross et al. (1989). These proposals may be tested by studies of proinsulin analogues obtained by site-directed mutagenesis.

Conclusions

The aromatic NMR resonances of human proinsulin have been assigned by comparison with the spectrum of human insulin and verified by chemical modification and partial proteolysis. The two histidine, three phenylalanine, and four tyrosine residues are observed to be in distinct local environments and to provide sensitive markers for comparative studies of split proinsulin analogues as models for prohormone processing intermediates. The environments of the tyrosine residues have also been investigated by use of photochemically induced dynamic nuclear polarization (photo-CIDNP).

Comparison of the observed features of proinsulin and insulin leads to the following five conclusions. (i) NMR features (chemical shifts, nuclear Overhauser enhancements, and resonance line widths) of human proinsulin correspond in part to those of human insulin, indicating that the insulin moiety is folded appropriately in the prohormone. The connecting peptide is largely unstructured. (ii) Insulin-specific amide resonances are broadened in the proinsulin monomer by intermediate exchange, reflecting a distribution of conformational substates related by motions on the millisecond time scale. Since a similar pattern of broadening is observed in the spectrum of the insulin monomer, the connecting peptide does not damp the motions involved. (iii) Dimerization of proinsulin is substantially weaker (under acidic conditions) than that of insulin. (iv) TyrB16 and TyrB26 are less accessible to photo-CIDNP enhancement in proinsulin than in insulin. This is likely to reflect either shielding by the connecting peptide or nonlocal stabilization of the B-chain surface in proinsulin. This difference in accessibility is reduced by cleavage at either the BC or CA junctions. (v) Differences are observed between the internal packing of proinsulin and that of insulin in the neighborhood of PheB24 and TyrA19 (i.e., the hydrophobic core). Comparative studies of split proinsulin analogues demonstrate that these differences are due to tethering of the N-terminus of the A-chain (CA junction) rather than tethering of the C-terminus of the B-chain (BC junction). We propose that a stable local structure exists at the CA junction and may provide a recognition element for type II proinsulin endopeptidase. In the future this hypothesis may be tested by site-directed mutagenesis of the CA junction.

Added in Proof

A preliminary crystal structure of bovine proinsulin has recently been obtained under acidic conditions by T. Blundell and co-workers (A. Cleasby, J. Murray-Rust, H. Driessen, S. P. Wood, I. J. Tickle, and T. L. Blundell, manuscript in preparation). Consistent with the present NMR results, the connecting peptide appears to be largely disordered. The N-terminal residues of the A-chain retain a helical conformation, and adjoining connecting-peptide residues may be in a position to interact with TyrA19 in accord with the pattern of NMR perturbations.

Acknowledgments

NMR spectra were obtained at the MIT Francis Bitter National Magnet Laboratory (NIH RR-00995), the NMR Facility of the University of Wisconsin at Madison (RR-02301 and NSF Grant PCM-8415048), the Harvard Medical School NMR Facility, the Laboratory of Chemical Physics (National Institutes of Health), and Varian, Inc. (Palo Alto, CA). We thank Dr. Ad Bax (NIH), Dr. W. M. Westler and Dr. B. Stockman (University of Wisconsin), and Dr. George Gray (Varian, Inc.) for helpful discussion and instrument time; Dr. Ron Chance (Lilly Research Laboratories) for insulin analogues; D. N. Brems for communication of results prior to publication; Michael Beckage (Eli Lilly) for assistance with CD measurements; Dr. H. T. Keutmann (Massachusetts General Hospital) for advice regarding peptide chemistry; Dr. D. Nguyen and Prof. M. Karplus for molecular mechanics calculations; E. K. O'Shea for technical assistance in the early stages of the work; and Dr. C. Rhodes, Prof. D. F. Steiner (University of Chicago), and Prof. A. Efstradiatis (Columbia College of Physicians and Surgeons) for discussion regarding prohormone processing. M.A.W. and S.E.S. thank Prof. Eugene Braunwald and Prof. John Potts, Jr., for their encouragement.

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.

Anasari, A., Berendze, J., Bowne, S. F., Frauenfelder, H., Iben, I. E. T., Sauke, T. B., Shyamsun, E., & Young, R. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 5000–5004.

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, B 319*, 389–456.

Bentley, G., Dodson, E., Dodson, G., Hodgkin, D., & Mercola, D. (1976) *Nature (London) 261*, 166–168.

Berliner, L. J., & Kaptein, R. (1981) *Biochemistry 20*, 799–807.

Bi, R. C., Dauter, Z., Dodson, E. J., Dodson, G. G., Giordano, F., & Reynolds, C. D. (1984) *Biopolymers 32*, 391–395.
Blundell, T. L., & Humbel, R. E. (1980) *Nature (London) 287*, 781–787.
Blundell, T. L., Bedarkar, S., Rinderknecht, E., & Humbel, R. E. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 180–184.
Bradbury, J. H., & Brown, L. R. (1977) *Biochemistry 16*, 573–582.
Bradbury, J. H., & Ramesh, V. (1985) *Biochem. J. 229*, 731–737.
Bradbury, J. H., Ramesh, R., & Dodson, G. (1981) *J. Mol. Biol. 150*, 609–613.
Brems, D. N., Brown, P. L., Heckenlaible, L. A., & Frank, B. H. (1990) *Biochemistry* (submitted for publication).
Brown, H., Sanger, F., & Kitai, R. (1955) *Biochem. J. 60*, 556–565.
Casaretto, M., Spoden, M., Diaconescu, C., Gattner, H.-G., Zahn, H., Brandenberg, D., & Wollmer, A. (1987) *Biol. Chem. Hoppe-Seyler 368*, 709–716.
Chan, S. J., Emdin, S. O., Kwok, S. C. M., Kramer, J. M., Falkmer, S., & Steiner, D. F. (1983) *J. Biol. Chem. 256*, 7595–7602.
Chan, S. J., Episkopou, V., Zeitlin, S., Karathanasis, S. K., MacKrell, A., Steiner, D. F., & Efstratiadis, A. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 5046–5050.
Chance, R. E., Ellis, R. M., & Bromer, W. W. (1968) *Science 161*, 165–167.
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.
Danho, W. O., Gattner, H.-G., Nissen, D., & Zarn, H. (1975) *Hoppe-Seyler's Z. Physiol. Chem. 356*, 1406–1412.
Davidson, H. W., & Hutton, J. C. (1987) *Biochem. J. 245*, 575–582.
Davidson, H. W., Rhodes, C. J., & Hutton, J. C. (1988) *Nature 333*, 93–96.
Davis, D. G., & Bax, A. (1985) *J. Magn. Reson. 64*, 533–535.
Docherty, K., & Steiner, D. F. (1982) *Annu. Rev. Physiol. 44*, 625–638.
Docherty, K., & Hutton, J. C. (1983) *FEBS Lett. 162*, 137–141.
Docherty, K., Rhodes, C. J., Taylor, N. A., Shennan, K. I. J., & Hutton, J. C. (1989) *J. Biol. Chem. 31*, 18335–18339.
Dodson, E. J., Dodson, G. G., & Hodgkin, D. C. (1979) *Can. J. Biochem. 57*, 469–479.
Dodson, E. J., Dodson, G. G., Hubbard, R. E., & Reynolds, C. D. (1984) *Biopolymers 22*, 281–291.
Fisher, J. M., & Scheller, R. H. (1988) *J. Biol. Chem. 263*, 16515–16518.
Frank, B. H., & Veros, A. J. (1970) *Biochem. Biophys. Res. Commun. 38*, 284–288.
Frank, B. H., Pekar, A. H., & Veros, A. J. (1972) *Biochemistry 11*, 4926–4931.
Frank, B. H., Bruck, P. J., Hutchins, J. F., & Root, M. A. (1982) in *Neue Insulin* (Petersen, K.-G., Schleuter, K. J., & Kerp, L., Eds.) pp 45–50, Freiburger Graphische Betriebe, Freiburg, Germany.
Frank, B. H., Peavy, D. E., Hooker, C. S., and Duckworth, W. C. (1983) *Diabetes 32*, 705–711.
Gabbay, K. H., Deluca, K., Fisher, J. N., Mako, M. E., and Rubenstein, A. H. (1976) *N. Engl. J. Med. 294*, 911–915.
Gelin, B., & Karplus, M. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 2002–2006.
Goldman, J., & Carpenter, F. H. (1974) *Biochemistry 13*, 4566–4574.
Grant, P. T., Coombs, T. L., & Frank, B. H. (1972) *Biochem. J. 126*, 433–440.
Gross, D. J., Villa-Komaroff, L., Kahn, C. R., Weir, G. C., & Halban, P. A. (1989) *J. Biol. Chem. 264*, 21486–21490.
Hore, P. J. (1983) *J. Magn. Reson. 54*, 539–542.
Horuk, R., Blundell, T. R., Lazarus, N. R., Neville, R. W. J., Stone, D., & Wollmer, a. (1980) *Nature (London) 286*, 822–824.
Hua, Q. X., & Weiss, M. A. (1990) *Biochemistry* (in press).
Hua, Q.-X., Chen, Y.-J., Wang, C.-C., Wang, D.-C., & Roberts, G. C. K. (1989) *Biochim. Biophys. Acta 994*, 114–120.
Inouye, K., Watanabe, K., Tochino, T., Kanaya, T., Kobayashi, M., & Shigeta, Y. (1981) *Experientia 37*, 811–813.
Jeffrey, P. D., & Coates, J. H. (1966a) *Biochemistry 5*, 489–498.
Jeffrey, P. D., & Coates, J. H. (1966b) *Biochemistry 5*, 3820–3824.
Jinbi, D., Meizhen, L., Junming, Y., & Dongcai, L. (1987) *Sci. Sin. 30*, 55–65.
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, NY.
Karplus, M. (1987) *Ann. N.Y. Acad. Sci. 482*, 255–266.
Kline, A. D., & Justice, R. M., Jr. (1990) *Biochemistry 29*, 2906–2913.
Kruger, P., Strassburger, W., Wollmer, A., van Gunsteren, W. F., & Dodson, G. G. (1987) *Eur. Biophys. J. 14*, 449–459.
Markussen, J. (1985a) *Int. J. Pept. Protein Res. 25*, 431–434.
Markussen, J. (1985b) *Int. J. Pept. Protein Res. 26*, 70–77.
Muszkat, K. A., & Gilon, C. (1978) *Nature (London) 271*, 685–686.
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*, 12054–12058.
Nolan, C., Margoliash, E., Peterson, J. D., & Steiner, D. F. (1971) *J. Biol. Chem. 246*, 2780–2796.
Oyer, P. e., Cho, S., Peterson, J. D., & Steiner, D. F. (1971) *J. Biol. Chem. 246*, 1375–1386.
Palmieri, R., Lee, R. W.-K., & Dunn, M. F. (1988) *Biochemistry 27*, 3387–3397.
Parham (1988) *Nature (London) 333*, 500–503.
Paselk, R. A., & Levy, D. (1974) *Biochim. Biophys. Acta 359*, 215–221.
Peavy, D. E., Brunner, M. R., Duckworth, W. C., Hooker, C. S., & Frank, B. H. (1985) *J. Biol. Chem. 260*, 13989–13994.
Pekar, A. H., & Frank, B. H. (1972) *Biochemistry 11*, 4013–4016.
Peking Insulin Structure Group (1971) *Peking Rev. 40*, 11–16.
Powell, S. K., Orci, L., Craik, C. S., & Moore, H.-P. H. (1988) *J. Cell Biol. 106*, 1843–1851.
Rhodes, C. J., Zumbrunn, A., Bailyes, E. M., Shaw, E., & Hutton, J. C. (1989) *Biochem. J. 258*, 305–308.
Robbins, D. C., Blix, P. M., Rubenstein, A. H., Kanazawa, Y., Kosaka, K., & Tager, A. H. (1981) *Nature 291*, 679–681.

Robbins, D. C., Shoelson, S. E., Rubenstein, A. H., & Tager, H. S. (1984) *J. Clin. Invest. 73*, 714–719.
Rupley, J. A., Renthal, R. D., & Praissman, M. (1976) *Biochim. Biophys. Acta 140*, 185.
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.
Steiner, D. F. (1979) *Diabetes 27*, Suppl. 1, 145–148.
Steiner, D. F. (1984) *Harvey Lect. 78*, 191–228.
Steiner, D. F., & Oyer, P. E. (1967) *Proc. Natl. Acad. Sci. U.S.A. 57*, 473–480.
Steiner, D. F., Cunningham, D. D., Spigelman, L., & Aten, B. (1967) *Science 157*, 697–700.
Steiner, D. F., Cho, S., Oyer, P. E., Terris, S., Peterson, J. D., & Rubenstein, A. H. (1971) *J. Biol. Chem. 246*, 1365–1374.
Strickland, E. H., & Mercola, D. A. (1976) *Biochemistry 15*, 3875–3884.
Thim, L., Hansen, M. T., Norris, K., Hoegh, I., Boel, E., Forstrom, J., Ammerer, G., & Fil, N. P. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 6766–6770.
Weiss, M. A., Eliason, J. A., & States, D. J. (1984) *Proc. Natl. Acad. Sci. U.S.A. 82*, 6019–6023.
Weiss, M. A., Nguyen, D. T., Khait, I., Inouye, K., Frank, B. H., Beckage, M., O'Shea, E., Shoelson, S. E., Karplus, M., & Neuringer, L. J. (1989) *Biochemistry 28*, 9855–9873.
Williamson, K. L., & Williams, R. J. P. (1979) *Biochemistry 18*, 5966–5972.
Wodak, S. J., Alard, P., Delhaise, P., & Renngboog, S. C. (1984), *J. Mol. Biol. 151*, 317–322.
Wollmer, A., Fleischhauer, J., Strassburger, W., Thiele, H., Bradenbury, D., Dodson, G., & Mercola, D. (1979) *Biophys. J. 20*, 233–243.
Wood, S. P., Blundell, T. L., Wollmer, A., Lazarus, N. R., & Neville, R. W. J. (1975) *Eur. J. Biochem. 55*, 531–542.
Wuthrich, K., Wider, G., Wagner, G., & Braun, W. (1983) *J. Magn. Reson. 155*, 311.

Sequential ^{1}H NMR Assignments and Secondary Structure of Aponeocarzinostatin in Solution†

M. Lyndsay Remerowski,*,‡ Steffen J. Glaser,‡ Larry C. Sieker,§ T. S. Anantha Samy,∥ and Gary P. Drobny‡
Departments of Chemistry and of Biological Structure, University of Washington, Seattle, Washington 98195, and Department of Oncology, University of Miami, Miami, Florida 33101

Received March 26, 1990; Revised Manuscript Received June 13, 1990

ABSTRACT: Sequential assignments and secondary structural analysis have been accomplished for the 113-residue apoprotein of the antitumor drug neocarzinostatin (NCS) from *Streptomyces carzinostaticus*. A total of 98% of the main-chain and 77% of the side-chain resonances have been sequence specifically assigned by use of information from coherence transfer experiments and by sequential and interstrand NOEs. Because of the complexity of the NCS spectrum, several sequential assignment strategies were employed to complete the analysis. Apo-NCS consists of three antiparallel β-sheeted domains by NMR analysis. There is an extensive four-strand antiparallel β-sheet, and two two-stranded domains. One of the two-strand domains is contiguous, S72–N87, with chain reversal occurring through the region L77–R82. The other two-stranded domain has the section G16–A24 antiparallel with respect to the region S62–R70. This secondary structure is consistent with the crystal structure of holo-NCS at 2.8-Å resolution.

Neocarzinostatin (NCS)[1] is a protein antitumor drug first isolated from *Streptomyces carzinostaticus* culture medium 25 years ago (Ishida et al., 1965). NCS is perhaps the best studied of the *Streptomyces* anticancer agents. It has a relatively low toxicity (LD_{50} 10–30 mg/kg in mice and dogs) and broad in vivo activity against human leukemia, bladder cancer, liver metastasis, and sarcoma (Montgomery et al., 1981). Although NCS is 10–100 times less toxic than the *Streptomyces* antitumor antibiotics auromomycin (AUR) and actinoxanthin (AXN), the polypeptide sequence of NCS shows marked homologies with those of AUR and AXN suggesting that they are somehow related in structure and perhaps function (Goldberg et al., 1981).

A detailed molecular basis for the action of NCS is still lacking, but the last several years have seen an enormous expansion of information on the molecule. Holo-NCS has been separated into a polypeptide component (apo-NCS) of MW ca. 11 100 and a tightly bound chromophore (Chr, termed NCS-Chr) of composition $C_{35}H_{33}NO_{12}$ (MW 659). This non-protein chromophore has been found to consist of four

† This research was supported by NIH RO1 CA45643-01A1 and a NATO Fellowship from the Deutsche Akademische Austauschdienst (DAAD) to S.J.G.
‡ Department of Chemistry, University of Washington.
§ Department of Biological Structure, University of Washington.
∥ University of Miami.

[1] Abbreviations: NMR, nuclear magnetic resonance; NCS, neocarzinostatin; Chr, chromophore AUR, auromomycin; AXN, actinoxanthin; MCR, macromomycin; FABMS, fast atom bombardment mass spectroscopy; GCMS, gas chromatography–mass spectroscopy; MIR, multiple isomorphous replacement; 2D, two dimensional;EDTA, ethylenediaminetetraacetic acid; COSY, 2D correlated spectroscopy; DQF, double quantum filtered; TQF, triple quantum filtered; RELAY, 2D relayed coherence transfer spectroscopy; TOCSY, 2D total correlattion spectroscopy, NOE, nuclear Overhauser enhancement; NOESY, 2D NOE spectroscopy; TPPI, time-proportional phase incrementation; ppm, parts per million; d_{AB}, the NOE connectivity between protons A and B of the same (intraresidue) or different (interresidue) amino acids in a polypeptide (protons A and B are designated N for amide protons, α for $C^{\alpha}H$, β for $C^{\beta}H$, and δ for $C^{\delta}H$).

0006-2960/90/0429-8401$02.50/0 © 1990 American Chemical Society