A Protein Caught in a Kinetic Trap: Structures and Stabilities of Insulin Disulfide Isomers†

Qing-Xin Hua,‡ Wenhua Jia,‡ Bruce H. Frank,§,|| Nelson F. B. Phillips,‡ and Michael A. Weiss*,‡

Department of Biochemistry, Case Western Reserve University School of Medicine, 10900 Euclid Avenue, Cleveland, Ohio 44106, and Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, Indiana 46285

Received April 22, 2002; Revised Manuscript Received October 11, 2002

ABSTRACT: Proinsulin contains six cysteines whose specific pairing (A6−A11, A7−B7, and A20−B19) is a defining feature of the insulin fold. Pairing information is contained within A and B domains as demonstrated by studies of insulin chain recombination. Two insulin isomers containing non-native disulfide bridges ([A7−A11,A6−B7,A20−B19] and [A6−A7,A11−B7,A20−B19]), previously prepared by directed chemical synthesis, are metastable and biologically active. Remarkably, the same two isomers are preferentially formed from native insulin or proinsulin following disulfide reassortment in guanidine hydrochloride. The absence of other disulfide isomers suggests that the observed species exhibit greater relative stability and/or kinetic accessibility. The structure of the first isomer ([A7−A11,A6−B7,A20−B19], insulin-*swap*) has been described [Hua, Q. X., Gozani, S. N., Chance, R. E., Hoffmann, J. A., Frank, B. H., and Weiss, M. A. (1995) *Nat. Struct. Biol. 2*, 129−138]. Here, we demonstrate that the second isomer (insulin-*swap2*) is less ordered than the first. Nativelike elements of structure are retained in the B chain, whereas the A chain is largely disordered. Thermodynamic studies of guanidine denaturation demonstrate the instability of the isomers relative to native insulin ($\Delta\Delta G_u$ > 3 kcal/mol). In contrast, insulin-like growth factor I (IGF-I) and the corresponding isomer IGF-*swap*, formed as alternative products of a bifurcating folding pathway, exhibit similar cooperative unfolding transitions. The insulin isomers are similar in structure and stability to two-disulfide analogues whose partial folds provide models of oxidative folding intermediates. Each exhibits a nativelike B chain and less-ordered A chain. This general asymmetry is consistent with a hierarchical disulfide pathway in which nascent structure in the B chain provides a template for folding of the A chain. Structures of metastable disulfide isomers provide probes of the topography of an energy landscape.

The structure of a protein, its presumed thermodynamic ground state, is determined by its sequence (*1*). What are the rules relating sequence to structure? This question, being of broad relevance to biology and chemistry, has attracted long-standing interest (for reviews, see refs *2*−*4*). Insight into the informational content of protein sequences has been obtained from genetic analysis of allowed and disallowed sequences (*5*, *6*) and from development of knowledge-based algorithms for evaluating the fit between a given sequence and a known structural template (*7*−*9*). Complementary insight into structural features that distinguish the native state has been provided by construction of incorrectly folded models and their analysis using empirical force fields (*10*−*12*). This approach allows direct visualization of competing "threads" but is largely restricted to computer simulation. In this paper, we investigate metastable disulfide isomers of human insulin as model misfolded proteins (*13*). The results demonstrate that nativelike supersecondary structure is consistently maintained in the B chain but not in the A chain. Such asymmetry supports the hypothesis that in the hierarchical folding pathway of proinsulin (*14*) nascent structure in the B domain provides a template for folding of the A domain (*15*). Such partial folds are discussed in relation to energy landscapes (*16*) and possible mechanisms of insulin fibrillation (*17*).

Insulin is a globular protein containing two chains, designated A (21 residues) and B (30 residues). Its structure as a monomer in solution (*18*, *19*) resembles a crystallographic T-state protomer (Figure 1A) (*20*−*24*). The hormone is generated in vivo by proteolytic processing of a single-chain precursor, designated *proinsulin*, in which a connecting peptide joins the C-terminus of the B chain to the N-terminus of the A chain (*25*). The connecting peptide provides a topological tether but is without specific folding information (*26*). Air oxidation of isolated A and B chains preferentially yields covalent A−B heterodimers with native disulfide pairing (*27*), a reaction known as *insulin chain combination.* The absence of detectable disulfide isomers following chain combination demonstrates that proinsulin's folding information resides within isolated A and B chains (*28*). The efficiency of chain combination is limited by off-pathway events, such as the reversible formation of cyclic A chains and irreversible B chain aggregation. Although the

† This work was supported by a grant to M.A.W. from the National Institutes of Health (R01 DK054622).

* To whom correspondence should be addressed. Phone: (216) 368-5991. Fax: (216) 368-3419. E-mail: weiss@biochemistry.cwru.edu.

‡ Case Western Reserve University School of Medicine.

§ Eli Lilly and Co.

|| Since the completion of this work, B.H.F. has retired from Eli Lilly and Co.

10.1021/bi0202981 CCC: $22.00
Published on Web 11/20/2002

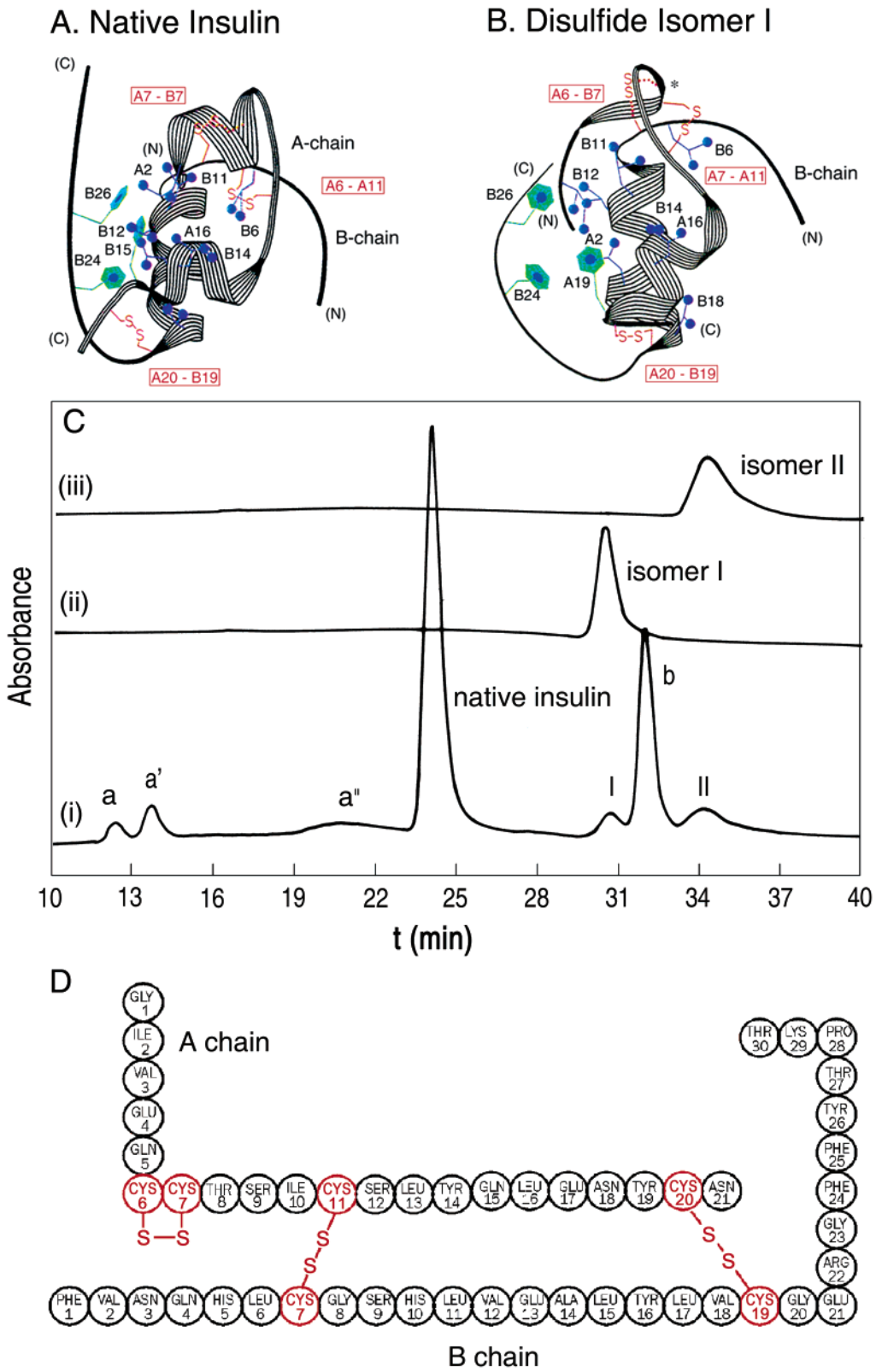

FIGURE 1: Structures of native insulin (A) and insulin-*swap* (B). Ribbon models are derived from NMR studies as described previously (*34*). In each panel, disulfide bridges (boxes) are shown in red. Selected side chains (blue) in the hydrophobic core are labeled. The structure of insulin-*swap* is remarkable for segmental unfolding of the A1−A8 α-helix and tilting of the A12−A18 helix in an otherwise nativelike globular domain. (C) Reverse-phase HPLC chromatograms showing (i) disulfide reassortment of native insulin, (ii) purified isomer I (insulin-*swap*), and (iii) purified isomer II (insulin-*swap2*). In chromatogram i, species a, a′, and a″ are isolated A chains containing internal disulfide bridges; species b is the isolated B7−B19 cyclic B chain. Isomers I and II are herein designated insulin-*swap* and insulin-*swap2*, respectively. Their differences in elution times and breadth of peaks relative to those of native insulin correlate with trends in CD-defined helix contents (Figure 3) and ^{1}H NMR chemical shift dispersion (Figure 5). (D) Primary structures and disulfide pairing schemes of insulin-*swap2*. Insulin-*swap2* contains an unusual A6−A7 vicinal disulfide bridge as well the A7−A11 pairing also found in insulin-*swap* and the native A20−B19 disulfide bridge.

mechanism of chain combination is not well characterized, studies of a single-chain proinsulin analogue suggest that disulfide pairing occurs via a preferred kinetic pathway (*14*). Chain combination and oxidative folding of single-chain precursor polypeptides have been widely employed to generate analogues, including novel insulins of therapeutic importance (*29*−*33*).

In a pioneering study, two disulfide isomers of human insulin were prepared by directed chemical synthesis (*13*). The isomers, which contain non-native pairings [A7−A11, A6−B7, and A20−B19 and A6−A7, A11−B7, and A20−B19 (herein designated insulin-*swap*[1] and insulin-*swap2*)], are metastable and function as insulin agonists.[2] The rate of disulfide rearrangement under acidic conditions was observed to be negligible. Remarkably, the same two isomers are preferentially formed from native insulin or proinsulin following disulfide reassortment in the presence of denaturants (Figure 1C; *34*). That other disulfide isomers are not detected suggests that the observed species exhibit greater relative stability and/or kinetic accessibility. One isomer (insulin-*swap*) has previously been shown to exhibit a partial fold (*34*). Although one of insulin's three canonical α-helices (comprising residues A1−A8) is not formed, nativelike secondary structure otherwise defines a globular domain and compact hydrophobic core (Figure 1B). Here, we extend these studies to the other isomer (insulin-*swap2*; disulfide pairing scheme shown in red in Figure 1D) and investigate its relative thermodynamic stability by guanidine denaturation. The second isomer is less organized than the first, precluding definition of a well-defined three-dimensional structure. Whereas the A chain is largely disordered, qualitative analysis of ^{1}H NMR nuclear Overhauser enhancements (NOEs) nonetheless demonstrates the maintenance of nativelike supersecondary structure in the B chain adjoining the cystine A20−B19 disulfide. Use of high-sensitivity cryogenic NMR probe technology demonstrates that these features are retained in dilute protein solutions and so characterize that misfolded monomer. The observed partial structure rationalizes the isomer's partial biological activity and immunological cross-reactivity with native insulin (*13*). Asymmetry between chains supports the hypothesis that the B domain of proinsulin provides a template for nascent folding of the A domain in a hierarchical pathway (*14*). Insulin-*swap* and insulin-*swap2* exemplify a polypeptide chain caught in a kinetic trap. Their structures probe the rugged topography of an energy landscape.

[1] Abbreviations: CD, circular dichroism; DPI, *des*-pentapeptide-[B26−B30]-insulin; DQF-COSY, double-quantum-filtered correlation spectroscopy; IGF-I, insulin-like growth factor I; insulin-*swap*, disulfide isomer with A6−B7, A7−A11, and A20−B19 pairings; insulin-*swap2*, disulfide isomer with A6−A7, A11−B7, and A20−B19 pairings; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser spectroscopy; rp-HPLC, reverse-phase high-performance liquid chromatography; TOCSY, total correlation spectroscopy; 2D, two-dimensional. Disulfide bonds are designated by residue number; e.g., A20−B19 indicates the disulfide bond between Cys^{A20} and Cys^{B19}. Amino acids are designated in the text with standard three-letter codes.

[2] Biological potencies of insulin isomers are qualitatively indistinguishable from those of native insulin, functioning in isolated adipocytes as full agonists in stimulation of glucose oxidation and inhibition of lipolysis (*13*). Potencies in cellular assays (14−28% relative to native insulin) are similar to percent activities in vivo in the rat hypoglycemia test and mouse convulsion assay (21−37%). The relative potencies of the two isomers are similar. Conversion of isomers to insulin in the presence of fat cells or in vivo has not been ruled out but seems unlikely given the similar results obtained under diverse assay conditions (*13*). No delay was observed in the onset of action as would be expected if slow isomerization to native insulin were required prior to assumption of an active structure.

MATERIALS AND METHODS

Protein Purification. Biosynthetic proteins [insulin, proinsulin, insulin-like growth factor I (IGF-I), and IGF-*swap*] were provided by Eli Lilly and Co. Isomers of insulin and proinsulin were generated from the native proteins (1 mM) by disulfide exchange in 10 mM dithiothreitol, 6 M guanidine hydrochloride, and 0.1 M Tris-HCl (pH 8.3). After 23 h in room air at 37 °C, the solution contained no protein SH groups (as assayed by reactivity to iodoacetate at pH 9.0), and the reactions were quenched by addition of 0.1 N HCl to a final pH of 3.1. Species with molecular masses of <12 kDa were desalted and separated from possible higher-order oligomers by size-exclusion chromatography (Sephadex G50 Superfine). Products were resolved by reverse-phase high-performance liquid chromatography (rp-HPLC; see below), desalted, lyophilized, and analyzed by electrospray ionization mass spectrometry and analytical HPLC in reference to native and non-native synthetic insulin standards (kindly provided by B. Riniker; Ciba-Geigy AG, Basel, Switzerland). The pairing schemes of proinsulin isomers were established by enzymatic cleavage of the connecting peptide with trypsin and carboxypeptidase B to yield the corresponding insulin isomers. Beginning from native insulin, ca. 70% of the starting material remained as native insulin; other products included insulin-*swap* (3−5% of the starting material), insulin-*swap2* (3−5%), and isolated A and B chains containing internal disulfide bonds (20−25%). Similar results were obtained with proinsulin except that isolated chains were not observed. Base-catalyzed reassortment of insulin or proinsulin in 6 M guanidine hydrochloride to the same two isomers was also observed over 1−4 weeks in the absence of a reducing agent in 1 M Tris-HCl (pH 9.5) at 4 °C. Quantitative rearrangement of either insulin isomer to native insulin was catalyzed by trace β-mercaptoethanol in 1 h at room temperature in 0.1 N ammonium hydroxide as described previously (*13*); the reaction is accompanied by formation of polymers. The purity of preparations used for NMR and circular dichroism (CD) studies was in each case >98% as evaluated by reverse-phase HPLC. Molecular masses of insulin and its disulfide isomers were indistinguishable by electrospray mass spectrometry; no contaminating molecular masses were observed.

HPLC Protocol. Products of disulfide reassortment were monitored by analytical rp-HPLC (Zorbax C8 150 Å column, 6 mm × 250 mm) at a column temperature of 40 °C. For gradient elution, we employed buffer A (20% acetonitrile in 200 mM ammonium sulfate and 50 mM sulfuric acid) and buffer B (40% acetonitrile in the same solution). The column was equilibrated and loaded in 70% buffer A and 30% buffer B. Proteins and peptides were eluted by a 30 min gradient from 30 to 70% B (Figure 1C). Products of disulfide reassortment were purified by semipreparative rp-HPLC (C18 column, 2.5 cm × 60 cm) by a similar protocol except that 0.5% trifluoroacetic acid was used instead of nonvolatile components (200 mM ammonium sulfate and 50 mM sulfuric acid). The respective fractions containing insulin or isomers were lyophilized.

Gel-Permeation Chromatography. A reduced level of self-association of insulin-*swap2* relative to that of native insulin at pH 8.1 and 32 °C was verified by HPLC gel-permeation chromatography. Whereas NMR studies were conducted in the absence of buffer, operation of the column requires a moderate ionic strength (5 mM Tris-HCl and 50 mM KCl), conditions that enhance the self-association of native insulin. Different concentrations of insulin or insulin-*swap2* were fractionated on a SynChropak GPCPEP column (250 mm × 4.6 mm, 5 μm as obtained from Eprogen, Darien, IL), using a Waters 1525 HPLC system equipped with a column heater. The column was maintained at 32 °C in a water bath and run at a flow rate of 0.5 mL/min prior to sample application. Samples (20 μL) were applied through an autosampler; peptide elution was monitored at 215 nm using a dual-lambda 2487 absorbance detector. Data acquisition and processing utilized the Waters Breeze HPLC software. Although gel-filtration studies are difficult to perform in 20% acetic acid, it is feasible to employ Sephadex G-50 (medium-grade, Pharmacia) in 5% acetic acid (HPLC-grade) in which insulin is partly monomer and dimer (at 200 μM insulin, the relative populations are 75 and 25%, respectively, as determined by NMR). Insulin and insulin-*swap2* (200 μM in 50 μL) were chromatographed separately on such a column (1 cm × 23 cm) at a flow rate of 0.5 mL/min at room temperature. Fractions (0.25 mL) were recovered in a fraction collector and eluted proteins detected by intrinsic tyrosine fluorescence using an Aviv ATF-105 spectrofluorimeter (Aviv Instruments, Lakewood, NJ). The excitation wavelength was 276 nm; emission spectra were obtained from 350 to 285 nm. Void and total column volumes were determined by fractionating blue dextran and xylene cyanol, respectively. The elution profile of the isomer is similar to that of native insulin, shifted somewhat to later elution times, consistent with a decreased level of dimerization.

NMR Spectroscopy. Initial NMR spectra were obtained at 500 and 600 MHz and 25 °C in 80% H_2O/20% deuterioacetic acid (pH 1.9) and 80% D_2O/20% deuterioacetic acid mixtures as described previously (*34*, *35*); the protein concentration was 1.5 mM. Resonance assignment (Supporting Information) was based on two-dimensional (2D) nuclear Overhauser spectroscopy (NOESY) (mixing times of 100 and 200 ms), total correlation spectroscopy (TOCSY) (mixing time of 55 ms), and double-quantum-filtered correlation spectroscopy (DQF-COSY) spectra. $^3J_{\alpha N}$ coupling constants in the folded moiety could not be inferred from DQF-COSY $H_\alpha-H_N$ cross-peaks due to conformational broadening (Supporting Information, Figure S8-A). Tables of chemical shifts are provided as Supporting Information. NMR assignments of nonexchangeable resonances were extended by analogy to spectra obtained at 600 MHz in D_2O at pD 7.7 (direct meter reading) at 32 °C in the absence of a cosolvent. This "bootstrap" procedure has previously been applied in studies of unstable insulin analogues containing pairwise substitution of cystine with alanine or serine (*15*, *36*). NMR analysis of amide resonances in an H_2O solution (pH 7.7−8.1; corresponding to pD 7.7) is not feasible due to rapid base-catalyzed amide proton exchange and conformational broadening (Supporting Information). Bootstrap extrapolation of assignments obtained in an organic cosolvent is justified by (a) insulin's native structure in 20% deuterioacetic acid (*18*, *37*), (b) the similarity of inter-residue nuclear Overhauser effects (NOEs) in insulin-*swap2* in spectra obtained in either 20% deuterioacetic acid or aqueous solution (pD 7.7), and (c) the consistency between NMR-defined helical segments and trends in CD-derived helix contents at neutral pH (see the Results). The solution structure of insulin-*swap* has previously been characterized in 20% deuterioacetic acid (*34*). The proteins are monomeric under conditions of study as inferred from gel-permeation chromatography (see above). To verify that the observed NMR features reflect the structural properties of the monomeric isomer, spectra were obtained at a protein concentration of 50 μM with an enhanced-sensitivity cryogenic probe at 600 MHz. Similar patterns of chemical shifts, resonance line widths, and NOEs were observed (Supporting Information).

Structural Analysis. Interpretation of NMR data focuses on the predominant conformation in solution. A small number of supernumerary NOE cross-peaks (<10 comprising at most two spin systems) are observed in 20% deuterioacetic acid. These are present in virgin samples and so do not represent degradation products. ROESY spectra indicate chemical exchange between major and minor states of Cys^{B7} (8.38 and 8.70 ppm, respectively); such exchange is not observed in native insulin. No other amide−amide ROESY cross-peaks are observed of the same sign as the diagonal (Supporting Information). The fractional occupancy of the minor state is <10%, leading to its exchange broadening relative to the major state (Supporting Information). We speculate that the minor state corresponds to an alternative configuration of the A7−B7 disulfide bridge.[3] We cannot exclude other mechanisms of structural heterogeneity as the TOCSY fingerprint spectrum contains several weak additional cross-peaks unassociated with detectable NOEs and unlinked to the major assignment pathway. Reverse-phase HPLC studies indicate that disulfide rearrangement does not occur during NMR data acquisition either in 20% acetic acid or at pD 7.7.

CD Spectroscopy. Spectra were obtained using an Aviv spectropolarimeter equipped with thermister-controlled temperature regulation. Samples were placed in a 1 mm path length quartz cuvette for wavelength scans and 1 cm quartz cuvette for guanidine denaturation studies. Spectra were obtained under acidic [0.01 N HCl (pH 2.0)] and neutral conditions [50 mM KCl and 10 mM potassium phosphate (pH 7.0)]; spectra are similar under these two conditions, in accord with past studies of insulin analogues (*38*). CD-detected guanidine unfolding curves at 222 nm (Figure 4A,B) were obtained in 50 mM KCl and 10 mM potassium phosphate (pH 7.0) at 4 °C with an automated titration unit as described previously (*15*). The protein concentration in guanidine studies was 5 μM. The reversibility of thermal transitions was evaluated by comparison of CD spectra and rp-HPLC chromatograms before and after heating (Supporting Information). Deconvolution of far-UV CD spectra to obtain estimates of fractional secondary structural composition (α-helix, β-sheet, turn, and random coil) was achieved by the SELCON method of Sreerama and Woody (*39*, *40*).

Thermodynamic Modeling. Guanidine denaturation data were fitted by a nonlinear least-squares method to a two-

[3] Among crystal structures of Zn−insulin hexamers, the cystine A7−B7 bridge exhibits a change in configuration associated with the T → R transition (*21*, *23*).

Table 1: Guanidine Denaturation[a]

analogue	ΔG_u	$\Delta\Delta G_u$	C_{mid} (M)	*m* (kcal mol^{-1} M^{-1})	potency (%)
(A) insulin analogues					
native insulin	4.4 ± 0.1	−	5.2 ± 0.1	0.84 ± 0.01	100
insulin-*swap*					
fit 1	1.3 ± 0.2	−3.1 ± 0.2	2.2 ± 0.2	0.58 ± 0.04	13−27
fit 2	1.4 ± 0.1	−3.0 ± 0.1	2.4 ± 0.2	0.57 ± 0.02	
fit 3	1.3 ± 0.1	−3.1 ± 0.2	2.6 ± 0.3	0.50 ± 0.05	
insulin-*swap2*					
fit 1	0.5 ± 0.2	−3.9 ± 0.3	1.2 ± 0.1	0.44 ± 0.03	13−27
fit 2	0.8 ± 0.1	−3.6 ± 0.1	2.0 ± 0.2	0.43 ± 0.02	
fit 3	0.9 ± 0.3	−3.5 ± 0.3	1.9 ± 0.2	0.46 ± 0.01	
(B) IGF analogues					
native IGF-1	2.5 ± 0.1	−	4.2 ± 0.1	0.58 ± 0.01	100
IGF-*swap*	3.3 ± 0.1	0.8 ± 0.2	5.1 ± 0.1	0.64 ± 0.01	10

[a] ΔG_u is the apparent change in free energy on denaturation in guanidine hydrochloride as extrapolated to zero denaturant concentration by a two-state model. $\Delta\Delta G_u$ is the difference in ΔG_u values relative to native insulin (A) or IGF-I (B). Uncertainties in two-state fitting parameters do not include possible systematic error due to non-two-state behavior. C_{mid} is defined as that concentration of guanidine hydrochloride at which 50% of the protein is unfolded. The *m* value provides the slope in plotting the unfolding free energy ΔG_u (G_u[HCl]) vs the molar concentration of the denaturant; this slope is proportional to the protein surface area exposed on unfolding. Insulin potencies (A) are expressed as the percent activity in isolated fat cell assays and *in vivo* rodent models (*13*; see footnote 2); IGF-I potencies (B) are defined as the percent affinity for the type I IGF receptor defined relative to the native ligand (100%; *49*). Fits 1−3 in the analysis of insulin isomers designate (1) simultaneous fitting of slopes of pre- and post-transition baselines, (2) an assumed horizontal pretransition baseline with simultaneous fitting of slope of the post-transition baseline only, and (3) assumed horizontal pre- and post-transition baselines without fitting of either slope, respectively (see Materials and Methods). For sigmoidal transitions (native insulin, IGF-I, and IGF-*swap*), the model employs simultaneous fitting baselines as in fit 1.

state model as described previously (*41*). Justification for the empirical use of a two-state model in the analysis of a partial fold has been given by Baldwin and co-workers[4] (*42*). In brief, CD data $\theta(x)$, where x indicates the concentration of the denaturant, were fitted by a nonlinear least-squares program according to

$$\theta(x) = \frac{\theta_A + \theta_B e^{-\Delta G_{H2O}^\circ - mx/(RT)}}{1 + e^{(-\Delta G_{H2O}^\circ - mx)/(RT)}} \quad (1)$$

where x is the concentration of guanidine and θ_A and θ_B are baseline values in the native and unfolded states, respectively. In "fit 1" (Table 1), these baselines were approximated by pre- and post-transition lines $\theta_A(x) = \theta_A^{H_2O} + m_A x$ and $\theta_B(x) = \theta_B^{H_2O} + m_B x$, respectively. Fitting CD data and baselines simultaneously circumvents artifacts associated with linear plots of ΔG as a function of denaturant concentration according to $\Delta G^\circ(x) = \Delta G_{H_2O}^\circ + m^\circ x$ (for a review, see ref *41*). Because the nonsigmoidal shapes of the isomers' unfolding transitions make definition of pre- and post-transition baselines uncertain, modeling was also performed assuming (a) that $\theta_A(x) = \theta_A^{H_2O}$ and fitting only the post-transition baseline (fit 2 in Table 1) and (b) that $\theta_A(x) = \theta_A^{H_2O}$ and $\theta_B(x) = \theta_B^{H_2O}$ without baseline fitting (fit 3); this protocol has been described previously (*15*). The *m* values obtained in fitting the isomer's unfolding curve are in each case significantly lower than the native *m* value (see Table 1). Such lower *m* values are likely to reflect a more exposed hydrophobic surface in the absence of denaturant and/or the existence of a native-state ensemble containing a distribution of incompletely folded forms with differing stabilities[5] (*43*).

RESULTS

Insulin-*swap* ([A7−A11,A6−B7,A20−B19]) and insulin-*swap2* ([A6−A7,A11−B7,A20−B19]) are generated at low yield from native insulin by disulfide reassortment under denaturing conditions (see Materials and Methods). A variety of conditions (pH, temperature, ratio of protein to reducing agent, and incubation time prior to acid quenching) were tested to optimize yield. HPLC chromatograms documenting the reassortment and purification of the two isomers are shown in Figure 1C. Proinsulin exhibits analogous disulfide reassortment (not shown). Prolonging the incubation time leads to a lower yield of insulin-*swap* and -*swap2* isomers as a more complex distribution of products (including polymers) evolves. The present reassortment reaction is thus not at equilibrium; the observed isomers represent kinetic traps accessible under denaturing conditions.[6] Polymeric products have not been characterized.

CD Studies Demonstrate an Unstable Partial Fold. CD spectra demonstrate that the isomers are partially folded but distinct in structure from native insulin. Spectra of the

[4] Rigorous application of the Baldwin−Luo formalism (*42*) to a molten globule or partial fold requires verification of concordance between probes of secondary and tertiary structure. The latter is typically provided by intrinsic tryptophan or tyrosine fluorescence. As insulin lacks tryptophan and its tyrosine fluorescence is uninformative, its present applicability is presumptive.

[5] Although partial folds would be expected to exhibit lower *m* values as observed here, a change in slope can also be associated with an underestimate of an analogue's stability as extrapolated to zero denaturant concentration. In the Baldwin−Luo formalism (*42*), an upper bound to the stability of the analogue is obtained by multiplying its $\langle C_{mid} \rangle$ value (an apparent value reflecting an average over an ensemble of distinct partial folds) by the native *m* value.

[6] Isomerization of insulin in 6 M guanidine hydrochloride occurs under conditions in which the three species are substantially unfolded (as indicated by CD spectra) and presumably exhibit similar residual stabilities. The low yield of isomers reflects kinetic control of the initial reassortment; an equilibrium has not been reached. If the initial reassortment were under thermodynamic control (i.e., if an equilibrium had been reached), then the relative populations of insulin (or proinsulin) to isomers I and II (Figure 1C) would imply a $\Delta\Delta G_u$ of ≈2 kcal/mol between native and alternative pairing schemes. Such a marked difference between the stabilities of respective unfolded-state ensembles in 6 M guanidine hydrochloride seems to be unphysical.

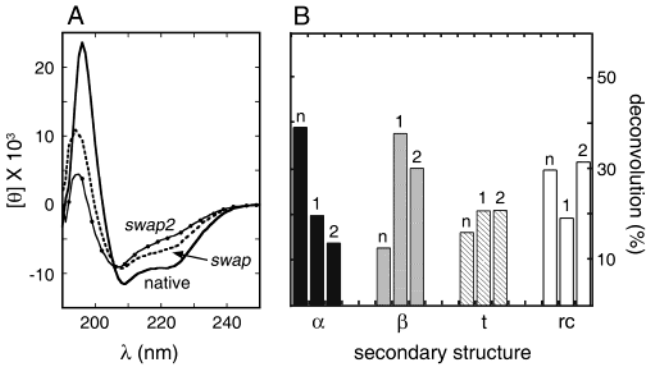

FIGURE 2: CD spectra of insulin and disulfide isomers (A) and spectral deconvolution (B). Far-UV spectra in panel A demonstrate attenuation of helix-specific bands at 222 and 195 nm in the following order: native insulin (—, highest helix content), insulin-*swap* (- - -), and insulin-*swap2* (●, lowest helix content). Deconvolution in panel B yields a histogram of secondary structures: (α, black bars) helix, (*β*, gray bars) *β*-sheet, (t, cross-hatched bars) turn, and (rc, white bars) random coil. Bars pertaining to insulin, insulin-*swap*, and insulin-*swap2* are labeled n, 1, and 2, respectively. Analysis of protein secondary structure was performed by the method of Woody and co-workers (*39*). Spectra were obtained in aqueous solution at pH 7.7 amd 32 °C at protein concentrations of 40 *μ*M.

isomers are remarkable for attenuation of helix-specific bands at 195, 208, and 222 nm (Figure 2A). Similar results are obtained at pH 7.7 and 32 °C (Figure 2A; chosen to correspond to the NMR conditions described below) and in 50 mM KCl and 10 mM potassium phosphate (pH 7.0) at 4 °C [Figure 3A; chosen to correspond to past studies of insulin analogues (*15*, *36*, *44*)]. Inferred α-helix contents are in each case less than that of native insulin; on the basis of SELCON deconvolution (*39*, *40*), estimated helix contents are 39 (native), 20 (insulin-*swap*), and 14% (insulin-*swap2*) (Figure 2B). The reduction in helix content is accompanied by an increase in predicted *β* and turn content (histogram in Figure 2B). The inferred helix content of native insulin is less than that expected on the basis of crystal structures (26 helical residues of 51 total residues or 52%). Despite the imprecision of the deconvolution method, the trend in inferred helix contents (native insulin > insulin-*swap* > insulin-*swap2*) is in accord with qualitative features of the spectra. These features do not resolve possible sites of perturbation in the protein nor distinguish between local unfolding versus global destabilization. CD spectra of proinsulin isomers are similar to those of the corresponding insulin isomers with addition of random-coil features attributed to the connecting peptide in accord with past studies of proinsulin (*45*−*47*).

CD spectra of insulin-*swap* and -*swap2* are shown at successive temperatures (4−70 °C) in panels B and C of Figure 3, respectively. Temperature-dependent features are reversible and unaccompanied by disulfide scrambling (Supporting Information). Whereas insulin-*swap* [like native insulin (*15*)] contains a subset of temperature-dependent structure, spectra of insulin-*swap2* are similar throughout this temperature range. The spectra of insulin, insulin-*swap*, and insulin-*swap2* exhibit similar nonrandom features at 70 °C (Figure 3E). This convergence is also seen in thermal melting curves at 222 nm (Figure 3D) in which the extent of thermal lability is greatest for native insulin (curve a) and least for insulin-*swap2* (curve c). Thermal unfolding of insulin would be non-two-state if melting of a labile substructure should precede that of a stable nucleus whose structure is analogous to that of insulin-*swap2* (see the Discussion).

Thermodynamic stabilities were estimated by CD-monitored guanidine denaturation (Figure 4A). Whereas native insulin exhibits a cooperative unfolding transition at 222 nm with a well-defined pretransition baseline (labeled *a* in Figure 4A), the isomers are each sensitive to low concentrations of guanidine hydrochloride (curves *b* and *c*). Visual inspection suggests that the isomers are each markedly less stable than native insulin and that the unfolding transition of insulin-*swap* is somewhat more sigmoidal in character than that of insulin-*swap2*. In neither case are pretransition baselines well defined. These broad transitions are similar to those observed in studies of two-disulfide insulin analogues (*15*, *36*). Despite their nonclassical shapes, the unfolding curves are consistent with a two-state model (eq 1 in Materials and Methods). This formalism (*41*), implemented with a variety of fitting procedures (see Materials and Methods; Table 1 and its footnote), would indicate that the isomers are less stable than native insulin ($\Delta\Delta G_u$ > 3 kcal/mol), in accord with their kinetic conversion to native insulin under appropriate conditions (*13*). Such fitting further suggests that insulin-*swap2* is less stable than insulin-*swap*, in qualitative accord with relative CD-defined helix contents. Quantitative interpretation of guanidine denaturation transitions is nonetheless limited by the assumptions of two-state modeling (*43*) and formal inequivalence of respective denatured states.[7]

Respective disulfide pairing schemes of insulin and insulin-*swap* are the same as those of the homologues IGF-I and IGF-*swap* (Figure 4C; *48*, *49*). Unlike insulin and proinsulin, however, the two forms of IGF-I are nearly equally populated at equilibrium, demonstrating the existence of alternative ground states (*49*). The structure of IGF-*swap* closely resembles that of insulin-*swap* (*50*). A corresponding "IGF-*swap2*" has not been observed. As a control for the guanidine denaturation method, we have characterized the respective unfolding transitions of native IGF-I and IGF-*swap* (designated n and swap in Figure 4B). Each exhibits a classical sigmoidal transition with well-defined pretransition baseline. A marked contrast is thus observed between the nativelike transitions of native insulin, native IGF-I, and IGF-*swap*, on one hand, and the broad and unstable transitions of insulin-*swap* and insulin-*swap2*, on the other hand. The stability of IGF-I is lower than that of insulin ($\Delta\Delta G_u = 1.9 \pm 0.2$ kcal/mol), in accord with observations that insulin is able to maintain its disulfide bridges at the oxidation−reduction potential of the blood stream whereas the isolated IGF-I monomer cannot (*51*).

NMR Studies Demonstrate Asymmetric Perturbations. To probe the structural origins of the isomer's altered CD spectrum and instability, insulin-*swap2* was investigated by ^{1}H NMR spectroscopy. Initial studies were conducted in 20% deuterioacetic acid at 25 °C in relation to native insulin (*35*) and insulin-*swap* (*34*). Spectra are unaffected by dilution of

[7] The guanidine denaturation method, unlike thiol-catalyzed disulfide rearrangement (*51*), employs distinct unfolded reference states (i.e., oxidized disulfide isomers in a saturated concentration of guanidine hydrochloride). Estimates of $\Delta\Delta G_u$ assume that the various unfolded states exhibit similar baseline stabilities. It is possible, however, that different unfolded-state ensembles have different (low) stabilities due to altered global topology (*69*) or residual local interactions (*70*) enforced by distinct disulfide connectivities. See also footnote 4.

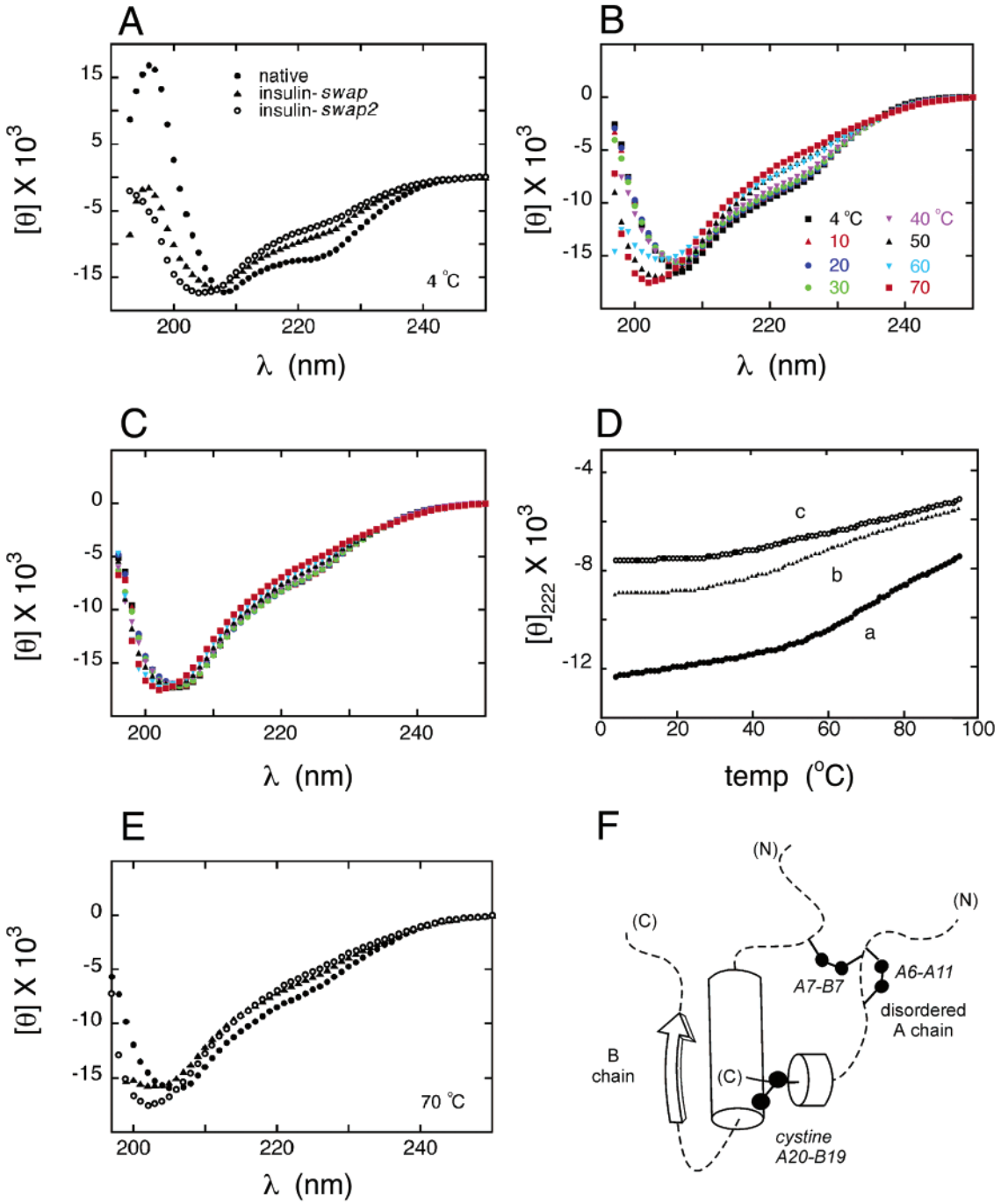

FIGURE 3: CD studies of insulin isomers. (A) Comparison of far-UV CD spectra of native insulin (●), insulin-*swap* (▲), and insulin-*swap2* (○) at 4 °C in 10 mM potassium phosphate (pH 7.4) and 50 mM KCl. (B and C) CD spectra of insulin-*swap* (B) and insulin-*swap2* (C) as a function of temperature. Both panels utilize the same temperature designations: 4 (black squares), 10 (red triangles), 20 (blue circles), 30 (light green circles), 40 (purple triangles), 50 (black triangles), 60 (light blue triangles), and 70 °C (red squares). (D) The mean residue ellipticity at 222 nm is shown as a function of temperature for native insulin (a), insulin-*swap* (b), and insulin-*swap2* (c). (E) Far-UV CD spectra of native insulin (●), insulin-*swap* (▲), and insulin-*swap2* (○) at 70 °C in 10 mM potassium phosphate (pH 7.4) and 50 mM KCl. (F) Proposed model of the partial insulin fold at 70 °C prior to the onset of fibrillation. Cylinders and the arrow indicate temperature-stable substructure (B9−B26 and A16−A20). Dashed lines indicate disordered regions. The protein's three disulfide bridges are indicated in schematic form by balls (sulfur atoms).

the protein to 50 μM as observed with enhanced-sensitivity cryogenic probe technology (Supporting Information). One-dimensional ^{1}H NMR spectra of native insulin, insulin-*swap*, and insulin-*swap2* exhibit the progressive loss of chemical shift dispersion (Figure 5). This trend is in qualitative accord with the respective stabilities and CD-defined α-helix contents in aqueous solution at neutral pH. Essentially complete sequential assignment of insulin-*swap2* was obtained by standard methods (Supporting Information). Motional narrowing is observed involving resonances in the N-terminal segments of both chains (A3−A5 and B2−B6) and C-terminal B chain segment (B27−B30); such narrowing is prominent in DQF-COSY fingerprint spectra due to antiphase cancellation of non-narrowed cross-peaks[8] (*52*) (Supporting Information). Two amide resonances (A2 and A10) are unobserved, presumably due to conformational

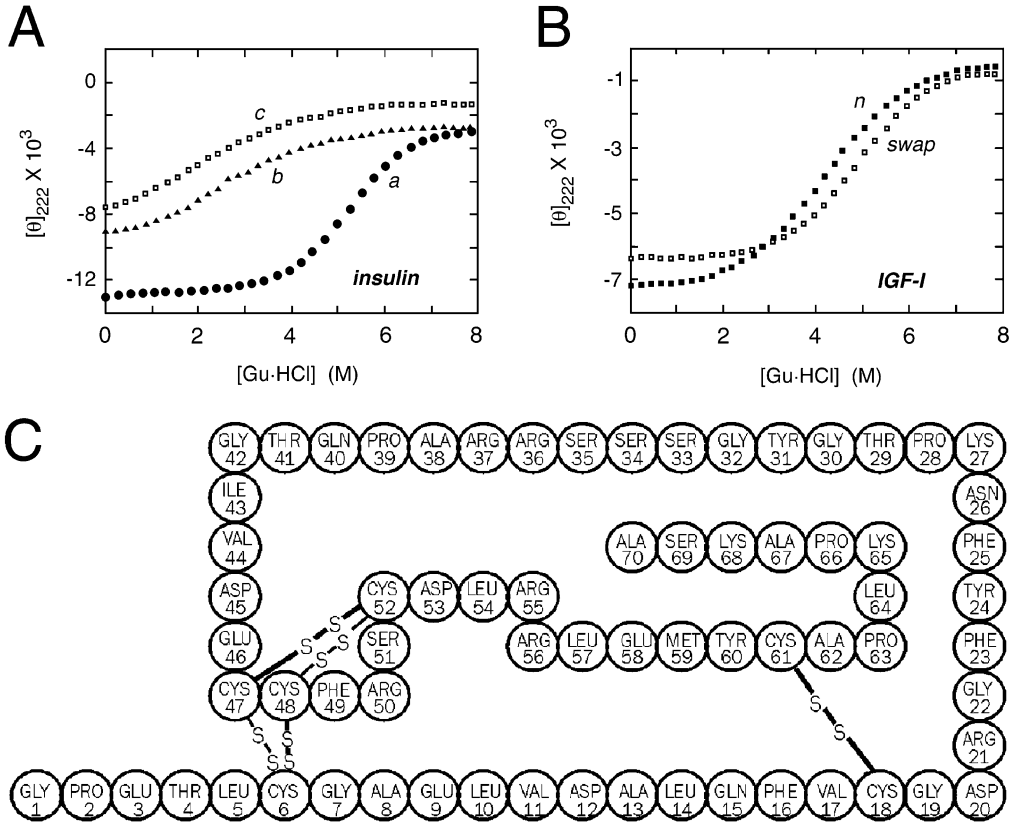

FIGURE 4: Contrasting thermodynamic features of insulin and IGF-I: guanidine unfolding transitions of insulin (A) and IGF-I (B) and disulfide pairing scheme of IGF-I (C). (A) Studies of native insulin (a), insulin-*swap* (b), and insulin-*swap2* (c). Guanidine unfolding transitions were monitored by CD ellipticity at 222 nm in panels A and B. (B) Control unfolding studies of native IGF-I and IGF-*swap*. Thermodynamic parameters inferred from two-state fitting (*41*) are given in Table 1. (C) Disulfide pairing schemes of IGF-I. Pairings in native IGF-I are designated with thick lines connecting sulfur atoms (residues 6−48, 18−61, and 47−52), whereas pairings in IGF-*swap* (6−47, 18−61, and 48−52) are denoted with thin lines.

broadening. Nonexchangeable side chain resonances are well-defined, with the exception of Gly^{B20} (whose H_α resonances are presumably overlapping near the diagonal of the 2D spectrum).

The upfield region of the aliphatic ^1H NMR spectrum of insulin-*swap2* is shown in Figure 6A-d in relation to the spectrum of an engineered monomer (DKP-insulin; *53*) in the same cosolvent (Figure 6A-a) and the native insulin dimer in dilute HCl at pH 2 (Figure 6A-b).[9] Whereas spectra of the native insulin monomer and dimer exhibit upfield-shifted methyl resonances from both A and B chains (Ile^{A2}, Ile^{A10}, and Leu^{B15}), the upfield spectrum of insulin-*swap2* contains only the methyl resonances of Leu^{B15}. The magnitude of B15 secondary shifts is attenuated relative to such shifts in the native dimer or monomer. Previous analyses of aromatic ring current shifts in insulin have ascribed the upfield shift of Leu^{B15} to Phe^{B24} (*54*). The near-random-coil chemical shifts

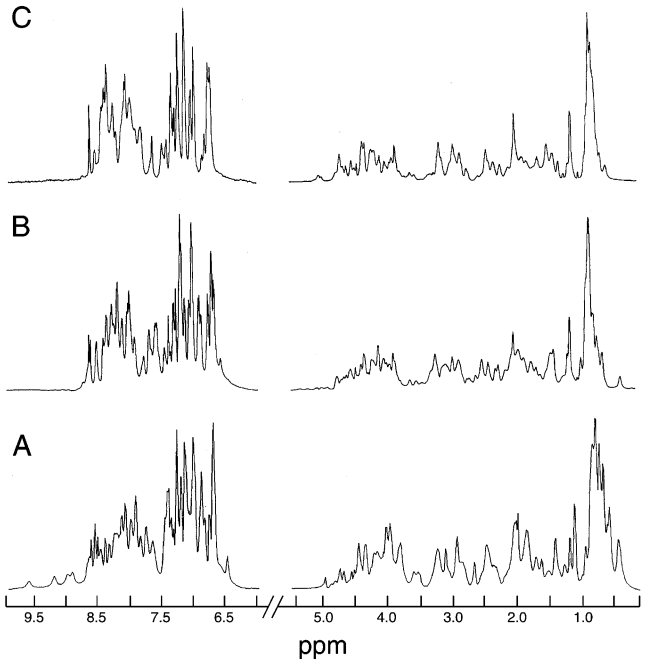

FIGURE 5: ^1H NMR spectra of native insulin (A), insulin-*swap* (B), and insulin-*swap2* (C) in 20% deuterioacetic acid (pH 1.9) at 25 °C. Chemical shift dispersion is greatest in the spectrum of native insulin and smallest in the spectrum of insulin-*swap2*. The protein concentration was in each case 1 mM. The spectrometer frequency was 500 MHz. Cryoprobe-enhanced ^1H NMR spectra at a protein concentration of 50 μM are provided as Supporting Information.

[8] Antiphase cancellation of DQF-COSY fingerprint cross-peaks (*52*) is accentuated in α-helical regions due to smaller values of $^3J_{HN\alpha}$ coupling constants. This effect is insufficient in itself, however, to account for the observed attenuation of such signals as indicated by past NMR studies of DPI (*71*) and other small helical proteins wherein such cross-peaks are well-defined. The extent of antiphase cancellation is not affected by dilution of the protein to 50 μM (Supporting Information).

[9] The crystal structure of native insulin at pH 2.0 has recently been determined by G. G. Dodson and co-workers to be a dimer (*72*). Its structure closely resembles the native dimer observed in the classical T_6 Zn hexamer (*20, 22*).

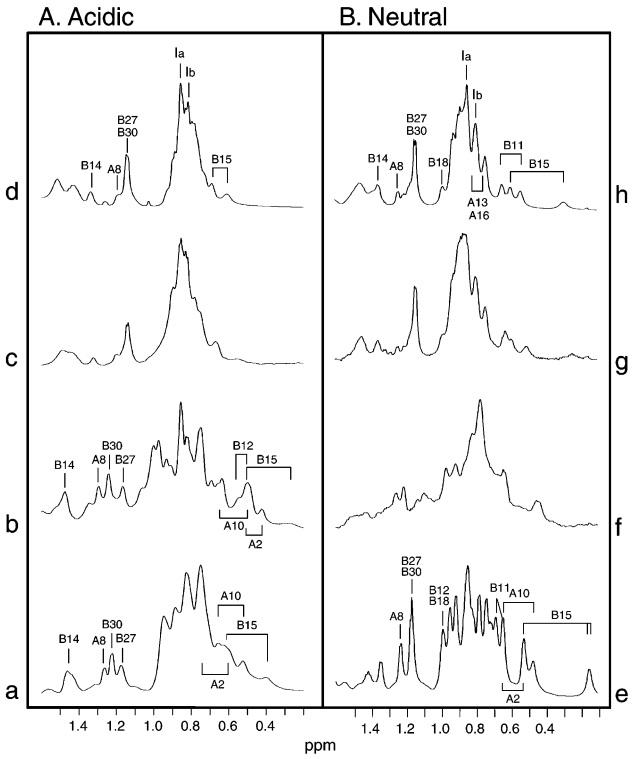

FIGURE 6: Reduced chemical shift dispersion of the insulin isomer. One-dimensional ^{1}H NMR aliphatic spectra of native insulin and insulin-*swap2* under acidic conditions (A) or at neutral pH (B): (a) 1 mM DKP-insulin monomer in 20% deuterioacetic acid at 25 °C, (b) native insulin dimer at pH 2 in 0.01 M HCl at 25 °C, (c) 0.2 mM insulin-*swap2* under the conditions of spectrum *b*, (d) 0.8 mM insulin-*swap2* under the conditions of spectrum *a*, (e) 1 mM DKP-insulin monomer in aqueous solution at pD 7.7 and 32 °C, (f) 0.2 mM native insulin aggregate at pD 7.0 and 25 °C, (g) 0.2 mM insulin-*swap2* at pD 7.0 and 25 °C, and (h) 0.5 mM insulin-*swap*2 at pD 7.7 and 32 °C. Assignments of spectra *a*, *d*, and *e* are based on complete sequential assignment; assignments of spectra *b* and *h* are based on analogous NOE patterns of spin systems in D_2O (see Figure 8). The two isoleucine spin systems of insulin-*swap2* are distinguishable but not assigned; their two δ-CH_3 resonances are labeled Ia and Ib in panels *d* and *h*.

of Ile^{A2} and Ile^{A10} (not individually assigned; δ-CH_3 resonances are labeled Ia and Ib in Figure 6A-d) suggest that neighboring ring currents (Ile^{A2}−Tyr^{A19} and Ile^{A10}−His^{B5}) have been perturbed. The ^{1}H NMR spectrum of isulin-*swap2* in 20% deuterioacetic acid (pH 1.9) is similar to (but somewhat less broad than) that obtained in dilute HCl solution (pH 2.0; Figure 6A-c). This similarity suggests that addition of the cosolvent prevents self-association of the isomer without introducing significant additional perturbations.

High-resolution spectra of insulin-*swap* may also be obtained in D_2O at 32 °C (Figure 6B-h). Fortuitously, whereas native insulin self-associates at neutral pH to yield broad NMR resonances (Figure 6B-f), insulin's classical self-association surfaces (*22*) are presumably perturbed in the isomer, leading to NMR line widths similar to those of an engineered insulin monomer under these conditions (Figure 6B-e). Similar spectra are observed at isomer concentrations of 1 mM and 50 μM. A decreased level of self-association under these conditions is independently verified by HPLC gel-permeation chromatography (Supporting Information). Assignments are extended by analogy as 2D NMR spectra obtained under the two solution conditions exhibit corresponding patterns of spin systems (Figures 7B and 8). Trends in chemical shifts are similar to those observed under acidic conditions. For both DKP-insulin (spectra *a* and *e* in Figure 6) and insulin-*swap2* (spectra *d* and *h*), dispersion of chemical shifts is more pronounced near neutral pH than in 20% deuterioacetic acid, especially in the magnitude of the upfield shift of Leu^{B15}. Nonetheless, spectral comparison at pD 7.7 (spectra *e* and *h* in Figure 6B) demonstrates an overall attenuation of chemical shift dispersion in the isomer similar to that observed in 20% deuterioacetic acid. The aliphatic portion of the ^{1}H NMR spectrum of insulin-*swap2* is in each case remarkable for retention of upfield-shifted B chain resonances but loss of upfield-shifted A chain resonances. The spectrum of insulin-*swap2* under CD conditions [50 mM KCl and 10 mM potassium phosphate (pH 7.0) at 25 °C] exhibits line broadening indicative of partial aggregation at this protein concentration (Figure 6B-g). The envelope of resonances nonetheless resembles that observed at pD 7.7 and 32 °C (Figure 6B-h).

Analysis of inter-residue NOEs in 80% H_2O and 20% deuterioacetic acid permits the observation of helical-associated contacts d_{NN} (contacts between successive amide protons); limited resolution in the fingerprint region precludes analysis of the corresponding $d_{\alpha N(i,i+3)}$ cross-peaks (contacts between H_α of residue *i* and the amide proton of residue $i+3$). The pattern of corroborative helix-related $d_{\alpha\beta(i,i+3)}$ contacts is similar in 20% deuterioacetic acid and in aqueous solution at pD 7.7. These contacts are summarized in Table 2 and outlined in Wüthrich format in the Supporting Information. Long-range inter-residue NOEs (Figures 7B and 8D) are also observed. These involve a subset of nativelike contacts and are similar under the two solution conditions (Table 3). Use of cryogenic probe technology demonstrates that these NOEs are retained at a protein concentration of 50 μM (Supporting Information) and hence arise in the monomer. Although insufficient NOEs are observed (or resolved) to permit calculation of a precise three-dimensional model by distance geometry, the following qualitative features are discerned.

(i) Maintenance of a Central B Chain α*-Helix (B9*−*B19).* Analysis of secondary structure reveals a nested pattern of $d_{\alpha\beta(i,i+3)}$ NOEs under both solution conditions spanning the B9−B18 segment. Similarly, d_{NN} NOEs are observed in 20% deuterioacetic acid involving the B9−B10, B10−B11, B14−B15, B17−B18, and B18−B19 steps (Supporting Information). The B11−B14 and B15−B16 steps are unclear because of either limited resolution near the diagonal or conformational broadening. Additional d_{NN} NOEs are observed at the B4−B5 (weak), B5−B6, and B6−B7 steps (i.e., the segment N-terminal to the non-native B7−A11 disulfide bridge). No $d_{\alpha\beta(i,i+3)}$ connectivities are observed in this region. It is possible that one or more turns occur between B4 and B9.

(ii) Maintenance of B Chain Supersecondary Structure. In native insulin, the B chain central α-helix is followed by a β-turn (residues B20−B23) and β-strand (B24−B28), which packs back across the α-helix (B9−B19) (Figure 7A). In insulin-*swap2*, a d_{NN} contact consistent with a turn is observed between residues B22 and B23; the expected B21 H_α−B23 H_N NOE may exist but is obscured by resonance overlap. Evidence for nativelike strand and helix packing is provided by side chain NOEs from the aromatic rings of Phe^{B24} and Tyr^{B26} to methyl resonances in the α-helix (Val^{B12} and Leu^{B15}; Figures 7B and 8D). Subtle differences in side

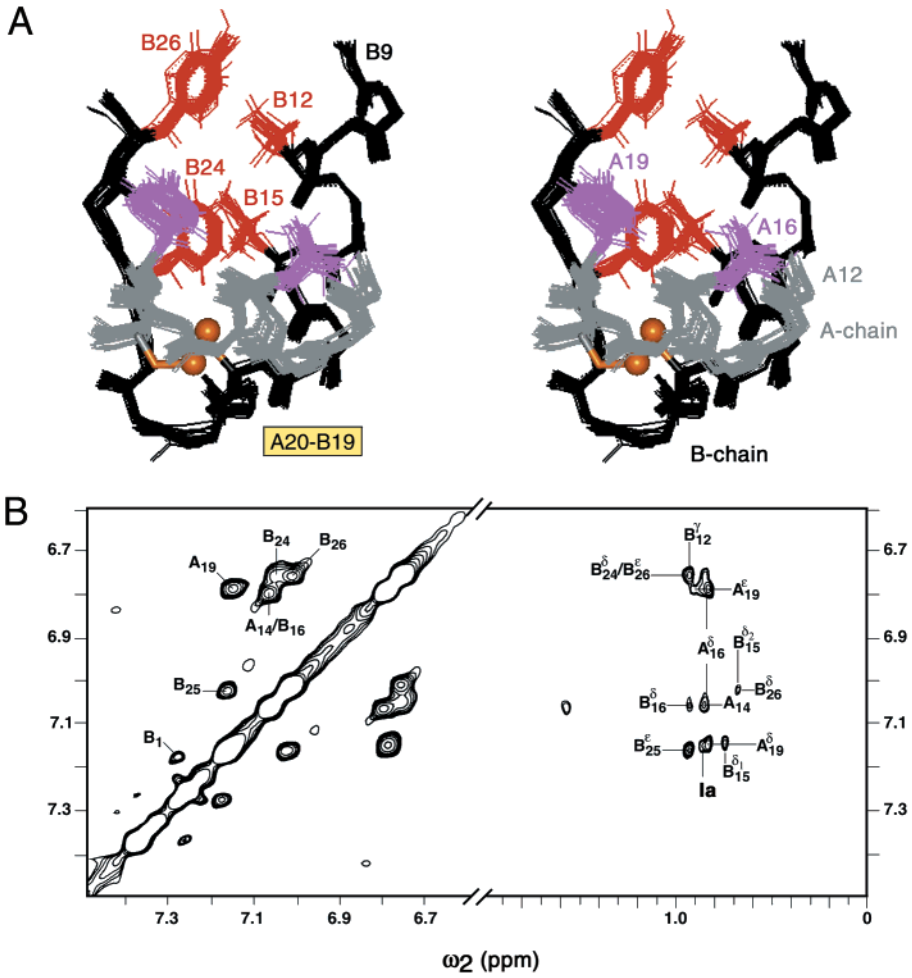

FIGURE 7: Long-range interactions of B chain supersecondary structure. (A) Superposition of crystal structures (residues A12−A20 and B9−B26) showing key side chain contacts among Leu^{A16}, Tyr^{A19}, Val^{B12}, Leu^{B15}, Phe^{B24}, and Tyr^{B26}. Main chain atoms of the A and B chains are shown in gray and black, respectively; side chains of the A and B chains are shown in lilac and red, respectively. The A20−B19 disulfide bridge is shown as yellow-orange balls (sulfur atoms) and sticks. The spread of the conformations reflects structural variation among crystal forms. Coordinates were obtained from the Protein Data Bank (entries 4INS, 1APH, 1BPH, 1CPH, 1DPH, 1TRZ, 1TYL, 1TYM, 1ZNI, 1LPH, 1G7A, 1EV6, and 1ZNJ). (B) NOESY spectrum of insulin-*swap2* in an 80% D_2O/20% deuterioacetic acid mixture showing long-range contacts between aromatic and aliphatic side chains. Of particular interest are contacts characteristic of nativelike B chain supersecondary structure (from Val^{B12} to Phe^{B24} and/or Tyr^{B26} and from Leu^{B15} to Tyr^{B26}) and nativelike A16−A19 and B15−A19 contacts. Aromatic resonances of Phe^{B24} H_γ and Tyr^{B26} H_ϵ overlap at 6.75 ppm. The assignment of a possible Ile^{A2}−Tyr^{A19} interaction (Ia) is provisional due to resonance overlap.

chain packing are observed between the two solution conditions (Table 3), but the observed NOEs are in each case consistent with maintenance of nativelike strand and helix packing. No nativelike contacts are observed under either solution condition between His^{B5} and Ile^{A10}, indicating that the structure of the mobile N-terminal portion of the B chain differs from that of native insulin. These NOEs are also absent in the spectrum of insulin-*swap*.

(iii) A Nascent A Chain C-Terminal α*-Helix*. Strings of $d_{\alpha\beta(i,i+3)}$ NOEs are absent in the N-terminal segment of the A chain. Assignments are largely defined and in principle would allow resolution of the majority of such helix-related contacts. Evidence suggesting a nascent C-terminal helix in the A chain is provided by an incomplete set of helix-related NOEs; these include A12−A15 $d_{\alpha\beta(i,i+3)}$, A12−A16 $d_{\alpha\gamma(i,i+4)}$ (not seen in native insulin), A15−A18 $d_{\alpha N(i,i+3)}$, a side chain (i, i + 3) contact between the aromatic ring of Tyr^{A19} and methyl resonances of Leu^{A16} (Figure 7B), and a string of d_{NN} connectivities between A13 and A19 (Supporting Information, Figure S8-B). An additional A15−A18 $d_{\alpha\beta(i,i+3)}$ is observed in D_2O (pD 7.7) not present in 20% deuterioacetic acid (Table 2). An A13−A16 $d_{\alpha N(i,i+3)}$ NOE may exist but is obscured by overlap under either solution condition. It is possible that this segment's helical preference is associated with side chain disorder in a molten ensemble of conformations. These features are shared by insulin-*swap* and insulin analogues lacking one disulfide bridge (either the cystine A6−A11 or A7−B7 bridge; *15*, *18*, *36*). A substantial loss of α-helical structure in the A chain is in accord with the attenuated CD spectra of these analogues. Long-range NOEs between side chains in the A chain, if present, are difficult to resolve.[10]

DISCUSSION

Insulin's three disulfide bridges (A6−A11, A7−B7, and A20−B19) play critical roles in its structure, stability, and

[10] An NOE is observed between the aromatic ring of Tyr^{A19} and a methyl resonance near 0.9 ppm. This contact is consistent with a nativelike contact involving Ile^{A2} (Figures 7B and 8D), but resonance overlap precludes a definitive assignment. Conformational broadening of A2 resonances (unlike those of A3−A5) suggests the presence of alternative interactions.

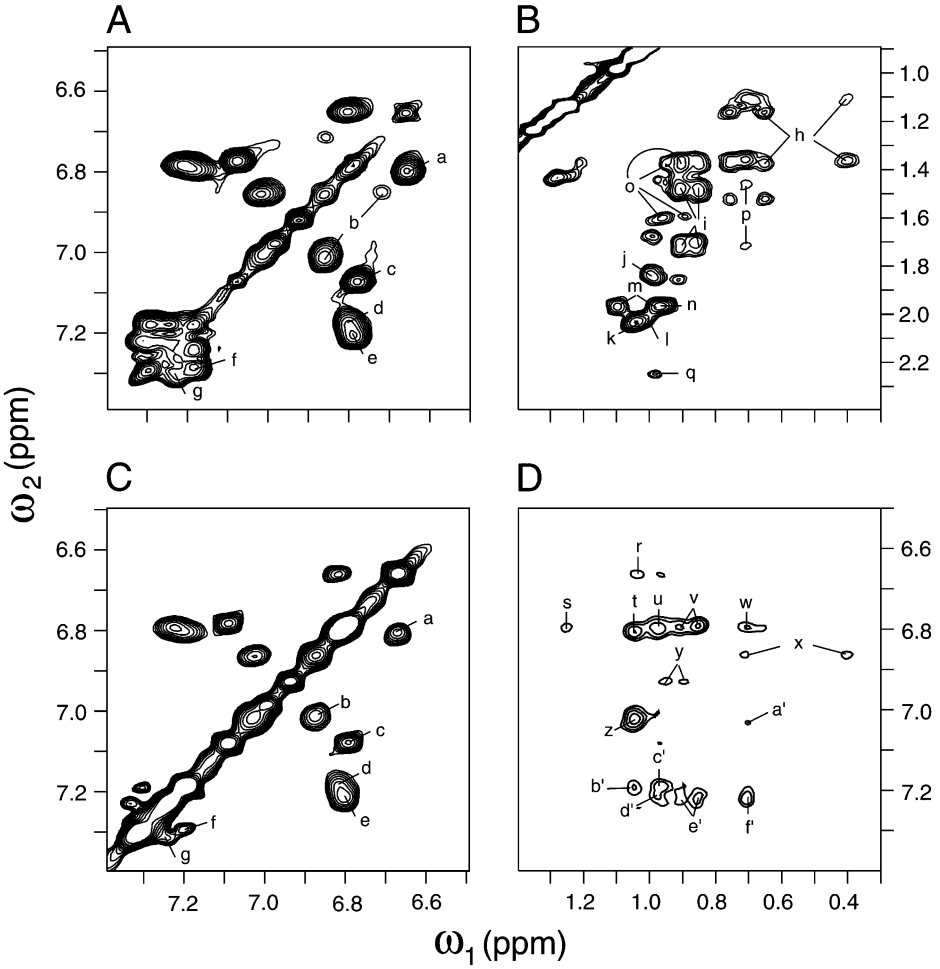

FIGURE 8: Structural relationships in insulin-*swap2* in aqueous solution (pD 7.7 and 32 °C). (A and B) NOESY spectrum showing aromatic (A) and methyl-related (B) NOEs. (C) TOCSY spectrum showing aromatic spin systems. (D) NOEs between aromatic protons and methyl resonances. The NOESY and TOCSY mixing times were 200 and 55 ms, respectively. Assignments are as follows: (A and C) (a) B26 H_ϵ−B26 H_δ, (b) B24 H_ϵ−B24 H_δ and B24 H_ϵ−B24 H_ζ, (c) A14 H_ϵ−A14 H_δ, (d) B16 H_δ−B16 H_ϵ, (e) A19 H_δ−A19 H_ϵ, (f) B1 H_δ−B1 H_ϵ, and (g) B25 H_δ−B25 H_ϵ and (B and D) (h) B15 H_β−B15 $\delta_{1,2}$ CH_3 and B15 H_γ−B15 $\delta_{1,2}$ CH_3, (i) A16 H_β−A16 $\delta_{1,2}$ CH_3 and A16 H_γ−A16 $\delta_{1,2}$ CH_3, (j) Ia H_β−Ia γ' CH_3, (k) B12 H_β−B12 $\gamma_{1,2}$ CH_3, (l) B2 H_β−B2 $\gamma_{1,2}$ CH_3, (m) B18 H_β−B18 $\gamma_{1,2}$ CH_3, (n) A3 H_β−A3 $\gamma_{1,2}$ CH_3, (o) B6 H_β−B6 $\delta_{1,2}$ CH_3 and B6 H_γ−B6 $\delta_{1,2}$ CH_3, (p) A16 H_β−B15 δ_1 CH_3 and A16 H_γ−B15 δ_1 CH_3, (q) A17 H_γ−B18 $\gamma_{1,2}$ CH_3, (r) B26 H_ϵ−B12 $\gamma_{1,2}$ CH_3, (s) A19 H_ϵ−B15 $H_{\beta1}$, (t) B26 H_δ−B12 $\gamma_{1,2}$ CH_3, (u) A19 H_ϵ−Ia δ CH_3, (v) A19 H_ϵ−A16 $\delta_{1,2}$ CH_3, (w) A19 H_ϵ−B16 δ_1 CH_3, (x) B24 H_ϵ−B15 $\delta_{1,2}$ CH_3, (y) B5 H_δ−B6 $\delta_{1,2}$ CH_3, (z) B24 H_δ−B12 $\gamma_{1,2}$ CH_3, (a′) B24 H_δ−B15 δ_1 CH_3, (b′) B1 H_δ−B2 $\gamma_{1,2}$ CH_3, (c′) B1 H_δ−B18 $\gamma_{1,2}$ CH_3, (d′) A19 H_δ−Ia δ CH_3, (e′) A19 H_δ−A16 $\delta_{1,2}$ CH_3, and (f′) A19 H_δ−B15 $\delta_{1,2}$ CH_3. The pattern of inter-residue NOEs is retained at an isomer concentration of 50 μM as observed with a high-sensitivity cryogenic probe (Supporting Information).

Table 2: Helix-Associated NOEs[a] in Insulin-*swap2*

(A) present in both 20% acetic acid and aqueous solution (pD 7.7)
$d_{\alpha\beta(i,i+3)}$, B9−B12, B10−B13, B13−B16, B14−B17,[b] B15−B18
$d_{\alpha\gamma(i,i+3)}$, B9−B12 and B15−B18
noncanonical (i, $i+3$) contacts between A16 δ-CH_3 and A19 $H_{\delta,\epsilon}$
(B) present only in 20% acetic acid
$d_{\alpha\beta(i,i+3)}$, B12−B15 and related NOEs[c]
noncanonical (i, $i+3$) contact[d] between A16 δ-CH_3 and A19 H_α
(C) present only in aqueous solution (pD 7.7)
$d_{\alpha\beta(i,i+3)}$, A15−A18 and B11−B14

[a] This table pertains to spectra observed in D_2O. No non-native helix-associated NOEs are observed; only a subset of native contacts is maintained. An A12−A16 $d_{\alpha\gamma(i,i+4)}$ contact is also observed in 20% deuterioacetic acid but not observed in the spectrum of native insulin. [b] The cross-peak is unresolved from an alternative possible assignment that is inconsistent with nativelike structure. [c] Helix-related contacts are also observed between H_α of Val^{B12} and the methine and methyl resonances of Leu^{B15}. [d] The δ_1 and δ_2 methyl resonances of Leu^{A16} have similar chemical shifts in 20% deuterioacetic acid.

Table 3: Long-Range NOEs[a] between Side Chains in Insulin-*swap2*

(A) present in both 20% acetic acid and aqueous solution (pD 7.7)[b]
Leu^{A16} $\delta_{1,2}$ CH_3−Leu^{B15} H_α, Tyr^{A19} H_δ−Leu^{B15} δ_1 CH_3,
Tyr^{A19} $H_{\beta1,2}$−Leu^{B15} δ_1 CH_3, Tyr^{A19} $H_{\delta,\epsilon}$−isoleucine[c] δ-CH_3
(B) present only in 20% acetic acid
Leu^{B15} δ_2 CH_3−Tyr^{B26} H_δ
(C) in aqueous solution (pD 7.7) and possibly in 20% acetic acid
Glu^{A17} H_γ−Val^{B18} γ_2 CH_3, Leu^{B6} δ_2 CH_3−Leu^{B11} H_α
(D) present only in aqueous solution (pD 7.7)[d]
Leu^{A16} $H_{\alpha,\beta,\gamma}$−Leu^{B15} δ_1 CH_3, Tyr^{A19} H_ϵ−Leu^{B15} H_β and δ_1 CH_3,
Val^{B12} γ CH_3−Tyr^{B26} H_δ, Val^{B12} δ CH_3−Phe^{B24} H_δ,
Leu^{B15} δ_1 CH_3−Phe^{B24} H_δ, Leu^{B6} δ CH_3−His^{B10} H_β

[a] This table pertains to spectra observed in D_2O. No non-native long-range NOEs are observed; only a subset of native contacts is maintained. [b] The δ_1 and δ_2 methyl resonances of Leu^{A16} have similar chemical shifts in 20% deuterioacetic acid. [c] The two isoleucine spin systems (Ile^{A2} and Ile^{A10}) are classified as Ia and Ib but not individually assigned (see Figures 6 and 7B). [d] The γ_1 and γ_2 methyl resonances of Val^{B12} have similar chemical shifts under both conditions.

function (*15*, *18*, *36*, *55*). We have previously described the structure of insulin-*swap* as a model of a protein caught in a kinetic trap (*34*). This disulfide isomer ([A7−A11,A6−B7,A20−B19] pairing scheme), originally prepared by total

protein synthesis (*13*), is accessible to kinetic disulfide reassortment under denaturing conditions. In this paper, a second accessible disulfide isomer ([A6−A7,A11−B7,A20−B19] pairing scheme) has been investigated by CD and NMR spectroscopy. CD spectra indicate an attenuated helix content in the following order: insulin-*swap2* (lowest helix content) < insulin-*swap* < native insulin (highest helix content). The limited chemical shift dispersion observed in the ^{1}H NMR spectrum of insulin-*swap2* indicates that the spatial organization of side chains (relative to aromatic rings currents) is not well maintained (*54*).

Regions of substructure populated by the majority of conformers in solution have been identified by 2D NMR spectroscopy. A bootstrap strategy has been applied whereby sequential assignment in an acidic cosolvent is extended to aqueous solution near neutral pH. The remarkable sensitivity of cryogenic probe technology enabled 2D NMR spectra of dilute protein solutions to be obtained (50 μM), thus verifying monomer-specific structural features. Although the secondary structure of the B chain of insulin-*swap2* is similar to that of native insulin and insulin-*swap*, the A chain is substantially less ordered. A partial hydrophobic core with nativelike interactions is nonetheless maintained in the neighborhood of the shared A20−B19 disulfide bridge. This partial structure and potential induced fit of disordered regions rationalize the isomer's biological activity as an insulin agonist (*13*). Thermodynamic studies by CD-detected guanidine denaturation demonstrate that insulin-*swap* and insulin-*swap2* are each markedly less stable than native insulin. Such instability is in qualitative accord with the isomers' lower α-helix content (Figure 2B) and attenuated ^{1}H NMR chemical shift dispersion (Figure 6). Similar instability has been observed among two-disulfide analogues of insulin (*15*, *36*). Of the two isomers, insulin-*swap2* is the less stable and less organized. These features rationalize its later rp-HPLC elution time (Figure 1C); greater exposure of hydrophobic surfaces would enhance the isomer's affinity for C8 and C18 aliphatic matrices.

The A and B domains of IGF-I are homologous to those of insulin and yet exhibit distinct disulfide-pairing properties (*48*, *49*). Native IGF-I exists in equilibrium with IGF-*swap*; a putative IGF-*swap2* has not been detected. The present guanidine unfolding studies indicate that native insulin and IGF-I exhibit different stabilities (relative to their respective unfolded states); insulin is significantly more stable. The lower stability of IGF-I to guanidine denaturation is consistent with the instability of its disulfide bridges *in vitro* at a redox potential similar to that of blood (*51*). The greater intrinsic stability of insulin is in accord with biological requirements; whereas IGF-I is transported as a complex with specific binding proteins, insulin circulates in the blood stream as a monomer and must maintain its disulfide bridges to be active. That insulin-*swap* (and so presumably proinsulin-*swap*) is destabilized relative to IGF-*swap* is also consistent with biological requirements; whereas nascent disulfide pairing by reduced proinsulin is thought to be affected within the monomeric polypeptide chain, the equilibrium between native IGF-I and IGF-*swap* in the endoplasmic reticulum is pulled toward the native pairing scheme by specific IGF-binding proteins (*56*). Such a coupled equilibrium in a heterodimeric system has apparently relaxed the selection pressure on the IGF-I sequence to encode a unique ground state. Interestingly, application of two-state fitting also predicts that IGF-*swap* is more stable under these conditions (50 mM KCl and 10 mM potassium phosphate at pH 7.0 and 4 °C) than is native IGF-I ($\Delta\Delta G_u = 0.8 \pm 0.1$ kcal/mol; Table 1), in accord with a shift in its unfolding curve to the right. (Because of a difference in the *m* values, however, the apparent difference in ΔG_u may be an overestimate. Correction for *m* value change yields a lower bound on $\Delta\Delta G_u$ of 0.3 kcal/mol.) It seems remarkable that the non-native isomer can be more stable than native IGF-I under some conditions. In contrast, disulfide exchange at higher temperatures and pHs demonstrates that native IGF-I is more populated than IGF-*swap*, making possible a robust estimate of $\Delta\Delta G$ of −0.4 to 0.7 kcal/mol (the negative sign reflects the opposite relative stabilities; *51*).

Implications for Protein Folding. A correspondence is observed between the partial fold of insulin-*swap2* and those of two-disulfide analogues of insulin sharing a cystine A20−B19 bridge (*15*, *36*). The latter analogues represent peptide models (*57*) of populated proinsulin folding intermediates (*14*). Each contains nativelike B chain supersecondary structure and a C-terminal A chain α-helix, the packing of which is proposed to provide a specific folding nucleus (*15*). Preferred pathways of disulfide pairing (*14*) and associated structures have motivated the hypothesis that nascent folding of the B domain provides a template for subsequent folding of the A domain. Stepwise organization of the the polypeptide chain with successive disulfide pairing suggests a hierarchical process (*3*, *4*). We imagine that diffusion and collision of nascent B chain and C-terminal A chain helices (a framework mechanism) leads to an extended disulfide-stabilized nucleus polarized with respect to frayed N-terminal segments in each chain. Nativelike B chain supersecondary structure and the nonlocal A20−B19 disulfide bridge would ensure that proinsulin's transition state exhibits a nativelike topology. Structures of insulin-*swap* and -*swap2* are in accord with this proposal in that an invariant B chain template is mismatched with aberrant N-terminal A chain folds. It is not known whether these isomers are the most stable of the possible alternative folds or only the most kinetically accessible from the ground state. Our attempts to obtain additional isomers through prolonged reassortment reactions were confounded by formation of polymers.

We imagine that the energy landscape of proinsulin has a funnel shape, in accord with the "new view" of protein folding (*16*; Figure 9A). Preferential trajectories down the funnel are populated (*58*), in accord with the experimental characterization of a preferred disulfide pathway (*14*, *15*). Disulfide intermediates occupy a range of molten-globule states (shaded region in Figure 9A) from which non-native structures may be trapped by disulfide mispairing (traps I and II in Figure 9A). The cystine A20−B19 bridge forms early in the kinetic pathway of disulfide pairing, rationalizing its maintenance in the two accessible isomers (*15*, *49*). We cannot exclude the possibility that non-native isomers lacking the A20−B19 bridge exist but cannot be reached by trajectories accessible within the funnel. In the future, this issue may be addressed through directed synthesis of inaccessible isomers. Such studies may illuminate why the sequence of proinsulin specifies a unique disulfide-pairing scheme whereas that of IGF-I encodes two distinct ground

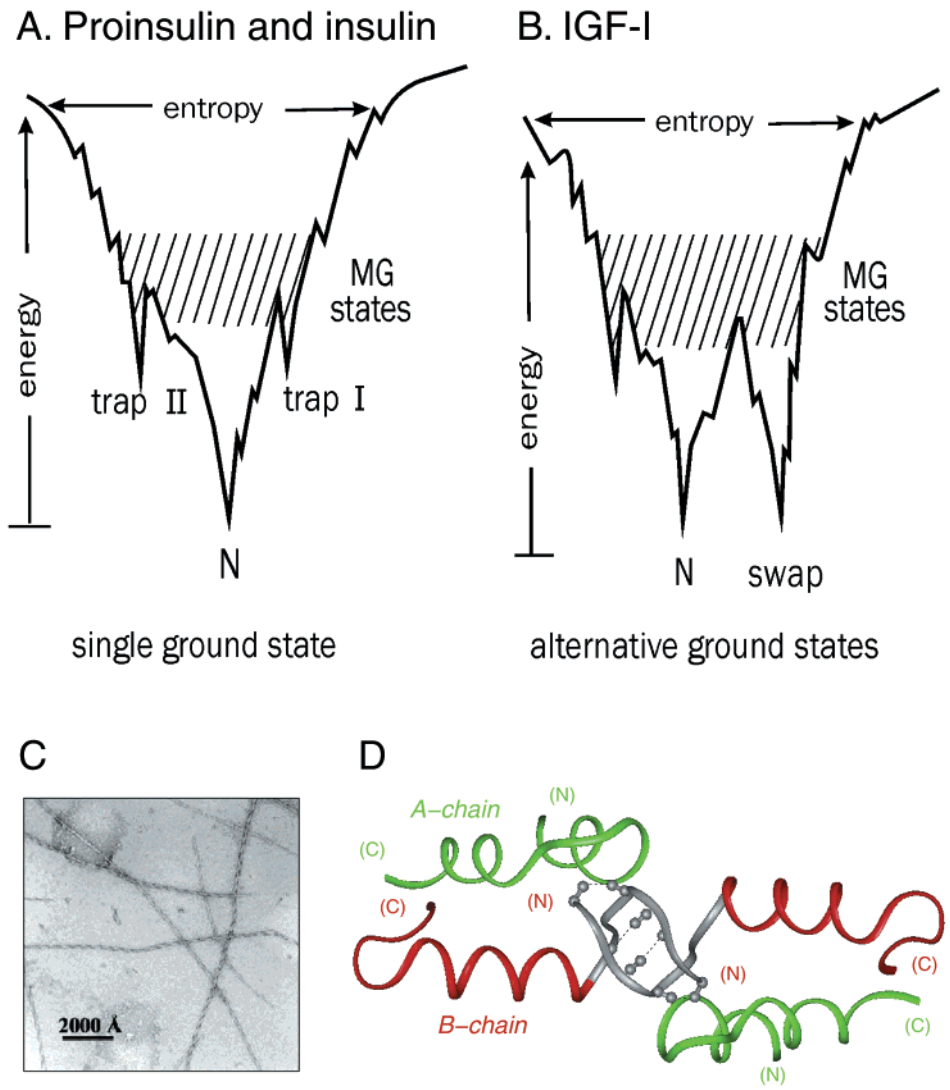

FIGURE 9: Protein folding landscapes and fibrillation. Comparison of funnel-like energy landscapes for proinsulin and insulin (A) and IGF-I (B). Whereas IGF-I exhibits two distinct ground states (native and *swap*), insulin and proinsulin exhibit a single minimum with higher-energy potential kinetic traps. These alternative folds are not ordinarily detected during insulin chain combination or oxidative folding of proinsulin. (C) Electron micrograph of insulin fibrils obtained from ref *68*. (D) The crystal structure of *des*-pentapeptide[B26−B30]-insulin contains a novel antiparallel β-sheet as a lattice contact between monomers (*17*). The N-terminal segment of the B chain is displaced from its native contact against the A chain to contact an adjoining A chain. This β-sheet is proposed as a model of an interaction in the cross-β insulin fibril (*17*) or within an amyloidogenic nucleus (*63*).

states[11] (native IGF-I and IGF-*swap*, Figure 9B; *48*, *49*).

Implications for Misfolding and Fibrillation. An unexpected feature of insulin-*swap2* is its apparent thermal stability as monitored by far-ultraviolet CD spectroscopy. Such stability stands in contrast to the isomer's exquisite sensitivity to guanidine denaturation. Whereas native insulin and (to a lesser extent) insulin-*swap* exhibit attenuated α-helix content at elevated temperatures (*15*), the CD spectrum of insulin-*swap2* is similar at high and low temperatures. This spectrum resembles those of insulin and insulin-*swap* at 70 °C. How can an unstable partial fold resist thermal denaturation? To resolve this apparent paradox, we propose that the structure of native insulin consists of two substructures: temperature-stable B chain supersecondary structure (residues B9−B26 and the adjoining A16−A20 hydrophobic cluster) and temperature-labile substructure elsewhere. This proposal suggests that the contrasting CD thermal features of insulin, insulin-*swap*, and insulin-*swap2* reflect differences in the extent to which the latter regions are folded at low temperatures. Since NMR studies demonstrate that the three species contain similar B chain structure and the adjoining A16−A20 hydrophobic cluster at room temperature (the presumed temperature-stable substructure), it is likely that insulin at 70 °C retains an analogous partial fold (Figure 3F). This same substructure is proposed to form a specific folding nucleus in the kinetic pathway of disulfide pairing (*14*, *15*, *59*).

Proteins ordinarily exhibit cooperative ("all or none") unfolding transitions. Such cooperativity is an evolved property of polypeptide sequences proposed to hinder their otherwise universal propensity to form cross-β fibrils (*60*). Studies of insulin analogues have demonstrated segmental

[11] Recent studies of chimeric single-chain insulin precursor polypeptides have demonstrated that the B domain is responsible for this difference (*73*). Polypeptides containing the insulin B domain and IGF-I A domain fold with unique disulfide pairing, whereas sequences containing the IGF-I B domain and insulin A domain form native and *swap* isomers in equilibrium.

unfolding (*37*, *61*), suggesting that the monomer is not highly cooperative. The structures of insulin isomers highlight this noncooperativity in that a nativelike B chain substructure is compatible with non-native A chain configurations. The insulin monomer is also highly susceptible to fibrillation, which for many decades has been a key issue in its clinical formulation (*31*, *62*). Fibrillation is enhanced by partial unfolding under a variety of stresses, including temperature (*17*). An example of an insulin fibril formed at 70 °C is shown in Figure 9C (*63*). The structure of insulin at 70 °C would be of interest in relation to the mechanism of fibrillation (*63*−*67*); the proposed partial fold (Figure 3F) would render accessible the N-terminal segments of the A and B chains for initiation of formation of an intermolecular β-sheet. Although the structure of an insulin fibril is not known, the participation of the B1−B8 segment in an intermolecular antiparallel β-sheet has in fact been observed as a lattice contact in the crystal structure of a monomeric insulin fragment [*des*-pentapeptide[B26−B30]-insulin (DPI)] (Figure 9D). This crystal contact is proposed as a model of non-native interactions in either the mature fibril or amyloidogenic intermediate[12] (*17*). It would be intriguing if insulin's susceptibility to fibrillation should arise as a side consequence of limited cooperativity.

This study has addressed how alternative disulfide pairing schemes are accommodated in nativelike partial folds of low stability. It would be of future interest to investigate whether insulin's disulfide isomers are also susceptible to fibrillation and, if so, how alternative disulfide bridges can be accommodated (or not) in an amyloidogenic nucleus or cross-β polymer. Would a particular pairing scheme be incompatible with long-range order and hence yield an amorphous precipitate? To our knowledge, there are presently no examples of a structural switch between ordered cross-β assembly and disordered aggregation based on a change in disulfide-directed chain topology rather than sequence.

CONCLUDING REMARKS

The landscape paradigm of protein folding rationalizes how a correct fold forms with high probability following hydrophobic collapse. Alternative metastable folds may exist at early stages of this process or as long-lived kinetic traps. This study exploits the kinetic accessibility of non-native disulfide isomers under denaturing conditions for obtaining models of a globular protein caught in a kinetic trap. The stabilities of these isomers are marginal under native conditions. Nonetheless, their structures exhibit some features of conventional native states, including well-defined elements of secondary structure and partial desolvation of nonpolar residues. Insulin-*swap* and insulin-*swap2* are less organized than native insulin. The incompatibility of the insulin sequence (especially in the A chain) with these alternative folds provides a physical realization of a template-threading procedure (*7*, *8*). The predominant structural features observed in NMR studies coincide with regions of nascent structure proposed to provide a specific folding nucleus in native oxidative folding (*59*). This nucleus is polarized, containing major elements of B chain supersecondary structure and a portion of the C-terminal A chain α-helix. Recapitulation of this nucleus in the metastable structures of insulin-*swap* and insulin-*swap2* may account for their kinetic accessibility from the native state. This perspective suggests that these traps exist in the walls of the native funnel and so do not represent rugged features of the outer landscape. Analysis of disulfide-trapped non-native states may provide a general method for probing the topography of an energy surface within the molten zone of hydrophobic collapse.

[12] Interestingly, the B1−B8 segment in the DPI crystal structure is displaced from its native T-state configuration adjoining the surface of the A chain (Figure 1A). The displaced segment contacts the A chain of an adjoining molecule as a rudimentary example of a "domain-swapped" dimer (*74*−*76*). Domain swapping occurs in diverse crystal structures and is proposed as a mechanism of amyloidogenesis (*75*), including in self-association of the prion protein PrP^c (*77*).

ACKNOWLEDGMENT

We thank A. Pekar and J. A. Hoffman (Lilly Research Laboratories) for assistance with HPLC separations and analysis of disulfide reassortment products, C. Anklin (Bruker Biospin, Inc.) for cryoprobe-assisted NMR spectroscopy, R. E. Chance, R. DiMarchi, G. G. Dodson, E. Dodson, C. M. Dobson, and J. A. Miller for discussion, D. Jones and K. Hallenga for assistance with NMR spectroscopy, E. Collins for preparation of the manuscript, and E. Blout, M. Karplus, and the late L. J. Neuringer for encouragement. This study is a contribution from the Cleveland Center for Structural Biology.

SUPPORTING INFORMATION AVAILABLE

Eleven figures and five tables providing results of gel-permeation chromatography, thermal reversibility of CD studies, cryoprobe-assisted NMR spectroscopy, DQF-COSY spectra, and chemical shift information. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Anfinsen, C. B. (1973) Principles that govern the folding of protein chains, *Science 181*, 223−230.
2. Balbach, J., Forge, V., van Nuland, N. A., Winder, S. L., Hore, P. J., and Dobson, C. M. (1995) Following protein folding in real time using NMR spectroscopy, *Nat. Struct. Biol. 2*, 865−870.
3. Baldwin, R. L., and Rose, G. D. (1999) Is protein folding hierarchic? I. Local structure and peptide folding, *Trends Biochem. Sci. 24*, 26−33.
4. Baldwin, R. L., and Rose, G. D. (1999) Is protein folding hierarchic? II. Folding intermediates and transition states, *Trends Biochem. Sci. 24*, 77−83.
5. Bowie, J. U., Reidhaar-Olson, J. F., Lim, W. A., and Sauer, R. T. (1990) Deciphering the message in protein sequences: tolerance to amino acid substitutions, *Science 247*, 1306−1310.
6. Hecht, M. H. (1994) De novo design of β-sheet proteins, *Proc. Natl. Acad. Sci. U.S.A. 91*, 8729−8730.
7. Bowie, J. U., Luthy, R., and Eisenberg, D. (1991) A method to identify protein sequences that fold into a known three-dimensional structure, *Science 253*, 164−170.
8. Sali, A., and Blundell, T. L. (1993) Comparative protein modeling by satisfaction of spatial restraints, *J. Mol. Biol. 234*, 779−815.
9. Marchler-Bauer, A., and Bryant, S. H. (1997) A measure of success in fold recognition, *Trends Biochem. Sci. 22*, 236−240.
10. Novotny, J., Bruccoleri, R., and Karplus, M. (1984) An analysis of incorrectly folded protein models. Implications for structure predictions, *J. Mol. Biol. 177*, 787−818.
11. Beck, C., Siemens, X., and Weaver, D. L. (2001) Diffusion-collision model study of misfolding in a four-helix bundle protein, *Biophys. J. 81*, 3105−3115.

12. Kinjo, A. R., Kidera, A., Nakamura, H., and Nishikawa, K. (2001) Physicochemical evaluation of protein folds predicted by threading, *Eur. Biophys. J. 30*, 1−10.
13. Sieber, P. S., Eisler, K., Kamber, B., Riniker, B., Rittel, W., Marki, F., and deGasparo, M. (1978) Synthesis and biological activity of two disulphide bond isomers of human insulin: [A7-A11, A6-B7-cystine] and [A6-A7, A11-B7-cystine] insulin (human), *Hoppe-Seyler's Z. Physiol. Chem. 359*, 113−123.
14. Qiao, Z. S., Guo, Z. Y., and Feng, Y. M. (2001) Putative disulfide-forming pathway of porcine insulin precursor during its refolding in vitro, *Biochemistry 40*, 2662−2668.
15. Hua, Q.-X., Nakagawa, S. H., Jia, W., Hu, S. Q., Chu, Y.-C., Katsoyannis, P. G., and Weiss, M. A. (2001) Hierarchical protein folding: asymmetric unfolding of an insulin analogue lacking the A7-B7 interchain disulfide bridge, *Biochemistry 40*, 12299−12311.
16. Onuchic, J. N., Socci, N. D., Luthey-Schulten, Z., and Wolynes, P. G. (1996) Protein folding funnels: the nature of the transition state ensemble, *Folding Des. 1*, 441−450.
17. Brange, J., Dodson, G. G., Edwards, D. J., Holden, P. H., and Whittingham, J. L. (1997) A model of insulin fibrils derived from the X-ray crystal structure of a monomeric insulin (*des*pentapeptide insulin), *Proteins 27*, 507−516.
18. Hua, Q. X., Hu, S. Q., Frank, B. H., Jia, W., Chu, Y. C., Wang, S. H., Burke, G. T., Katsoyannis, P. G., and Weiss, M. A. (1996) Mapping the functional surface of insulin by design: structure and function of a novel A-chain analogue, *J. Mol. Biol. 264*, 390−403.
19. Olsen, H. B., Ludvigsen, S., and Kaarsholm, N. C. (1996) Solution structure of an engineered insulin monomer at neutral pH, *Biochemistry 35*, 8836−8845.
20. Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. C., Mercola, D. A., and Vijayan, M. (1971) Atomic positions in rhombohedral 2-zinc insulin crystals, *Nature 231*, 506−511.
21. Bentley, G., Dodson, E., Dodson, G., Hodgkin, D., and Mercola, D. (1976) Structure of insulin in 4-zinc insulin, *Nature 261*, 166−168.
22. Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M., Hubbard, R. E., Isaacs, N. W., and Reynolds, C. D. (1988) The structure of 2Zn pig insulin crystals at 1.5 Å resolution, *Philos. Trans. R. Soc. London 319*, 369−456.
23. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Smith, G. D., Sparks, C., and Swenson, D. (1989) Phenol stabilizes more helix in a new symmetrical zinc insulin hexamer, *Nature 338*, 594−596.
24. Badger, J., Harris, M. R., Reynolds, C. D., Evans, A. C., Dodson, E. J., Dodson, G. G., and North, A. C. (1991) Structure of the pig insulin dimer in the cubic crystal, *Acta Crystallogr. B47*, 127−136.
25. Steiner, D. F. (1967) Evidence for a precursor in the biosynthesis of insulin, *Trans. N.Y. Acad. Sci. 30*, 60−68.
26. Heath, W. F., Belagaje, R. M., Brooke, G. S., Chance, R. E., Hoffmann, J. A., Long, H. B., Reams, S. G., Roundtree, C., Shaw, W. N., Slieker, L. J., et al. (1992) (A-C-B) human proinsulin, a novel insulin agonist and intermediate in the synthesis of biosynthetic human insulin, *J. Biol. Chem. 267*, 419−425.
27. Katsoyannis, P. G., and Tometsko, A. (1966) Insulin synthesis by recombination of A and B chains: a highly efficient method, *Proc. Natl. Acad. Sci. U.S.A. 55*, 1554−1561.
28. Tang, J. G., and Tsou, C. L. (1990) The insulin A and B chains contain structural information for the formation of the native molecule. Studies with protein disulphide-isomerase, *Biochem. J. 268*, 429−435.
29. Brange, J., Ribel, U., Hansen, J. F., Dodson, G., Hansen, M. T., Havelund, S., Melberg, S. G., Norris, F., Norris, K., and Snel, L. (1988) Monomeric insulins obtained by protein engineering and their medical implications, *Nature 333*, 679−682.
30. Brems, D. N., Alter, L. A., Beckage, M. J., Chance, R. E., DiMarchi, R. D., Green, L. K., Long, H. B., Pekar, A. H., Shields, J. E., and Frank, B. H. (1992) Altering the association properties of insulin by amino acid replacement, *Protein Eng. 5*, 527−533.
31. Brange, J. (1997) The new era of biotech insulin analogues, *Diabetologia 40*, S48−S53.
32. Lee, H. C., Kim, S. J., Kim, K. S., Shin, H. C., and Yoon, J. W. (2000) Remission in models of type 1 diabetes by gene therapy using a single-chain insulin analogue, *Nature 408*, 483−488.
33. DeFelippis, M. R., Chance, R. E., and Frank, B. H. (2001) Insulin self-association and the relationship to pharmacokinetics and pharmacodynamics, *Crit. Rev. Ther. Drug Carrier Syst. 18*, 201−264.
34. Hua, Q. X., Gozani, S. N., Chance, R. E., Hoffmann, J. A., Frank, B. H., and Weiss, M. A. (1995) Structure of a protein in a kinetic trap, *Nat. Struct. Biol. 2*, 129−138.
35. Hua, Q. X., and Weiss, M. A. (1991) Comparative 2D NMR studies of human insulin and *des*-pentapeptide insulin: sequential resonance assignment and implications for protein dynamics and receptor recognition, *Biochemistry 30*, 5505−5515.
36. Weiss, M. A., Hua, Q.-X., Jia, W., Chu, Y.-C., Wang, R.-Y., and Katsoyannis, P. G. (2000) Hierarchiacal protein "un-design": insulin's intrachain disulfide bridge tethers a recognition α-helix, *Biochemistry 39*, 15429−15440.
37. Hua, Q. X., Shoelson, S. E., Kochoyan, M., and Weiss, M. A. (1991) Receptor binding redefined by a structural switch in a mutant human insulin, *Nature 354*, 238−241.
38. Shoelson, S. E., Lu, Z. X., Parlautan, L., Lynch, C. S., and Weiss, M. A. (1992) Mutations at the dimer, hexamer, and receptor-binding surfaces of insulin independently affect insulin-insulin and insulin-receptor interactions, *Biochemistry 31*, 1757−1767.
39. Sreerama, N., and Woody, R. W. (1993) A self-consistent method for the analysis of protein secondary structure from circular dichroism, *Anal. Biochem. 209*, 32−44.
40. Sreerama, N., and Woody, R. W. (2000) Estimation of protein secondary structure from CD spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set, *Anal. Biochem. 287*, 252−260.
41. Sosnick, T. R., Fang, X., and Shelton, V. M. (2000) Application of circular dichroism to study RNA folding transitions, *Methods Enzymol. 317*, 393−409.
42. Luo, Y., and Baldwin, R. L. (2001) How Ala → Gly mutations in different helices affect the stability of the apomyoglobin molten globule, *Biochemistry 40*, 5283−5289.
43. Pace, C. N., and Shaw, K. L. (2000) Linear extrapolation method of analyzing solvent denaturation curves, *Proteins* (Suppl. 4), 1−7.
44. Weiss, M. A., Hua, Q. X., Jia, W., Nakagawa, S. H., Chu, Y. C., and Katsoyannis, P. G. (2002) in *Insulin & Related Proteins: Structure to Function and Pharmacology* (Dieken, M. L., Federwisch, M., and De Meyts, P., Eds.) pp 103−119, Kluwar Academic Publishers, Dordrecht, The Netherlands.
45. Frank, B. H., Pekar, A. H., and Veros, A. J. (1972) Insulin and proinsulin conformation in solution, *Diabetes 21*, 486−491.
46. Weiss, M. A., Frank, B. H., Khait, I., Pekar, A., Heiney, R., Shoelson, S. E., and Neuringer, L. J. (1990) NMR and photo-CIDNP studies of human proinsulin and prohormone processing intermediates with application to endopeptidase recognition, *Biochemistry 29*, 8389−8401.
47. Brems, D. N., Brown, P. L., Heckenlaible, L. A., and Frank, B. H. (1990) Equilibrium denaturation of insulin and proinsulin, *Biochemistry 29*, 9289−9293.
48. Hober, S., Forsberg, G., Palm, G., Hartmanis, M., and Nilsson, B. (1992) Disulfide exchange folding of insulin-like growth factor I, *Biochemistry 31*, 1749−1756.
49. Miller, J. A., Narhi, L. O., Hua, Q. X., Rosenfeld, R., Arakawa, T., Rohde, M., Prestrelski, S., Lauren, S., Stoney, K. S., Tsai, L., and Weiss, M. A. (1993) Oxidative refolding of insulin-like growth factor 1 yields two products of similar thermodynamic stability: a bifurcating protein-folding pathway, *Biochemistry 32*, 5203−5213.
50. Sato, A., Koyama, S., Yamada, H., Suzuki, S., Tamura, K., Kobayashi, M., Niwa, M., Yasuda, T., Kyogoku, Y., and Kobayashi, Y. (2000) Three-dimensional solution structure of a disulfide bond isomer of the human insulin-like growth factor-I, *J. Pept. Res. 56*, 218−230.
51. Hober, S., Lundstrom Ljung, J., Uhlen, M., and Nilsson, B. (1999) Insulin-like growth factors I and II are unable to form and maintain their native disulfides under in vivo redox conditions, *FEBS Lett. 443*, 271−276.
52. Weiss, M. A., Eliason, J. L., and States, D. J. (1984) Dynamic filtering by two-dimensional ^{1}H NMR with application to phage lambda repressor, *Proc. Natl. Acad. Sci. U.S.A. 81*, 6019−6023.
53. Weiss, M. A., Hua, Q. X., Lynch, C. S., Frank, B. H., and Shoelson, S. E. (1991) Heteronuclear 2D NMR studies of an engineered insulin monomer: assignment and characterization of the receptor-binding surface by selective ^{2}H and ^{13}C labeling with application to protein design, *Biochemistry 30*, 7373−7389.

54. Jacoby, E., Hua, Q. X., Stern, A. S., Frank, B. H., and Weiss, M. A. (1996) Structure and dynamics of a protein assembly. ^{1}H NMR studies of the 36 kDa R_6 insulin hexamer, *J. Mol. Biol. 258*, 136−157.
55. Guo, Z.-Y., and Feng, Y.-M. (2001) Effects of cysteine to serine substitutions in the two intra-A-chain disulfide bonds of insulin, *Biol. Chem. 382*, 443−448.
56. Hober, S., Hansson, A., Uhlen, M., and Nilsson, B. (1994) Folding of insulin-like growth factor I is thermodynamically controlled by insulin-like growth factor binding protein, *Biochemistry 33*, 6758−6761.
57. Oas, T. G., and Kim, P. S. (1988) A peptide model of a protein folding intermediate, *Nature 336*, 42−48.
58. Lazaridis, T., and Karplus, M. (1997) "New view" of protein folding reconciled with the old through multiple unfolding simulations, *Science 278*, 1928−1931.
59. Hua, Q.-X., Chu, Y.-C., Jia, W., Phillips, N. F. B., Wang, R.-Y., Katsoyannis, P. G., and Weiss, M. A. (2002) Mechanism of insulin chain combination. Asymmetric roles of A-chain α-helices in disulfide pairing, *J. Biol. Chem.* (in press).
60. Chamberlain, A. K., Receveur, V., Spencer, A., Redfield, C., and Dobson, C. M. (2001) Characterization of the structure and dynamics of amyloidogenic variants of human lysozyme by NMR, *Protein Sci. 10*, 2525−2530.
61. Xu, B., Hua, Q. X., Nakagawa, S. H., Jia, W., Chu, Y. C., Katsoyannis, P. G., and Weiss, M. A. (2002) A cavity-forming mutation in insulin induces segmental unfolding of a surrounding α-helix, *Protein Sci. 11*, 104−116.
62. Shnek, D. R., Hostettler, D. L., Bell, M. A., Olinger, J. M., and Frank, B. H. (1998) Physical stress testing of insulin suspensions and solutions, *J. Pharm. Sci. 87*, 1459−1465.
63. Bouchard, M., Zurdo, J., Nettleton, E. J., Dobson, C. M., and Robinson, C. V. (2000) Formation of insulin amyloid fibrils followed by FTIR simultaneously with CD and electron microscopy, *Protein Sci. 9*, 1960−1967.
64. Jimenez, J. L., Guijarro, J. I., Orlova, E., Zurdo, J., Dobson, C. M., Sunde, M., and Saibil, H. R. (1999) Cryo-electron microscopy structure of an SH3 amyloid fibril and model of the molecular packing, *EMBO J. 18*, 815−821.
65. Nielsen, L., Frokjaer, S., Brange, J., Uversky, V. N., and Fink, A. L. (2001) Probing the mechanism of insulin fibril formation with insulin mutants. *Biochemistry 40*, 8397−8409.
66. Nielsen, L., Khurana, R., Coats, A., Frokjaer, S., Brange, J., Vyas, S., Uversky, V. N., and Fink, A. L. (2001) Effect of environmental factors on the kinetics of insulin fibril formation: elucidation of the molecular mechanism, *Biochemistry 40*, 6036−6046.
67. Nielsen, L., Frokjaer, S., Carpenter, J. F., and Brange, J. (2001) Studies of the structure of insulin fibrils by Fourier transform infrared (FTIR) spectroscopy and electron microscopy, *J. Pharm. Sci. 90*, 29−37.
68. Nettleton, E. J., Tito, P., Sunde, M., Bouchard, M., Dobson, C. M., and Robinson, C. V. (2000) Characterization of the oligomeric states of insulin in self-assembly and amyloid fibril formation by mass spectrometry, *Biophys. J. 79*, 1053−1065.
69. Shortle, D., and Ackerman, M. S. (2001) Persistence of native-like topology in a denatured protein in 8 M urea, *Science 293*, 487−489.
70. Neri, D., Billeter, M., Wider, G., and Wuthrich, K. (1992) NMR determination of residual structure in a urea-denatured protein, the 434-repressor, *Science 257*, 1559−1563.
71. Hua, Q. X., and Weiss, M. A. (1990) Toward the solution structure of human insulin: sequential 2D ^{1}H NMR assignment of a *des*-pentapeptide analogue and comparison with crystal structure, *Biochemistry 29*, 10545−10555.
72. Whittingham, J. L., Scott, D. J., Chance, K., Wilson, A., Finch, J., Brange, J., and Dodson, G. G. (2002) Insulin at pH 2: Structural analysis of the conditions promoting insulin fibre formation, *J. Mol. Biol. 318*, 479−490.
73. Guo, Z. Y., Shen, L., and Feng, Y. M. (2002) The different folding behavior of insulin and insulin-like growth factor 1 is mainly controlled by their B-chain/domain, *Biochemistry 41*, 1556−1567.
74. Schlunegger, M. P., Bennett, M. J., and Eisenberg, D. (1997) Oligomer formation by 3D domain swapping: a model for protein assembly and misassembly, *Adv. Protein Chem. 50*, 61−122.
75. Liu, Y., Gotte, G., Libonati, M., and Eisenberg, D. (2001) A domain-swapped RNase A dimer with implications for amyloid formation, *Nat. Struct. Biol. 8*, 211−214.
76. Ogihara, N. L., Ghirlanda, G., Bryson, J. W., Gingery, M., DeGrado, W. F., and Eisenberg, D. (2001) Design of three-dimensional domain-swapped dimers and fibrous oligomers, *Proc. Natl. Acad. Sci. U.S.A. 98*, 1404−1409.
77. Knaus, K. J., Morillas, M., Swietnicki, W., Malone, M., Surewicz, W. K., and Yee, V. C. (2001) Crystal structure of the human prion protein reveals a mechanism for oligomerization, *Nat. Struct. Biol. 8*, 770−774.

BI0202981