Equilibrium Intermediates in the Denaturation of Human Insulin and Two Monomeric Insulin Analogs

Rohn L. Millican and David N. Brems*

Lilly Research Laboratories, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, Indiana 46285

Received June 17, 1993; Revised Manuscript Received October 29, 1993®

ABSTRACT: The equilibrium denaturation of human insulin in a monomer-inducing solvent and of two monomeric insulin analogs, lys^{B28}pro^{B29} insulin and asp^{B10}des^{B28-30} insulin, was reexamined [Brems, D. N., Brown, P. L., Heckenlaible, L. A., & Frank, B. H. (1990) *Biochemistry 29,* 9289–9293] by circular dichroism (CD) at additional wavelengths in the near-UV region. Previous denaturation studies were limited by the solubility of guanidine hydrochloride being only slightly greater than the level of denaturant required to fully unfold human insulin; therefore, only a few data points were available for construction of the post-transitional baseline. In the present study, we report the use of an unfolded mimic created by enzymatic digestion of insulin to confirm the slope of the post-transitional baseline. Evidence for equilibrium unfolding intermediates for each of these insulins was indicated by noncoincidence of the denaturation transitions as monitored by tyrosine and helical-dependent CD bands (270 and 224 nm, respectively). Additional evidence for intermediates through multiphasic denaturation transitions was obtained at a wavelength likely related to disulfide conformation, 251 nm. The results suggest that for each of the insulins, at least two intermediates are significantly populated. An unfolding model is proposed in which the conformation of the least stable intermediate is slightly unfolded only in the C-terminal segment of the B chain. A second more stable intermediate retains minimal secondary structure while containing localized structure proximal to one or more of the disulfide groups. The presence of equilibrium intermediates has important implications for the folding pathway of insulin, in pharmaceutical applications such as formulation stability, and for conformational transitions that accompany receptor binding.

The structure of insulin has been the target of numerous studies originating from the first primary structure determination of a protein (Sanger & Tuppy, 1951). Secondary and tertiary structural details were later determined by Blundell and co-workers (1972) and since then, multiple crystal forms of insulin studied by X-ray diffraction have revealed not one but several stable conformations (Derewenda et al., 1989). Conclusions from these crystallography studies indicate that insulin's conformation is flexible, dynamic, and dependent on crystal packing forces. Solution structure of human insulin determined by NMR extends these crystallographic observations to the solution state. Insulin's solution structure most closely resembles the hexameric 2-zinc crystal form of porcine insulin with flexibility in the N- and C-terminus of the B chain readily observed (Hua et al., 1991). Solution and crystal structure analyses reveal that human/pork insulin has two helical stretches in the A chain at A$^{1-8}$ and A$^{13-19}$ in addition to a central B chain helix at B$^{9-19}$. A β-turn located at residues B$^{20-23}$ orients the C-terminus of the B chain over the hydrophobic surface of the B-chain helix. These C-terminal residues, B$^{24-28}$, also participate in an antiparallel β-sheet structure between two monomers at the dimer interface. The N-terminus of the B chain (B$^{1-9}$) can exist in an α-helix or extended conformation depending on the presence or absence of specific ligands (Derewenda et al., 1989).

An insulin analog resulting from a peptide cross-link between the N-terminus of the A chain and the C-terminus of the B chain (lys^{B29} cross-linked gly^{A1}) was determined to be biologically inactive yet surprisingly found to retain a crystal structure near identical to native insulin (Derewenda et al., 1991). On this basis, the known structures of insulin are considered to reflect a biologically inactive conformation. To bind receptor it is believed that the C-terminal residues of the B chain must unfold and move away from the N-terminus of the A chain exposing part of the hydrophobic core. In this way, the flexibility and dynamic nature of the insulin structure are inherently linked to biological activity (Baker et al., 1990; Derewenda et al., 1991; Hua et al., 1993).

The equilibrium denaturation of an insulin analog with a unique spectroscopic probe (tryptophan) engineered at residue B^{28} was previously determined (Bryant et al., 1992). A comparison of the transitions detected by circular dichroism (CD)[1] at 224 nm and tryptophan fluorescence demonstrated evidence of an equilibrium intermediate. This intermediate was characterized as having the C-terminal segment of the B chain unfolded while the secondary structure remained intact. This preferential susceptibility to denaturant perturbation in the C-terminal region of the B chain is consistent with the hypothesis that this segment is quite flexible and highly dynamic with regard to its inherent motion.

Insulin's self-association is well documented. In metal-free solutions, insulin associates to dimer and higher ordered aggregates through noncovalent interactions (Jeffrey et al., 1976). Specific association of insulin to hexamers is inducible by limited concentrations of divalent metal ions (Goldman & Carpenter, 1974). *In vivo,* this metal-dependent association is believed to provide considerable stability within the

* Author to whom correspondence should be addressed.

® Abstract published in *Advance ACS Abstracts,* January 15, 1994.

[1] Abbreviations: CD, circular dichroism; UV, ultraviolet; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*'-2-enthanesulfonic acid; GdnHCl, guanidine hydrochloride; HPLC, high-pressure liquid chromatography; TFA, trifluoroacetic acid; ΔH, enthalpy change; N ↔ U, native ↔ unfolded; NMR, nuclear magnetic resonance; analogs of human insulin are denoted by the replacement amino acid followed in superscript by its chain location (A chain or B chain) and sequence position; LC–MS, liquid chromatography–mass spectroscopy.

0006-2960/94/0433-1116$04.50/0 © 1994 American Chemical Society

secretory/storage granules of the pancreatic islet cells (Greider et al., 1969). Modification of insulin's self-association can be readily altered through solvent conditions (Brems et al., 1990; Bryant et al., 1992) or by select mutations at residues critical to association (Brange et al., 1991; Brems et al., 1992b). Several such mutants have been synthesized that are monomeric and equipotent to wild-type insulin (Brange et al., 1991; Brems et al., 1992b). Monomeric insulin analogs are being developed for commercial use as rapid-acting preparations due to increased absorption rates from the subcutaneous injection site (Brange et al., 1991; DiMarchi et al., 1992). This increased absorption of monomeric analogs has been attributed to the absence of a rate-limiting dissociation step.

We previously examined the GdnHCl-induced equilibrium denaturation of human insulin in 20% ethanol (Brems et al., 1990). Transitions detected by circular dichroism at 224 and 275 nm for human insulin were reversible, sigmoidal in shape, and coincident. From these results it was concluded that denaturation of insulin could be explained by a simple two-state process, native ↔ unfolded, without significantly populated equilibrium intermediates. In this report we reexamine the equilibrium denaturation of human insulin and extend the studies to include detection at additional wavelengths in the near-UV CD spectra. Previous studies were limited by the solubility of GdnHCl being only slightly greater than the level of denaturant required to fully unfold human insulin. In the present study we report the use of a denatured insulin mimic to confirm the slope of the post-transitional baseline thereby reducing the inherent variability in data analysis. The current results, utilizing additional probes of folding and modified data analysis, demonstrate a more complicated folding pathway that includes the presence of at least two well-populated intermediates. Detection of an intermediate in the unfolding of Asp^{B9} insulin was also recently determined by noncoincidence of the CD transitions at 270 and 224 nm (Kaarsholm et al., 1993).

EXPERIMENTAL PROCEDURES

Materials

Biosynthetic human insulin, $lys^{B28}pro^{B29}$ human insulin, and $asp^{B10}des^{B28-30}$ human insulin were prepared by Eli Lilly and Company in a zinc-free lyophilized or sodium crystal form. *Staphylococcus aureus* V8 strain protease and ultrapure, 99.99% guanidine hydrochloride (GdnHCl) was purchased from ICN Biochemicals. Anhydrous absolute ethanol (200 proof) was supplied from Quantum Chemicals. Thermolysin (protease type X) and HEPES (*N*-(2-hydroxyethyl)piperazine-*N*'-2-ethanesulfonic acid), titration purity >99.5%, was purchased from Sigma Chemical Company. Acetonitrile, ChromAR HPLC grade, was obtained from Mallinckrodt Specialty Chemicals Company. All other reagents were of analytical grade or better, and deionized double-distilled water, Milli-Q grade, was used throughout this study.

Methods

Equilibrium Denaturation by Circular Dichroism. A stock solution of insulin or an insulin analog at 6 mg/mL in 20 mM HEPES, pH 7.5, was diluted to a fixed protein concentration of 1 mg/mL by a combination of 8.4 M GdnHCl and 20 mM HEPES, pH 7.5, and 20 mM HEPES, pH 7.5, in various ratios to yield final GdnHCl concentrations between 0 to 7.4 M. A separate stock at 6 mg/mL insulin in 8.4 M GdnHCl and 20 mM HEPES, pH 7.5, was prepared and diluted as above for samples between 7.4 and 8.4 M GdnHCl (lower GdnHCl concentrations were also prepared from this stock to assess reversibility). Ethanol (20%, v/v) was added to the denaturation solutions for human insulin and $lys^{B28}pro^{B29}$ insulin (as a control) to disrupt higher order association. No ethanol was included in samples above 7.4 M GdnHCl due to reduced solubility of GdnHCl.

Refractive index was used to determine GdnHCl concentrations on a Bausch & Lomb Abbe 3-L refractometer by the method of Nozaki (1972). Protein concentrations were determined by UV absorption on an AVIV model 14DS double-beam spectrophotometer using an extinction coefficient at 276 nm of 1.12 $(mg/mL)^{-1}$ cm^{-1} for $asp^{B10}des^{B28-30}$ insulin or 1.05 $(mg/mL)^{-1}$ cm^{-1} for human insulin and $lys^{B28}pro^{B29}$ insulin [unpublished program *ProteinPP* written by Jim Shields, Eli Lilly & Company; the calculated results for human insulin agree with the experimentally determined value of Frank et al. (1968)].

Equilibrium denaturation data were recorded on an AVIV model 62DS circular dichroism spectrophotometer calibrated with (1*S*)-(+)-10-camphorsulfonic acid at 290 nm. In the near-UV CD range, data were collected at room temperature in a 1.0-cm cell using a bandwidth of 1 nm averaged for 60 s. The far-UV CD data were collected in an equivalent manner in a 0.05-cm cell. The raw data were converted to mean residue ellipticity using a mean residue weight of 113.88 for both human insulin and $lys^{B28}pro^{B29}$ insulin or 113.73 for $asp^{B10}des^{B28-30}$ insulin according to the following formula:

$$[\theta]_\lambda = (\theta)(\mathrm{MRW})/10[\mathrm{C}](L)$$

where θ is the observed rotation in millidegrees, MRW is the mean residual weight in daltons, C is the protein concentration in mg/mL, L is the cell path length in centimeters, and $[\theta]_\lambda$ is mean residue ellipticity at a particular wavelength (λ) and has the units of deg $cm^2/dmol$ (Adler, 1973).

Staphylococcus aureus V8 Strain Protease Digest. Lys^{B28}-pro^{B29} insulin dissolved at 10 mg/mL in 20 mM HEPES, pH 7.5 containing 880 units/mL (Kunitz, 1947) *Staphylococcus aureus* V8 strain protease was incubated for 8 h at room temperature. The reaction was quenched by addition of an equal volume of 8.4 M GdnHCl, 20 mM HEPES, pH 7.5. Stability of this sample toward further digestion or chemical degradation over the course of the study was determined by analytical reversed-phase HPLC analysis under the following conditions: The quenched sample was injected on a Shimadzu LC-610 system with an Altech Spherisorb ODS-2, 3-um (4.6 × 150 mm) reversed-phase column thermostated to 40 °C. The resulting three peptide fragments were detected at 220 nm (Grau, 1985) and eluted with a mobile phase consisting of 10% and 40% (v/v) acetonitrile (A and B buffer, respectively) in 0.125 M ammonium sulfate, pH 2.3. A linear gradient from 10% B to 70% B over 60 min was run. To assess secondary structure by CD, preparative purification of the three peptide fragments was conducted on a Vydac Protein and Peptide C_{18} (21.2 × 250 mm) column under similar HPLC conditions.

The resulting quenched peptide mixture from the enzymatic digestion of $lys^{B28}pro^{B29}$ insulin was further analyzed by CD-detected equilibrium denaturation (see Equilibrium Denaturation by Circular Dichroism). A blank consisting of protease and buffer at each GdnHCl concentration was subtracted from the results.

Thermolysin Digest. $Lys^{B28}pro^{B29}$ insulin dissolved at 5 mg/mL in 20 mM 4-methylmorpholine, 0.5 mM $ZnCl_2$, containing 1.6 mg/mL thermolysin by weight, pH 7.7 was allowed to digest for 24 h at room temperature. Samples

were analyzed by analytical reversed-phase HPLC to ensure complete digestion of intact insulin on a Zorbax RX-C8 (4.6 × 150 mm) column thermostated to 40 °C. The mobile-phase consisted of 0% and 50% (v/v) acetonitrile in 0.1% TFA (A and B buffer, respectively). Samples were injected on a column equilibrated in 14% B buffer and eluted with a mobile-phase composition adjusted at 0.25% B per minute for 20 min followed by a 0.5% B per minute linear adjustment for 40 min. Between injections the column was washed by adjusting to 100% B over 15 min and holding for 5 min. Peaks detected at 214 nm were identified by liquid chromatography mass spectroscopy (LC-MS) and amino acid analysis (personal communication with David Sharknas, Eli Lilly & Co.).

The digested peptide mixture was analyzed by CD-detected equilibrium denaturation (see Equilibrium Denaturation by Circular Dichroism). A blank consisting of protease and buffer at each GdnHCl concentration was subtracted from the results.

Biological Activity. The *in vivo* biological properties of insulin and the insulin analogs were determined by quantifying the percent change in blood glucose in fasted normal male Sprague–Dawley rats following subcutaneous administration of the test compound. Dose–response curves were constructed from experiments at five different concentrations of the compound using the maximal response (usually occurring 1 or 2 h post-injection). Five test animals and five control animals were used at each concentration. The subcutaneous dose (microgram of insulin/kilogram of animal weight) that yielded the half-maximal hypoglycemic response (ED_{50}) was determined from each dose–response curve and was taken as a measure of the *in vivo* potency of the particular test compound.

RESULTS

Biological Activity. The rat hypoglycemic effect elicited by human insulin, $lys^{B28}pro^{B29}$ insulin, and $asp^{B10}des^{B28-30}$ insulin were statistically equivalent on the basis of ED_{50} values of 7.75 ± 0.14, 7.20 ± 0.30, and 7.70 ± 0.40 μg/kg, respectively.

Determination of Post-Transitional Baseline. Analysis of the post-transitional baseline in an equilibrium denaturation study was accomplished by monitoring the GdnHCl solvent interaction with an unfolded insulin mimic. This mimic was produced through an enzymatic digest of insulin with either *Staphylococcus aureus* V8 strain protease or thermolysin, which resulted in unfolding of the structure. Enzymatic digest of $lys^{B28}pro^{B29}$ insulin with *S. aureus* V8 strain protease (Grau, 1985) hydrolyzed three peptide bonds on the C-terminal side of glutamic acids A^{17}, B^{13}, and B^{21} (glu^{A4} was not cut under these conditions). HPLC purification of the resulting fragments (≥95% purity by analytical reversed phase, see Experimental Procedures) and subsequent CD spectra analysis qualitatively confirmed the unfolded conformation of all three peptides (data not shown). Stability of this digest with regard to cleavage at glu^{A4} or chemical degradation was not significant over the course of the study as determined by analytical HPLC. Thermolysin digestion of $lys^{B28}pro^{B29}$ insulin resulted in 14 major peptide peaks, which were identified by LC-MS and amino acid analysis. The relative population of each peptide, based on peak area absorbance at 214 nm, ranged from 2 to 13%. Fragments ranged from 1 to 16 residues in length with an area weighted average of seven residues per peptide.

Equilibrium Denaturation. Respectively, Figures 1–4 illustrate the GdnHCl equilibrium denaturation of human insulin in 20% ethanol, $lys^{B28}pro^{B29}$ insulin in 20% ethanol, $lys^{B28}pro^{B29}$ insulin excluding 20% ethanol, and $asp^{B10}des^{B28-30}$

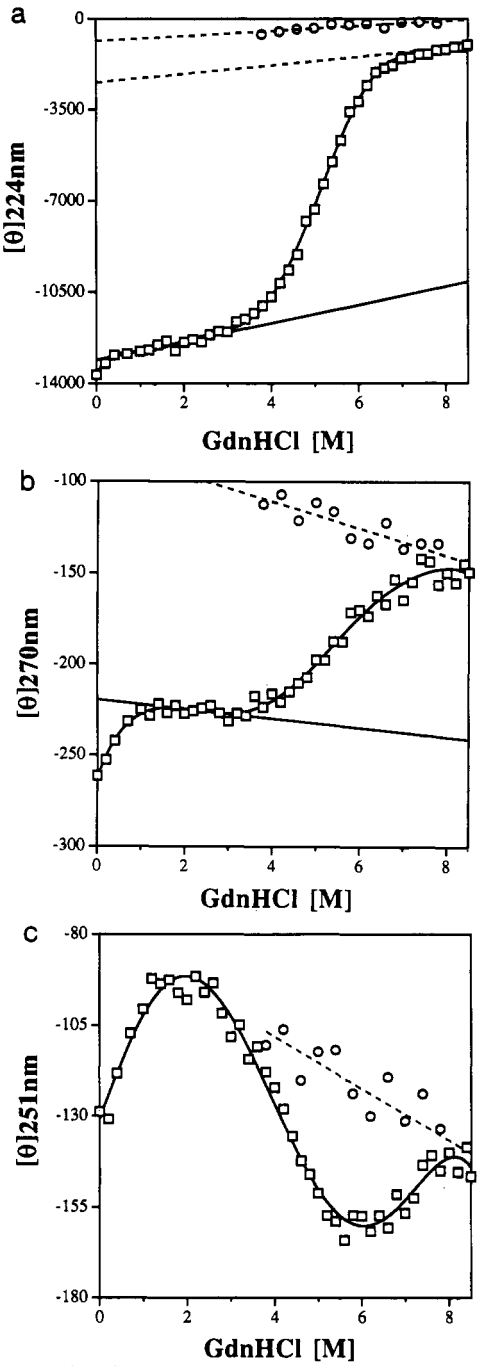


FIGURE 1: GdnHCl-induced equilibrium denaturation of human insulin in 20% ethanol, 20 mM HEPES, pH 7.5, at 1 mg/mL. Square symbols indicate protein denaturation data; open circles represent solvent interaction with *S. aureus* V8 digested insulin mimic. Pre- and post-transitional baselines are represented by the linear solid and broken lines, respectively. The curved solid line represents the best polynomial fit to the data. Transitions detected by CD at (a) 224, (b) 270, and (c) 251 nm.

insulin excluding 20% ethanol as monitored by CD at 270, 251, and 224 nm. Multiple wavelengths were selected to monitor different conformationally dependent chromophores. Circular dichroism at 270 nm principally monitors tyrosine environment and reflects overall tertiary structure, 251 nm most likely reflects the conformation of the disulfide bonds, and 224 nm represents helical content and is representative of secondary structure. Equilibrium denaturation detected at 270 and 224 nm demonstrates transitions that are generally smooth, sigmoidal, and reversible. The pre- and post-transition baselines are indicated by the linear solid and broken lines, respectively, in Figures 1–4. For human insulin detected at

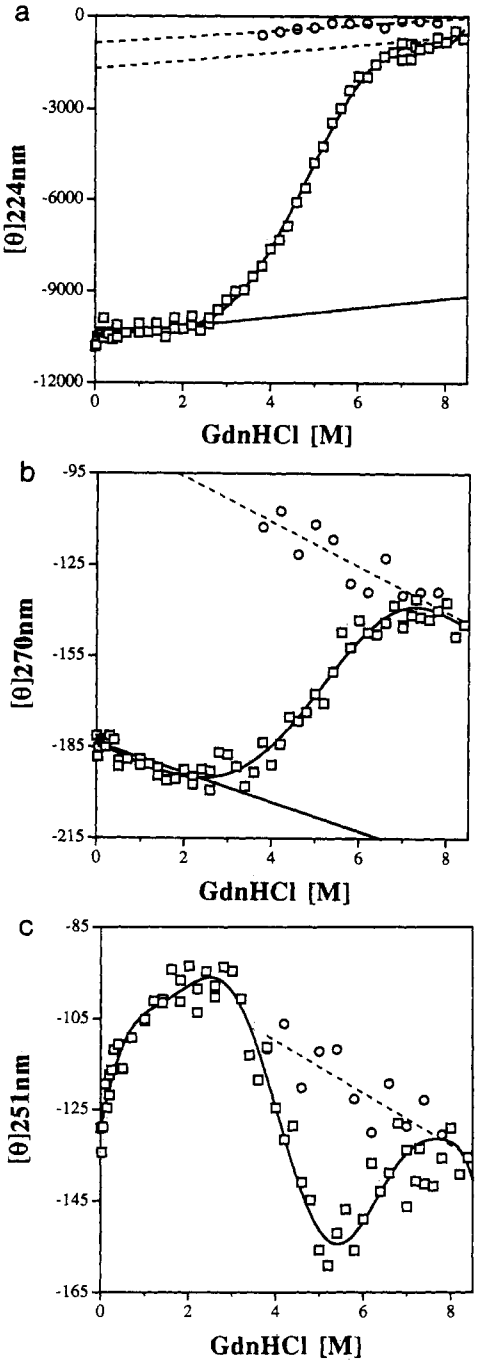


FIGURE 2: GdnHCl-induced equilibrium denaturation of lys^{B28}pro^{B29} insulin in 20% ethanol, 20 mM HEPES, pH 7.5, at 1 mg/mL. Square symbols indicate protein denaturation data; open circles represent solvent interaction with *S. aureus* V8 digested insulin mimic. Pre- and post-transitional baselines are represented by the linear solid and broken lines, respectively. The curved solid line represents the best polynomial fit to the data. Transitions detected by CD at (a) 224, (b) 270, and (c) 251 nm.

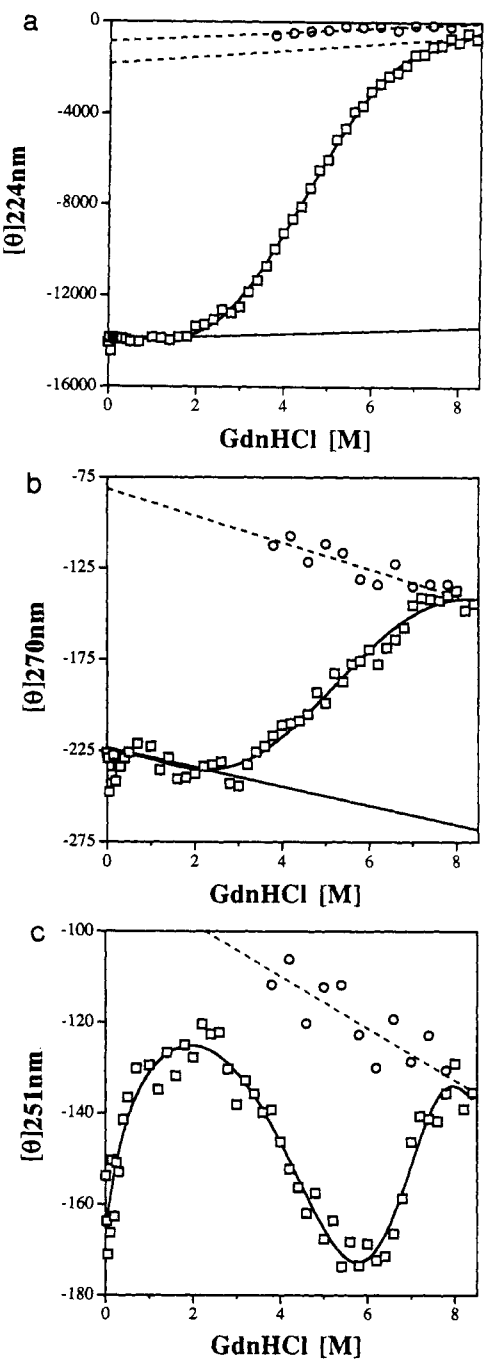


FIGURE 3: GdnHCl-induced equilibrium denaturation of lys^{B28}pro^{B29} insulin in 20 mM HEPES, pH 7.5, at 1 mg/mL. Square symbols indicate protein denaturation data; open circles represent solvent interaction with *S. aureus* V8 digested insulin mimic. Pre- and post-transitional baselines are represented by the linear solid and broken lines, respectively. The curved solid line represents the best polynomial fit to the data. Transitions detected by CD at (a) 224, (b) 270, and (c) 251 nm.

270 nm, deviation from the pre-transition baseline in the data from 0 to 1 M GdnHCl was due to a small amount of residual association that persists even in 20% ethanol (Bryant et al., 1992). This is consistent with the results obtained from the monomeric insulin analogs that show no deviation in the data from the pre-transition baseline.

Denaturant levels that approached the solubility of GdnHCl were required to completely unfold these insulins. Therefore, only a few data points (3–4) were available for determination of the post-transition baseline. To increase the reliability of this line, we utilized the enzymatically cleaved lys^{B28}pro^{B29} insulin denaturation data to reflect solvent interaction with an unfolded insulin. With this peptide mixture, any CD signal change that occurred with varying denaturant concentration was reflective of solvent interaction with the peptides and not due to conformational changes.

The apparent equilibrium unfolding constant at each denaturant concentration in the transition region was calculated by the method of Pace (1975), and they are presented in Figure 5 for human insulin at 224 and 270 nm. Plots for the other analogs or solvent conditions are not shown but resulted in similar profiles. For all the insulins investigated, the transitions detected at 224 and 270 nm are not coincident. The transitions detected at 270 nm are displaced to higher

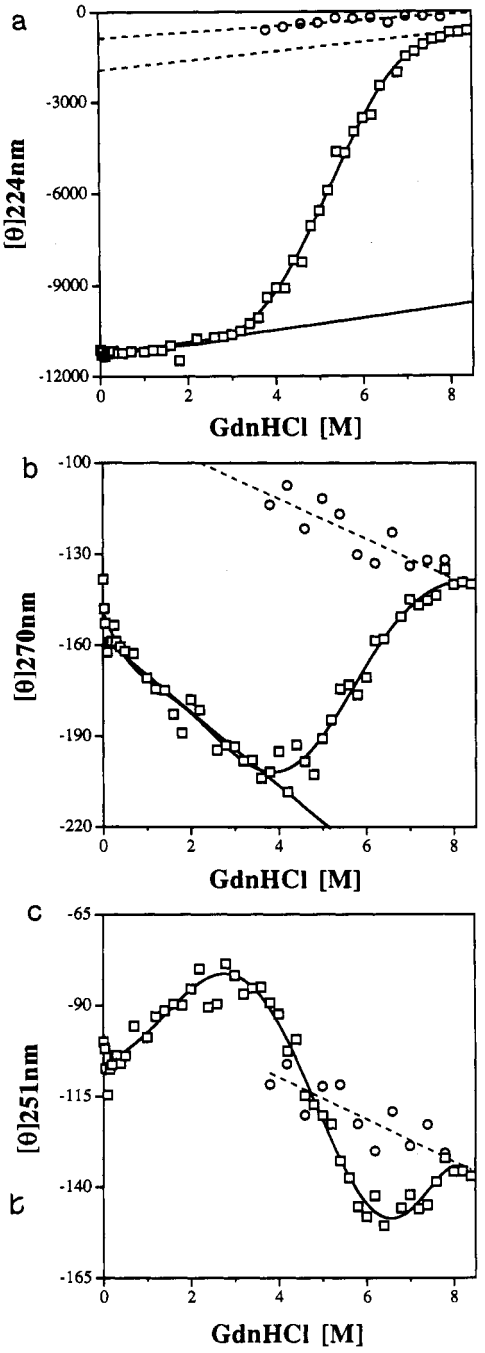


FIGURE 4: GdnHCl-induced equilibrium denaturation of asp^{B10}-des^{B28-30} insulin in 20 mM HEPES, pH 7.5, at 1 mg/mL. Square symbols indicate protein denaturation data; open circles represent solvent interaction with *S. aureus* V8 digested insulin mimic. Pre- and post-transitional baselines are represented by the linear solid and broken lines, respectively. The curved solid line represents the best polynomial fit to the data. Transitions detected by CD at (a) 224, (b) 270, and (c) 251 nm.

denaturant concentrations and/or have more shallow slopes than the transitions determined at 224 nm. Noncoincidence of the curves was obtained regardless of whether the post-transitional baselines were constructed from the raw insulin data or from the digested insulin mimic (comparison not shown), with the latter providing more reliable analysis.

In contrast to the single smooth transitions observed at 270 and 224 nm, the data at 251 nm resulted in multiple transitions for all three insulins. Three reversible, apparent transitions exist: a transition occurring between 0 and 2.5 M GdnHCl, a second transition from 2.5 to 6 M GdnHCl, and a third between 6 to 8 M GdnHCl. The linear broken lines in Figures

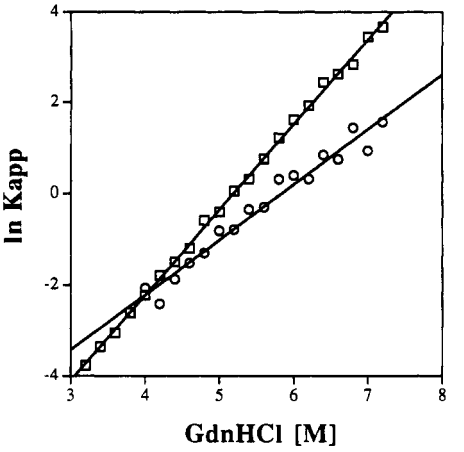


FIGURE 5: A representative comparison of denaturant versus ln K_{app} for human insulin in 20% ethanol. Square symbols represent 224-nm CD data; circles represent 270-nm CD data. Solid lines represent a linear least-square fit to the data.

1c–4c represent the 251-nm denatured baseline as determined for *S. aureus* V8 cleaved lys^{B28}pro^{B29} insulin. Due to the overlap of transitions, the apparent equilibrium constants for unfolding were not calculated.

DISCUSSION

Evidence for Intermediates. There are multiple lines of evidence reported in the literature to indicate deviation from a simple two-state equilibrium unfolding model of a protein, N ↔ U. Evidence for equilibrium unfolding intermediates is indicated in the following cases (Privalov, 1979; Ghélis & Yon, 1982; Kim & Baldwin, 1982): if analysis of the denaturation transition, as monitored by different probes of protein structure, results in noncoincidence of the curves; or, if the denaturation transition displays multiple plateaus or distinct shoulders; or, if the ΔH determined from calorimetric measurements is not equivalent to a van't Hoff determination of enthalpy.

In the present study, we have demonstrated that insulin and the two monomeric insulin analogs deviate from a two-state unfolding model on two lines of evidence: noncoincidence of denaturation transitions as monitored by multiple structural probes, and multiphasic transitions as determined by a single probe. Multiphasic and noncoincident transitions may also result if denaturation is accompanied by changes in a protein's association state. This is of practical concern since the self-association of insulin is well documented and results in spectrophotometric changes in the near- and far-UV CD (Goldman & Carpenter, 1974; Pocker & Biswas, 1980; Melberg & Johnson, 1990). Analytical equilibrium ultracentrifugation was used to characterize the noncovalent association state of insulin and the insulin analogs used in this study (Brems et al., 1992b; Bryant et al., 1992). Human insulin displays a strong concentration-dependent increase in the apparent molecular weight reflective of association to dimers and higher order aggregates. Addition of 20% (v/v) ethanol significantly disrupts this association but does not significantly perturb secondary structure nor alter receptor binding capacity (Brems et al., 1990). Specific mutation of residues involved in the dimer or hexamer interface can also disrupt association properties. Two such mutants used in this study, lys^{B28}pro^{B29} insulin and asp^{B10}des^{B28-30} insulin, are essentially monomeric at 1 mg/mL even in the absence of 20% ethanol (Brems et al., 1992b; and unpublished results on asp^{B10}des^{B28-30}). Thus, under the conditions of this study,

human insulin retains a small degree of association and the two insulin analogs are monomeric. The small residual association of human insulin in 20% ethanol is further disrupted by the addition of low concentrations of GdnHCl, as evident by the change in ellipticity at 270 nm over 0 to 1 M GdnHCl (Figure 1b). For insulin, the CD at 270 nm is dependent on both folding and self-association (Strickland, 1976; Brems et al., 1992b), with the disruption of self-association occurring at much lower denaturant concentration than disruption of secondary structure. Therefore the residual self-association of human insulin in 20% ethanol does not interfere with the observed monomeric denaturation transition.

To compare the denaturation transitions determined by CD at different wavelengths, an accurate determination of the pre- and post-transitional baselines is required. Determination of the pre-transition baselines was straightforward. However, determination of the post-transition baselines was ambiguous due to overlap with the solubility limit of GdnHCl. At these high denaturant concentrations, the signal-to-noise ratio increased due to the minimal CD signal of a random coil protein and increased absorbance from the denaturant. This contributed to a low confidence in the constructed post-transitional baselines. To more accurately assess this baseline, an unfolded mimic of insulin was developed. This unfolded mimic was produced through an enzymatic digest of lys^{B28}-pro^{B29} insulin resulting in unfolding of the structure. Since the digested insulin has essentially the same primary structure as insulin, except at the cleavage sites, the interaction of the denaturant with either random coil protein or peptide fragments is expected to be identical. We can follow the denaturation of the unfolded insulin mimic over a large GdnHCl range and use this data to construct the post-transitional baseline for intact insulin. As illustrated in Figures 1–4 by the linear broken lines, the baselines determined from the *S. aureus* V8 digested insulin mimic align very nicely with the last 3–4 data points of intact insulins' denaturation profiles and is consistent with a line that would be constructed from the insulin's raw data alone. For any of the denaturation transitions monitored by the near-UV CD, this *S. aureus* V8 digested insulin serves as an excellent source for the denatured baseline. Denaturation transitions detected by the far-UV CD (224 nm) demonstrated that the digested insulin has a slope consistent with, but slightly offset from, the post-transition baseline expected from the denatured intact insulin. The slight offset between the digested insulin and the intact insulin baselines is probably due to the fewer peptide bonds present in the digested insulin because the far-UV CD is significantly composed of contributions from peptide bonds (Johnson, 1985).

The slope of the baseline determined from the *S. aureus* V8 digest of insulin could be compromised by residual structure or association of the peptides resulting in a CD signal change with varying denaturant concentrations. CD spectrum analysis of the isolated *S. aureus* V8 digested fragments, which ranged from 9 to 30 amino acids in length (Grau, 1985), produced spectra characteristic of random coil conformation for all three peptides (data not shown). Association of this peptide mixture, derived from a monomeric insulin analog, is unlikely especially at the high denaturant concentrations (4–8 M GdnHCl) used in baseline construction. To stringently test this statement, an alternate enzymatic digest of $lys^{B28}pro^{B29}$ insulin with thermolysin was conducted to produce a smaller and more diverse composition of fragments. Smaller fragments are less likely to associate or have ordered structure due to specific sequence requirements (Wu & Yang, 1981). A limited

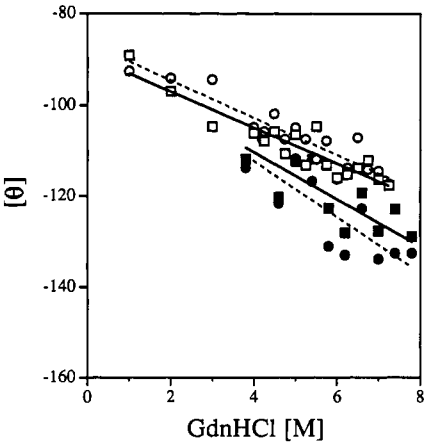
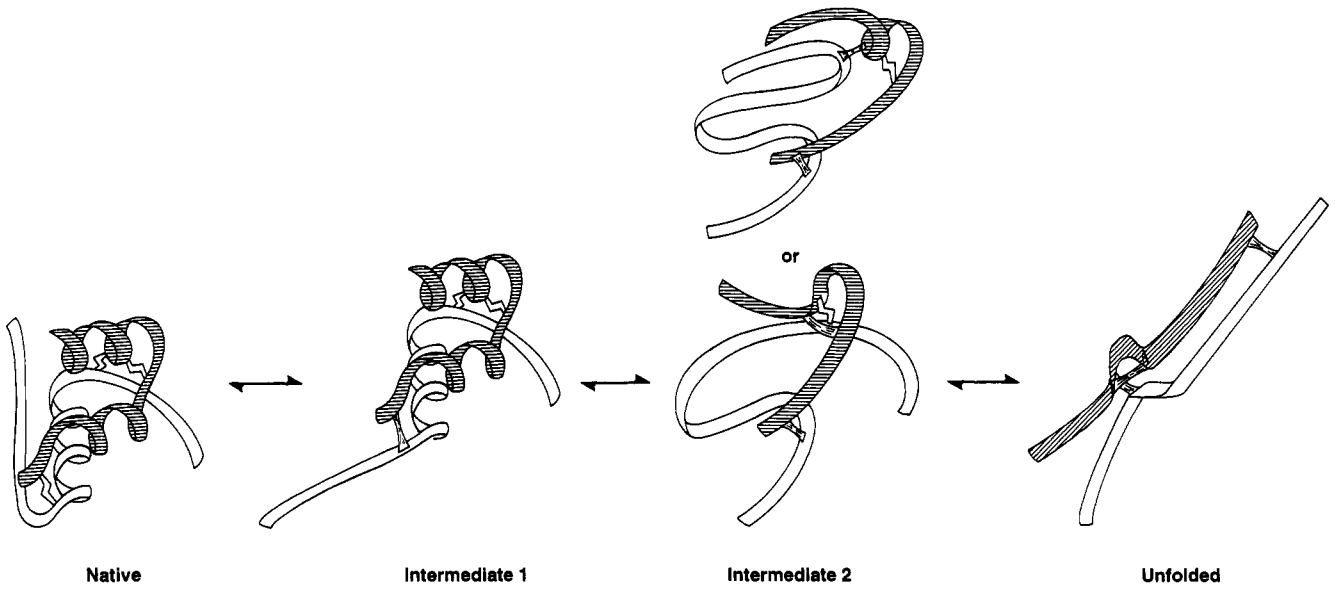


FIGURE 6: Comparison of GdnHCl solvent interaction with *S. aureus* V8 or thermolysin digested $lys^{B28}pro^{B29}$ insulin at 270 nm and 251 nm. Circles represent detection at 270 nm; squares represents detection at 251 nm; solid and broken lines represent linear regression fit to the data at 251 and 270 nm, respectively. Thermolysin data is represented by the open symbols and *S. aureus* V8 data by the filled symbols.

thermolysin digest of $lys^{B28}pro^{B29}$ insulin, which hydrolyzes on the amino side of hydrophobic residues (Matsubara et al., 1966; Beynon et al., 1989) resulted in 14 major peptide peaks, each having a relative HPLC peak area between 2–13% at 214 nm. These fragments ranged from 1 to 16 residues in length with an area weighted average of seven residues per peptide as compared to 17 for the *S. aureus* V8 digested fragments. The effect of increasing denaturant on the thermolysin digest at 270 and 251 nm by circular dichroism is compared to the *S. aureus* V8 digest in Figure 6 and resulted in near identical slopes. It is unlikely that multiple peptide digests that vary in length and composition would associate or have ordered structure to the same extent and have similar CD signals with increasing denaturant levels. A third unfolded insulin mimic was produced through disulfide bond reduction of $lys^{B28}pro^{B29}$ insulin with a 10-fold molar ratio of dithioerythritol to insulin, pH 8.0 for 1 h at room temperature. The far-UV CD spectrum was also consistent with a random coil conformation (data not shown). The GdnHCl interaction with the equal molar mixture of A and B chain insulin also resulted in a slope consistent to the other two mimics at 270 nm (data not shown). The magnitude of the signal at 270 nm was smaller than the other mimics due to disruption of the disulfide chromophore's contribution to the 270-nm signal (Beychok, 1965; Ettinger & Timasheff, 1971). All three mimics derived from a monomeric insulin analog, representing a diversity of peptides, resulted in near identical baseline slopes with increasing GdnHCl concentrations as measured by near-UV CD at 270 nm. This supports the validity of the *S. aureus* V8 digested insulin mimic as a model for solvent interactions on an unfolded insulin.

The denaturation transition for each analog was analyzed according to an assumed two-state unfolding mechanism. The apparent equilibrium constant for unfolding, $K_{app} = F_u/F_f$, where F_u = fraction unfolded and F_f = fraction folded at a denaturant concentration, was calculated by extrapolation of the pre- and post-transition baseline into the transition region (Pace et al., 1989). A representative comparison of denaturant versus ln K_{app} for human insulin detected at various wavelengths is shown in Figure 5. The CD-determined GdnHCl denaturation data at 270 and 224 nm resulted in noncoincidence for both human insulin and the two monomeric analogs. To assess the effect of the 20% ethanol that was included to limit self-association of human insulin, the denaturation of

Scheme 1



lys^{B28}pro^{B29} insulin was determined with and without the addition of ethanol. In either case, noncoincidence of the denaturation transitions as measured by different probes was observed (data not shown). Inclusion of 20% ethanol increased the Gibb's free energy of unfolding for lys^{B28}pro^{B29} insulin and may be attributed to a general solvent effect (Ettinger & Timasheff, 1971) or could be due to selective ethanol stabilization of a conformational state.

The noncoincidence of the denaturation transitions is due to the presence of at least one well-populated equilibrium intermediate that has a significantly different and detectable conformation than the native or unfolded state. In the determination of noncoincidence, two probes were selected to represent different structural features. The near-UV CD probe at 270 nm is an indicator of tertiary structure that results from electronic transition moment-to-moment coupling of a tyrosine residue with a proximal aromatic side chain; i.e., tyrosine, phenylalanine, or histidine (Strickland, 1974). The far-UV CD probe at 224 nm is representative of secondary structure, in particular the strong α-helix $n \rightarrow \pi^*$ amide signal and to a lesser extent the weaker $n \rightarrow \pi^*$ amide signal resulting from β-sheet, β-turn, and random coil structures (Johnson, 1985).

We have previously reported that the GdnHCl equilibrium denaturation of human insulin in 20% ethanol as monitored by circular dichroism at 223 and 275 nm was coincident (Brems et al., 1990). Determination of the post-transitional baseline was limited to a few data points due to the solubility limit of GdnHCl being only slightly greater than the level of denaturant required to fully unfold human insulin. In addition, the signal-to-noise ratio for these data points was large due to the small optical rotation of a randomly coiled protein and increased absorbance from the high denaturant concentrations. In the earlier report, a post-transitional baseline with a positive slope was constructed at 275 nm that was consistent with the limited data. Replicate collection of this denaturation data at 270 nm (Figure 1b) resulted in a negative slope being consistent with the data and was confirmed through the solvent interactions with an enzymatically digested insulin mimic. The discrepancy between the previous and current reports is due to the low confidence in the post-transitional baseline slope resulting from inherently variable and restricted data points in the previous work. This limitation was abolished in the present study through the use of an unfolded insulin mimic.

An alternate, independent line of evidence for equilibrium unfolding intermediates resulted from multiple transitions observed by CD at 251 nm. Although not extensively characterized, disulfide bonds have been reported to produce significant CD signals near the 250-nm region modulated by vicinal interactions, disulfide dihedral angles, and C–S–S bond angles (Casey & Martin, 1972; Kahn, 1972; Strickland, 1974; Bewley, 1977). Due to the high percentage of cystine bonds present in insulin (6%), the CD signal at 251 nm is likely dominated by disulfide contributions (Kruger et al., 1990). Three denaturation transitions were detected by CD at 251 nm (Figures 1c–4c). Changes in the 251-nm CD signal occur between 0 to 2.5, 2.5 to 6, and from 6 to 8 M. The unfolded post-transition baseline can be ascertained by the denatured insulin mimic, but other baselines cannot be determined with certainty, and therefore the related transitions cannot be separated and analyzed directly. The two apparent transitions from 0 to 2.5 and 6 to 8 M GdnHCl appear to be unique compared to the other CD-detected transitions. The transition from 0 to 2.5 M GdnHCl is not thought to be due to any residual self-association since it is present in the two monomeric insulin analogs. The middle of the three transitions detected at 251 nm occurs over a similar denaturant range as the transitions observed at 224 and 270 nm.

Proposed Structures of Intermediates. From the denaturation data obtained by various probes, we speculate on the possible structural characteristics of the unfolding intermediates. Scheme 1 proposes a possible folding pathway model that is consistent with the denaturation results. In Scheme 1, the A chain is illustrated by the cross-hatched structure with the N-terminus of the A chain and the C-terminus of the B chain located in the upper left-hand side of the native insulin structure. To account for the three distinct and noncoincident transitions at 251 nm, at least two stable intermediates are required. Alternate intermediates may be populated, but additional structural details will be required before further distinctions can be drawn. Proposed intermediate 1 retains complete native-like helical structure with local interactions around one or more disulfides being perturbed. This intermediate represents the denaturation results in which the C-terminal segment of the B chain is perturbed, resulting in a conformational alteration of the A^{20}–B^{19} disulfide bond and is depicted by the 251-nm CD transition occurring over 0–2.5 M GdnHCl (Figures 1c–4c). The structure of the C-terminal

segment of the B chain is speculated to be the most labile to denaturation based on numerous X-ray diffraction studies of different crystal forms of insulin (Derewenda et al., 1989) and NMR structural studies (Hua et al., 1991; Hua & Weiss, 1991) that have shown this segment to be highly flexible and easily displaced from its native conformation. Fluorescence-detected equilibrium denaturation of a monomeric insulin analog with a unique tryptophan at B^{28} demonstrated that the C-terminal segment of the B chain preferentially unfolds at a lower denaturant concentration than secondary structure (Bryant et al., 1992).

The denaturation transition occurring between 6 to 8 M GdnHCl in the 251-nm CD results indicates the presence of another intermediate species (intermediate 2, Scheme 1) that is well separated from the transition corresponding to intermediate 1. In the high GdnHCl concentration range of 6 to 8 M, only a small percentage of secondary structure remains, as indicated by 224-nm CD. The transition detected only by CD at 251 nm suggests there is residual localized structure around one or more of the disulfides. This second intermediate is illustrated in Scheme 1 with localized structure around either disulfide A^7–B^7 or A^6–A^{11}. The denaturation of intermediate 1 to intermediate 2 results in the major unfolding of the 2° and 3° structure as observed by CD-detected transitions at 224 and 270 nm and additional changes in the 251 nm band (Figures 1–4). The noncoincidence of the 224- and 270-nm detected transitions do not imply the presence of an additional intermediate but can be explained by the overlapping differential denaturation by the retained partial structures in intermediates 1 and 2.

Implications for Insulin Folding Intermediates. The identification and characterization of insulin equilibrium intermediates have important ramifications for the mechanism of protein folding and structure–function relationships. Kinetic folding studies will be required to determine the correct sequence of intermediates in the folding pathway of insulin. The order and placement of intermediates in Scheme 1 are based on the hypothesis that interactions that are most resistant to denaturation are the first to form in the folding pathway. For example, intermediate 2 represents a localized structure that can form in the absence of other long-range 2° and 3° interactions. Such regions of local structure are possible sites of nucleation at very early stages of folding and could be important in restricting conformational space and for guiding folding in the direction of the native state. Interest in characterizing nonrandom regions of structure in otherwise denatured states is growing rapidly (Shortle, 1993). The characterization of denatured states provides important insights into areas of structure prediction, transmembrane transport, intracellular protein turnover, and chaperone function.

The X-ray crystal structure of a completely inactive cross-linked insulin has been shown to be nearly superimposable to the structure of insulin, resulting in the conclusion that the determined three-dimensional crystal structure of insulin is not the conformation of the biologically active state (Derewenda et al., 1991). Restricted mobility at the C-terminus of the B chain, due to cross-linking with the N-terminus of the A chain, was proposed to account for complete loss of biological activity. A similar conclusion was obtained through the solution structure determination of a natural insulin mutant, insulin Los Angeles, that results in type I diabetes mellitus (Hua et al., 1993). Close structural similarities between this ser^{B24} mutant and human insulin are readily observed through NMR although significant biological activity is lost. An alternate mutation of the B^{24} residue to glycine results in retention of near-full biological activity while solution studies indicate that the C-terminus of the B chain is displaced from the native conformation exposing part of the underlying hydrophobic core. These and other studies have linked the dynamic nature of the insulin structure to the requirement for a conformational change in order to bind receptor. The proposed structure of intermediate 1, resembling the gly^{B24} insulin solution structure, may represent the biologically active state of human insulin.

Unfolding intermediates have important implications to the pharmaceutical formulation of insulin products. All commercial formulations of insulin possess a limited shelf-life of approximately 2 years due to chemical and physical degradation. We have previously established a direct correlation between chemical stability of human insulin and numerous insulin analogs to their conformational stability (Brems et al., 1992a). The population of transient, partially unfolded intermediates (intermediates 1 and 2, Scheme 1) could drastically alter the susceptibility of certain unfolded regions of insulin's structure to chemical degradation. The relative population and stability of intermediates 1 and 2 among various insulin analogs may alter the extent to which the formulation chemical stability would be dependent on the stability and relative population of one or more of these intermediates. Additional insulin mutations specifically directed toward stabilizing or destabilizing these intermediates may result in increased shelf-life of commercial insulin products.

Physical instability of insulin resulting in precipitation (fibril or gel formation) has long been a curse to the pharmaceutical industry and has limited development of insulin infusion devices (Brange, 1987). Insulin gelation/fibrillation is a nucleation-driven process with the rate-limiting step being the partial denaturation and aggregation of intermediate states (Sluzky et al., 1991; Sluzky et al., 1992). Raman spectroscopy (Yu et al., 1974) and low-angle X-ray diffraction studies (Burke & Rougvie, 1972) on insulin fibrils also indicate a perturbed native structure with increased β-sheet. Intermediate structures 1 and 2 (Scheme 1) may play a pivotal role in the gelation/fibrillation process. Understanding the role of intermediate structures in the physical stability of insulin may have a significant impact on improving the future quality and elegance of commercial formulations.

ACKNOWLEDGMENT

We thank Mr. Allen Pekar for providing analytical equilibrium ultracentrifugation data for $asp^{B10}des^{B28-30}$ insulin and Dr. Walter Shaw for biological activity measurements.

REFERENCES

Adler, A. J., Greenfield, N. J., & Fasman, G. D. (1973) *Methods Enzymol. 27*, 675–735.

Baker, E. N., Blundell, T. L., Cutfield, G. S., Cutfield, S. M., Dodson, E. J., Boelens, R., Ganadu, M. L., Verheyden, P., & Daptein, R. (1990) *Eur. J. Biochem. 191*, 147–153.

Bewley, T. A. (1977) *Biochemistry 16*, 209–215.

Beychok, S. (1965) *Proc. Natl. Acad. Sci. U.S.A. 53*, 999–1006.

Beynon, R. J., & Bond, J. S., Eds. (1989) *Proteolytic enzymes, a practical approach*, p 240, IRL Press, Oxford.

Blundell, T. L., Dodson, G., Hodgkin, D., & Mercola, D. (1972) *Adv. Protein Chem. 26*, 279–402.

Brange, J. (1987) *Galenics of Insulin*, pp 1–101, Springer Verlag, New York.

Brange, J., Dodson, G. G., & Xiao, B. (1991) *Curr. Opin. Struct. Biol. 1*, 934–940.

Brems, D. N., Brown, P. L., Heckenlaible, L. A., & Frank, B. H. (1990) *Biochemistry 29*, 9289–9293.
Brems, D. N., Brown, P. L., Bryant, C., Chance, R. E., Green, K. L., Long, H. B., Miller, A. A., Millican, R. L., Shields, J. E., & Frank, B. H. (1992a) *Protein Eng. 5*, 519–525.
Brems, D. N., Alter, L. A., Beckage, M. J., Chance, R. E., DiMarchi, R. D., Green, K. L., Long, H. B., Pekar, A. H., Shields, J. E., & Frank, B. H. (1992b) *Protein Eng. 5*, 527–533.
Bryant, C., Strohl, M., Green, K. L., Alter, L. A., Pekar, A. H., Chance, R. E., & Brems, D. N. (1992) *Biochemistry 31*, 5692–5698.
Burke, M. J., & Rougvie, M. A. (1972) *Biochemistry 11*, 2435–2439.
Casey, J. P., & Martin, R. B. (1972) *J. Am. Chem. Soc. 94*, 6141–6151.
Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Smith, G. D., Sparks, C., & Swenson, D. (1989) *Nature 338*, 594–596.
Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Bing, X., & Markussen, J. (1991) *J. Mol. Biol. 220*, 425–433.
DiMarchi, R. D., Mayer, J. P., Fan, L., Brems, D. N., Frank, B. H., Green, L. K., Hoffmann, J. A., Howey, D. C., Long, H. B., Shaw, W. N., Shields, J. E., Slieker, L. J., Su, K. S. E., Sundell, K. L., & Chance, R. E. (1992) *Peptides: Chemistry and Biology*, Proceedings of the Twelfth American Peptide Symposium (Smith, J. E., & Rivier, J. E., Eds.) pp 26–28, ESCOM, Leiden.
Ettinger, M. J., & Timasheff, S. N. (1971) *Biochemistry 10*, 831–840.
Frank, B. H., & Veros, A. J. (1968) *Biochem. Biophys. Res. Commum. 32*, 155–160.
Ghélis, C., & Yon, J. (1982) *Protein Folding*, pp 298–330, Academic Press, New York.
Goldman, J., & Carpenter, F. H. (1974) *Biochemistry 13*, 4566–4574.
Grau, U. (1985) *Diabetes 34*, 1174–1180.
Greider, M. H., Howell, S. L., & Lacy, P. E. (1969) *J. Cell. Biol. 41*, 162–166.
Hua, Q. X., & Weiss, M. A. (1991) *Biochemistry 30*, 5505–5515.
Hua, Q. X., Shoelson, S. E., Kochoyan, M., & Weiss, M. A. (1991) *Nature 354*, 238–241.
Hua, Q. X., Shoelson, S. E., Inouye, K., & Weiss, M. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 582–586.
Jeffrey, P. D., Milthorpe, B. K., & Nichol, L. W. (1976) *Biochemistry 15*, 4660–4665.
Johnson, C. W., Jr. (1985) *Methods Biochem. Anal. 31*, 62–160.
Kaarsholm, N. C., Norris, K., Jørgensen, R. J., Mikkelsen, J., Ludvigsen, S., Olsen, O. H., Sørensen, A. R., & Havelund, S. (1993) *Biochemistry 32*, 10773–10778.
Kahn, P. C. (1972) Thesis, Columbia University, New York.
Kim, P. S., & Baldwin, R. L. (1982) *Ann. Rev. Biochem. 51*, 459–489.
Kruger, P., Gilge, G., Cabuk, Y., & Wollmer, A. (1990) *Biol. Chem. Hoppe-Seyler 371*, 669–673.
Kunitz, M. (1947) *J. Gen. Physiol. 30*, 291–310.
Matsubara, H., Sasaki, R., Singer, A., & Jukes, T. H. (1966) *Arch. Biochem. Biophys. 115*, 324–331.
Melberg, S., & Johnson, C., Jr. (1990) *Proteins: Struc., Funct., Genet. 8*, 280–286.
Nozaki, Y. (1972) *Methods Enzymol. 26*, 43–50.
Pace, C. N. (1975) *CRC Crit. Rev. Biochem. 3*, 1–100.
Pace, C. N. (1989) *Methods Enzymol. 131*, 266–280.
Pocker, Y., & Biswas, S. (1980) *Biochemistry 19*, 5043–5049.
Privalov, P. L. (1979) *Adv. Protein Chem. 33*, 167–241.
Sanger, F., & Tuppy, H. (1951) *Biochem. J. 49*, 463–490.
Shortle, D. (1993) *Current Opin. Struct. Biol. 3*, 66–74.
Sluzky, V., Tamada, J. A., Klibanov, A. M., & Langer, R. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9377–9381.
Sluzky, V., Klibanov, A. M., & Langer, R. (1992) *Biotechnol. Bioeng. 40*, 895–903.
Strickland, E. H. (1974) *CRC Crit. Rev. Biochem. 2*, 113–175.
Strickland, E. H. (1976) *Biochemistry 15*, 3875–3884.
Wu, Chuen-Shang, & Yang, J. T. (1981) *Biochem. Biophys. Acta 667*, 285–293.
Yu, N., Jo, B. H., Chang, R. C. C., & Huber, J. D. (1974) *Arch. Biochem. Biophys. 160*, 614–622.