An Achilles’ Heel in an Amyloidogenic Protein and Its Repair

INSULIN FIBRILLATION AND THERAPEUTIC DESIGN

Received for publication, September 18, 2009, and in revised form, January 19, 2010. Published, JBC Papers in Press, January 27, 2010, DOI 10.1074/jbc.M109.067850

Yanwu Yang,‡ Aneta Petkova,‡ Kun Huang,‡ Bin Xu,‡ Qing-xin Hua,‡ I-Ju Ye,‡ Ying-Chi Chu,‡ Shi-Quan Hu,‡ Nelson B. Phillips,‡ Jonathan Whittaker,‡ Faramarz Ismail-Beigi,‡ Robert B. Mackin,** Panayotis G. Katsoyannis,‡ Robert Tycko,§ and Michael A. Weiss**

From the Departments of ‡Biochemistry and §Medicine, Case Western Reserve University, Cleveland, Ohio 44106, the †Laboratory of Chemical Physics, NIDDK, National Institutes of Health, Bethesda, Maryland 20892-0520, the ‡Department of Pharmacology and Biological Chemistry, Mt. Sinai School of Medicine, New York University, New York, New York 10029, and the **Department of Biomedical Sciences, Creighton University School of Medicine, Omaha, Nebraska 68178

Insulin fibrillation provides a model for a broad class of amyloidogenic diseases. Conformational distortion of the native monomer leads to aggregation-coupled misfolding. Whereas β-cells are protected from proteotoxicity by hexamer assembly, fibrillation limits the storage and use of insulin at elevated temperatures. Here, we have investigated conformational distortions of an engineered insulin monomer in relation to the structure of an insulin fibril. Anomalous 13C NMR chemical shifts and rapid 15N-detected 1H–2H amide-proton exchange were observed in one of the three classical α-helices (residues A1–A8) of the hormone, suggesting a conformational equilibrium between locally folded and unfolded A-chain segments. Whereas hexamer assembly resolves these anomalies in accordance with its protective role, solid-state 13C NMR studies suggest that the A-chain segment participates in a fibril-specific β-sheet. Accordingly, we investigated whether helicogenic substitutions in the A1–A8 segment might delay fibrillation. Simultaneous substitution of three β-branched residues (IleA2 → Leu, ValA3 → Leu, and ThrA8 → His) yielded an analog with reduced thermodynamic stability but marked resistance to fibrillation. Whereas amide-proton exchange in the A1–A8 segment remained rapid, 13C chemical shifts exhibited a more helical pattern. This analog is essentially without activity, however, as IleA2 and ValA3 define conserved receptor contacts. To obtain active analogs, substitutions were restricted to A8. These analogs exhibit high bisemil-binding affinity; representative potency in a rodent model of diabetes mellitus was similar to wild-type insulin. Although 13C chemical shifts remain anomalous, significant protection from fibrillation is retained. Together, our studies define an “Achilles’ heel” in a globular protein whose repair may enhance the stability of pharmaceutical formulations and broaden their therapeutic deployment in the developing world.

Insulin, a hormone central to the regulation of metabolism (Fig. 1), is a prototypical amyloidogenic protein. In pathological amyloid deposits, such proteins undergo non-native cross-β-assembly to form linear polymers (fibrils) (1), a hallmark of diseases of toxic misfolding (2). Fibrillation of insulin when stored above room temperature has long complicated its use in the treatment of diabetes mellitus (DM) (3). Why is insulin susceptible to fibrillation, and how may such degradation be prevented? These questions are of both basic and applied interest. In this study, we employ 13C and 15N NMR spectroscopy to identify a dynamic anomaly in an engineered insulin monomer (4). Molecular repair of an unstable polypeptide segment (its “Achilles’ heel”) provides evidence of a link to the mechanism of fibrillation and yields novel analogs that may enhance the stability, efficacy, and safety of insulin replacement therapy in the treatment of DM in the developed and developing world.

The mechanism of insulin fibrillation is outlined in Fig. 2. Of signal importance is conformational distortion of the monomer (Fig. 2, purple rhombus). Whereas native insulin (Fig. 2, triangle) may be sequestered within storage hexamers (left), non-native aggregation of distorted monomers leads to successive formation of amyloidogenic nuclei, protofibrils, and fibrils (right). The fact that susceptibility to fibrillation is conserved among vertebrate insulins (5) may be a side consequence of the mechanism of action of the hormone (5); induced fit on recep-

---

8 This work was supported, in whole or in part, by National Institutes of Health Grants DK52085 (to R. B. M.) and DK079233, DK04949, and DK054662 (to M. A. W.). This work was also supported by Grant 0530067N from the American Heart Association (to Y. Y.) and a grant from the American Diabetes Association (to N. B. P.). This paper is a contribution from the Cleveland Center for Membrane and Structural Biology. M. A. W. has equity in and consults to Thermalin Diabetes, Inc. (Cleveland, OH).

9 The on-line version of this article (available at http://www.jbc.org) contains supplemental “Methods,” “Discussion,” Figs. 51–59, Tables 51–54, and additional references.

1 To whom correspondence may be addressed. Tel: 216-368-8703; Fax: 216-368-4458; E-mail: yanwu.yang@case.edu.

2 Supported in part by the Diabetes Association of Greater Cleveland and the Dietrich Diabetes Research Institute.

3 To whom correspondence may be addressed. Tel: 216-368-5991; Fax: 216-368-3419; E-mail: Michael.Weiss@case.edu.

4 The abbreviations used are: DM, diabetes mellitus; CSI, chemical shift index; DKP-insulin, insulin analog containing three substitutions in B-chain (AspB10, LysB28, and ProB29); KP-insulin, KP-insulin contains the substitutions ProB29 to lys and lys to pro; HSCC, heteronuclear single-quantum coherence; IGFR, insulin-like growth factor receptor; IR, insulin receptor; MAS, magic-angle spinning; NOE, nuclear Overhauser enhancements; NOEY, NOE spectroscopy; r.m.s.d., root-mean-square deviation; rp-HPLC, reverse-phase HPLC; ThT, thioflavin T; TOCSY, total correlation spectroscopy; SSP, secondary structural propensities.

5 A striking exception to the general susceptibility of vertebrate insulins to fibrillation is provided by synthetic single-chain insulin analogs (“mini-proinsulins”), in which the C terminus of the B-chain is connected to the N terminus of the A-chain (1, 7). Such peptide tethers (0–3 residues) are compatible with native-like structure (75) but markedly impair receptor binding (76, 77). Such tethers may prevent conformational distortion of both the N-terminal A-chain α-helix and the C-terminal B-chain β-strand.
Amyloid-related Achilles' Heel in Insulin

FIGURE 1. Ribbon model of human insulin. Conformation of an insulin monomer as observed in a crystallographic T state protomer (2-Zn molecule 1; Protein Data Bank accession code 4INS). The A-chain is shown in red (residues A1–A8) and light gray (A9–A21); the B-chain is dark gray. Analog design focuses on residues IleA2, ValA3, and ThrA8 (asterisk); active analogs have A8 substitutions (magenta); NMR studies employed an engineered insulin monomer (DKP-insulin) with substitutions of wild-type residues HisB10, ProB28, and LysB29 (green). Also shown are cystines A7–B7 (largely obscured at top), A6–A11 (middle), and A20–B19 (bottom); sulfur atoms are depicted as gold spheres.

FIGURE 2. Mechanism of insulin fibrillation. Proposed pathway of insulin fibrillation via partial unfolding of monomer (9). The native state is protected by classical self-assembly (far left). Disassembly leads to equilibrium between native and partially folded monomers (open triangle and purple rhombus, respectively). This partial fold may unfold completely as an off-pathway event (open circle) or aggregate to form a nucleus en route to a protofibril (far right). This figure was adapted from Fink and co-workers (9).

Our study has three parts. First, to map sites of conformational distortability in an insulin monomer, we sought to exploit the $^{13}$C NMR chemical shift index (CSI) (10, 11) as a probe of ensemble-averaged protein dynamics. Based on associations between secondary structure and trends among $^1$Hα, $^{13}$Cα, $^{13}$Cβ, and $^{13}$CO chemical shifts, the CSI provides a rapid and sensitive method to identify α-helices, β-stands, and disordered segments in a protein without analysis of nuclear Overhauser effects (NOEs) (10, 11). Application to an insulin monomer (12) uncovered a discrepancy between CSI- and NOE-based models. These results suggested that, despite maintenance of helix-related NOEs, the N-terminal segment of the A-chain (red in Fig. 1) undergoes conformational fluctuations between helical and nonhelical conformations in accordance with its rapid $^1$H–$^2$H amide proton exchange in D$_2$O (13). Second, to test whether this or other segments of insulin participate in a fibril-specific β-sheet, we employed $^{13}$C-based solid-state NMR to measure $^{13}$Cα, $^{13}$Cβ, and $^{13}$CO chemical shifts at selected sites in a powder sample (14). The probes suggested that on fibrillation the fluctuating A1–A8 α-helix undergoes an α → β transition as part of a global change in protein conformation. The solid-state NMR results are in accordance with prior evidence of an overall α → β transition obtained by infra-red and Raman spectroscopy (7, 15–17).

In the third and final part of this study, we sought to repair the A1–A8 segment as a test of its relationship to fibrillation and as a potential route to novel insulin analogs of clinical interest. Because this segment contains unfavorable β-branched residues at three positions (IleA2, ValA3, and ThrA8), an analog was prepared containing only non-β-branched residues at these sites to augment α-helical propensity (18–20). Together, the substitutions (IleA2 → Leu, ValA3 → Leu, and ThrA8 → His) impair thermodynamic stability (presumably due to perturbed tertiary packing by LeuA2 and LeuA3) and yet extend the lag time prior to fibrillation by 5-fold. Such extension is associated with a more helical pattern of $^{13}$Cα chemical shifts in the A1–A8 segment. Although these findings suggest that the intrinsic helical propensity of fluctuating α-helix can indeed modulate amyloidogenicity, the affinity of the analog for the insulin receptor (IR) is reduced by 100-fold, precluding its clinical utility. To obtain potential therapeutic analogs, substitutions were therefore restricted to the solvent-exposed position A8 at which site a variety of non-β-
Amyloid-related Achilles’ Heel in Insulin

branched substitutions are well tolerated (21). Models were provided by rapid-acting insulin analogs in current clinical use (22–25). We have found that the A8 substitutions extend the lag time by between 2- and 3-fold. Although such protection is less marked than that of the above analog, the A8-modified insulin analogs exhibit high receptor-binding affinities. Such an analog exhibits native hypoglycemic potency in a rodent model of type 1 DM.

Our study is of dual basic and translational significance. From a biophysical perspective, it is remarkable that the anomalous $^{13}$C-based CSI signature of a globular protein correlates with both classical patterns of structural variability among crystal forms (26, 27) and its propensity for fibrillation. Furthermore, application of $^{13}$C NMR methods combined with protein engineering implies that a relationship exists between the intrinsic $\alpha$-helical propensity of a peptide segment and fibrillation lag time. From a therapeutic perspective, mutagenesis of a fluctuating $\alpha$-helix may augment the stability of rapid-acting insulin formulations, a potential benefit of particular relevance to continuous pump infusion (28). Because fibrillation-associated occlusion of the delivery system can cause metabolic decompensation leading to diabetic ketoacidosis (29), A8-modified rapid-acting insulin analogs may enhance the efficacy and safety of such devices. Whereas pump-based therapy is a technology of affluent societies, in the developing world, thermal fibrillation poses a barrier to the storage and use of insulin for treatment of DM. Together, these results suggest that joint $^{13}$C chemical shift mapping by solution- and solid-state $^{13}$C NMR may provide a general approach to detect sites of conformational distortibility in amyloidogenic proteins as a foundation for their rational repair and therapeutic application.

MATERIALS AND METHODS

**Bacterial Expression of DKP-Proinsulin**—DKP-human proinsulin, containing substitutions His$^{B10}$ → Asp, Pro$^{B28}$ → Lys, and Lys$^{B29}$ → Pro, was expressed in *Escherichia coli* and purified as described previously (30). The sites of substitution in the B-chain are illustrated in Fig. 1 (green). The protein was isotopically labeled by fermentation in M9 minimal medium containing $[^{13}$C]$\alpha$-glucose as the sole carbon source and $[^{15}$N]$\alpha$-ammonium chloride as the sole nitrogen source.

**Proteolysis of DKP-Proinsulin**—Conversion of DKP-proinsulin to a two-chain analog was accomplished by tryptic digestion. The isotopically labeled proinsulin (1 mg/ml) was dissolved in 0.01 M sodium chloride as the sole nitrogen source.

To 70% over 60 min; peaks were pooled and identified by mass spectrometry. Whereas proinsulin yielded a symmetrical HPLC peak with retention time of 53 min, on trypsin digestion this fraction progressively disappeared, and two new peaks appeared with elution times of 32 and 45 min; each exhibited a small shoulder (supplemental Fig. S1). Mass spectrometry indicated that the C-peptide and related fragments eluted at 32 min, whereas a two-chain insulin analog eluted at 45 min (supplemental Fig. S2). This analog (52 residues) retains Arg$^{B31}$ and so differs from DKP-insulin by one residue. NMR analysis indicated that the additional residue is disordered.

**Chemical Synthesis of Insulin Analogs**—Variant B-chains containing substitutions Asp$^{B28}$ or Lys$^{B28}$,Pro$^{B29}$ were obtained by oxidative sulfitolysis of insulin aspart (NovoLog®; Novo-Nordisk A/S, Bagsværd Denmark) or insulin lispro (Humalog®; Lilly) as described previously (30). Unlabeled or isotopically labeled DKP B-chains and the labeled wild-type A-chain were prepared by N-tert-butoxycarbonyl-based manual peptide synthesis as described previously (5); uniformly ($^{13}$C,$^{15}$N)-enriched N-tert-butoxycarbonyl amino acids (Ala, Gly, Ile, Leu, and Val) were obtained from Cambridge Isotope Laboratories, Inc. (Woburn, MA). Variant A-chains containing A8 substitutions were prepared by automated Fmoc (9-fluorenylmethoxycarbonyl)-based synthesis and derivatized by S-sulfonation. Insulin chain combination was effected as described previously (5). Purification of insulin analogs was accomplished by cation-exchange chromatography with carboxymethyl cellulose (CM-52, Whatman) (24 × 1.2 cm) using acetate buffer (pH 3.3) and a NaCl gradient. Final purification was by rp-HPLC using a C18 column (25 × 1 cm) with a 0.1% trifluoroacetic acid/acetonitrile solvent system. Combination of 20 mg of A-chain and 10 mg of the variant B-chain typically resulted in a final yield of 2 mg similar to that of wild-type chain combination. The purified analogs were observed to be a single component on analytical rp-HPLC (C18 column, 25 × 0.46 cm) using an acetonitrile elution gradient in 0.1% trifluoroacetic acid. Matrix-assisted laser desorption-ionization mass spectra of the products were in accordance with predicted values.

**Solution NMR Spectroscopy**—Samples were prepared in nitrogen-purged H$_2$O solutions (7% D$_2$O) in 50 mM sodium phosphate (pH 7.1 or 7.8) and 50 mM NaCl in a 300-μL Shigemi NMR tube; the protein concentration was ~0.2 mM. Spectra were obtained between pH 7 and 8 due to progressive aggregation as the pH is reduced below 7, leading to macroscopic precipitation below pH 5.5. Restriction of NMR conditions to pH values >7 is associated with base-catalyzed amide proton exchange. Reverse-phase HPLC excluded detectable disulfide rearrangement during the course of data acquisition. NMR spectra were acquired at 25°C and 700 MHz using a Bruker AVANCE spectrometer equipped with triple-resonance cryo-probe. Methods for heteronuclear resonance assignment and analysis of amide proton exchange are provided as supplemental material.

**NMR Resonance Assignments**—Homonuclear resonance assignment of [Leu$^{A2}$,Leu$^{A3}$,His$^{A8}$]DKP-insulin was obtained by standard two-dimensional methods as described previously (31). Heteronuclear main-chain ($^1$H, $^{13}$C, and $^{15}$N) NMR resonance assignments were obtained by combined use of through-
bond correlation spectra (HNCACB and CBCA(CO)NH). $^{13}$Cα and $^{13}$Cβ chemical shifts are provided in supplemental Table S2. In HNCACB spectra, amide resonances are correlated with intra-residue and preceding inter-residue $^{13}$Cα and $^{13}$Cβ resonances, whereas in CBCA(CO)NH spectra, a subset of connectivities is established only to the inter-residue $^{13}$Cα and $^{13}$Cβ resonances. Such correlations are illustrated in supplemental Fig. S3. Connectivities are broken at β-turn residues B7–B9 and B21 due to rapid amide solvent exchange.  

Side-chain spin systems were assigned by analysis of a series of isotope-directed TOCSY spectra (C(CO)NH-TOCSY, H(CCO)NH-TOCSY, and HCCH-TOCSY). Spin system classification was in each case consistent with main-chain assignments. Gaps or ambiguities in analysis of main-chain (through bond) correlation spectra were overcome by analysis of (through-space) NOEs as follows: three-dimensional $^{15}$N-$^{13}$C NOESY and $^{13}$C-$^{13}$C NOESY spectra as described in supplemental “Methods” (see supplemental Fig. S4).

Secondary Structural Propensities—Secondary structural propensities (SSP) were calculated using $^{13}$Cα and $^{13}$Cβ chemical shifts using the program SSP (32). For display, SSP scores were color-coded on a ribbon model of insulin using the program GRASP (33).

Solid-state NMR Spectroscopy—Two-dimensional solid-state $^{13}$C-$^{13}$C NMR spectra of DKP-insulin fibrils (with uniform ($^{15}$N,$^{13}$C) labeling at specific sites (sample A, Val$^{A3}$, Ile$^{A10}$, and Leu$^{A16}$; sample B, Val$^{B2}$, Gly$^{B8}$, and Ala$^{B14}$ in the B-chain) were obtained at a $^{13}$C NMR frequency of 100.83 MHz and room temperature, using a Varian Infinity spectrometer and a Varian 3.2-mm magic angle spinning (MAS) probe. The finite pulse radio frequency-driven recoupling technique (34–36) was used for $^{13}$C-$^{13}$C polarization transfer, with 23.3–23.6-kHz MAS frequencies, 12.72–12.88 $\mu$s $^{13}$C π pulses, and a 1.356–1.373-ms mixing period. 512 complex points were acquired in the $t_2$ dimension, with a $t_1$ increment of 18.88 $\mu$s, a 1.5-s recycle delay, and 160–320 scans per free-induction decay. 1H decoupling fields in the $t_1$ and $t_2$ dimensions were 106 kHz, with two-pulse phase modulation (37). Continuous-wave decoupling at 100 kHz without modulation was applied in the mixing period. Data were processed with NMRPipe software (38). $^{13}$C chemical shifts in solid-state NMR spectra were defined relative to tetramethylsilane, based on an external adamantane reference. Analysis of site-specific secondary structure based on $^{13}$C chemical shifts of uniformly $^{15}$N-$^{13}$C-labeled residues in the fibrillar samples followed the approach used in studies of β-amyloid fibrils (14, 34, 39), where β-strand and non-β-strand segments were identified from secondary shifts relative to tabulated random-coil values (40).

Analysis of Crystal Structures—Thirteen T state crystallographic protomers (obtained from Protein Data Bank entries 4INS, 1APH, 1EVE, 1MPP, 3MTH, 1TRZ, 1TYL, 1TYM, 2INS, 1ZN, 2TCI, 1G7A, and 1BPH) were employed for analysis of conformational variability. Atoms omitted in deposited coordinates were built using InsightII (Molecular Simulations, Inc.). Structures were aligned based on the main-chain atoms of residues A1–A21 and B9–B28; segmental root-mean-square deviations (r.m.s.d.) were calculated as described in supplemental Table S1 and displayed using GRASP (33).

Preparation of Fibrils—Fibrillation was monitored by thioflavin T (ThT) fluorescence on gentle agitation in phosphate-buffered saline (pH 7.4) at 37 °C. Samples were prepared immediately prior to fibrillation. Proteins were made 60 $\mu$M in degassed buffer with 0.01% sodium azide at pH 7.4 and incubated in pre-sterilized glass vials with airtight sealed caps. Vials were rocked at 60 rpm. Aliquots were withdrawn at successive times and added to a ThT solution for fluorescence assay. Samples for solid-state NMR studies were prepared from DKP-insulin as described previously (17) by gentle agitation at 37 °C in 250 mM sodium chloride and 10 mM sodium phosphate (pH 7.4) in a solution containing 60% ethanol (added to expedite fibrillation). In each case the presence of linear fibrils was verified by transmission electron microscopy as described previously (17).

Fluorescence Spectroscopy—ThT was made 1 mM in double-distilled water and stored at 4 °C in the dark. To monitor fibrillation, 5-$\mu$l aliquots obtained at the indicated time points were mixed with 3 ml of ThT assay buffer (5 $\mu$l ThT in 50 mM Tris-HCl (pH 7.5) and 100 mM NaCl). Fluorescence measurements were performed using an Aviv spectrofluorometer in 1-cm quartz cuvettes. Emission spectra were collected from 470 to 500 nm following excitation at 450 nm. The fibrillation lag time is defined as the time required to observe 2-fold enhancement of ThT emission.

Circular Dichroism—Far-UV CD was employed to monitor denaturant-induced unfolding as a probe of thermodynamic stability. Spectra were recorded using an Aviv spectropolarimeter (Aviv Biomedical, Inc., Lakewood, NJ) in 10 mM potassium phosphate and 50 mM KCl at pH 7.4. The protein concentration was ∼25–50 $\mu$M. Guanidine denaturation was monitored by ellipticity at 222 nm and fitted by nonlinear least square regression to a two-state model to extract apparent free energies ($\Delta G_f$) (7). Data were acquired at 25 °C.

Receptor Binding Assay—Studies employed a FLAG epitope-tagged holoreceptor (either IR isoform B or the homologous type 1 insulin-like growth factor receptor (IGFR)) transiently expressed in human 293 PEAK cells (American Tissue Culture Collection CRl-2828) (41, 42). The protocol is illustrated in supplemental Fig. S5. Receptors were partially purified by wheat germ agglutinin chromatography and further fractionated on to polyethylene 96-well plates (Nunc Maxisorb) coated with an anti-FLAG monoclonal antibody (FLAG M2 immunoglobulin G; Sigma). Plates were incubated overnight at 4 °C with the immunoglobulin (100 $\mu$l/well of 40 $\mu$g/ml in phosphate-buffered saline). This protocol circumvents confounding binding of insulin or insulin analogs to endogenous IGFR; 293 PEAK cells do not express detectable endogenous IR. Additional experimental details are provided as supplemental material. Relative activity is defined as the ratio of specific dissociation constants as determined by competitive displacement of bound $^{125}$I-Tyr$^{14}$ human insulin (in the case of IR) or $^{125}$I-Tyr$^{31}$ human IGF-I (in the case of IGFR). In all assays the percentage of tracer bound in the absence of competing ligand was <15% to avoid ligand-depletion artifacts (41). To obtain high

$^7$ In $^{1}$H-$^{15}$N HSQC spectra, the cross-peak of residue B7 is anomalously broad at both pH 7.8 and 7.1; cross-peaks of B8 and B21 are observable only at the lower pH values, whereas B9 is absent under either condition.
Amyloid-related Achilles’ Heel in Insulin

FIGURE 3. Chemical shift index of DKP-insulin. Histograms show 13C NMR secondary shifts by residue in B-chain (left) and A-chain (right); sequences are provided at top. Upper histogram refers to Cα shifts, and lower histogram refers to Cβ shifts. Positive (negative) values indicate that the observed chemical shift is downfield (upfield) of expected random-coil frequency (10, 11). Canonical secondary structure of insulin is indicated at bottom: helix (spiral) or β-strand (arrow). CSI values in the A1–A8 segment (asterisk) are attenuated relative to α-helices 1 and 3. The sequence of insulin is shown with B-chain at left and A-chain at right to correspond to the N- to C-terminal order of B and A domains in proinsulin.

affinity analog-receptor dissociation constants, binding data were analyzed by nonlinear regression analysis using the model of Wang (43), which employs an exact mathematical expression to describe competitive binding of two different ligands to a protein molecule. Control studies using untransfected PEAK cell lysates demonstrated that this microtiter plate antibody-capture assay has negligible endogenous background.

Rodent Assay—Male Lewis rats (mean body mass ~300 g) were rendered diabetic by treatment with streptozotocin as described previously (44). To test the in vivo potency of a representative insulin analog in relation to wild-type human insulin, protein solutions containing wild-type human insulin, a representative A8-modified insulin analog ([His4,Asp28]human insulin), or buffer alone (protein-free sterile diluent obtained from Lilly; composed of 16 mg of glycercin, 1.6 mg of meta-cresol, 0.65 mg of phenol, and 3.8 mg of sodium phosphate (pH 7.4 per ml)) were injected subcutaneously, and the resulting changes in blood glucose were monitored by serial measurements using a clinical glucometer (Hypoguard Advance Micro-Draw meter). To ensure uniformity of formulation, insulin analogs were each re-purified by rp-HPLC, dried to powder, dissolved in diluent at the same maximum protein concentration (300 µg/ml), and re-quantified by analytical C4 rp-HPLC; dilutions were made using the above buffer. Rats were injected subcutaneously at time t = 0 with 20 µg of insulin in 100 µl of buffer per 300-g rat. This dose corresponds to ~67 µg/kg body weight, which corresponds in international units to 2 IU/kg body weight. Dose-response studies of wild-type insulin indicated that at this dose a near-maximal rate of glucose disposal during the 1st h following injection was achieved. Blood was obtained from clamped tip of the tail at time 0 and every 10 min up to 90 min. The efficacy of insulin action to reduce blood glucose concentration was calculated using the change in concentration over time (using least mean squares and initial region of linear fall) divided by the concentration of insulin injected.

RESULTS

Our results are presented in three parts. We first describe heteronuclear NMR studies to identify an anomaly in the 13C chemical shifts of an engineered insulin monomer. We next extend these findings by solid-state NMR to insoluble insulin fibrils containing selective isotopic labels in the A- or B-chains. Finally, based on these results, we describe the design of fibrillation-resistant insulin analogs. Together, these findings provide evidence for a relationship between the intrinsic α-helical propensity of a peptide segment and the amyloidogenicity of the intact protein.

Solution NMR Chemical Shift Mapping

Chemical Shift Index and Core-related NOEs—Secondary chemical shifts of Cα- and Cβ-carbons (the differences between observed and random-coil values) are sensitive to secondary structure (10, 11). CSI values are summarized by the histogram in Fig. 3 in relation to canonical elements of secondary structure (bottom). Positive (or negative) values correspond to resonances that are downfield (or upfield) of expected random-coil values. Whereas CSI values are in accordance with maintenance of α-helices 1 and 3 (residues B9–B19 and A13–A19, respectively) and the C-terminal B-chain β-strand (B24–B28), the A1–A8 segment exhibits marked attenuation of CSI values (Fig. 3, arrow and asterisks).

Heteronuclear resonance assignment also enabled application of an algorithm to convert chemical shift data to a single residue-specific SSP score (32). This score predicts at each residue the percentage of stably folded conformers with canonical secondary structure (α-helical or β-strand) in an ensemble. Limiting SSP scores of 1 or ~1 imply formation of a fully formed α-helix or β-strand, respectively; intermediate scores (~0.5) imply 50% local occupancy; no information is provided regarding the nature of alternative noncanonical conformations. The SSP score of each residue in DKP-insulin (supplemental Table S2) is mapped onto the structure of an insulin monomer in Fig. 4A. Ribbon color in Fig. 4A ranges from violet (highest β-strand character) to white (no predominant secondary structure) to blue (highest α-helical character). Mean segment-specific SSP scores are ~20% (N-terminal A-chain segment), 75% (C-terminal A-chain α-helix), and 100% (central B-chain α-helix). Strikingly, these segmental scores correlate with structural variability among crystallographic protomers as probed by main-chain atomic r.m.s.d. (mapped onto the structure in Fig. 4B; see “Discussion”). An alignment of crystallographic protomers is shown in Fig. 4C. With respect to strands, the N-ter-
Although its SSP score implies that the N-terminal A-chain segment undergoes significant main-chain conformational fluctuations, the core structure of insulin is not disrupted as indicated by maintenance of strong NOEs from the side chain of IleA2 to the core side chain of TyrA19. Native-like contacts are likewise maintained between the side chains of ValA3 and TyrB26 within an inter-chain crevice. The N-terminal A1–A8 segment also retains a subset of canonical α-helical NOEs. Such NOEs, exemplified by an (i, i + 3) contact between A2-Hα and A5-Hβ, are weaker than corresponding NOEs in other helices (31). Such attenuation presumably reflects nonhelical local conformations; the signal-to-noise ratio is further reduced by line broadening in this segment, itself a consequence of conformational exchange (45). We imagine that core side chains in the B-chain and C-terminal A-chain α-helix provide a nonpolar template for folding of the molten N-terminal A-chain segment as an amphipathic helix (see “Discussion”).

**Amide Proton Exchange**—Base-catalyzed amide exchange in DKP-insulin is too rapid to investigate by the standard D2O exchange method. We therefore employed the heteronuclear solvent presaturation 1H-15N HSQC method at 25 °C in H2O. Base-line HSQC spectra of DKP-insulin at pH 7.1 or 7.8 (Fig. 5A) exhibit favorable chemical shift dispersion in both 1H- and 15N dimensions; the inset shows the weak cross-peak of CysA11 (green dashed circle), observable only at the lower pH following prolonged acquisition (14 h). Although spectra acquired at pH 7.8 (Fig. 5A, red) or pH 7.1 (black) exhibit only slight changes in cross-peak positions, six additional cross-peaks (residues A11, B8, B21, B26, B27, and B28) are observed at the lower pH. Although for residues B26–B28 (lying at the edge of a β-strand), this pH effect is presumably due to less efficient base-catalyzed exchange, and for residues A11 and B8 (participating in chain reversals) the change in pH may also be affecting the extent of conformational broadening as these amide resonances exhibit unusually large 1H line widths.

---

**Figure 4. Structure of insulin and relationship between conformational variability and chemical shift mapping.** A, color-coded ribbon model of insulin T state indicating SSP score derived from 13C-secondary chemical shift (32). B, ribbon model of insulin T state color-coded according to main-chain r.m.s.d. values among collection of T state protomers aligned according to the main-chain atoms of residues B9–B24 and A1–A21. Red indicates greatest variability and blue indicates least variability. Quantitative scales in A and B are shown above each structure. C, superposition of crystallographic protomers (15 T states and 15 R (or Rf) states). The structures were aligned according to the main-chain atoms of residues B9–B24 and A12–A21. The A1–A8 α-helix in each protomer is shown in red; the variable secondary structure of the N-terminal segment of the B-chain is shown in green (extended in T state; right) or blue (extended α-helix in R state; left). D, ribbon model of wild-type insulin dimer in crystallographic T and R states. The color code is as in C. Structures in C were obtained from the following entries in the Protein Data Bank: T states: 4INS, 1APH, 1CPH, 1DPH, 1TRZ, 1TYL, 1TYM, 2INS, 1ZNI, 1LPH, 1G7A, 1MSO; R states: 1EV6, 1DNJ, 1TRZ, 1ZNI, 1LPH.

---

8 The crystallographic T → R transition, an allosteric reorganization of insulin hexamers (64, 80), is remarkable for a change in the secondary structure of the B1–B8 segment from extended (T state) to α-helix (R state). Segmental reorganization is coupled to changes in the conformation of GlyB8, core packing, and handedness of cystine A7–B7. The latter sulfur atoms are exposed in the T state but buried in a nonpolar crevice in the R state. The superscript Rf refers to N-terminal fraying of the B1–B19 α-helix. The T → R transition has been analyzed as a model for the long range transmission of conformational change (26).

9 Raising the temperature from 25 to 40 °C, which would be expected to accelerate amide proton exchange, leads to partial narrowing of A11 and B8 1HN resonances and their detectability in 1H-15N HSQC spectra.
Additional $^1$H-$^15$N HSQC spectra were acquired following 2 s of presaturation at either the solvent frequency or a control downfield of protein resonances (Fig. 5B). Disproportionate attenuation due to solvent presaturation is observed at multiple sites. These include sites both predicted and not predicted by the structure of DKP-insulin to be engaged in hydrogen bonding. Of its three α-helices, only the N-terminal A-chain segment lacks observable cross-peaks following solvent presaturation. Comparison of this limiting observation with predicted intrinsic rates of solvent exchange indicates that protection factors in the A1–A8 segment are very low (for example, <2.8 at position A4 relative to a predicted global protection factor base line of 880 ± 75; see supplemental material). Because the global stability of DKP-insulin (as inferred from guanidine denaturation experiments) (46) is 4.4 kcal/mol at 25 °C, amide proton exchange must be mediated by local or subglobal conformational fluctuations (13). The majority of residues in other α-helices (but not all) exhibit reduced cross-peak intensities (relative to the control spectrum), implying intermediate protection factors generally less than 200 and thus exchange via a subglobal mechanism. These observations imply that a large fraction of hydrogen bonds in the native state undergo transient breakage leading to amide proton exchange. Absence of protected amide protons in the A1–A8 segment and a broad range of lifetimes in the other helical segments have previously been observed at pH 1.9 in 20% deuterio-acetic and 80% D$_2$O (47).

**Solid-state NMR Chemical Shift Mapping**

Uniformly enriched ($^{13}$C, $^{15}$N) amino acids were incorporated into DKP-insulin at three positions in the A-chain (Val$^A_3$, Ile$^A_{10}$, and Leu$^A_{16}$; sample A) or B-chain (Val$^B_2$, Gly$^B_8$, and Ala$^B_{14}$; sample B). These residues were chosen for their distinctive spin systems, interesting locations, and economy of labeling. In native insulin, these sites lie either in an extended and flexible N-terminal segment (Val$^B_2$), noncanonical turns (Gly$^B_8$ and Ile$^{A_{10}}$), or α-helix (Val$^{A_{13}}$, Leu$^{A_{16}}$, and Ala$^{B_{14}}$). Whereas previous optical spectroscopic studies of insulin fibrils demonstrated a predominance of β-sheet (7, 15–17), such probes could not resolve residue-specific conformations. By contrast, two-dimensional solid-state $^{13}$C-$^{13}$C NMR MAS corre-

![Figure 5](image-url)
lation spectroscopy (14, 34, 39) can provide evidence for β-strand-associated isotropic chemical shifts relative to random-coil values (40).

Solution 13C NMR spectra are shown at left in Fig. 6; two-dimensional 13C-13C cross-polarization MAS spectra are shown at right. A-chain spin systems in sample A are shown in Fig. 6A (top), and B-chain spin systems in sample B are shown in Fig. 6B (bottom). Resonance assignment in the native state is in accordance with the general heteronuclear analysis of DKP-insulin (part I) and was verified by analysis of 1H-13C HSQC and 1H-13C HMQC-TOCSY spectra of samples A and B. Resonance assignment in the fibrillar powder was obtained by inspection of the two-dimensional 13C-13C MAS chemical shift correlation spectra (supplemental Table S3). In the fibril, residues A3, A10, A16, B2, and B8 exhibit β-sheet-associated Cα and Cβ chemical shift signatures. The chemical shifts of AlaB14
by contrast retain α-helical positions. Nonuniformity of A- or B-chain β-strands (or both) would be in accordance with the incongruent spacing between inter-chain disulfide bridges A7–B7 and A20–B19; CysA7 and CysA20 are separated by an even number of residues, and CysB7 and CysB19 are separated by an odd number. The 13C NMR solid-state MAS line widths at the labeled sites are in general similar to those observed in well ordered peptide fibrils (14, 34) with additional broadening observed at positions B2 and A16 (supplemental Fig. S6).

by contrast retain α-helical positions. Nonuniformity of A- or B-chain β-strands (or both) would be in accordance with the incongruent spacing between inter-chain disulfide bridges A7–B7 and A20–B19; CysA7 and CysA20 are separated by an even number of residues, and CysB7 and CysB19 are separated by an odd number. The 13C NMR solid-state MAS line widths at the labeled sites are in general similar to those observed in well ordered peptide fibrils (14, 34) with additional broadening observed at positions B2 and A16 (supplemental Fig. S6).

Design of Fibrillation-resistant Analogs

The sequence of the N-terminal segment of the A-chain (GIVEQCCT) contains three β-branched residues (underlined; magenta in Fig. 1) and so exhibits low intrinsic α-helical propensity. The side chain of IleA2 packs near the side chains of LeuA16, TyrA19, LeuB11, and LeuB15 in the hydrophobic core; the side chain of ValA3 projects into an inter-chain crevice near ValB12 and TyrB26, and the side chain of ThrA8 projects into solvent (Fig. 1, asterisk). To test whether simultaneous substitution of these residues by non-β-branched side chains would delay fibrillation by a mechanism unrelated to global thermodynamic stability, an analog of DKP-insulin was prepared with substitutions IleA2 → Leu, ValA3 → Leu, and ThrA8 → His. The analog is less stable than DKP-insulin (ΔΔGμ = 0.5 ± 0.2 kcal/mol; supplemental Fig. S7) and yet exhibits a lag time prior to fibrillation 5-fold longer than that of DKP-insulin as follows: 65 ± 9 versus 12.4 ± 2.5 days (DKP-insulin exhibits a base-line lag time longer than that of wild-type insulin (below) due to the stabilizing AspB10 substitution). Fibrillation was provoked by gentle agitation at 37 °C under neutral pH conditions as a model for the physical degradation of pharmaceutical formulations in a pump or challenged region of the developing world. The pattern of 13C chemical shifts, obtained at natural abundance by analysis of its two-dimensional 1H-13C HSQC spectrum in D2O (Fig. 7A), exhibits a more α-helical pattern (supplemental Table S4). Despite the reduced thermodynamic stability of the analog, NOESY spectra exhibit a native-like pattern of inter-residue contacts, including the side chains of LeuA2 and LeuA3 (red and green, respectively). Assignments are as shown. The vertical axis (ω1) contains aromatic resonances, and the horizontal axis (ω2) aliphatic resonances.

by contrast retain α-helical positions. Nonuniformity of A- or B-chain β-strands (or both) would be in accordance with the incongruent spacing between inter-chain disulfide bridges A7–B7 and A20–B19; CysA7 and CysA20 are separated by an even number of residues, and CysB7 and CysB19 are separated by an odd number. The 13C NMR solid-state MAS line widths at the labeled sites are in general similar to those observed in well ordered peptide fibrils (14, 34) with additional broadening observed at positions B2 and A16 (supplemental Fig. S6).

I-J. Ye, K. Huang, Q. X. Hua, S. Q. Hu, and M. A. Weiss, manuscript in preparation.
Insulin analogs containing substitutions LeuA₂, LeuA³, and HisA₈ would be without clinical benefit due to low activity; binding of the above DKP-insulin variant is 100-fold lower than that of DKP-insulin (supplemental Fig. 58). Such impaired affinity is in accordance with previous studies showing that IleE² and ValA³ are singly required for receptor binding (48).11 Because ThrA₈ is by contrast not conserved and because diverse side chains at A₈ are readily accommodated on the surface of insulin (21), we investigated whether non-β-branched substitutions only at A₈ would protect, at least partially, insulin from fibrillation.12 Our studies employed rapid-acting insulin analogs in current clinical use. We chose His and Lys due to their natural occurrence among vertebrate insulin sequences12 and previous structural characterization of HisA₈-insulin and LysA₈-insulin as analogs with enhanced receptor-binding affinity (21, 49). In isolated peptides, His and Lys exhibit greater intrinsic α-helical propensity and C-capping potential than does Thr (20).

We chose to test A₈ substitutions in the context of AspB²₈-insulin (insulin aspart; the active component of NovoLog®, Novo-Nordisk A/S) and [LysB²₈,ProB²₉]insulin (KP-insulin, also designated insulin lispro, the active component of Humalog®, Lilly) (23, 50). These rapid-acting analogs are less stable than wild-type insulin, but their pharmacokinetic properties are advantageous for pump-based continuous infusion in patients with DM (24, 51, 52). Insulin solutions in external pumps may experience agitation at elevated ambient temperatures, conditions conducive to fibrillation. The stability of insulin within pumps and catheters is of critical concern as fibril plugs can occlude delivery, exposing patients with type 1 DM to fibrillation.12 Our studies employed rapid-acting insulin analogs as well as insulin lispro and [LysB²₈,ProB²₉]insulin form fibrils more quickly (1.7 ± 0.3 days, respectively) than does native human insulin (3.5 ± 0.6 days; column 4 in Table 1). Whereas the analogs are monomeric under conditions of study, the delay in wild-type fibrillation may be due in part to its partial dimerization and in part to intrinsic difference between corresponding monomers. Strikingly, the A₈-modified analogs exhibit prolonged lag times (~6 days). The lag time of TrpA₈-modified insulin lispro is also prolonged but to a lesser extent (4.2 ± 0.3 days). Natural abundance 1H-13C HSQC spectrum of a representative analog (LysA₈-KP-insulin) revealed no significant changes in A₁–A₈ SSP scores (data not shown).

Analogues were characterized with respect to receptor-binding affinity and thermodynamic stability (columns 2 and 3 in Table 1) as illustrated in Fig. 8A (CD-detected guanidine denaturation studies) and Fig. 8B (IR competitive-displacement binding assay; isoform B). The A₈-modified analogs each exhibit enhanced affinities and stabilities relative to respective parent analogs. Although the stability of LysA₈,KP-insulin (ΔGₙ₂₀ ± 0.1 kcal/mol) is similar to that of wild-type insulin (3.3 ± 0.1 kcal/mol) under conditions in which both are monomeric, the fibrillation lag time of the LysA₈ analog is prolonged by almost 2-fold. The A₈ substitutions do not augment cross-binding to IGFR (data not shown).

The hypoglycemic potency of a representative A₈-modified rapid-acting analog ([HisA₈,AspB²₈]insulin) was tested in a streptozotocin-induced rat model of DM. Five rats were treated with wild-type insulin, and five rats with the analog at a standard dose of 20 mg per rat (mean body weight ~300 g). Following subcutaneous injection, the rate of fall in blood glucose concentration was indistinguishable (Fig. 8, C and D), 243 ± 19 mg/dl/h (wild type) and 272 ± 27 mg/dl/h (analog). This difference is not statistically significant. In control studies injection of buffer alone did not cause a reduction in blood sugar (Fig. 8C, top trace).

### DISCUSSION

Insulin provides a classical model for studies of protein folding and assembly (27). Stored in β-cells as zinc hexamers, insulin dissociates in the bloodstream to function as a zinc-free monomer (54). Amino acid substitutions have widely been investigated for possible favorable effects on the pharmacokinetics of insulin action following subcutaneous injection. Such substitutions (such as those in NovoLog® and Humalog®) can be associated with more rapid fibrillation and poorer physical stability. Indeed, in a series of 10 analogs tested by Fink and co-workers (9) for susceptibility to fibrillation, each was found to be more susceptible to fibrillation (at pH 7.4 and 37 °C) than human insulin. Because the substitutions were at diverse sites

### TABLE 1

| Protein          | Activity* | Stability* | Fibr. lag time* |
|------------------|-----------|------------|-----------------|
| Insulin          | 1.0       | 3.3        | 3.5 ± 0.6       |
| KP-insulin⁴      | 0.9       | 2.8        | 2.6 ± 0.3       |
| HisA₈,KP-insulin | 2.9       | 3.9        | 6.2 ± 0.8       |
| LysA₈,KP-insulin | 3.1       | 3.2        | 6.3 ± 0.6       |
| AspB²₈-insulin⁴  | 1.2       | 2.9        | 1.7 ± 0.3       |
| HisA₈,AspB²₈-insulin | 3.4   | 3.7        | 5.8 ± 0.3       |

*Activity is defined as relative affinity of the analog at room temperature for an immobilized C-terminal (FLAG), epitope-tagged and detergent-solubilized IR (isoform B). Values shown have uncertainties of ±15% based on estimates of means ± S.E. The hormone-receptor dissociation constant under these conditions is 0.06 ± 0.009 nM. The stability (kcal/mol) is defined as the free energy of unfolding (ΔGₙ₂₀) at 25 °C as inferred (in the absence of zinc ions) from guanidine denaturation studies by application of a two-state model (74). Uncertainties are ±0.1 kcal/mol.

Fibrillation (Fibr. lag time) is defined by the onset of amyloid-enhanced fluorescence of ThT as described (7). Proteins were made 60 μM zinc-free in phosphate-buffered saline (pH 7.4) and gently rocked at 37 °C in a glass vial in the presence of an air-water interface. KP indicates substitutions ProB²₈ → Lys and LysB²₉ → Pro as employed in the rapid-acting insulin product Humalog® (Lilly).

AspB²₈-insulin (also known as insulin aspart) is the active component of the rapid-acting insulin product NovoLog® (Novo-Nordisk A/S).
Amyloid-related Achilles’ Heel in Insulin

SSP Scores and Crystallographic Data Base—Crystal structures of insulin (26, 27) contain canonical architectural elements (i.e. globular packing of \( \alpha \)-helices, \( \beta \)-turns, and \( \beta \)-sheets to define polar surfaces and a hydrophobic core). A complementary view has emerged from NMR (31, 55). Although conventional NMR-derived models (12, 31, 55) are similar to crystallographic \( T \) states (right-hand protomer in Fig. 4A) (27), this study provides evidence for unequal segment-specific helical stabilities. An anomalous pattern of \(^{13}\)C chemical shifts in the native state provides evidence of a breakdown of the cooperativity of folding in the \( \alpha \)-helical domain (56).

The first part of our study exploited an empirical correlation between secondary structure and the \(^{13}\)C CSI, an NOE-independent structural probe (10, 11), to investigate conformational variability. The \(^{13}\)Ca secondary shift is most sensitive to helical structure, whereas the \(^{13}\)CB secondary shift is most sensitive to the folding of a \( \beta \)-strand. A generalization of the index has led to the development of the SSP algorithm (32), which at a given position predicts the percentage of conformers in an ensemble that are helical or strand. Correlation was observed between SSP and conformational variability among crystallographic protomers.

The crystallographic data base provides information regarding sites of mobility and conformational change. Although each individual structure is conventionally organized, comparison of multiple protomers reveals local and nonlocal differences, including variation in main-chain dihedral angles, rigid-body displacement of secondary structural elements, alternate secondary structures, and propagated changes in core packing. Evidence for the flexibility of the monomer in solution was previously obtained from Raman spectroscopy (17). In the monomer, \( \beta \)-sheet- and \( \alpha \)-helix-specific vibrational bands (amide I) are anomalously broadened relative to corresponding bands in insulin dimers and hexamers, suggesting that conformational excursions in the monomer are dampened on assembly. Complementary evidence for assembly-dependent stabilization has been provided by CD (57). Unlike the monomer, the phenol-stabilized \( R \_h \) zinc insulin hexamer exhibits spectroscopic features characteristic of well organized globular proteins (including standard Raman line widths (17), CD amplitudes (58), \(^{1}\)H NMR line widths (59), and \(^{13}\)C NMR chemical shifts (60)). The solution structure of an \( R \_h \) hexamer is essentially identical to corresponding crystal structures (60, 61).

FIGURE 8. Stability and activity of a representative insulin analog. A, protein denaturation studies. CD-detected protein denaturation studies demonstrate that AspB28-insulin (solid line) is less stable than wild-type insulin (dashed line), whereas [HisA8,AspB28]insulin (dotted line) is more stable than wild type (wt). In each case fractional change in ellipticity is plotted as a function of concentration of guanidine HCl at 25 °C. Analysis by a two-state model (74) yields estimates of \( \Delta G \), (see Table 1), B, receptor-binding assay. Competitive displacement of \(^{125}\)I-labeled insulin from the insulin receptor (isoform B; IR-B) in the presence of varying concentrations of wild-type human insulin (○), AspB28 analog (△), or A8-modified analog ([HisA8,AspB28]insulin; ●). C, initial rate of decline of blood glucose following subcutaneous injection of wild-type human insulin (○), the A8-modified analog ([HisA8,AspB28]insulin; △), or control diluent injection (×). Human insulin or the insulin analog was given at dose 20 \( \mu \)g per rat (~300 g body weight). Mean initial values of blood glucose (\( n = 5 \) for each group) were similar in sets of rats treated with the analog or wild-type insulin. Error bars represent S.D. Control studies were performed at a different date using rats of similar body mass and degree of hyperglycemia. D, histogram showing initial rate of decline (mean ± S.D.) among the two groups of diabetic rats (\( n = 5 \) for each group). Although the \( \text{in vitro} \) IR-binding affinity of [HisA8,AspB28]human insulin is higher than that of human insulin (Table 1), their biological potencies in this rat assay are indistinguishable.

and presumably associated with a wide range of changes in thermodynamic stability, no structural criteria or rules emerged for the screening or design of fibrillation-resistant insulin analogs. If obtained, such analogs would be of interest in relation to the safety and efficacy of insulin replacement therapy in DM.

This study sought to augment the resistance of insulin to fibrillation based on foundational biophysical principles. Our approach built on the crystallographic data base of insulin structures to investigate sites of conformational variability in an engineered monomer as probed by anomalous CSI attenuation and amide proton exchange in solution. These results focused attention of the A1–A8 segment, next shown by solid-state and amide proton exchange in solution. These results focused structures to investigate sites of conformational variability in an approach built on the crystallographic data base of insulin fibrillation based on foundational biophysical principles. Our

initial rate of decline (mean ± S.D.) among the two groups of diabetic rats (\( n = 5 \) for each group). Although the \( \text{in vitro} \) IR-binding affinity of [HisA8,AspB28]human insulin is higher than that of human insulin (Table 1), their biological potencies in this rat assay are indistinguishable.

and presumably associated with a wide range of changes in thermodynamic stability, no structural criteria or rules emerged for the screening or design of fibrillation-resistant insulin analogs. If obtained, such analogs would be of interest in relation to the safety and efficacy of insulin replacement therapy in DM.

This study sought to augment the resistance of insulin to fibrillation based on foundational biophysical principles. Our approach built on the crystallographic data base of insulin structures to investigate sites of conformational variability in an engineered monomer as probed by anomalous CSI attenuation and amide proton exchange in solution. These results focused attention of the A1–A8 segment, next shown by solid-state and amide proton exchange in solution. These results focused structures to investigate sites of conformational variability in an approach built on the crystallographic data base of insulin fibrillation based on foundational biophysical principles. Our
We imagine that the A1–A8 segment (Fig. 9A, yellow ribbon) flexibly packs against nonpolar template provided by the underlying core of the insulin (Fig. 9B, green surface). Quantitative insight is provided by analysis of segment-specific variation in degree of structural uniformity (or nonuniformity) among α-helices. Because an insulin monomer in solution resembles the crystallographic T state (12, 30, 55), we have focused on a collection of 13 such protomers; representative T state structures were extracted from T6 and T27R3R3 zinc hexamers (27, 62) and the T2 dimer (63). This collection exhibits significant structural variation. Global alignment based on the main-chain atoms of A1–A21 and B9–B28 yields r.m.s.d. of 0.46 Å for aligned atoms and 1.20 Å for side chains. In this alignment, the three constituent α-helices (residues A1–A8, A12–A18, and B9–B19) differ in overall position. A segment-specific measure of structural uniformity is provided by the mean main-chain r.m.s.d. within a given helix as follows: low for the central B-chain α-helix (0.25 ± 0.10 Å), intermediate for the C-terminal A-chain helix (0.47 ± 0.33 Å), and high for the N-terminal A-chain helix (0.64 ± 0.55 Å).

The correlation between NMR-based SSP scores and crystallography-based r.m.s.d. values among α-helices is shown in schematic form in Fig. 4A and B. An SSP score near zero (white in Fig. 4A) indicates significant conformational averaging in solution, corresponding to regions of marked conformational variability among crystal structures (Fig. 4B, red). Outside of the helices, the isolated C-terminal B-chain β-strand in DKP-insulin (Fig. 4A, violet) correlates with the uniformity of this segment in crystal structures (Fig. 4B, blue); this uniformity is in part imposed by the dimer-related anti-parallel β-sheet (Fig. 4D). Discordance between SSP and crystallographic r.m.s.d. values is observed at the N-terminal B-chain segment, whose positioning in crystal structures is influenced by lattice packing. More marked reorganization of the N-terminal segment occurs in the TR transition (Fig. 4D) (62, 64), excluded from the present analysis.

Crystal structures also exhibit variability in residue-specific helical geometry. Whereas the above segment-specific r.m.s.d. values in principle reflect both local differences and global structural variation (i.e. rigid-body rotation or displacement of helices), a strictly local measure of variation is provided by main-chain (φ, ψ) dihedral angles. For any one residue, the collection of T state structures yields 13 (φ, ψ) values, permitting calculation of their mean values and standard deviation. A segment-specific measure of local conformational variability is provided by the average of the standard deviations (S.D.) for residues within a given helix. The corresponding values (S.D.φ and S.D.ψ) are 4.82° and 5.33° for the central B-chain α-helix, 6.41° and 8.31° for the C-terminal A-chain helix, and 8.96° and 9.95° for the N-terminal A-chain helix. This measure of (φ, ψ) variability thus yields the same rank order of conformational variability as segment-specific r.m.s.d. values (supplemental Table S2). A similar trend is observed among main-chain thermal B-factors of the T₆ hexamer refined at a resolution of 1.9 Å: mean segmental values are 9.4 Å for the central B-chain α-helix, 15.2 Å for the C-terminal A-chain helix, and 19.8 Å for the N-terminal A-chain helix (27).

Conformational variability among crystal structures also affects rigid-body positioning of the A1–A8 α-helix and local main-chain conformations without disruption of main-chain hydrogen bonds or lengthening of key (i, i + 3) and (i, i + 4) inter-proton distances beyond the 5-Å cut-off for observable NOEs. Although such NOEs are retained in DKP-insulin, it is possible that the range of fluctuations in solution exceeds that seen among crystallographic protomers, thereby rationalizing rapid amide proton exchange in this segment. Helix-related hydrogen bonds are likewise maintained in the C-terminal A-chain segment despite its intermediate level of variability (25% attenuation of ¹³C-SSP score relative to the central B-chain α-helix).

SSP analysis of the insulin monomer is reminiscent of the partial folds of protein-folding intermediates (45). Unlike conventional ground states (supplemental Fig. S9), such intermediates are characterized by broad free-energy landscapes. By
Fibrillation of globular proteins is mediated by aggregation of partially unfolded intermediates (68). In the case of insulin, a variety of evidence supports such a mechanism (6, 8, 69). Under classical amyloidogenic conditions (pH 2 and 60 °C) (9, 70), a novel partial fold (Fig. 10, center) has been observed by NMR in which the N-terminal segments of the A- and B-chains detach from the core; unfolding of the N-terminal α-helix of the A-chain exposes a hydrophobic surface formed by native-like packing of the remaining α-helices (6). The C-terminal segment of the B-chain, although not well ordered, remains tethered to this partial core (6). Segmental flexibility of the N-terminal A-chain segment may contribute to formation of amyloid, either by itself mediating non-native aggregation or by exposing underlying nonpolar surfaces that do so. It is possible that direct and indirect mechanisms are each operative as both insulin chains participate in cross-β-assembly (present solid-state NMR results; see also Ref. 71), and each isolated chain is itself amyloidogenic (72). Although our studies focused on only three sites in each chain, the β-sheet-associated $^{13}$Ca and $^{13}$CB chemical shifts of Val$^\text{A3}$, Ile$^\text{A10}$, and Leu$^\text{A16}$ suggest an extensive α → β transition, including in part or in whole the A1–A8 segment. Enhancing the predicted helical propensity of this segment by replacement of its three native β-branched side chains reduces the stability of the insulin monomer (presumably via nonoptimal core packing involving Leu$^\text{A2}$ and Leu$^\text{A3}$) but protects the analog from fibrillation by a factor of 5 in lag time. Such protection is associated with a more α-helical pattern of $^{13}$Ca chemical shifts. This association was not significant in a less well protected A8 analog. Although our perspective focuses on conformational distortions in the native state, it is also possible that protective substitutions destabilize non-native oligomeric intermediates, the protofibril, or mature fibril.

Although β-branched residues Ile$^\text{A2}$ and Val$^\text{A3}$ are required for biological activity, optimization of the C-cap of the unstable A1–A8 α-helix can also delay fibrillation without loss of activity (Table 1). It is likely that the A8 modified α-helix is stabilized by removal of a β-branched residue, by favorable (i, i + 4) side-chain interactions with a Glu$^\text{A4}$, and by (in the case of Lys$^\text{A9}$) electrostatic compensation of the helical dipole. The lag time of the A1–B30 tether in mini-proinsulins would hinder both direct and indirect contributions of segmental unfolding of the N-terminal A-chain segment to such aggregation (7).
the TrpA8 analog, intermediate between ThrA8 and HisA8 or LysA8, suggests the removal of the β-branch is itself advantageous in accordance with model peptide studies. Because patients with type 1 DM managed by continuous insulin infusion routinely replace the insulin in the pump reservoir every 3 days, this improvement may enhance the safety of such devices at summer temperatures. With current technology, external pumps exposed to direct sunlight (as can occur during summertime outdoor activities) can expose the insulin solution with the reservoir to temperatures significantly higher than ambient. Risk of occlusion leading to ketoacidosis can be magnified by delays in automated pump detectors of reduced flow (which can sound only after onset of hyperglycemia) and further exacerbated in childhood and adolescence if the unsupervised patient chooses to ignore or disable the alarm. The degree of protection from thermal fibrillation provided by A8-modified insulin analogs, although less marked than that of the “β-branchless” LeuA2,LeuA3,HisA8 biophysical model, may be sufficient to make existing pump alarm systems robust to the lifestyle demands of adolescence.

Although pump-based insulin therapy is a technology of affluent societies, a pandemic of DM is emerging in the developing world. Indigent populations are at risk in regions exposed to seasonal temperatures >35°C and lacking consistent access to electricity and refrigeration. Mechanism-based studies of insulin fibrillation and further development of ultra-stable insulin analog formulations may be of humanitarian use in such regions.14

Concluding Remarks—This study is of both technical and mechanistic interest. Segment flexibility of an engineered insulin monomer highlights a limitation of conventional protocols for calculation of NMR structures. Such protocols ordinarily require that each model obey each restraint simultaneously. Reported solution structures of insulin and insulin analogs (including from this laboratory) thus exhibit three α-helices with similar apparent precisions (7, 31, 55). Such idealized ensembles are likely to underestimate conformational sampling. Newly described approaches to characterize the extent of fluctuations in proteins (73) may be informative. It would also be of interest to investigate whether motions in insulin are correlated and, if so, might provide a mechanism of allostery among crystallographic hexamers (26). The present results also suggest that further application of solid-state NMR methods to insulin fibrils would be feasible. In the future, systematic measurements of main-chain dihedral angles and distances between labeled sites could enable construction of a three-dimensional model.

The mechanism of fibrillation by disulfide cross-linked globular proteins poses an important general problem. Despite the broad biological utility of segmental flexibility in proteins, an associated breakdown in cooperativity may lead to an amyloidogenic partial fold (2). Segmental unfolding is likely to contribute to the exquisite susceptibility of insulin to fibrillation. The anomalous pattern of 13C NMR chemical shifts thus suggested a strategy to protect the monomer based on a general principle of peptide chemistry. We imagine that the likelihood of formation of an amyloidogenic nucleus is increased by segmental unfolding to enable non-native interactions by β-strand conformations. This model predicts that the risk of fibrillation would be decreased by optimization of intrinsic α-helical propensities (18–20) as demonstrated by the inactive [LeuA2,LeuA3,HisA8]DKP-insulin analog.

Application to clinical insulin analogs may prevent occlusion-related metabolic decompensation during continuous insulin infusion (25) and facilitate the transport, storage, and use of insulin in the developing world. Rapid-acting A8-modified insulin analogs thus exemplify the promise of foundational biophysics to enhance protein therapeutics. Comparative 13C NMR chemical shift mapping in the native state and in a fibril may provide a general method to identify amyloid-related “Achilles’ heels” in globular proteins.

Acknowledgments—Solid-state NMR measurements were supported by the Intramural Research Program of the NIDDK, National Institutes of Health. We thank Drs. P. Carey, G. M. Clore, G. G. Dodson, R. Levy, D. F. Steiner, and B. Sykes for discussion; and we thank M. Choquette for technical assistance in generating the labeled DKP-proinsulin. M. A. W. thanks Drs. D. Jacobson (Case Western Reserve University Mandel School of Applied Social Sciences), M. R. Miah (Director, Southern Illinois University Carbondale School of Social Work), M. O. Rahman (Provost, Independent University of Bangladesh), and B. M. Chowdhury (Vice Chancellor, Independent University of Bangladesh) for sponsorship of a student-faculty trip to study poverty, micro-finance, and medical needs in the developing world.

REFERENCES
1. Brange, J., Andersen, L., Laursen, E. D., Meyn, G., and Rasmussen, E. (1991) J. Pharm. Sci. 86, 517–525
2. Dobson, C. M. (2003) Nat. Rev. Drug Discov. 2, 154–160
3. Brange, J. (1987) Galenics of Insulin: The Physicochemical and Pharmacological Aspects of Insulin and Insulin Preparations, pp. 52–68, Springer-Verlag, Berlin
4. Weiss, M. A., Hua, Q. X., Lynch, C. S., Frank, B. H., and Shoelson, S. E. (1991) Biochemistry 30, 7373–7389
5. Hua, Q. X., Xu, B., Huang, K., Hu, S. Q., Nakagawa, S., Jia, W., Wang, S., Whitaker, J., Katsoyannis, P. G., and Weiss, M. A. (2009) J. Biol. Chem. 284, 14586–14596
6. Hua, Q. X., and Weiss, M. A. (2004) J. Biol. Chem. 279, 21449–21460
7. Huang, K., Dong, J., Phillips, N. B., Carey, P. R., and Weiss, M. A. (2005) J. Biol. Chem. 280, 42345–42355
8. Ahmad, A., Millett, I. S., Doniach, S., Uversky, V. N., and Fink, A. L. (2003) Biochemistry 42, 11404–11416
9. Nielsen, L., Frokjaer, S., Brange, J., Uversky, V. N., and Fink, A. L. (2001) Biochemistry 40, 8397–8409
10. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) Biochemistry 31, 1647–1651
11. Wishart, D. S., and Sykes, B. D. (1994) J. Biomol. NMR 4, 171–180
12. Hua, Q. X., Shoelson, S. E., Kochoyan, M., and Weiss, M. A. (1991) Nature 354, 238–241
R. Y., Whittaker, J., Katsoyannis, P. G., and Weiss, M. A. (2009) J. Biol. Chem. 284, 14597–14608
83. Wan, Z. L., Xu, B., Chu, Y. C., Katsoyannis, P. G., and Weiss, M. A. (2003) Biochemistry 42, 12770–12783
84. Wan, Z. L., Huang, K., Xu, B., Hu, S. Q., Wang, S., Chu, Y. C., Katsoyannis, P. G., and Weiss, M. A. (2005) Biochemistry 44, 5000–5016
85. Kobayashi, M., Ohgaku, S., Iwasaki, M., Maegawa, H., Shigeta, Y., and Inouye, K. (1982) Biochem. Biophys. Res. Commun. 107, 329–336
86. Nourse, A., Treacy, G. B., Shaw, D. C., and Jeffrey, P. D. (1996) Biol. Chem. Hoppe-Seyler 377, 147–153
87. Yunus, M. (1998) Banker to the Poor: Micro-lending and the Battle against World Poverty, Aurum Press Ltd., London, UK
88. Hua, Q. X., Nakagawa, S. H., Jia, W., Huang, K., Phillips, N. B., Hu, S. Q., and Weiss, M. A. (2008) J. Biol. Chem. 283, 14703–14716