# Structure-Specific Effects of Protein Topology on Cross-$\beta$ Assembly: Studies of Insulin Fibrillation†

Kun Huang,‡ Nakul C. Maiti,‡ Nelson B. Phillips, Paul R. Carey, and Michael A. Weiss*

*Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106*

*Received May 3, 2006; Revised Manuscript Received June 8, 2006*

ABSTRACT: Systemic amyloidoses, an important class of protein misfolding diseases, are often due to fibrillation of disulfide-cross-linked globular proteins otherwise unrelated in sequence or structure. Although cross-$\beta$ assembly is regarded as a universal property of polypeptides, it is not understood how such amyloids accommodate diverse disulfide connectivities. Does amyloidogenicity depend on protein topology? A model is provided by insulin, a two-chain protein containing three disulfide bridges. The importance of chain topology is demonstrated by mini-proinsulin (MP), a single-chain analogue in which the C-terminus of the B chain (residue B30) is tethered to the N-terminus of the A chain (A1). The B30-A1 tether impedes the fiber-specific $\alpha \rightarrow \beta$ transition, leading to slow formation of a structurally nonuniform amorphous precipitate. Conversely, fibrillation is robust to interchange of disulfide bridges. Whereas native insulin exhibits pairings [A6-A11, A7-B7, and A20-B19], metastable isomers with alternative pairings [A6-B7, A7-A11, A20-B19] or [A6-A7, A11-B7, A20-B1] readily undergo fibrillation with essentially identical $\alpha \rightarrow \beta$ transitions. Respective pairing schemes are in each case retained. Isomeric fibrils and the amorphous MP precipitate are each able to seed the fibrillation of wild-type insulin, suggesting a structural correspondence between respective nuclei or modes of assembly. Together, our results demonstrate that effects of polypeptide topology on amyloidogenicity depend on structural context. Although the native structures and stabilities of single-chain insulin analogues are similar to those of wild-type insulin, the interchain tether constrains the extent of conformational distortion at elevated temperature, retards initial non-native aggregation, and is apparently incompatible with the mature structure of an insulin protofilament. We speculate that the general danger of fibrillation has imposed a constraint in protein evolution, selecting for topologies unfavorable to amyloid formation.

A major frontier of molecular medicine is defined by amyloid-associated diseases, including prion-related encephalopathies, neurodegenerative diseases, and systemic amyloidoses (*1*). Although these pathological processes involve only a small subset of endogenous proteins, fibrillation is in vitro a generic property of polypeptides (*2*−*4*). X-ray diffraction studies indicate that such fibrils arise from cross-$\beta$ assembly; the constitutive $\beta$-strands are perpendicular to the long axis of the fibril (*5*, *6*). Whereas some amyloidogenic species in vivo are peptides or natively unfolded (*1*), others are well-defined globular proteins. Examples of clinical importance include immunoglobulin light chains, $\beta_2$-microglobulin, and human lysozyme variants (*7*−*9*). The native states of these proteins each exhibit a classical hydrophobic core stabilized by multiple disulfide bridges. Such bridges are generally maintained in the fibril (*10*, *11*). Given the generality of peptide cross-$\beta$ assembly (*3*), a seeming paradox is posed by the variable positioning of cystines within sequences of globular proteins.

A range of models of protein fibrils (as distinct from peptide fibrils) have been proposed that differ in the extent to which native structural elements are retained or reorganized (*12*−*16*). Ribonuclease fibrils, for example, are envisaged to undergo elongation through domain swapping within an essentially nativelike assembly (*11*). Formation of lysozyme fibrils by contrast requires substantial protein unfolding and subsequent conversion of $\alpha$-helix to $\beta$-sheet (*17*). An attractive model for the study of protein fibrillation is provided by insulin, a small globular protein containing two chains, designated A (21 residues) and B (30 residues). Insulin is the biosynthetic product of a larger single-chain precursor polypeptide, proinsulin, in which a 35-residue domain connects the C-terminal residue of the B chain to the N-terminal residue of the A chain (Figure 1A; *18*). Oxidative folding of proinsulin within the endoplasmic reticulum of the $\beta$ cell directs specific disulfide pairing (cystines A6-A11, A7-B7, and A20-B19; orange lines in Figure 1A). The connecting peptide contains flanking pairs of basic residues (filled black circles in Figure 1A), sites of cleavage in the biosynthesis of insulin (*19*). Although the crystal structure of proinsulin has not been determined, a variety of evidence suggests the presence of a well-ordered insulin moiety and disordered connecting domain (dashed blue line in Figure 1B).

The structure of insulin has been extensively investigated. The native state contains representative elements of larger globular proteins: $\alpha$-helices, $\beta$-strands, $\beta$-turns, a noncanonical chain reversal, and a hydrophobic core (Figure 1C; *20*). As in classical amyloids, insulin fibrillation is associated

† This work was supported in part by grants from the National Institutes of Health to M.A.W. (DK0697674 and DK074176).

* Address correspondence to this author. Tel: (216) 368-5991. Fax: (216) 368-3419. E-mail: maw21@case.edu.

‡ These authors contributed equally.

10.1021/bi060879g CCC: $33.50 
Published on Web 08/04/2006

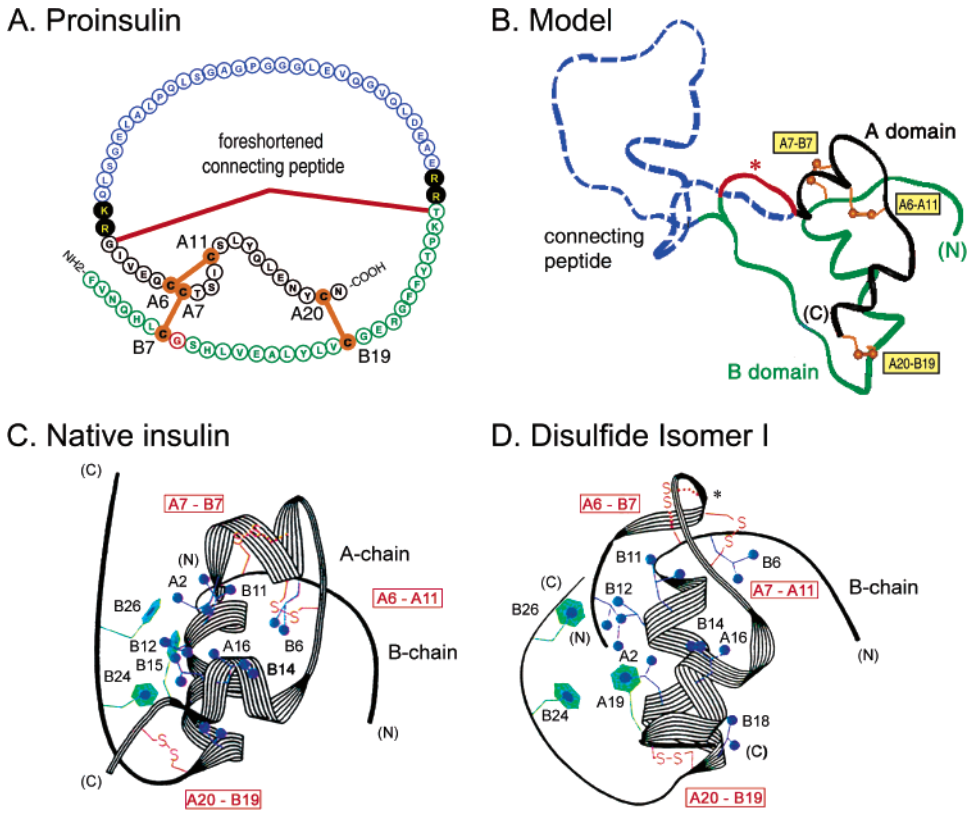


FIGURE 1: Proinsulin and insulin disulfide isomers. (A) Sequence of human proinsulin: A and B chains are represented by black and green open circles, respectively, with cystines as filled orange circles (native pairing). The connecting domain is represented by blue open circles (C-peptide) and filled circles (dibasic cleavage sites). Alternative B30-A1 tether in MP is shown in red. (B) Structural model of proinsulin, which consists of an insulin-like moiety and disordered connecting domain (dashed line). Coloring scheme as in panel A. Cystines are labeled in yellow boxes. (C and D) Ribbon models of native insulin and insulin-*swap* (*32*). Cystines are labeled in red boxes.

with a substantial $\alpha \rightarrow \beta$ transition (*21*, *22*) with retention of native disulfide pairing (*23*, *24*). An engineered tether between the C-terminus of the B chain and the N-terminus of the A chain (the junctions of the connecting domain in human proinsulin; red lines in Figure 1A,B) facilitates folding (*25*) but does not otherwise affect the structure of insulin (*26*, *27*). Human proinsulin (HPI)[1] is markedly less susceptible to fibrillation than is human insulin despite their similar thermodynamic stabilities (*28*). Although HPI fibrils may be obtained at acidic pH at elevated temperature, the lag time is prolonged by 15-fold; at neutral pH HPI forms a $\beta$-sheet-rich amorphous precipitate, indicating that either initial cross-$\beta$ assembly or its orderly propagation is unfavorable (*28*). Because parallel models of an insulin protofilament envisage wide separation of the C-terminus of the B chain from the N-terminus of the A chain (*24*), the delayed time course of HPI fibrillation may reflect unfavorable conformational constraints imposed by the connecting tether, hindering a generic mode of peptide cross-$\beta$ assembly. The high solubility of the connecting peptide and its intrinsic resistance to fibrillation may also contribute to the low amyloidogenicity of HPI (*28*).

In this paper we investigate whether the amyloidogenicity of insulin can be modulated by changes in chain topology. The essential idea is that spatial constraints due to chain connectivity may alter (a) the range of distorted conformations initially populated in solution under amyloidogenic conditions, (b) the susceptibility of such distorted monomers to non-native aggregation, (c) structures of insoluble assemblies or nascent protofilaments (*24*), or (d) their possible modes of growth (*28*). In these studies chain topology is constrained or perturbed in two different ways. First, the proximity between chain termini B30 and A1 is tethered by foreshortening the connecting peptide from 35 residues (in HPI) to 2 residues [in mini-proinsulin (MP); asterisk in Figure 1B; *29*]. Although essentially without biological activity (*25*−*27*, *29*), the native states of MP and related single-chain analogues exhibit enhanced efficiency of refolding after reduction (*25*) and resistance to fibrillation (*30*). The structure of MP (*26*, *27*) and its mode of zinc-mediated hexameric assembly (*26*) are essentially identical to that of wild-type insulin. A second strategy to alter chain topology is based on disulfide rearrangement. Whereas the native structure of insulin requires specific pairing [A6-A11, A7-B7, A20-B19] (highlighted in orange in Figure 1A; *31*), under partially denaturing conditions two disulfide isomers accumulate following thiol-catalyzed rearrangement (*32*). Designated insulin-*swap* and insulin-*swap2*, these species contain respective pairings [A7-A11, A6-B7, A20-B19] and [A6-A7, B7-A11, A20-B19]. The isomers exhibit low but significant biological activity (*33*) with partial folds contain-

[1] Abbreviations: ANS, 1-anilino-8-naphthalenesulfonate; CD, circular dichroism; DTT, dithiothreitol; HEWL, hen egg white lysozyme; HPI, human proinsulin; HPLC, high-performance liquid chromatography; NMR, nuclear magnetic resonance; MP, mini-proinsulin; PAGE, polyacrylamide gel electrophoresis; PPII, polyproline II; rp-HPLC, reverse-phase high-performance liquid chromatography; TFA, trifluoroacetic acid; ThT, thioflavin T. Amino acids are designated by standard one- and three-letter codes.

ing a subset of nativelike structural elements (Figure 1D; *32*, *34*).

Our results demonstrate marked differences in how changes in protein topology affect amyloidogenicity. In accord with previous studies (*28*, *30*) imposition of a tight B30-A1 tether blocks fibrillation. This blockage is a result of several factors: the B30-A1 tether protects the distorted monomer at elevated temperature from aberrant aggregation whereas prolonged incubation leads to amorphous precipitation rather than fibrillation. Whereas classical fibrils exhibit ultra-narrow Raman bandwidths indicative of structural uniformity (*35*), MP precipitates exhibit broad bands indicative of structural heterogeneity. Further, whereas insulin fibrils exhibit a near-complete $\alpha \rightarrow \beta$ transition (*35*−*38*), deconvolution of the amide I Raman region suggests that in MP precipitates this transition is incomplete, suggesting a less compact structure with higher residual α-helix content. In contrast, interchange of disulfide bridges is well tolerated: disulfide isomers insulin-*swap* and *swap*2 form classical fibrils in which non-native pairings are retained within classical $\beta$-sheet-rich structures. In cross-seeding experiments the isomeric fibrils each promote the fibrillation of wild-type insulin, suggesting similar modes of assembly. Surprisingly, an amorphous MP aggregate, despite its incomplete $\alpha \rightarrow \beta$ transition, can also accelerate insulin fibrillation. Such cross-seeding suggests that the local structure of an MP aggregate can mimic that of wild-type amyloidogenic intermediates. Although less well organized than wild-type fibrils, MP aggregates nevertheless provide an analogous amyloidogenic growth surface for insulin monomers (*39*, *40*), accelerating formation of non-native oligomers whose altered conformations foreshadow the classical $\alpha \rightarrow \beta$ transition characteristic of mature fibrils (*41*).

Human insulin illustrates a general and poorly understood frontier of protein chemistry: the competition between folding and misfolding of a polypeptide in a rugged energy landscape. Here, the complementary properties of single-chain analogues and disulfide isomers demonstrate that whereas protein fibrillation is robust to some changes in protein topology, others can constrain accessible modes of non-native aggregation and cross-$\beta$ assembly. Although fibrillation may be a universal feature of linear polypeptides (*3*), the nonlinear connectivities of disulfide-cross-linked globular proteins may preclude simple models of cross-$\beta$ assembly. Engineered changes in topology may thus provide an approach to protect proteins from fibrillation in pharmaceutical and industrial applications. These considerations suggest that protection from amyloidogenicity may have provided a hidden constraint underlying the evolution of disulfide bridges in globular proteins.

## MATERIALS AND METHODS

*Materials.* Human insulin, insulin-*swap*, and insulin-*swap2* were kindly provided by Eli Lilly and Co. (Indianapolis, IN). MP (containing a dipeptide linker AK between B30 and A1; *29*, *42*) was the gift of Y.-m. Feng (Shanghai, China). 1-Anilino-8-naphthalenesulfonate (ANS), hen egg white lysozyme (HEWL), and thioflavin T (ThT) were obtained from Sigma (St. Louis, MO). Other chemicals (analytical grade) were obtained from Fisher Chemicals (Pittsburgh, PA).

*Preparation of Fibrils.* Solutions of proteins, prepared immediately prior to fibrillation, were made 60 $\mu$M in 0.01 N HCl [0.01 N HCl (pH 2.0) and 150 mM NaCl] and incubated in Eppendorf tubes at 65 °C without agitation. HEWL fibrils were obtained from a 10 mg/mL protein at pH 2 at 65 °C for 1 week without agitation (*43*). At indicated intervals aliquots were withdrawn and either added to a ThT solution for fluorescence assay or stored at −20 °C for EM.

*Fluorescence Spectroscopy.* ThT was made 1 mM in doubly distilled water and stored at 4 °C in the dark. To monitor fibrillation, 10 $\mu$L aliquots obtained at indicated time points were mixed with 3 mL of ThT assay buffer [5 $\mu$M ThT in 50 mM Tris-HCl (pH 7.5) and 100 mM NaCl]. Fluorescence measurements were performed using an Aviv spectrofluorometer (Aviv Biomedical, Lakewood, NJ) in 1 cm quartz cuvettes. Emission spectra were acquired from 470 to 500 nm following excitation at 450 nm; the integration time was 1 s. ThT in buffer without protein was used as baseline. The fibrillation lag time is defined as described (*28*). For ANS assay, 10 $\mu$L aliquots obtained at indicated time points were mixed with 200 $\mu$L of ANS assay buffer [1 $\mu$M ANS in 50 mM Tris-HCl (pH 7.5)]. Fluorescence measurements were performed in 3 mm quartz cuvettes. Emission spectra were acquired from 450 to 600 nm following excitation at 350 nm; ANS in buffer without protein was used as baseline.

*Cross-Seeding Studies.* To obtain a distribution of seeds, preformed protein fibrils were sonicated on ice with a Sonic Dismembrator (model 550; Fisher Scientific, Pittsburgh, PA) for a total of 30 s by a series of ten 3 s pulses. Sonicated seeds (1% w/w) were added to a freshly prepared insulin solution; fibrillation kinetics and fibril morphologies were monitored by ThT and transmission EM, respectively. The time course of MP amorphous precipitation (not associated with an increase in ThT fluorescence) was monitored by serial EM studies: detection of proteinaceous deposits correlates with onset of macroscopic precipitation in the parent vial.

*Transmission Electron Microscopy.* Aliquots (10 $\mu$L) were deposited on Formvar-coated 400 mesh copper grids (Electron Microscopy Sciences, Hatfield, PA) for 5 min. Excess solution was adsorbed to filter paper. Grids were washed three times with distilled water and three times with filtered 1% uranyl acetate for negative staining. Stained grids were allowed to dry for 20 min at room temperature. Specimens were observed with a JEOL 1200EX transmission EM operating with an accelerating voltage of 80 kV.

*Circular Dichroism.* Samples were dissolved in 0.01 N HCl (pH 2.0) at a protein concentration of 55 $\mu$M. To remove particulate matter and protein aggregates, samples were filtered (0.22 $\mu$M; Satorius, Goetlingen, Germany). Spectra, acquired with an Aviv spectropolarimeter (Aviv Biomedical, Inc., Lakewood, NJ), were normalized by mean residue ellipticity. The Selcon3 deconvolution algorithm was used to estimate fractional secondary structure composition (*44*); its comparison to two other algorithms is provided in Supporting Information. Errors in secondary structure composition were estimated on the basis of the mean error in observed ellipticities. Precision of measurement was ca. 1% between 200 and 250 nm, increasing at shorter wavelengths to 5% at 195 nm. To obtain mean errors in inferred secondary structure composition, errors in ellipticity were used to

simulate a collection of 10 noise-enhanced spectra as follows. At each wavelength the mean error in ellipticity was either added to or subtracted from the observed ellipticity according to the results of a pseudorandom (+1, −1) number generator. Each simulated spectrum was analyzed by the Selcon3 algorithm, giving rise to a distribution of estimates of secondary structure composition. Error bars in Figure 3C represent the standard deviations of these distributions.

*Native Gel Electrophoresis.* Native (nondenaturing) polyacrylamide gel electrophoresis (PAGE) was performed at a constant voltage (120 V) with a Criterion Bio-Rad electrophoresis system using a Tris-HCl polyacrylamide gel with 4% stacking gel and 10−20% gradient resolving gel. Samples were prepared in the absence or presence of sodium dodecyl sulfate (SDS). Gels were stained with Gelcode Blue (Pierce, Rockford, IL). Band intensities (native and deamidation products) were calculated with the program Quantity One (Bio-Rad, Hercules, CA).

*Raman Microscopy and Spectroscopy.* Raman spectra of solid samples were recorded using a Raman microscope system (*45*). In brief, samples were placed on aluminum foil; 20 $\mu$L of the suspended fibril solution was directly applied onto this surface, air-dried, and viewed via a long focal length objective and a video charge-coupled device (CCD) camera. A halogen lamp was used for imaging the sample and as an aid to focusing the laser. The microscope was operated in the nonconfocal mode to maximize light throughput. Approximately 90 mW of 647.1 nm laser excitation from a krypton laser was focused through the microscope using a 20× objective to form the focal volume and to generate Raman scattering; the data acquisition time was 1−2 min. Spectra displayed in figures are averages of three or more individual spectra.

*Band Fitting of Raman Spectrum.* Changes in secondary structure were investigated by analyzing the asymmetric amide I Raman band (*46*−*49*). The extended amide I region (1575−1720 $cm^{-1}$) was fitted assuming three component bands that represent the different secondary structural conformations: an α-helical band at 1655 $cm^{-1}$ with spectral window 1650−1660 $cm^{-1}$, organized $\beta$-sheet component band near 1673 $cm^{-1}$ (1670−1675 $cm^{-1}$), and composite contributions from loose $\beta$-strands, PPII, and disordered structure near 1685 $cm^{-1}$ (1680−1690 $cm^{-1}$) (*50*). Accurate fitting of experimental profile required an additional feature near 1637 $cm^{-1}$ (1630−1645 $cm^{-1}$) whose origin is uncertain; possible sources are disordered structure (*48*) and/or vibronic coupling (*48*, *51*, *52*). Contribution from $\beta$-sheet structure is also possible in this range as suggested by IR studies (*53*). Although interpretation of such spectral decomposition is an area of continuing refinement (*54*), the present analysis using four major component bands provides a robust and semiquantitative probe of changes in secondary structure. Component bands were fitted by 15−40 $cm^{-1}$ bandwidth features; Gaussian and Lorentzian functions were employed as models of homogeneous and heterogeneous broadening, respectively. Fitting was performed using the Levenberg−Marquardt nonlinear least-squares method as implemented in the CurveFit.Ab routine of GRAMS/32 (*55*). Three bands at ∼1585, ∼1604, and ∼1615 $cm^{-1}$, due to ring modes of Phe and Tyr, were included. The baseline was taken as a linear function as described (*56*, *57*). The standard error for peak positions is $<3$ $cm^{-1}$ and for peak widths $<5$ $cm^{-1}$ for well-defined components; this typically introduces an uncertainty of ∼10% in the area measurement. For weak peaks the standard error is slightly higher.

*Fibril Renaturation.* Fibrils obtained from solutions of insulin or insulin isomers (as formed at pH 2) were centrifuged at 12000*g* for 10 min. Supernatants were in each case removed, and pellets were washed three times with distilled water containing 0.1% trifluroacetic acid (TFA). A solution of guanidine hydrochloride (8 M at pH 6; Pierce, Rockford, IL) was added to dissolve the pellet; guanidine hydrochloride was also added to the supernatants to a final concentration of 6 M. After 1 h at room temperature fibrils dissolved to yield a clear solution. Supernatants and dissolved pellets were each loaded onto an analytical rp-HPLC column (C8 column; Vydac, Inc., Hesperia, CA) and eluted with a methanol gradient (40%−80% methanol in 40 min) in 0.1% TFA. Control solutions of native human insulin and disulfide isomers (in their original globular states) were used as HPLC standards.

## RESULTS

Our results are presented in two parts. In part I we investigate the single-chain insulin analogue MP (*25*) in relation to insulin and HPI (Figure 1A,B). Raman spectroscopy is employed to characterize respective native states and misfolded aggregates. In part II we extend these studies to insulin disulfide isomers (*34*), which exhibit nativelike partial folds of low stability (Figure 1D; *32*). In each case cross-seeding experiments are utilized as a probe of structural compatibility between nuclei or sonicated protofilaments. The properties of MP and insulin isomers are interpreted in relation to the fibrillation of wild-type insulin and HPI (*28*).

Insulin fibrils form under a wide variety of conditions, including following agitation at neutral pH and 37 °C as can occur in insulin pumps (Figure 2A; *28*). Our studies employ classical conditions for insulin fibrillation: a quiescent solution of protein in 0.01 N HCl (pH 2) at 65 °C (*22*, *58*, *59*). Under these conditions insulin forms well-defined fibrils within 4 h (Figure 2B); at the same protein concentration HPI forms fibrils 20-fold more slowly despite its similar thermodynamic stability (Figure 2C; *28*). Fibrillation lag times and thermodynamic stabilities are summarized in Table 1. Raman spectra are in each case obtained in the solid state to enable direct comparison between native and non-native states. Raman microscopy offers a higher signal-to-noise ratio than that typically obtained in solution due to reduced background scattering (*60*).

*(I) An Interchain Tether Retards Aggregation and Blocks Fibrillation.* MP is refractory to fibrillation but slowly forms an amorphous precipitate over 2 weeks at pH 2 and 65 °C (*28*). EM photomicrographs are notable for the absence of linear structures (Figure 2D). Unlike insulin fibrils, which exhibit 30-fold enhanced ThT fluorescence relative to the native state, MP precipitates exhibit an increase of less than 2-fold. Such attenuated ThT fluorescence suggests a paucity of well-ordered cross-$\beta$ structure (*61*).

The B30-A1 tether in MP may affect one or more of several steps in the fibrillation pathway, ranging from initial formation of a distorted monomeric conformation to the orderly propagation of a cross-$\beta$ assembly. As a probe of the distorted monomer, successive far-UV CD spectra were

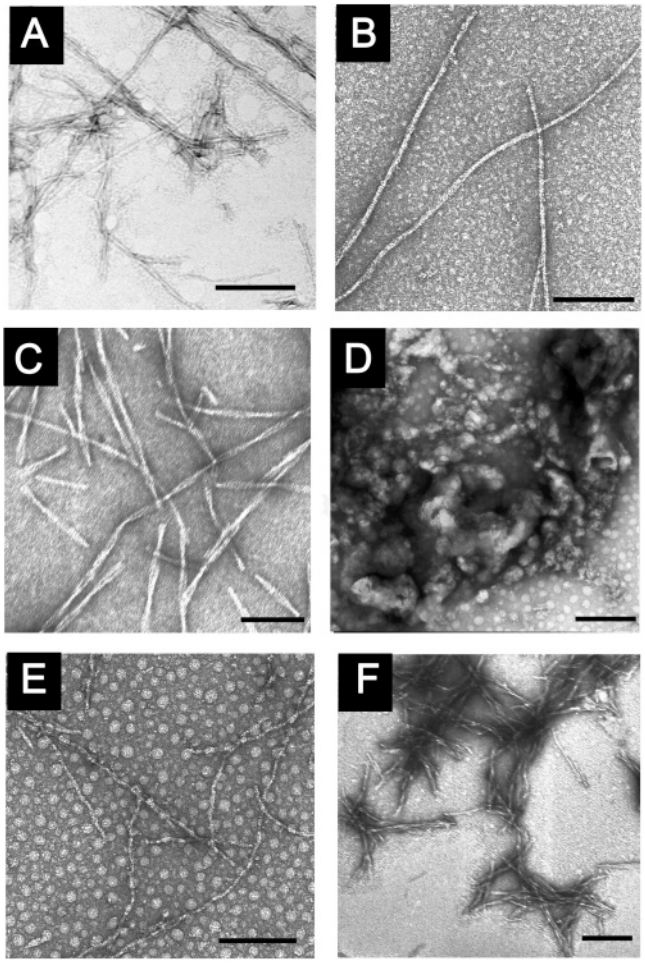


FIGURE 2: Morphology of fibrils and amorphous precipitates. EM images of insulin fibrils formed at neutral (A) or acidic pH (B); (C−F) fibrils or aggregate formed by HPI (C), MP (D), insulin-*swap* (E), and insulin-*swap2* (F). Linear well-defined fibrils formed in all samples except MP, in which only amorphous aggregates are observed. Scale bar equals 200 nm.

Table 1: Fibrillation Lag Times and Thermodynamic Stabilities

| species | lag time[a] (h) | $\Delta G_u$[b] (kcal/mol) | $C_{mid}$[c] (M) | $m$[d] (kcal mol$^{-1}$ M$^{-1}$) |
|---|---|---|---|---|
| insulin | 3.5 ± 0.4 (7)[e] | 4.4 ± 0.1 | 5.2 ± 0.1 | 0.84 ± 0.01 |
| proinsulin | 79 ± 14 (4) | 3.9 ± 0.1 | 4.7 ± 0.1 | 0.77 ± 0.02 |
| MP | >360 (4) | 3.1 ± 0.1 | 4.9 ± 0.1 | 0.63 ± 0.01 |
| *swap* | 1.5 ± 0.3 (3) | 1.3 ± 0.1 | 2.4 ± 0.2 | 0.55 ± 0.04 |
| *swap2* | <0.5 (3) | 0.7 ± 0.2 | 1.7 ± 0.2 | 0.44 ± 0.04 |

[a] Fibrillation lag times are defined by the change in ThT fluorescence (*28*). [b] Guanidine denaturation data were interpreted according to a two-state model (*107*); values were measured at 4 °C in phosphate buffer (pH 7.4). [c] $C_{mid}$ is defined as the denaturant concentration at which 50% of the protein is unfolded. [d] The $m$ value is the slope $d(\Delta G_u)/d(M)$. Statistical errors in fitting $\Delta G_u$ are in each case less than the values provided due to upward rounding to nearest tenth. [e] Number of experiment replicates.

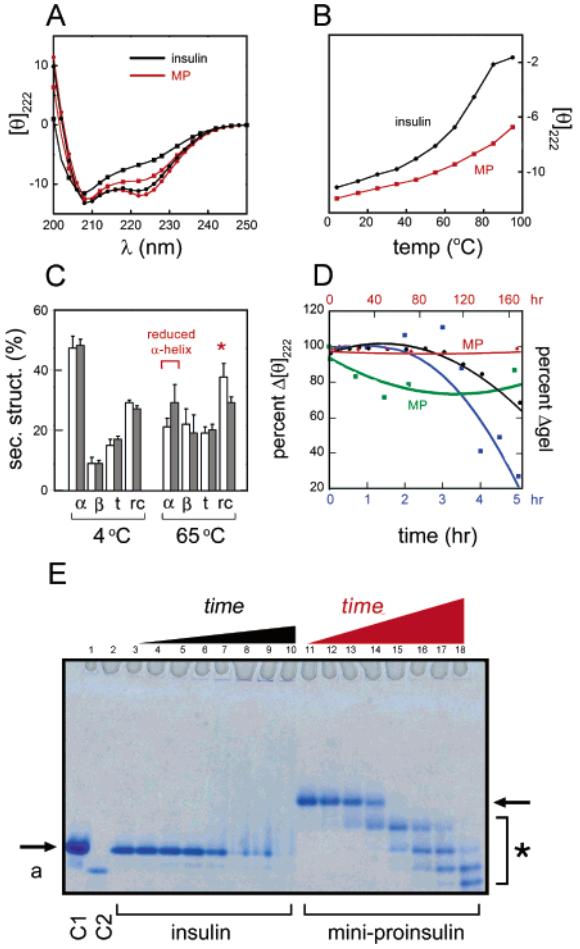


FIGURE 3: Secondary structure and kinetics of non-native aggregation. (A) Far-UV CD spectra of insulin at 4 and 65 °C (black circles and black squares, respectively) versus MP at 4 and 65 °C (red circles and squares). (B) Change in mean residue ellipticity of insulin (black circles) and MP (red squares) at 222 nm as a function of temperature. (C) Histogram showing fractional CD deconvolution of insulin (white) and MP (gray) at 4 °C (left) and 65 °C (right). Percent α-helix, $\beta$-sheet, turn, and random coil are respectively designated α, $\beta$, t, and rc. Reduced α-helix content at 65 °C (relative to 4 °C) is indicated; an asterisk highlights greater random-coil component in the spectrum of insulin at 65 °C than in MP at 65 °C. (D) Time-dependent non-native aggregation of insulin and MP as monitored by native gel electrophoresis and CD attenuation. Aggregation of MP is delayed relative to insulin; the upper horizontal time axis (red) pertains to MP whereas the lower horizontal axis pertains to insulin (blue). Gel assay: insulin (blue line and squares) and MP (green line and squares). CD attenuation assay: insulin (black line and circles) and MP (red line and circles). (E) Native PAGE electrophoresis illustrating time-dependent attenuation of monomeric bands, representing initial starting material (arrows) and deamidation products (asterisk at right). Samples were incubated at pH 2 and 65 °C for 5 h (insulin) or 7 days (MP). Lanes 3−10 correspond to insulin samples incubated for 0, 1, 2, 3, 3.5, 4, 4.5, and 5 h, respectively; native-state gel standards are provided by analogues KP-insulin (C1, lane 1; arrow at left) and DKP-insulin (C2, lane 2; labeled a at left). After 3 h aggregates accumulate as prominent bands in loading wells and smears (lanes 7−10). Lanes 11−18 correspond to MP samples incubated for 0, 1, 3, 7, 24, 50, 72, and 168 h. Deamidation products were confirmed by MALDI-TOF mass spectrometry. KP-insulin contains substitutions $Pro^{B28} \rightarrow Lys$ and $Lys^{B29} \rightarrow Pro$; DKP-insulin contains these substitutions and in addition $His^{B10} \rightarrow Asp$ as described (*76*).

obtained of wild-type insulin and MP at 4−65 °C and pH 2 (Figure 3A,B). At low temperature initial spectra are similar (filled black and red circles in Figure 3A), in each case consistent with the native structure of insulin. At elevated temperatures (>50 °C) the spectra each exhibit attenuation of mean residue ellipticity at 196, 208, and 222 nm ($[\theta]_{196}$, $[\theta]_{208}$, and $[\theta]_{222}$) consistent with incremental loss of α-helix content. This perturbation is significantly more marked in the spectrum of insulin at 65 °C (black squares in Figure 3A) than in the corresponding spectrum of MP (red squares). Deconvolution of these spectra delineates possible structural similarities and differences. Fractional secondary structural

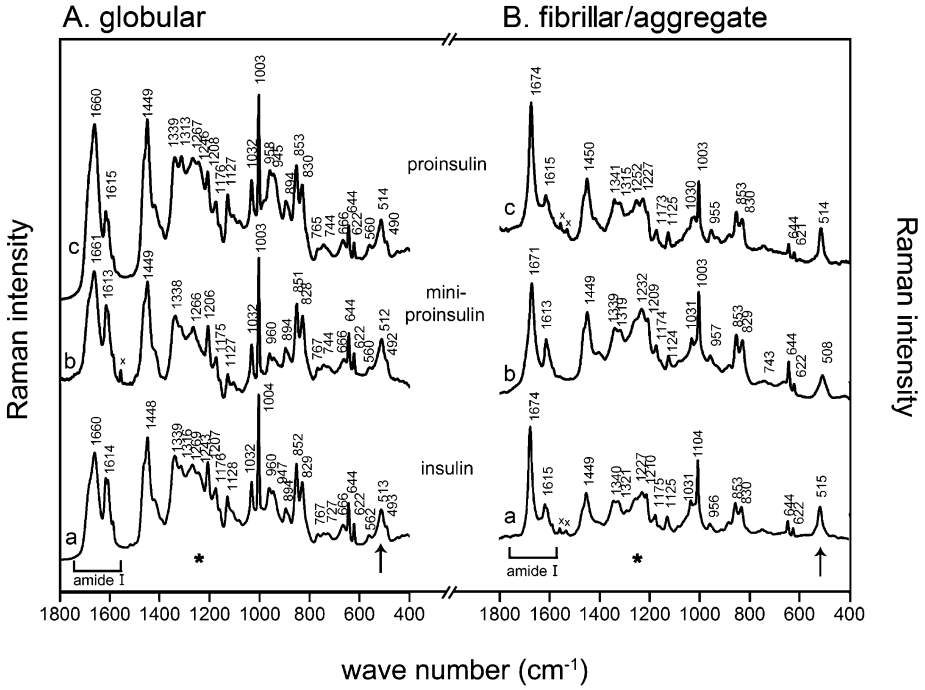


FIGURE 4: Raman studies of MP in relation to insulin and proinsulin. Solid-state spectra of analogues in native (A) and misfolded (B) states: (A) native globular domains, insulin (bottom panel), MP (middle), and HPI (top); (B) insulin and HPI fibrils (bottom and top panels) and MP aggregates (middle panel) as indicated. Arrows indicate S−S bands; brackets indicate the amide I region; asterisks indicate the amide III region; x indicates artifacts due to the laser plasma line.

compositions are shown by the histogram in Figure 3C (white columns pertain to insulin and shaded columns to MP). The apparent helix content of insulin decreases from 46(±4)% at 4 °C to 21(±3)% at 65 °C; the corresponding values for MP are 48(±2)% and 30(±5)%, respectively. The slight difference in inferred helix content between the proteins at 4 °C (within the statistical error of the Selcon3 program but evident in the original CD spectra, Figure 3A), likely to reflect damping of conformational fluctuations (*62*), is thus accentuated at 65 °C in accord with qualitative inspection of the spectra (above). Whereas at 4 °C the random-coil content of insulin [28(±2)%] is only slightly larger than that of MP [26(±1)%], this difference is likewise magnified at 65 °C [respective values of 37(±5)% and 28(±2)%].

Additional insight is provided by analysis of the temperature dependence of these CD changes between low and high temperatures. Two temperature regimes are apparent. Whereas the fractional change in $[\theta]_{222}$/°C between 4 and 40 °C is similar in the two cases, insulin exhibits a more marked attenuation of ellipticity between 50 and 70 °C (black line in Figure 3B). By contrast, in MP the thermal dependence of $[\theta]_{222}$ continues to be linear with a slope similar to that observed at low temperatures (red line). Together, these data indicate that the distorted monomeric conformations of insulin and MP differ at high temperature. Whereas both proteins exhibit gradual and progressive attenuation of mean helix content with increasing temperature, the B30-A1 tether constrains a supervening unfolding transition characteristic of insulin at high temperature. Because the solution structure of insulin at elevated temperature exhibits a nativelike partial fold with segmental unfolding of the A1-A8 α-helix (*63*), we envisage that the B30-A1 tether in MP partially constrains the unfolding of this helix (see Discussion).

CD may also be used to monitor the time course of nonnative aggregation. Within a few hours at 65 °C the CD spectrum of insulin exhibits an overall attenuation without significant distortion; little or no change is observed in the ratio of ellipticities at helix-sensitive wavelengths. By contrast, the amplitude of the CD spectrum of MP remains little changed at 65 °C for >7 days. Native PAGE of successive aliquots of insulin (lanes 3−10 in Figure 3E), cooled to room temperature, demonstrates that insulin's progressive spectral attenuation at 65 °C (black line in Figure 3D) correlates in large part with a decrease in the fraction of the protein able to enter the gel and run with a mobility characteristic of the native monomer or deamidated monomer (species that run faster than native insulin); the blue line in Figure 3D represents the combined intensity of the native and deamidated PAGE bands. Also present in the gel at later time points (lanes 7−10) are poorly resolved "tails" of staining above the native band, representing a broad distribution of oligomers. The similar trends in CD intensity and monomeric PAGE bands (blue and black lines in Figure 3D) suggest that the CD attenuation is due to progressive loss of α-helical conformations in solution; the aggregates either are sufficiently large as to scatter the incident light (rather than rotate its plane of polarization) such that their contribution to the observed spectrum is negligible or, if CD-visible, are altered in conformation such that their far-UV CD helical signature is negligible. The rapid transition of insulin under these conditions from a solution of distorted monomers to gel-impermeable aggregates occurs without accumulation of a single predominant small oligomer.

The conformation of MP at 65 °C, less distorted than that of insulin despite the similar thermodynamic stabilities of the two proteins at 25 °C, is refractory to aggregation for several days. By contrast to insulin, there is no attenuation of CD ellipticity between 0 and 160 h (red line in Figure 3D); PAGE indicates a predominance of the native monomer (arrow at right in Figure 3E) with progressive accumulation

Table 2: Curve Fitting Analysis of the Amide I Raman Band: Primary Contribution to Raman Intensity from Different Component Bands[a,b]

| | α-helix | | | organized β-sheet | | | loose β-strand, PPII + disordered | | | disordered/vibronic coupling bands | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| species | peak[c] ($cm^{-1}$) | % *A* | width ($cm^{-1}$) | peak ($cm^{-1}$) | % *A* | width ($cm^{-1}$) | peak ($cm^{-1}$) | % *A* | width ($cm^{-1}$) | peak ($cm^{-1}$) | % *A* | width ($cm^{-1}$) |
| (a) Globular State | | | | | | | | | | | | |
| insulin | 1660 | 51 | 29 | 1675 | 4 | 40 | 1684 | 27 | 27 | 1645 | 17 | 37 |
| MP | 1660 | 49 | 30 | 1670 | 8 | 27 | 1686 | 28 | 28 | 1642 | 15 | 40 |
| HPI | 1659 | 42 | 27 | 1670 | 7 | 29 | 1684 | 29 | 30 | 1644 | 22 | 37 |
| *swap* | 1659 | 46 | 30 | 1672 | 4 | 20 | 1684 | 36 | 31 | 1633 | 14 | 36 |
| *swap2* | 1658 | 46 | 31 | 1673 | 13 | 23 | 1687 | 31 | 31 | 1639 | 9 | 30 |
| (b) Misfolded State | | | | | | | | | | | | |
| insulin | 1660 | 14 | 22 | 1674 | 70 | 16 | 1683 | 5 | 30 | 1634 | 12 | 26 |
| MP | 1659 | 34 | 38 | 1671 | 49 | 19 | 1687 | 16 | 28 | 1637 | 1 | 24 |
| HPI | 1660 | 10 | 21 | 1674 | 61 | 17 | 1685 | 13 | 35 | 1633 | 16 | 21 |
| *swap* | 1660 | 17 | 32 | 1674 | 69 | 19 | 1690 | 12 | 13 | 1632 | 7 | 23 |
| *swap2* | 1660 | 14 | 23 | 1673 | 63 | 16 | 1688 | 8 | 25 | 1630 | 15 | 39 |

[a] Percent component areas (% *A*) are calculated considering the total area under four component bands as 100%. Standard errors for well-defined components introduce an uncertainty of ~10% in the area measurement. [b] Width is the bandwidth of each fitted component. [c] Wavenumber indicates fitted peak position.

of deamidated monomeric species (asterisk; lanes 13−18). The combined intensity of the native and deamidated bands remains nearly constant (green line in Figure 3D). In further contrast to the PAGE analysis of wild-type insulin, the MP lanes do not contain detectable protein in the well or as higher molecular weight tails. These results indicate that the B30-A1 tether not only protects the monomer from severe thermal distortion (relative to insulin) but also retards the subsequent aggregation of such distorted monomers that ordinarily occur en route to insulin fibrillation.

The physical state of MP precipitates was investigated by Raman microscopy in relation to its native state and classical insulin and HPI fibrils. A foundation is provided by control studies of insulin and MP in their respective native states. As expected on the basis of prior crystallographic and NMR studies (*26*, *27*), the Raman spectrum of MP (middle panel of Figure 4A; as obtained from a powder lyophilized from a solution of native MP) is similar to that of native insulin (bottom panel in Figure 4A). In each case amide I and III regions (1610−1690 and 1230−1300 $cm^{-1}$) indicate a predominance of α-helix; an extended β-strand-related feature appears as a small shoulder near 1685 $cm^{-1}$. Overall spectral complexity and bandwidths are typical of globular domains (*35*). Spectra of MP and HPI (top panel in Figure 4A) differ due to the substantial random-coil signature of the connecting peptide (*28*).

The physical state of MP precipitates was likewise investigated by Raman microscopy in relation to its native state and classical insulin and HPI fibrils. Comparison of MP spectra as a native powder or amorphous precipitate (middle panels of Figure 4A,B, respectively) provides evidence of a significant change in conformation. Analysis of amide I and III bands indicates a partial conversion of α-helix to β-sheet and random coil with retention of significant α-helical content (Table 2). Whereas the Raman spectra of bona fide fibrils exhibit marked simplification due to loss of spectral dispersion and narrow bandwidths (top and bottom panels of Figure 4B), the spectrum of MP precipitates in part retains heterogeneous band positions and widths (middle panel of Figure 4B). This distinction is exemplified by the amide I region, whose decomposition provides a semiquantitative estimate of secondary structure (Figure 5; *35*). Results of component analysis and integrated areas from a band-fitting procedure (see Materials and Methods) are given in Table 2. Respective native states of insulin and MP give rise to similar contributions from these component bands (bottom and middle panels of Figure 5A). In each case the major contribution (~50%) is due to a component at 1660 $cm^{-1}$, assigned to α-helix. About 44% of the amide I band intensities are attributed to nonhelical components. The combined intensity of bands at 1684 and 1642 $cm^{-1}$ represents loose β-strands, polyproline II conformations (PPII), and other disordered structures. The contribution from well-organized β-sheet at 1675 $cm^{-1}$ is in each case minimal. As expected, native HPI by percentage has less α-helical content (~42%) and a larger nonhelical fraction (~55%) due to the presence of the connecting domain (top panel of Figure 5A and Table 2).

In accord with past studies (*35*) the amide I bands of insulin and HPI fibrils are extremely narrow (respective widths 18 and 22 $cm^{-1}$ at ~1674 $cm^{-1}$; Figure 5B). Deconvolution suggests in each case a predominance of well-organized β-sheet (~70% for insulin fibrils and ~61% for HPI fibrils; Figure 5B and Table 2).[2] Such fibrillation-associated changes in the amide I band provide a Raman signature of classical cross-β assembly. This signature is incomplete in MP precipitates (middle panel of Figure 5B): the amorphous aggregate exhibits a broader bandwidth (30 $cm^{-1}$; Table S1 in Supporting Information) and larger contribution from α-helix (component band at 1659 $cm^{-1}$; Table 2), suggesting that the physical state of the polypeptide in the aggregate is not as compact or uniform as bona fide fibrils. Because the well-organized β-sheet component in MP precipitates remains narrow, the increase in total amide I width is due primarily to its residual α-helix content. A similar contrast is observed between the broad composite disulfide band in MP precipitates (in Figure 4B near 508 $cm^{-1}$, bandwidth 32 $cm^{-1}$, similar to that of MP or insulin in their native states) and the narrowed disulfide band of insulin and HPI fibrils (respective widths 19 and 18 $cm^{-1}$; arrow in Figure 4B). Although it is very difficult to draw

[2] The present deconvolution protocol differs from that employed in a previous study (*35*) (see Materials and Methods) as inclusion of an additional fitting component improves the fit and enables better resolution of the potential non-cross-β contribution to the structure of an insulin fibril.

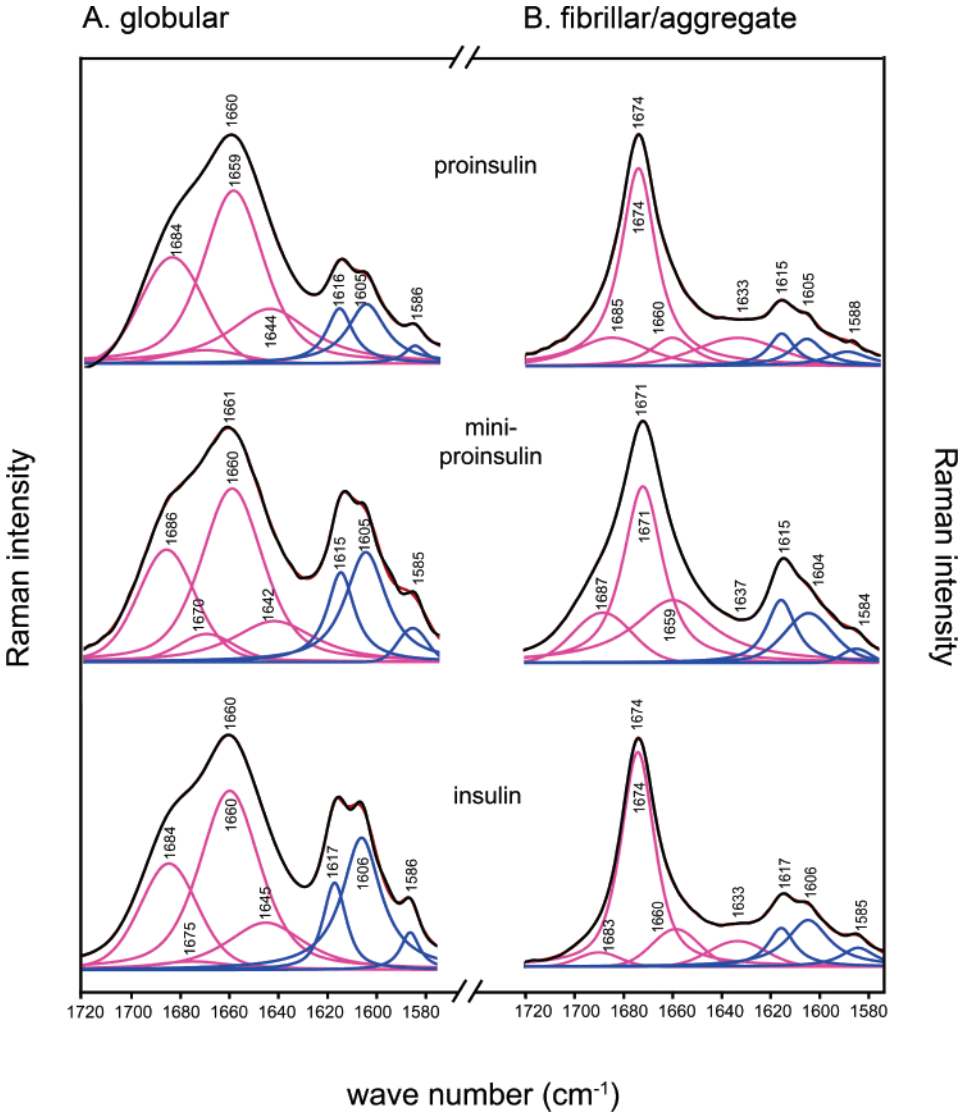


FIGURE 5: Expansion of amide I Raman region with Gaussian deconvolution: (A) native states, insulin (bottom), MP (middle), and HPI (top); (B) insulin and HPI fibrils (bottom and top) and MP precipitates (middle). In each case experimental and fitted spectra are shown in black and red (nearly superimposed as the top trace in each panel). Four component bands that represent total amide I band intensity are in magenta, and component bands assigned to aromatic residues are in blue.

quantitative conclusions, in MP aggregate the appearance of an S−S feature at 508 $cm^{-1}$ suggests a small change in its average geometry in the disulfide linkage compared to insulin and proinsulin (*64*, *65*). In addition, the broader bandwidth of the 508 $cm^{-1}$ band in the spectrum of MP aggregates indicates a greater spread in the geometries of disulfide linkages.

A complementary probe of protein structure is provided by the fluorescence of ANS, whose emission intensity increases on binding to exposed hydrophobic surfaces or within nonpolar pockets. Whereas wild-type insulin exhibits a ca. 10-fold increase in ANS fluorescence on fibrillation (Supporting Information), the increase in MP-associated fluorescence is less than 50%; i.e., the increase in ANS fluorescence on binding to MP precipitates is more than 7-fold lower than that observed on binding to insulin fibrils. These contrasting results suggest a major structural difference between insulin fibrils and MP precipitates: the incomplete MP Raman signature is associated with reduced binding of ANS to hydrophobic sites. Nevertheless, sonicated insulin fibrils and sonicated MP precipitates exhibit similar abilities to act as seeds to promote the fibrillation of freshly prepared insulin solutions (Figure 6). Such seeding is not a generic property of cross-$\beta$ assemblies as no acceleration was observed in the presence of sonicated fibrils prepared from HEWL (filled magenta circles in Figure 6). Interestingly, amorphous aggregates of HPI formed at neutral pH, likewise rich in $\beta$-structure (*28*), can also act as seeds to accelerate insulin fibrillation (data not shown). Cross-seeding between native insulin and MP precipitates suggests that a subset of MP structures within the precipitate resembles oligomeric amyloidogenic intermediates in the pathway of insulin fibrillation, which can provide a complementary growth surface for insulin fibril. Thus, whereas the term "amorphous" is appropriate on the mesoscopic scale of EM, we suggest that at the atomic level MP precipitates exhibit nonrandom conformational preferences even in the absence of coherent long-range order. Sonicated insulin fibrils are neither able to induce canonical fibril formation in MP nor able to efficiently accelerate its amorphous precipitation. Such precipitation was observed in an MP sample containing sonicated insulin fibrils after 13 days, similar to that of

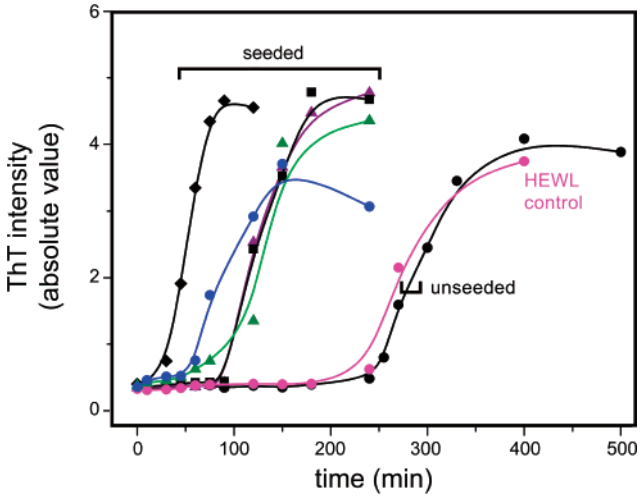


FIGURE 6: Cross-seeding experiments. Fibrillation kinetic profiles of wild-type human insulin without seeding (black circles at far right) and in the presence of 1% (w/w) wild-type insulin seeds (black diamonds at far left), HPI seeds (green triangles), MP aggregates (blue circles), insulin-*swap* seeds (purple triangles), insulin-*swap2* seeds (black squares), and HEWL seeds (magenta circles). Fibrillation was monitored by enhancement of ThT fluorescence (vertical axis).

nonseeded samples (15 days; the small difference is within scatter of the data). MP aggregates obtained from either condition exhibit similar EM morphologies and low ThT emission.

*(II) Insulin Fibrillation Is Robust to Disulfide Interchange.* Insulin-*swap* and insulin-*swap2* each exhibit nativelike partial folds containing two α-helices; the N-terminal segments of the A and B chains are not well ordered (*32*, *34*). Raman spectra of the isomers as freshly prepared powders are consistent with these models: amide III bands provide evidence for a mixture of helical and disordered structure (asterisk in Figure 7A). In insulin-*swap* these bands appear near 1247 (shoulder) and 1262 $cm^{-1}$ (top panel of Figure 7A) whereas in the spectrum of insulin-*swap2* (middle panel) corresponding bands appear near 1248 (shoulder) and 1265 $cm^{-1}$. Further insight is provided by respective amide I bands. Loss of α-helix content is indicated by the shift of the band maximum from 1660 $cm^{-1}$ (insulin) to near 1664 $cm^{-1}$ (insulin-*swap* and insulin-*swap2*); as in native insulin (bottom panel of Figure 7A), a $\beta$-strand signature appears in each case as a small shoulder near 1684 $cm^{-1}$. The amide I regions of isomers in their globular states exhibit greater intensity at higher frequency (near 1684 $cm^{-1}$) than does the amide I region of native insulin. In the isomers the peak maximum appears at 1664 $cm^{-1}$ (shifted from 1660 $cm^{-1}$ in native insulin), which may indicate the presence of a larger contribution from disordered or nonhelical structures. Deconvolution likewise indicates an attenuation of α-helical components and an increased contribution from the features near 1684 $cm^{-1}$ (Figure 8A and Table 2). Surprisingly, despite disulfide interchange, the C−S and S−S stretch modes of cystines are similar in the isomers and native insulin: the C−S stretch modes appear as broad features near 666 and 725 $cm^{-1}$ whereas the S−S stretch modes exhibit two overlapping bands at ∼512 and 493 $cm^{-1}$ (490−562 $cm^{-1}$; arrow in Figure 7A). Insulin, insulin-*swap*, and insulin-*swap2* nonetheless exhibit different S−S bandwidths (28, 28, and 34 $cm^{-1}$, respectively), consistent with expected differences in disulfide environments. Although conformational analysis of the C−S bonds was not attempted due to difficulties in assignment of these bands to a particular conformation, each globular species exhibits a similar band pattern with only subtle differences in intensity distribution among component bands (at 767, 742, and 727 $cm^{-1}$).

Insulin-*swap* and insulin-*swap2* rapidly form fibrils at 65 °C and pH 2.0 (Table 1 and Figure 2E,F): whereas the lag time for insulin is ∼4 h (*28*, *38*), the isomers exhibit respective lag times of 90 min and <30 min (Table 1). Such rapidity is presumably due to their marginal stabilities and initial partial folds with preexposed hydrophobic surfaces. EM photomicrographs in each case depict linear fibrils (Figure 2E,F). To investigate whether the isomers retain their respective pairing schemes during fibrillation, we depolymerized the fibrils with guanidine hydrochloride under acidic

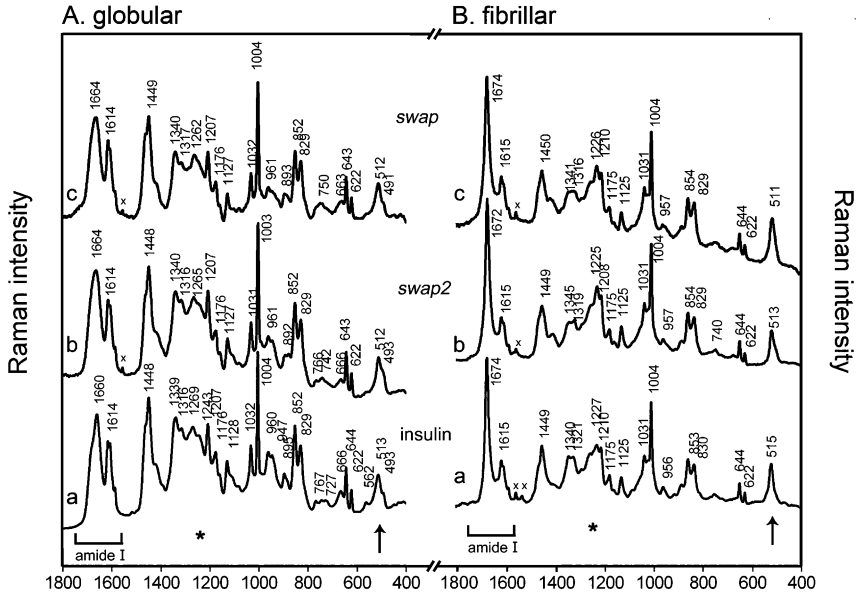


FIGURE 7: Raman studies of insulin disulfide isomers. Solid-state spectra of analogues in (A) globular states and (B) fibrils. In each panel insulin is shown at the bottom, insulin-*swap2* in the middle, and insulin-*swap* at the top. Arrows indicate S−S bands; brackets indicate the amide I region; asterisks indicate the amide III region; x indicates artifacts due to the laser plasma line.

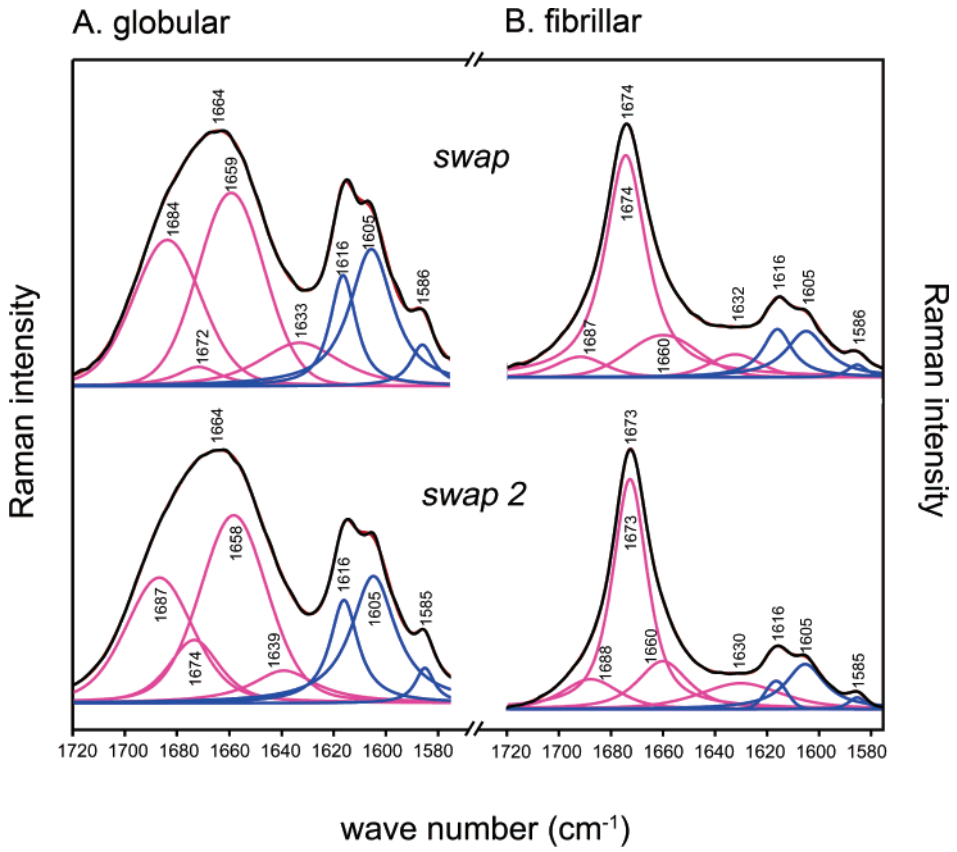


FIGURE 8: Expansion of the amide I Raman region of insulin isomers with Gaussian deconvolution: (A) disulfide isomers *swap* and *swap2* in the globular state and (B) corresponding isomeric fibrils. In each case experimental and fitted spectra are shown in black and red (nearly superimposed as the top trace in each panel). The four component bands that represent total amide I band intensity are shown in magenta, and component bands assigned to aromatic residues are in blue.

conditions (to avoid base-catalyzed disulfide rearrangement) and analyzed the products by rp-HPLC. Since wild-type insulin and the isomers exhibit different rp-HPLC retention times (*32*), respective HPLC profiles provide a fingerprint of disulfide rearrangement (Supporting Information). These fingerprints, and hence respective disulfide pairing schemes, are in each case retained following fibrillation and depolymerization.[3]

The Raman spectra of the isomeric fibrils strongly resemble the spectrum of classical insulin fibrils (Figure 7B). In each case the spectra exhibit an overall simplification due to more uniform and well-ordered residue environments. Of particular interest, the extent of the canonical α- to *β*-structure transitions (as monitored in the amide I region) is similar in each case to that observed in control studies of wild-type insulin. Deconvolution corroborates fitting by an intense component at the *β* position (Figure 8B). Further, fibril-specific amide I bandwidths [18 $cm^{-1}$ (insulin), 22 $cm^{-1}$ (insulin-*swap*), and 19 $cm^{-1}$ (insulin-*swap2*)] are in each case significantly narrower than the corresponding widths in respective globular states (Figure 8A). Differences between wild-type and isomeric fibrils are nonetheless observed in the S−S stretch (arrow in Figure 7B). This stretch in fibrils comprised of wild-type insulin or insulin-*swap* is relatively narrow and symmetric. Modest differences are observed in peak position and bandwidth: wild-type insulin fibrils (515 $cm^{-1}$ with width 19 $cm^{-1}$; bottom panel of Figure 7B) versus insulin-*swap* (511 $cm^{-1}$ with width 24 $cm^{-1}$; top panel of Figure 7B). The S−S stretch of insulin-*swap2* is by contrast less symmetrical, as the major band at 513 $cm^{-1}$ adjoins a shoulder at 495 $cm^{-1}$ (middle panel of Figure 7B). We presume that these variations reflect different modes of local structural accommodation of disulfide constraints within otherwise canonical cross-*β* assemblies. Structural similarity among wild-type and isomeric fibrils, maintained despite their inequivalent disulfide pairing schemes, is suggested by cross-seeding studies. Sonicated isomeric fibrils efficiently function as seeds to accelerate the fibrillation of native insulin. As in studies of MP precipitates, no acceleration was observed in the presence of sonicated HEWL fibrils (Figure 6).

## DISCUSSION

Insulin is a disulfide-cross-linked globular domain and as such provides a model for nonneuropathic systemic amyloidoses, human diseases associated with widespread deposition of amyloid (*66*−*68*). Such conditions cause serious morbidity due to organ infiltration and are often fatal. Autosomal dominant syndromes have been identified with mutations in diverse globular proteins (Table 3). The discovery of the mutant lysozyme syndrome is of particular interest (*69*) because lysozyme has been so extensively characterized, facilitating comparative studies of wild-type and disease-associated variants (*9*, *70*−*73*). Stabilized by four disulfide bridges, lysozyme contains two domains (α and

[3] The extent of chemical modification [deamidation of asparagine and glutamine (*114*, *115*)] in the isomeric fibrils varied, as indicated in each case by shoulder HPLC peaks that elute immediately following the major peaks (Supporting Information).

Table 3: Amyloidogenic Globular Proteins in Human Disease

| protein | length | S−S bridges | structure | syndrome[a] | ref |
|---|---|---|---|---|---|
| apolipoprotein AI | 243 | 0 | α-helical | HSA | *108* |
| $\beta_2$-microglobulin | 100 | 1 | $\beta$-sheet | HAA | *109* |
| fibrinogen A α-chain | 57 | 0 | $\beta$-sheet | HRA | *110* |
| light chain | 110 | 1 | $\beta$-sheet | PSA | *111* |
| lysozyme | 130 | 4 | α/$\beta$ | HSA | *79, 112* |
| cystatin C | 146 | 2 | α/$\beta$ | HBCCAA | *113* |

[a] Abbreviations: HSA, hereditary systemic amyloidosis; HAA, hemodialysis-associated amyloidosis; HRA, hereditary renal amyloidosis; PSA, primary systemic amyloidosis; HBCCAA, hereditary brain cystatin C amyloid angiopathy.

$\beta$) and folds in two steps: folding of the α-domain precedes that of the $\beta$-domain (*9*). Folding of the amyloidogenic variants is less cooperative than that of the wild-type protein (*70*). The presence of partially folded species in equilibrium with the predominant native state is proposed to underlie amyloidogenesis (*9*). As in insulin fibrils, X-ray diffraction studies demonstrate the presence of meridional reflections at 4.6−4.8 Å characteristic of a cross-$\beta$ core structure (*5*). How such protein fibrils accommodate multiple disulfide bridges within (or adjoining) a cross-$\beta$ assembly is not understood.

The present study provides complementary examples of how the connectivity of a polypeptide may be altered. MP is constrained by a short tether between the C-terminal residue of the B chain (residue B30; see Figure 1A) and N-terminal residue of the A chain (A1). The B30-A1 tether is compatible with native structure and stability but impairs induced fit on receptor binding (*26*, *27*). Disulfide rearrangement yields metastable isomers insulin-*swap* (*34*) and insulin-*swap2* (*32*). These partial folds exhibit altered connectivities within the A chain and between it and the N-terminal segment of the B chain. Together, these modifications have contrasting effects on fibrillation and so demonstrate structure-specific modulation of amyloidogenicity by changes in protein topology. We discuss our results in relation to possible mechanisms of fibrillation (*23*, *24*, *74*).

*Insulin Disulfide Isomers*. The native structure, stability, and function of insulin require a specific disulfide pairing scheme (A6-A11, A7-B7, and A20-B19; Figure 1B; *31*). Although disulfide isomers are not ordinarily observed in chemical syntheses (*75*, *76*), studies of the expression and folding of native and variant proinsulins in mammalian secretory cells suggest that mispairing can occur in the endoplasmic reticulum (*77*, *78*). Two disulfide isomers of insulin have been trapped following disulfide interchange under partially denaturing conditions; these exhibit non-native pairings [A7-A11, A6-B7, A20-B19] (insulin-*swap*) and [A6-A7, B7-A11, A20-B19] (insulin-*swap2*; *32*, *34*). Although of low thermodynamic stability, the isomers retain nativelike partial folds (*32*) and reduced but significant biological activity (*33*). Here, we have demonstrated that insulin fibrillation is robust to such disulfide rearrangement. The swapped isomers readily form fibrils under conditions favorable to classical insulin fibrillation. Further, Raman microscopic studies of the isomeric fibrils provide evidence for a formation of a similar $\beta$-sheet-rich structure. We thus propose that the overall organization of insulin chains in a protofilament can accommodate variation in local structure at sites cross-linked by the rearranged disulfide bridges. Wild-type insulin and the accessible disulfide isomers share the A20-B19 disulfide bridge and thus an overall topological feature: the two chains in an extended conformation would be roughly parallel.

It is well established that fibrillation rates of amyloidogenic protein can be markedly enhanced by the addition of small amount of its own preformed fibril (seeding; *40*, *43*). Although the isomeric insulin fibrils retain non-native disulfide pairing, they can effectively function as seeds to promote the fibrillation of wild-type insulin. Cross-seeding by distinct polypeptide chains has been extensively investigated using a variety of candidate sequences (ranging from mutants of the same protein to heterologous proteins unrelated in sequence or native structure). Although a broad range of results has been reported, an important recent study demonstrated a correlation between cross-seeding efficiency and extent of sequence homology (*79*). This correlation is only approximate as cross-seeding of homologous proteins can fail; an example is provided by studies of islet amyloid polypeptide (IAPP; amylin) and the Alzheimer's related A-$\beta$ peptide (*80*). Conversely, the fibrillation rate of α-synuclein, a natively unfolded protein associated with Parkinson's disease, is accelerated by nonhomologous seeds obtained from either GroES of *Escherichia coli*, HEWL, or bovine insulin (*81*). The structural basis and limitations of cross-seeding are of medical importance in relation to prion strains (*82*−*84*). Promiscuous cross-seeding between native insulin and isomeric fibrils, as shown here, demonstrates that such variant fibrils are not "strains." To our knowledge, these results provide the first example of cross-seeding by a protein with identical sequence but different disulfide pairing. We suggest that such cross-seeding reflects a structural correspondence among the isomeric fibrils: despite local disulfide rearrangement, a common global mode of canonical cross-$\beta$ assembly is in each case retained. Although this hypothesis is consistent with the available Raman data, its testing will require high-resolution methods to compare wild-type and isomeric fibrils at atomic dimensions.

Disulfide rearrangement of insulin and proinsulin is restricted to the *swap* and *swap2* pairing schemes (*32*, *34*). It is not known whether the exclusion of other possible isomers is due to kinetic barriers to further rearrangement or to the profound instability of such folds. Hence, although insulin fibrillation is robust to the particular rearrangements investigated here, our study was necessarily limited by the inaccessibility of other pairing schemes (*32*, *34*). The disulfide isomer [A20-B7, A6-B19, A7-A11], if synthesizable for example, might by contrast be found to be incompatible with fibrillation (as distinct from amorphous precipitation) due to inversion of the A- and B-chain orientation. Thus, our results do not exclude whether the architecture of a protein fibril may in general be constrained by the pattern of disulfide pairing. Indeed, the protection afforded by the B30-A1 tether in MP (see below) provides proof of principle that certain disulfide bridges might likewise be incompatible with fibrillation. We speculate that some secreted human proteins not found to participate in pathological amyloid deposits (i.e., not in Table 3) are fortuitously protected by the spacing of their disulfide bridges relative to potential sites of cross-$\beta$ assembly. Resistance to fibrillation might even

provide a constraint contributing to the evolution of protein sequences.

*Models of an Insulin Protofilament.* Models of protein fibrils can be provided by constituent amyloidogenic peptides (*85*, *86*). Although the relationship between such peptide fibrils and the parent protein fibrils is unclear, detailed solid-state NMR studies of peptide fibrils have been described (*86*, *87*). Peptide dissection of insulin has led to the identification of amylodogenic peptide segments within both the A and B chains (*88*). Indeed, the two isolated chains can each form fibrils. Cross-seeding has recently been observed under acidic conditions between isolated B-chain fibrils and wild-type insulin, suggesting that aspects of B-chain cross-$\beta$ assembly may resemble the intact insulin protofilament under these conditions (*89*). These observations suggest a simple model of an insulin protofilament comprising two layers, one sheet formed by the cross-$\beta$ assembly of the A chain and another by cross-$\beta$ assembly of the B chain (Supporting Information). To accommodate discrepant cryo-EM constraints, however, Saibil and colleagues have proposed a novel model in which each chain unfolds to form a U-shaped polypeptide (Supporting Information). In this model one molecule of insulin would occupy two $\beta$-strand layers along the fiber axis, thus rationalizing the observed electron density (0.5−0.6 molecule per layer). Although this model also contains two layers, a key difference is that both chains contribute to each layer. As above, retention of specific disulfide pairing (A7-B7 and A20-B19) implies that the two chains within a protofibril are parallel.

The present results do not exclude either of the above models. Given the nonuniform central bends proposed in either case to accommodate the A6-A11 disulfide bridge, for example, one may readily imagine that alternative pairing [A6-B7, A7-A11, A20-B19] (insulin-*swap*) could be accommodated within either model with only local adjustments; by contrast, accommodation of pairing [A6-A7, B7-A11, A20-B19] (insulin-*swap2*) would appear to realign the respective N-terminal segments of the two chains. It would be intriguing if structural relationships among the wild-type and isomeric insulin protofilaments should resemble those observed in A-$\beta$ fibrils (*90*−*93*). In particular, structural differences between wild-type and protomeric insulin fibrils might be analogous to those observed among distinct and self-propagating forms of A-$\beta$ fibrils (*94*). It would be of future interest to test contrasting features of the above models through isotope-directed NMR studies.

*Mini-proinsulin and Topological Resistance to Fibrillation.* Fibrillation of globular proteins is thought to be mediated by the aggregation of partially unfolded intermediates (*1*). In the case of insulin fibrillation a variety of evidence supports such a mechanism (*59*, *95*, *96*). In the storage granule of the $\beta$ cell as in pharmaceutical formulations insulin is protected from fibrillation through $Zn^{2+}$-mediated self-assembly: the susceptible monomer is sequestered within classical hexamers (Figure 9A, left). Following disassembly of such native oligomers, the insulin monomer may undergo conformational distortion, in turn permitting non-native aggregation to form a putative amyloidogenic nucleus (Figure 9A, middle panels). The nature of this nucleus, its structure, and mechanism of cross-$\beta$ assembly to yield mature fibrils (Figure 9A, right) are not well understood. Extensive site-directed mutagenesis of insulin has been undertaken in an effort to identify specific side chains involved in initial non-native aggregation and fibril growth (*97*). Although effects of mutations on lag time and rate of elongation have been described (*97*), it is difficult to resolve direct effects of substitutions on specific steps in the pathway from indirect effects on structure, stability, or dynamics of the native protein or intermediate species.

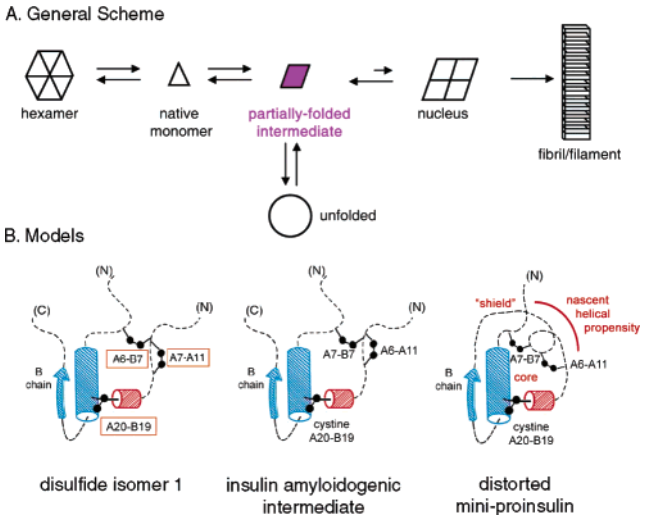


FIGURE 9: Mechanism of fibrillation and putative amyloidogenic partial folds. (A) Proposed pathway of insulin fibrillation via partial unfolding of the monomer. Whereas the native state is protected by classic self-assembly (far left), disassembly leads to an equilibrium between native and partially folded monomers (open triangle and magenta trapezoid, respectively). The partial fold may unfold completely as an off-pathway event (open circle) or aggregate to form a nucleus en route to protofilament assembly (far right). (B) Cylinder models of insulin-related partial folds. Panels: left, insulin disulfide isomer at 25 °C (insulin-*swap*; *32*, *34*); middle, putative insulin amyloidogenic intermediate (at 60 °C prior to onset of fibrillation; *59*); right, model of MP partial fold at 65 °C. In each case A- and B-chain helices are represented by red and blue cylinders, respectively; sulfur atoms in disulfide bridges are shown as black balls. Whereas the tertiary structure of insulin-*swap* (left) resembles that of insulin at elevated temperature (middle), we suggest that MP (right) is less distorted: the B30-A1 tether shields part of the hydrophobic core and augments the helical propensity of the N-terminal A-chain segment. In the left-hand panel rearranged disulfide bridges are boxed.

Insight into the conformational distortions of the insulin monomer at pH 2 and 60 °C has been obtained from heteronuclear NMR studies (*59*). Although the protein is not well ordered under such conditions, a novel partial fold is observed in which the N-terminal segments of the A and B chains detach from the core (Figure 9B, middle). Unfolding of the N-terminal α-helix of the A chain exposes a hydrophobic surface formed by nativelike packing of the remaining α-helices. The C-terminal segment of the B chain, although flexible, remains tethered to this partial core (*59*). Subsequent $^{1}H-^{15}N$ NMR monitoring of this distorted monomer reveals progressive loss of cross-peak amplitude without changes in line width or chemical shift: no "footprint" is obtained of sites of aberrant protein−protein interaction. These observations (analogous to the progressive attenuation of CD spectra reported here) indicate (a) that downstream non-native aggregates are either too large or too rare to give rise to a detectable NMR signal and (b) that such aggregation is either irreversible or too slow to permit detectable NMR exchange phenomena. Fink and colleagues

have nevertheless shown that in 20% acetic acid (pH 1.9) an insulin monomer undergoes a detectable sequence of multiple aggregation-coupled changes in conformation en route to formation of fibrils (*41*). The use of multiple probes (CD, FTIR, fluorescence, light scattering, and hydrogen exchange) enabled resolution of two major subpopulations of soluble oligomeric intermediates, each with non-native $\beta$-rich structures.

The general scheme of insulin fibrillation (Figure 9A) provides a framework for understanding why MP and other single-chain insulin analogues are refractory to fibrillation. Such resistance was first described by Brange and co-workers in 1987 (*30*). The particular MP investigated at that time (a 50-residue *des*-B30 polypeptide in which residue B29 was connected directly to A1 by a peptide bond) is more tightly constrained than that characterized here (a 53-residue polypeptide in which a tripeptide is interposed between B29 and A1, hence allowing more play between chains). Solutions of the 50-mer remained clear during the course of these prior studies, precluding characterization of possible precipitates. Brange and co-workers proposed that fibrillation was prevented by the restricted conformational mobility of MP: in particular, the original B29-A1 tether constrained the C-terminal B-chain $\beta$-strand from detaching from the $\alpha$-helical subdomain, thus limiting exposure of underlying hydrophobic surfaces (*30*). The present results suggest that such B-A tethers may also constrain segmental unfolding of the N-terminal A-chain $\alpha$-helix, a prominent feature of insulin at elevated temperature (*63*). CD studies demonstrate that the B30-A1 tether partially protects MP from loss of $\alpha$-helix content at high temperatures. We speculate that such protection is due to maintenance of $\alpha$-helix-like conformations in the N-terminal A-chain segment (Figure 9B, right). Whereas the segmental unfolding of the native N-terminal A-chain $\alpha$-helix in wild-type insulin at temperatures $>50$ °C is anomalous (indicating a breakdown in the cooperativity of folding), analysis of the temperature dependence of the CD spectrum at helix-sensitive wavelengths suggests that the B30-A1 tether does not protect MP from baseline thermal destabilization as extrapolated from lower temperatures (4−40 °C). In the future it would be of interest to investigate the distorted conformation of MP at 65 °C by heteronuclear NMR approaches. Differences between this conformation and that of wild-type insulin may rationalize why MP resists non-native aggregation under these conditions. We speculate that the B30-A1 tether and partial helical propensity of the A1-A8 segment shield underlying hydrophobic surfaces in the core (Figure 9B, right) that would otherwise mediate such aggregation.

MP is unable to form canonical linear fibrils and instead forms an amorphous precipitate. Raman studies indicate a significant but incomplete $\alpha \rightarrow \beta$ transition relative to wild-type insulin or HPI. This resistance to fibrillation may reflect a structural incompatibility between the B30-A1 tether length and the respective positions of these chain termini in a protofilament or amyloidogenic nucleus. Such incompatibility is suggested by both models described above. Extension of such models to key soluble oligomeric intermediates would be speculative as the structures of such intermediates have not to date been characterized. Because amorphous precipitation of MP is >5-fold slower than the fibrillation of HPI, it is likely that failure of MP to form fibrils is not due to a more efficient (and competing) kinetic pathway leading to an amorphous precipitate. We imagine that MP aggregates resemble one (or both) intermediate species identified by Fink and colleagues (*41*). This analogy is supported by observation in each case of weak ANS fluorescence enhancement and by cross-seeding between native insulin and sonicated MP precipitates. Interestingly, sonicated insulin fibrils, competent to seed native insulin, are unable to accelerate the amorphous precipitation of MP.

Although mechanisms of amyloid nucleation and seeding are not well understood, Lansbury and co-workers have proposed that seeding is "extremely discriminating" since it relies on a complementary sequence between the growth face of the nuclei/seed and the monomer (*39*, *98*). MP amorphous aggregates may thus provide a structurally discriminating growth surface recognized by distorted insulin monomers. Consistent with this mechanism, amorphous aggregates of HPI (as formed in PBS buffer at pH 7.4; *28*) can also seed insulin fibrillation; despite their amorphous EM appearance, these aggregates exhibit a Raman $\beta$-signature similar to that of insulin fibrils. The inability of sonicated insulin fibrils to accelerate the amorphous precipitation of MP may reflect structural differences between mature fibrils and soluble oligomeric intermediates. In particular, the proposed steric incompatibility between the B30-A1 tether in MP and the architecture of a mature fibril does not imply that this tether cannot be accommodated in at least a subset of populated amyloidogenic intermediates, which provide at least part of the fibril growth surface recognized by distorted insulin monomers. Such reasoning suggests that amorphous precipitation of MP reflects an interruption in the pathway of insulin fibrillation, leading to irregular aggregation of conformationally altered oligomers enriched in $\beta$-sheet content but retaining a subset of native $\alpha$-helical structure.

The properties of MP suggest that it may in principle be possible to design cross-links within proteins to hinder or prevent fibrillation. Although MP itself is without biological activity, previous protein design studies have demonstrated how disulfide bridges can be introduced with maintenance of activity (*99*−*103*). Such cross-links often enhance the thermal stability of the native state, presumably via entropic effects in the unfolded state. Whereas such enhanced stability may in itself delay fibrillation (*104*), the present considerations suggest that disulfide bridges at key sites (relative to the structure of a protofilament) may impede the wild-type mode of cross-$\beta$ assembly through changes in chain topology. Proteins refractory to fibrillation would be of broad interest in pharmaceutical formulations and as industrial catalysts. In addition to such applications, it would be of fundamental interest to investigate whether globular proteins can form more than one type of protofilament, so that hindering a single mode of fibrillation by such design would not preclude alternative modes of misfolding and cross-$\beta$ assembly. The existence of distinct prion strains (*82*−*84*) and propagated molecular differences among A-$\beta$ fibrils (*94*) suggest that "designing out" fibrillation may require characterization of multiple pathways of misfolding and assembly.

## CONCLUDING REMARKS

Insulin provides a prototype for the fibrillation of a globular protein. Because the structure of a protein fibril at

atomic resolution is not well understood, models have been proposed on the basis of low-resolution methods (*24*, *105*, *106*). Major classes of models differ regarding such fundamental features as whether the core cross-$\beta$ assembly is parallel or antiparallel, which polypeptide segments participate in this core, and what is their hydrogen-bonding registry. Given the diversity of globular proteins susceptible to amyloidogenesis, it is also possible that the extent of unfolded or non-$\beta$ segmental structures in such fibrils will be observed to vary. We envisage that disulfide pairing within globular proteins provides a fundamental constraint, favoring or excluding otherwise competing modes of cross-$\beta$ assembly. Distinguishing these and other general molecular features of diverse protein fibrils constitutes an outstanding unsolved problem in protein chemistry. The present results highlight the potential importance of covalent protein topology in mechanisms of conformation distortion, non-native aggregation, and cross-$\beta$ assembly.

## ACKNOWLEDGMENT

We thank W. Jia for assistance with CD studies, R. Tycko for discussion and communication of results prior to publication, W. Surewicz for assistance with fibrillation assays, B. Frank and Y.-m. Feng for materials, J. Brange, B. Frank, G. G. Dodson, H. Havel, Q.-X. Hua, and J. Whittingham for advice, and S. Price for preparation of the manuscript and figures. This article is a contribution from the Cleveland Center for Structural Biology.

## SUPPORTING INFORMATION AVAILABLE

Five figures showing renaturation of isomeric fibrils, additional cross-seeding studies, EM morphologies of cross-seeding products, ANS assays, and model of insulin protofilament and three tables providing spectral parameters in Raman analysis and additional CD deconvolution studies. This material is available free of charge via the Internet at http://pubs.acs.org.

## REFERENCES

1. Uversky, V. N., and Fink, A. L. (2004) Conformational constraints for amyloid fibrillation: the importance of being unfolded, *Biochim. Biophys. Acta 1698*, 131−153.
2. Sunde, M., and Blake, C. C. (1998) From the globular to the fibrous state: protein sutrcture and structural conversion in amyloid formation, *Q., Rev. Biophys. 31*, 1−39.
3. Dobson, C. M. (1999) Protein misfolding, evolution and disease, *Trends Biochem. Sci. 24*, 329−332.
4. Dobson, C. M. (2003) Protein folding and misfolding, *Nature 426*, 884−980.
5. Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B., and Blake, C. C. (1997) Common core structure of amyloid fibrils by synchrotron X-ray diffraction, *J. Mol. Biol. 273*, 729−739.
6. Serpell, L. C., Sunde, M., Benson, M. D., Tennent, G., Pepys, M. B., and Fraser, P. E. (2000) The protofilament substructure of amyloid fibrils, *J. Mol. Biol. 301*, 553−563.
7. Lachmann, H. J., Gallimore, R., Gillmore, J. D., Carr-Smith, H. D., Bradwell, A. R., Pepys, M. B., and Hawkins, P. N. (2003) Outcome in systemic AL amyloidosis in relation to changes in concentration of circulating free immunoglobulin light chains following chemotherapy, *Br. J. Haemotol. 122*, 78−84.
8. McParland, V. J., Kalverda, A. P., Homans, S. W., and Radford, S. E. (2002) Structural properties of an amyloid precursor of beta-(2)-microglobulin, *Nat. Struct. Biol. 9*, 326−331.
9. Canet, D., Sunde, M., Last, A. M., Miranker, A., Spencer, C., Robinson, C. V., and Dobson, C. M. (1999) Mechanistic studies of the folding of human lysozyme and the origin of amyloidogenic behavior in disease-related variants, *Biochemistry 18*, 6419−6427.
10. Ivanova, M. I., Sawaya, M. R., Gingery, M., Attinger, A., and Eisenberg, D. (2004) An amyloid-forming segment of $\beta_2$-microglobulin suggests a molecular model for the fibril, *Proc. Natl. Acad. Sci. U.S.A. 101*, 10584−10589.
11. Sambashivan, S., Liu, Y., Sawaya, M. R., Gingery, M., and Eisenberg, D. (2005) Amyloid-like fibrils of ribonuclease A with three-dimensional domain-swapped and native-like structure, *Nature 437*, 266−269.
12. Brange, J., Dodson, G. G., Edwards, D. J., Holden, P. H., and Whittingham, J. L. (1997) A model of insulin fibrils derived from the x-ray crystal structure of a monomeric insulin (*des*pentapeptide insulin), *Proteins 27*, 507−516.
13. Liu, Y., Gotte, G., Libonati, M., and Eisenberg, D. (2001) A domain-swapped RNase A dimer with implications for amyloid formation, *Nat. Struct. Biol. 8*, 211−214.
14. Janowski, R., Kozak, M., Jankowska, E., Grzonka, Z., Grubb, A., Abrahamson, M., and Jaskolski, M. (2001) Human cystatin C, an amyloidogenic protein, dimerizes through three-dimensional domain swapping, *Nat. Struct. Biol. 8*, 316−320.
15. Perutz, M. F., Finch, J. T., Berriman, J., and Lesk, A. (2002) Amyloid fibers are water-filled nanotubes, *Proc. Natl. Acad. Sci. U.S.A. 99*, 5591−5595.
16. Blake, C. C., Serpell, L. C., Sunde, M., Sandgren, O., and Lundgren, E. (1996) A molecular model of the amyloid fibril, *Ciba Found. Symp. 199*, 6−15; discussion 15−21, 40−46.
17. Yonezawa, Y., Tanaka, S., Kubota, T., Wakabayashi, K., Yutani, K., and Fujiwara, S. (2002) An Insight into the pathway of the amyloid fibril formation of hen egg white lysozyme obtained from a small-angle X-ray and neutron scattering study, *J. Mol. Biol. 323*, 237−251.
18. Steiner, D. F., and Clark, J. L. (1968) The spontaneous reoxidation of reduced beef and rat proinsulins. *Proc. Natl. Acad. Sci. U.S.A. 60*, 622−629.
19. Dodson, G., and Steiner, D. (1998) The role of assembly in insulin's biosynthesis, *Curr. Opin. Struct. Biol. 8*, 189−194.
20. Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M., Hubbard, R. E., Isaacs, N. W., and Reynolds, C. D. (1988) The structure of 2Zn pig insulin crystals at 1.5 Å resolution, *Philos. Trans. R. Soc. London 319*, 369−456.
21. Waugh, D. F. (1944) The linkage of corpuscular protein molecules. I. A fibrous modification on insulin, *J. Am. Chem. Soc. 66*, 663.
22. Waugh, D. F. (1946) A fibrous modification of insulin. I. The heat precipitate of insulin, *J. Am. Chem. Soc. 68*, 247−250.
23. Nettleton, E. J. (1998) Ph.D. Thesis, Oxford University, Oxford.
24. Jimenez, J. L., Nettleton, E. J., Bouchard, M., Robinson, C. V., Dobson, C. M., and Saibil, H. R. (2002) The protofilament structure of insulin amyloid fibrils, *Proc. Natl. Acad. Sci. U.S.A. 99*, 9196−9201.
25. Markussen, J., Jorgensen, K. H., Sorensen, A. R., and Thim, L. (1985) Single chain *des*-(B30) insulin. Intramolecular crosslinking of insulin by trypsin catalyzed transpeptidation, *Int. J. Pept. Protein Res. 26*, 70−77.
26. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Bing, X., and Markussen, J. (1991) X-ray analysis of the single chain B29-A1 peptide-linked insulin molecule. A completely inactive analogue, *J. Mol. Biol. 220*, 425−433.
27. Hua, Q.-X., Hu, S. Q., Jia, W., Chu, Y.-C., Burke, G. T., Wang, S. H., Wang, R. Y., Katsoyannis, P. G., and Weiss, M. A. (1998) Mini-proinsulin and mini-IGF-I: homologous protein sequences encoding non-homologous structures, *J. Mol. Biol. 277*, 103−118.
28. Huang, K., Dong, J., Phillips, N. B., Carey, P. R., and Weiss, M. A. (2005) Proinsulin is refractory to protein fibrillation. Topological protection of a precursor protein from cross-beta assembly, *J. Biol. Chem. 280*, 42345−42355.
29. Qiao, Z. S., Guo, Z. Y., and Feng, Y. M. (2001) Putative disulfide-forming pathway of porcine insulin precursor during its refolding in vitro, *Biochemistry 40*, 2662−2668.
30. Brange, J., Hansen, J. F., Havelund, S., and Melberg, S. G. (1987) Studies of insulin fibrillation process, in *Advanced models for the therapy of insulin-dependent diabeses* (Brunetti, P., and Waldhäusl, W. K., Eds.) pp 85−90, Raven Press, New York.
31. Conlon, J. M. (2001) Evolution of the insulin molecule: insights into structure−activity and phylogenetic relationships, *Peptides 22*, 1183−1193.

32. Hua, Q. X., Jia, W., Frank, B. H., Phillips, N. F., and Weiss, M. A. (2002) A protein caught in a kinetic trap: structures and stabilities of insulin disulfide isomers, *Biochemistry 41*, 14700−14715.
33. Sieber, P. S., Eisler, K., Kamber, B., Riniker, B., Rittel, W., Marki, F., and deGasparo, M. (1978) Synthesis and biological activity of two disulphide bond isomers of human insulin: [A7-A11, A6-B7-cystine]- and [A6-A7, A11-B7-cystine] insulin (human), *Hoppe-Seyler's Z. Physiol. Chem. 359*, 113−123.
34. Hua, Q. X., Gozani, S. N., Chance, R. E., Hoffmann, J. A., Frank, B. H., and Weiss, M. A. (1995) Structure of a protein in a kinetic trap, *Nature Struct. Biol. 2*, 129−138.
35. Dong, J., Wan, Z., Popov, M., Carey, P. R., and Weiss, M. A. (2003) Insulin assembly damps conformational fluctuations: Raman analysis of amide I line widths in native states and fibrils, *J. Mol. Biol. 330*, 431−442.
36. Burke, M. J., and Rougvie, M. A. (1972) Cross-$\beta$ protein structures. I. Insulin fibrils, *Biochemistry 11*, 2435−2439.
37. Yu, N.-T., Jo, B. H., Chang, R. C. C., and Huber, J. D. (1974) Single-crystal Raman spectra of native insulin. Structures of insulin fibrils, glucagon fibrils, and intact calf lens, *Arch. Biochem. Biophys. 160*, 614−622.
38. Bouchard, M., Zurdo, J., Nettleton, E. J., Dobson, C. M., and Robinson, C. V. (2000) Formation of insulin amyloid fibrils followed by FTIR simultaneously with CD and electron microscopy, *Protein Sci. 9*, 1960−1967.
39. Jarrett, J. T., and Lansbury, P. T. J. (1992) Amyloid fibril formation requires a chemically discriminating nucleation event: studies of an amyloidogenic sequence from the bacterial protein OsmB, *Biochemistry 31*, 12345−12352.
40. Harper, J. D., and Lansbury, P. T. J. (1997) Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins, *Annu. Rev. Biochem. 66*, 385−407.
41. Ahmad, A., Uversky, V. N., Hong, D. P., and Fink, A. L. (2005) Early events in the fibrillation of monomeric insulin, *J. Biol. Chem. 280*, 42669−42675.
42. Kjeldsen, T. (2000) Yeast secretory expression of insulin precursors. *Appl. Microbiol. Biotechnol. 54*, 277−286.
43. Krebs, M. R., Wilkins, D. K., Chung, E. W., Pitkeathly, M. C., Chamberlain, A. K., Zurdo, J., Robinson, C. V., and Dobson, C. M. (2000) Formation and seeding of amyloid fibrils from wild-type hen lysozyme and a peptide fragment from the beta-domain, *J. Mol. Biol. 300*, 541−549.
44. Sreerama, N., and Woody, R. W. (2000) Estimation of protein secondary structure from CD spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set, *Anal. Biochem. 287*, 252−260.
45. Dong, J., Atwood, C. S., Anderson, V. E., Siedlak, S. L., Smith, M. A., Perry, G., and Carey, P. R. (2003) Metal binding and oxidation of amyloid-beta within isolated senile plaque cores: Raman microscopic evidence, *Biochemistry 42*, 2768−2773.
46. Chen, M. C., and Lord, R. C. (1974) Laser-excited Raman spectroscopy of biomolecules. VI. Polypeptides as conformational models, *J. Am. Chem. Soc. 96*, 4750−4752.
47. Miura, T., and Thomas, G. J., Jr. (1995) Raman spectroscopy of proteins and their assemblies, *Subcell. Biochem. 24*, 55−99.
48. Sane, S. U., Cramer, S. M., and Przybycien, T. M. (1999) A holistic approach to protein secondary structure characterization using amide I band Raman spectroscopy, *Anal. Biochem. 269*, 255−272.
49. Tuma, R. (2005) Raman spectroscopy of proteins: from peptides to large assemblies, *J. Raman Spectrosc. 36*, 307−319.
50. Dong, J., Wan, Z. L., Chu, Y.-C., Nakagawa, S. H., Katsoyannis, P. G., Weiss, M. A., and Carey, P. R. (2001) Isotope-edited Raman spectroscopy of proteins: a general strategy to probe individual peptide bonds with application to insulin, *J. Am. Chem. Soc. 123*, 7919−7920.
51. Miyazawa, T. (1960) Perturbation treatment of the characteristic vibrations of polypeptide chains in various configurations, *J. Chem. Phys. 32*, 1647−1652.
52. Krimm, S., and Yasuaki, A. (1972) Intermolecular interaction effects in the amide I vibrations of beta polypeptides, *Proc. Natl. Acad. Sci. U.S.A. 69*, 2788−2792.
53. Byler, D. M., and Susi, H. (1986) Examination of the secondary structure of proteins by deconvolved FTIR spectra, *Biopolymers 25*, 469−487.
54. Brauner, J. W., Flach, C. R., and Mendelsohn, R. (2005) A quantitative reconstruction of the amide I contour in the IR spectra of globular proteins: from structure to spectrum, *J. Am. Chem. Soc. 127*, 100−109.
55. Marquardt, D. W. (1963) An algorithm for least-squares estimation of nonlinear parameters, *J. Soc. Ind. Appl. Math. 11*, 431−441.
56. Maiti, N. C., Apetri, M. M., Zagorski, M. G., Carey, P. R., and Anderson, V. E. (2004) Raman spectroscopic characterization of secondary structure in natively unfolded proteins: alpha-synuclein, *J. Am. Chem. Soc. 126*, 2399−2408.
57. Apetri, M. M., Maiti, N. C., Zagorski, M. G., Carey, P. R., and Anderson, V. E. (2006) Secondary structure of alpha-synuclein oligomers: characterization by Raman and atomic force microscopy, *J. Mol. Biol. 355*, 63−71.
58. Nettleton, E. J., and Robinson, C. V. (1999) Probing conformations of amyloidogenic proteins by hydrogen exchange and mass spectronomy, *Methods Enzymol. 309*, 633−646.
59. Hua, Q. X., and Weiss, M. A. (2004) Mechanism of insulin fibrillation: the structure of insulin under amyloidogenic conditions resembles a protein-folding intermediate, *J. Biol. Chem. 279*, 21449−21460.
60. Altose, M. D., Zheng, Y., Dong, J., Palfey, B. A., and Carey, P. R. (2001) Comparing protein-ligand interactions in solution and single crystals by Raman spectroscopy, *Proc. Natl. Acad. Sci. U.S.A. 98*, 3006−3011.
61. Krebs, M. R., Bromley, E. H., and Donald, A. M. (2005) The binding of thioflavin-T to amyloid fibrils: localisation and implications, *J. Struct. Biol. 149*, 30−37.
62. Weiss, M. A., Nguyen, D. T., Khait, I., Inouye, K., Frank, B. H., Beckage, M., O'Shea, E., Shoelson, S. E., Karplus, M., and Neuringer, L. J. (1989) Two-dimensional NMR and photo-CIDNP studies of the insulin monomer: assignment of aromatic resonances with application to protein folding, structure, and dynamics, *Biochemistry 28*, 9855−9873.
63. Hua, Q. X., and Weiss, M. A. (2004) Mechanism of insulin fibrillation: structure of insulin under amyloidogenic conditions, *J. Biol. Chem. 279*, 21449−21460.
64. Sugeta, H., Go, A., and Miyazawa, T. (1973) Vibrational spectra and molecular conformations of dialkyl disulfides, *Bull. Chem. Soc. Jpn. 46*, 3407−3411.
65. Van Wart, H. E., and Scheraga, H. A. (1976) Raman spectra of strained disulfides. Effect of rotation about sulfur−sulfur bonds on sulfur−sulfur stretching frequencies, *J. Phys. Chem. 80*, 1823−1832.
66. Lanham, J. G., Meltzer, M. L., De Beer, F. C., Hughes, G. R., and Pepys, M. B. (1982) Familial amyloidosis of Ostertag, *Q. J. Med. 51*, 25−32.
67. Hawkins, P. N. (2003) Hereditary systemic amyloidosis with renal involvement, *J. Nephrol. 16*, 443−448.
68. Hirschfield, G. M., and Hawkins, P. N. (2003) Amyloidosis: new strategies for treatment, *Int. J. Biochem. Cell Biol. 35*, 1608−1613.
69. Pepys, M. B., Hawkins, P. N., Booth, D. R., Vigushin, D. M., Tennant, G. A., Soutar, A. K., Totty, N., Nguyen, O., Blake, C. C., Terry, C. I., Feest, A., Zalin, M., and Hsuan, J. J. (1993) Human lysozyme gene mutations cause hereditary systemic amyloidosis, *Nature 362*, 553−557.
70. Booth, D. R., Sunde, M., Bellotti, V., Robinson, C. V., Hutchinson, W. L., Fraser, P. E., Hawkins, P. N., Dobson, C. M., Radford, S. E., Blake, C. C., and Pepys, M. B. (1997) Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis, *Nature 385*, 787−793.
71. Morozova-Roche, L. A., Zurdo, J., Spencer, A., Noppe, W., Receveur, V., Archer, D. B., Joniau, M., and Dobson, C. M. (2000) Amyloid fibril formation and seeding by wild-type human lysozyme and its disease-related mutational variants, *J. Struct. Biol. 130*, 339−351.
72. Chamberlain, A. K., Receveur, V., Spencer, A., Redfield, C., and Dobson, C. M. (2001) Characterization of the structure and dynamics of amyloidogenic variants of human lysozyme by NMR, *Protein Sci. 10*, 2525−2530.
73. Moraitakis, G., and Goodfellow, J. M. (2003) Simulations of human lysozyme: probing the conformations triggering amyloidosis, *Biophys. J. 84*, 2149−2158.
74. Brange, J., Andersen, L., Laursen, E. D., Meyn, G., and Rasmussen, E. (1997) Toward understanding insulin fibrillation, *J. Pharm. Sci. 86*, 517−525.
75. Katsoyannis, P. G., and Tometsko, A. (1966) Insulin synthesis by recombination of A and B chains: a highly efficient method, *Proc. Natl. Acad. Sci. U.S.A. 55*, 1554−1561.

76. Hua, Q.-X., Chu, Y.-C., Jia, W., Phillips, N. F. B., Wang, R.-Y., Katsoyannis, P. G., and Weiss, M. A. (2002) Mechanism of insulin chain combination. Asymmetric roles of A-chain alpha-helices in disulfide pairing, *J. Biol. Chem. 277*, 43443−43453.
77. Zhang, B. Y., Liu, M., and Arvan, P. (2003) Behavior in the eukaryotic secretory pathway of insulin-containing fusion proteins and single-chain insulins bearing various B-chain mutations, *J. Biol. Chem. 278*, 3687−3693.
78. Liu, M., Li, Y., Cavener, D., and Arvan, P. (2005) Proinsulin disulfide maturation and misfolding in the endoplasmic reticulum, *J. Biol. Chem. 280*, 13209−13212.
79. Krebs, M. R., Morozova-Roche, L. A., Daniel, K., Robinson, C. V., and Dobson, C. M. (2004) Observation of sequence specificity in the seeding of protein amyloid fibrils, *Protein Sci. 13*, 1933−1938.
80. O'Nuallain, B., Williams, A. D., Westermark, P., and Wetzel, R. (2004) Seeding specificity in amyloid growth induced by heterologous fibrils, *J. Biol. Chem. 279*, 17490−17499.
81. Yagi, H., Kusaka, E., Hongo, K., Mizobata, T., and Kawata, Y. (2005) Amyloid fibril formation of alpha-synuclein is accelerated by preformed amyloid seeds of other proteins: Implications for the mechanism of transmissible conformational diseases, *J. Biol. Chem. 280*, 38609−38616.
82. Bessen, R. A., Kocisko, D. A., Raymond, G. J., Nandan, S., Lansbury, P. T., and Caughey, B. (1995) Non-genetic propagation of strain-specific properties of scrapie prion protein, *Nature 375*, 698−700.
83. Kretzschmar, H. A., and Feiden, W. (2002) Human prion diseases, *Pathologe 23*, 241−251.
84. Jones, E. M., and Surewicz, W. K. (2005) Fibril conformation as the basis of species- and strain-dependent seeding specificity of mammalian prion amyloids, *Cell 121*, 63−72.
85. Gustavsson, A., Engstrom, U., and Westermark, P. (1991) Normal transthyretin and synthetic transthyretin fragments form amyloid-like fibrils in vitro, *Biochem. Biophys. Res. Commun. 175*, 1159−1164.
86. Jaroniec, C. P., MacPhee, C. E., Astrof, N. S., Dobson, C. M., and Griffin, R. G. (2002) Molecular conformation of a peptide fragment of transthyretin in an amyloid fibril, *Proc. Natl. Acad. Sci. U.S.A. 99*, 16748−16753.
87. Jaroniec, C. P., MacPhee, C. E., Bajaj, V. S., McMahon, M. T., Dobson, C. M., and Griffin, R. G. (2004) High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy, *Proc. Natl. Acad. Sci. U.S.A. 101*, 711−716.
88. Ivanova, M. I., Thompson, M. J., and Eisenberg, D. (2006) A systematic screen of beta2-microglobulin and insulin for amyloid-like segments, *Proc. Natl. Acad. Sci. U.S.A. 103*, 4079−4082.
89. Hong, D. P., and Fink, A. L. (2005) Independent heterologous fibrillation of insulin and its B-chain peptide, *Biochemistry 44*, 16701−16709.
90. Petkova, A. T., Ishii, Y., Balbach, J. J., Antzutkin, O. N., Leapman, R. D., Delaglio, F., and Tycko, R. (2002) A structural model for Alzheimer's beta-amyloid fibrils based on experimental constraints from solid-state NMR, *Proc. Natl. Acad. Sci. U.S.A. 99*, 16742−16747.
91. Petkova, A. T., Buntkowsky, G., Dyda, F., Leapman, R. D., Yau, W. M., and Tycko, R. (2004) Solid-state NMR reveals a pH-dependent antiparallel beta-sheet registry in fibrils formed by a beta-amyloid peptide, *J. Mol. Biol. 335*, 247−260.
92. Sciarretta, K. L., Gordon, D. J., Petkova, A. T., Tycko, R., and Meredith, S. C. (2005) Abeta40-lactam(D23/K28) models a conformation highly favorable for nucleation of amyloid, *Biochemistry 44*, 6003−6014.
93. Petkova, A. T., Yau, W. M., and Tycko, R. (2006) Experimental constraints on quaternary structure in Alzheimer's beta-amyloid fibrils, *Biochemistry 45*, 498−512.
94. Petkova, A. T., Leapman, R. D., Guo, Z., Yau, W. M., Mattson, M. P., and Tycko, R. (2005) Self-propagating, molecular-level polymorphism in Alzheimer's beta-amyloid fibrils, *Science 307*, 262−265.
95. Ahmad, A., Millett, I. S., Doniach, S., Uversky, V. N., and Fink, A. L. (2003) Partially folded intermediates in insulin fibrillation, *Biochemistry 42*, 11404−11416.
96. Ahmad, A., Millett, I. S., Doniach, S., Uversky, V. N., and Fink, A. L. (2004) Stimulation of insulin fibrillation by urea-induced intermediates, *J. Biol. Chem. 279*, 14999−15013.
97. Nielsen, L., Frokjaer, S., Brange, J., Uversky, V. N., and Fink, A. L. (2001) Probing the mechanism of insulin fibril formation with insulin mutants, *Biochemistry 40*, 8397−8409.
98. Harper, J. D., Lieber, C. M., and Lansbury, P. T., Jr. (1997) Atomic force microscopic imaging of seeded fibril formation and fibril branching by the Alzheimer's disease amyloid-beta protein, *Chem. Biol. 4*, 951−959.
99. Sauer, R. T., Hehir, K., Stearman, R. S., Weiss, M. A., Jeitler-Nilsson, A., Suchanek, E. G., and Pabo, C. O. (1986) An engineered intersubunit disulfide enhances the stability and DNA binding of the N-terminal domain of lambda repressor, *Biochemistry 25*, 5992−5998.
100. Pabo, C. O., and Suchanek, E. G. (1986) Computer-aided model-building strategies for protein design, *Biochemistry 25*, 5987−5991.
101. Matsumura, M., and Matthews, B. W. (1991) Stabilization of functional proteins by introduction of multiple disulfide bonds, *Methods Enzymol. 202*, 336−356.
102. Clarke, J., Henrick, K., and Fersht, A. R. (1995) Disulfide mutants of barnase. I: Changes in stability and structure assessed by biophysical methods and X-ray crystallography, *J. Mol. Biol. 253*, 493−504.
103. Robinson, C. R., and Sauer, R. T. (2000) Striking stabilization of Arc repressor by an engineered disulfide bond, *Biochemistry 39*, 12494−12502.
104. Ramirez-Alvarado, M., and Regan, L. (2002) Does the location of a mutation determine the ability to form amyloid fibrils?, *J. Mol. Biol. 323*, 17−22.
105. Makin, O. S., and Serpell, L. C. (2005) Structures for amyloid fibrils, *FEBS J. 272*, 5950−5961.
106. Nelson, R., and Eisenberg, D. (2006) Recent atomic models of amyloid fibril structure, *Curr. Opin. Struct. Biol. 16*, 260−265.
107. Sosnick, T. R., Fang, X., and Shelton, V. M. (2000) Application of circular dichroism to study RNA folding transitions, *Methods Enzymol. 317*, 393−409.
108. Gillmore, J. D., Stangou, A. J., Tennent, G. A., Booth, D. R., O'Grady, J., Rela, M., Heaton, N. D., Wall, C. A., Keogh, J. A., and Hawkins, P. N. (2001) Clinical and biochemical outcome of hepatorenal transplantation for hereditary systemic amyloidosis associated with apolipoprotein AI Gly26Arg, *Transplantation 71*, 986−992.
109. Katou, H., Kanno, T., Hoshino, M., Hagihara, Y., Tanaka, H., Kawai, T., Hasegawa, K., Naiki, H., and Goto, Y. (2002) The role of disulfide bond in the amyloidogenic state of $\beta_2$-microglobulin studied by heteronuclear NMR, *Protein Sci. 11*, 2218−2229.
110. Benson, M. D., Liepnieks, J., Uemichi, T., Wheeler, G., and Correa, R. (1993) Hereditary renal amyloidosis associated with a mutant fibrinogen alpha-chain, *Nat. Genet. 3*, 252−255.
111. Kim, Y., Wall, J. S., Meyer, J., Murphy, C., Randolph, T. W., Manning, M. C., Solomon, A., and Carpenter, J. F. (2000) Thermodynamic modulation of light chain amyloid fibril formation, *J. Biol. Chem. 275*, 1570−1574.
112. Cao, A., Hu, D., and Lai, L. (2004) Formation of amyloid fibrils from fully reduced hen egg white lysozyme, *Protein Sci. 13*, 319−324.
113. Janowski, R., Kozak, M., Jankowska, E., Grzonka, Z., Grubb, A., Abrahamson, M., and Jaskolski, M. (2001) Human cystatin C, an amyloidogenic protein, dimerizes through three-dimensional domain swapping, *Nat. Struct. Biol. 8*, 316−320.
114. Brange, J., Langkjaer, L., Havelund, S., and Volund, A. (1992) Chemical stability of insulin. 1. Hydrolytic degradation during storage of pharmaceutical preparations, *Pharm. Res. 9*, 715−726.
115. Nilsson, M. R., and Dobson, C. M. (2003) Chemical modification of insulin in amyloid fibrils, *Protein Sci. 12*, 2637−2641.


BI060879G